How might a statistical cloud scheme be coupled to a mass-flux convection scheme?

Stephen A. Klein[1]
NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA

Robert Pincus and Cécile Hannay[2]
Climate Diagnostics Center, NOAA Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado, USA

Kuan-Man Xu
NASA Langley Research Center, Hampton, Virginia, USA

Received 14 May 2004; revised 29 September 2004; accepted 11 October 2004; published 21 April 2005.

[1] The coupling of statistical cloud schemes with mass-flux convection schemes is addressed. Source terms representing the impact of convection are derived within the framework of prognostic equations for the width and asymmetry of the probability distribution function of total water mixing ratio. The accuracy of these source terms is quantified by examining output from a cloud-resolving model simulation of deep convection. Practical suggestions for the inclusion of these source terms in large-scale models are offered.

Citation: Klein, S. A., R. Pincus, C. Hannay, and K.-M. Xu (2005), How might a statistical cloud scheme be coupled to a mass-flux convection scheme?, *J. Geophys. Res.*, *110*, D15S06, doi:10.1029/2004JD005017.

1. Introduction

[2] In recent years, increasing attention has been paid to treating the non-linear effects of radiation and precipitation microphysics in large-scale models. The traditional assumption that a cloud is horizontally homogeneous on the scale of a large-scale model grid box for purposes of radiation and precipitation leads to substantial biases due to the non-linear nature of these processes (e.g., *Cahalan et al.* [1994] for radiation; *Larson et al.* [2001] or *Pincus and Klein* [2000] for microphysics).

[3] Statistical cloud schemes provide an attractive framework to self-consistently predict the horizontal inhomogeneity because the probability distribution function (PDF) of total water contained in the scheme can be used to calculate the PDF of cloud condensate, from which the non-linear effects of radiation and precipitation may be self-consistently estimated. (While "assumed-PDF scheme" might be a better term than "statistical cloud scheme," the term "statistical cloud scheme" is ingrained in the literature and will be retained in this paper.) Statistical cloud schemes were originally developed in the context of boundary layer studies, so their extension to the full atmosphere is non-trivial [*Mellor*, 1977; *Sommeria and Deardorff*, 1977]. For example, in their seminal paper, *Sommeria and Deardorff* [1977, p. 345] state: "For large grid volumes, such as in a global circulation model, the assumption of Gaussian distributions on the subgrid scale, for even θ_l [liquid water potential temperature] and q_t [total water specific humidity, a quantity closely related to total water mixing ratio] would presumably be poor."

[4] Indeed, one problem not envisioned by the pioneers is that one would want to have a statistical cloud scheme in a model that also contained a separate mass-flux formulation for atmospheric convection. Originally, it was envisioned that the turbulence scheme when formulated with Reynolds averaged equations would be responsible for all the subgrid-scale transport within a grid box. However, it was quickly learned that Gaussian PDFs do not well represent the trade cumulus boundary layer, where intermittent convection leads to highly skewed distributions of total water [*Bougeault*, 1982]. While one can overcome these difficulties for shallow convection within a self-consistent statistical approach that treats all subgrid-scale transport [*Golaz et al.*, 2002], one may choose to represent subgrid-scale vertical transports with both a turbulence and a mass-flux convection scheme. As this represents the case for virtually all large-scale models which must treat both shallow and deep convection, it is worth improving the consistency between a statistical cloud scheme and a mass-flux convection scheme.

[5] If, within a large-scale model, a statistical cloud scheme is to co-exist with a mass-flux convection scheme, how should they be coupled? First attempts at using a statistical cloud scheme in a global model just ignored the coupling altogether. For example, models at the UK Met Office have essentially assumed a triangle distribution to a variable that is essentially the difference of r_t and r_s, the total water and saturation mixing ratios, respectively [*Smith*,

[1] Now at Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, California, USA.

[2] Now at National Center for Atmospheric Research, Boulder, Colorado, USA.

0148-0227/05/2004JD005017$09.00

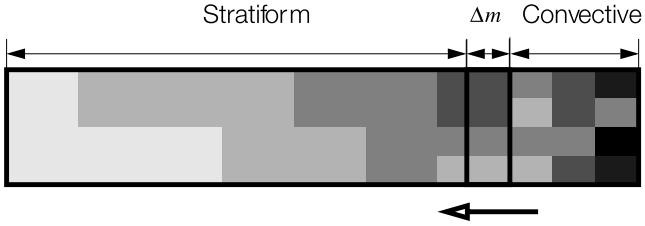

Figure 1. Schematic diagram illustrating the process of detrainment of mass from the convective part of a grid box to the stratiform part. The domain of the schematic is one single level from one single grid box of a large-scale model. The shading indicates the subgrid distribution of the total water mixing ratio r with darker shading indicating more abundant quantities of r. The scenario might represent a detrainment event in the upper troposphere where a mass Δm leaves the convective part and becomes part of the stratiform part, transferring high values of r from the convective part to the stratiform part. The variable amounts of r in the detrained mass indicate that the variance in the detrained air may be a significant source of variance for the stratiform part.

1990]. The width of the distribution when normalized by r_s, however, was either a global constant or a fixed function of pressure. With a time-invariant width to the distribution, this scheme is equivalent to a relative humidity threshold scheme [*Smith*, 1990]. Without any explicit connection to convection, imagine what happens to the parameterized clouds when updrafts detrain cloud condensate into clear air. According to this cloud scheme, clouds can only begin to occur when the relative humidity in the grid box exceeds a threshold. Thus, if the initial relative humidity is less than the threshold value, all cloud water detrained from convection instantaneously evaporates until the relative humidity of the grid box rises to the threshold value. How one might remedy this unnatural behavior in the context of a statistical cloud scheme is the subject of this paper.

[6] Two issues must be dealt with if one wishes to treat clouds in the environment of convection with a statistical approach. The first and relatively straightforward issue is to remove the conventional assumption of a symmetric shape to the r_t PDF, as PDFs tend to be highly skewed [*Xu and Randall*, 1996; *Bony and Emanuel*, 2001; *Tompkins*, 2002; *Larson et al.*, 2002]. The second and more difficult issue is the prediction of the shape of the r_t PDF, both its width and asymmetry. *Bony and Emanuel* [2001] use a simple coupling whereby the shape of the PDF is altered so that at every time step, the in-cloud value of cloud condensate equals the sum of that diagnosed by traditional large-scale saturation and that diagnosed from the Emanuel convection scheme. In the formulation of *Tompkins* [2002], prognostic equations for essentially the variance and skewness of the r_t PDF are added to the large-scale model, with ad-hoc source terms from the mass-flux convection scheme.

[7] In this work, physically based source terms for the impact of convection on the variance and skewness of the r_t PDF are proposed in section 2. These source terms are intended to replace the ad-hoc source terms given by *Tompkins* [2002]. In section 3, the suitability of these terms for parameterization is tested with output from a cloud-resolving model (CRM) simulation of deep atmospheric convection. The ultimate goal of this work is successful incorporation of these terms into a large-scale model. While this is not accomplished herein, practical suggestions to this end are offered in section 4.

2. Higher-Order Moment Source Terms From Convection

[8] In this section, equations to represent the impact of subgrid-scale convection on the variance of total water mixing ratio r are derived. (For clarity of reading the "t" subscript is dropped in the rest of this paper.) Note that the following derivation is for the impact of convection on the variance in the part of the grid-box exclusive of that containing the convective updrafts and downdrafts (i.e., the "stratiform" part of the grid box). Thus the forms of the equations will differ from those of *Lappen and Randall* [2001] who provide variance budget equations appropriate for the whole grid box, including convective regions.

[9] It is intended that the convective terms be one of the source/sink terms in the advective-diffusive equation for the time evolution of $\overline{r'^2}$, the variance of total water mixing ratio in the stratiform part of the grid box:

$$\frac{\partial}{\partial t}\overline{r'^2} + \vec{V} \bullet \overline{r'^2} = \sum_i S_i\left(\overline{r'^2}\right), \tag{1}$$

where the second term on the left-hand side is the advection of $\overline{r'^2}$ by the resolved-scale wind $\vec{V}$, and the summation on the right-hand side is over the sinks and sources of $\overline{r'^2}$ by subgrid-scale processes. The overbar symbol indicates the grid box mean and the prime symbol indicates the deviation from grid box mean. These subgrid-scale processes include vertical transport, precipitation, and convection (others might include horizontal transport). For example, the source term for vertical transport S_{vt} is:

$$S_{vt}\left(\overline{r'^2}\right) = -2\overline{w'r'}\frac{\partial r}{\partial z} - \frac{\partial}{\partial z}\overline{w'r'^2} - \varepsilon, \tag{2}$$

where w is the vertical velocity and ε is the subgrid-scale dissipation of variance within a grid box [*André et al.*, 1978]. For precipitation, the source term S_{pr} is:

$$S_{pr}\left(\overline{r'^2}\right) = -2\,\overline{G'_{pr}r'}, \tag{3}$$

where G_{pr} is the sink of total water due to stratiform precipitation processes. Note that (3) has the simple interpretation that if precipitation removal processes are greater in the portion of the grid box where r is larger, as is generally true, then precipitation reduces the variance of r within a grid box.

[10] The starting point of the derivation of the convective source terms is consideration of how the variance in the "stratiform" part of the grid box is impacted by the detrainment of mass from the "convective" part of the grid box (Figure 1). This detrainment event may correspond to mass detrainment from either a convective updraft or downdraft. If the ratio of the detrained mass Δm in a time step Δt to the sum of the detrained mass and the mass in the

stratiform part is defined as Δa, then $\overline{r'^2}^{n+1}$, the variance in r at the next time step $n+1$, can be written as:

$$\overline{r'^2}^{n+1} = \Delta a(1-\Delta a)(r_d - r^n)^2 + (1-\Delta a)\overline{r'^2}^n + \Delta a \overline{r_d'^2}, \quad (4)$$

where r^n and $\overline{r'^2}^n$ are the mean and variance of r at time step n in the stratiform part of the grid box before the detrainment event occurs (see Appendix A for an outline of the derivation of (4)). Other symbols in (4) are r_d and $\overline{r_d'^2}$, the mean and variance of r in the detrained mass. The subscript "d" indicates that the variable has the property of the mass detrained from the convective part of the grid box. Equation (4) can be recognized as the expression for the variance of a volume containing two parts, the stratiform mass before detrainment (i.e., the part of Figure 1 marked "stratiform") and the detrained mass Δm. Equation (4) states that the total variance has three contributions: the first is the variance due to the fact the stratiform part and the detrained mass may have different mean r, and the other two terms are the contributions to the total variance from r variance within each of the respective parts. If one subtracts $\overline{r'^2}^n$ from both sides of (4) and divides by the time step Δt, one has:

$$\frac{\overline{r'^2}^{n+1} - \overline{r'^2}^n}{\Delta t} = \frac{\Delta a(1-\Delta a)}{\Delta t}(r_d - r^n)^2 + \frac{\Delta a}{\Delta t}\left(\overline{r_d'^2} - \overline{r'^2}^n\right). \quad (5)$$

Taking the differential limit of very small Δt and also the limit of $\Delta a \ll 1$, one arrives at:

$$\frac{\partial \overline{r'^2}}{\partial t} = D(r_d - r)^2 + D\left(\overline{r_d'^2} - \overline{r'^2}\right), \quad (6)$$

where D is the mass detrainment rate from convection in units of inverse seconds. (One might be concerned regarding the assumption that $\Delta a \ll 1$. From analysis of the cloud-resolving model simulation discussed below, $\Delta a < 0.2$ for nearly all detrainment events over an interval of one hour, a typical time step to physics routines of large-scale models. If one does not wish to assume that $\Delta a \ll 1$, then (4) can be used with $\Delta a = \min(1, D\,\Delta t)$ instead of (6) for the increment of variance over a time step.) D is related to the convective mass flux M_c by: [*Yanai and Johnson*, 1993]

$$\frac{\partial M_c}{\partial p} = D - E. \quad (7)$$

In (7), E is the rate of entrainment of mass into the convective part and p is pressure. Note that M_c is defined in pressure coordinates and has units of Pa s^{-1}.

[11] Equation (6) states that the variance in the stratiform part of the grid box increases if the air being detrained from convection has a different value of total water mixing ratio r_d than that present in the stratiform environment r. Furthermore, even if two regions had the same mean value of r, the variance in the stratiform part could increase if the variance in the air being detrained from convection $\overline{r_d'^2}$ exceeds the variance already present in the stratiform part $\overline{r'^2}$.

[12] To complete the derivation of the convective source terms requires consideration of two other processes: entrainment of stratiform air into the convective region and compensating subsidence. Entrainment is straightforward and mirrors the derivation above exactly. That is, if some part of the stratiform air enters the convective part (i.e., entrainment), the variance in the stratiform part may change for two reasons: the value of mean total water mixing ratio of the air being entrained into the convective part r_e may be different than the mean in the stratiform part, and even if they are the same, the variance of r in the air being entrained $\overline{r_e'^2}$ may be different than that present in the stratiform part. Mathematically, this is expressed as:

$$\frac{\partial \overline{r'^2}}{\partial t} = -E(r_e - r)^2 - E\left(\overline{r_e'^2} - \overline{r'^2}\right). \quad (8)$$

The negative sign in front of both of these terms results from the fact that if convection entrains air with a different value of r than the mean value in the stratiform part then the variance of r in the stratiform part decreases. Likewise, if the variance of r in the air being entrained exceeds that of the stratiform part, the variance in the stratiform part will decrease. If one assumes that the properties of air entrained into the convective regions do not differ from the average properties in the stratiform environment, as is generally true of conventional convection parameterizations, then the entrainment terms are zero.

[13] The final process to consider is compensating subsidence. At levels of the atmosphere where there is no entrainment or detrainment but positive mass flux M_c, the air in the stratiform environment is being advected downward by the subsidence which compensates for the upward mass flux in the convective part. This term is computed as for any passive scalar in the mass-flux framework:

$$\frac{\partial \overline{r'^2}}{\partial t} = -M_c \frac{\partial \overline{r'^2}}{\partial p}. \quad (9)$$

Combining (6), (8), and (9), one has the total form of $S_{cv}(\overline{r'^2})$, the convective source term of variance:

$$S_{cv}\left(\overline{r'^2}\right) = D(r_d - r)^2 + D\left(\overline{r_d'^2} - \overline{r'^2}\right) - E(r_e - r)^2 - E\left(\overline{r_e'^2} - \overline{r'^2}\right) - M_c \frac{\partial \overline{r'^2}}{\partial p}. \quad (10)$$

As convection produces highly skewed total water PDFs, one might be interested in the source term from convection of $\overline{r'^3}$, the third moment of r. The derivation for $\overline{r'^3}$ is analogous to that presented above for $\overline{r'^2}$, so only the end result is given:

$$S_{cv}\left(\overline{r'^3}\right) = D(r_d - r)^3 + D\left(\overline{r_d'^3} - \overline{r'^3}\right) + 3D(r_d - r)\left(\overline{r_d'^2} - \overline{r'^2}\right) - E(r_e - r)^3 - E\left(\overline{r_e'^3} - \overline{r'^3}\right) - 3E(r_e - r)\left(\overline{r_e'^2} - \overline{r'^2}\right) - M_c \frac{\partial \overline{r'^3}}{\partial p}. \quad (11)$$

The form of the third moment convective source term is analogous to that for the variance source term except for the presence of two additional terms (those beginning with coefficient 3). The additional terms arise from the formula

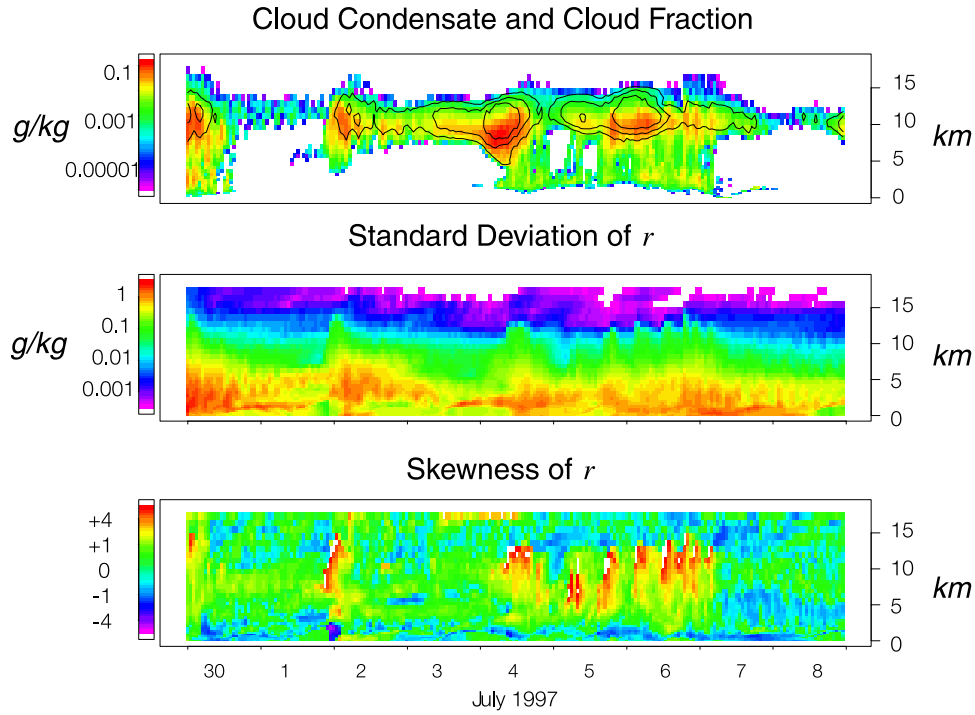

Figure 2. Selected time-height fields from a 9-day segment of the cloud-resolving model simulation. The top panel displays the domain averaged cloud condensate mixing ratio (in colors) and cloud fraction (in contours at levels 0.2, 0.5, and 0.8). The middle and bottom panels display the standard deviation and skewness of r, the total water mixing ratio. The displayed fields are computed every hour from the stratiform portion of the cloud-resolving model domain using the archived 5-min model snapshots. Whenever deep convection occurs, the variance and skewness of r increase in the middle and upper troposphere. Near the surface, downdrafts bring low values of r into the boundary layer imparting negative skewness to the r PDF.

for the third moment of a scalar in a volume containing two parts. In the case of detrainment, it states that the third moment of r increases if the detrained mass has both a higher mean r and a higher level of $\overline{r'^2}$ than present in the stratiform environment.

3. Examination of a Cloud-Resolving Model Simulation

[14] The appropriateness of the source terms from convection is tested by examination of detailed output from a 29 day simulation of summertime deep convection at the Southern Great Plains site of the Atmospheric Radiation Measurement program. The simulation period, 19 June to 17 July 1997, occurs during an intensive observing period whose data were used to calculate the model forcing fields according to the variational analysis method of *Zhang et al.* [2001]. The simulation is performed by the 2-dimensional UCLA-CSU cloud-resolving model with horizontal resolution of 2 km and a total domain size of 512 km [*Xu et al.*, 2002]. Other characteristics of this model simulation include 34 vertical levels beneath 20 km on a stretched grid, a 5 category bulk microphysics scheme, a third-order turbulence closure scheme, and interactive radiation. Information on the performance of this CRM relative to observations and other CRMs for this period is provided by *Xu et al.* [2002]. A detailed comparison of the cirrus clouds produced by this CRM integration with radar observations is provided by *Luo et al.* [2003]. From this integration, full model fields saved at intervals of 5 min are examined.

[15] To calculate the source terms from convection, the domain of the cloud-resolving model is divided into convective and stratiform parts according to the algorithm of *Xu* [1995]. To determine the rates of mass detrainment or entrainment and the properties of air undergoing entrainment or detrainment, the definitions of *Siebesma* [1998] are used. Details regarding these calculations are presented in Appendix B.

[16] Figure 2 displays time-height cross-sections from a 9 day segment in the middle of the integration. The upper panel displays the 1 hour mean cloud condensate mixing ratio (the sum of the model's cloud liquid and ice) and cloud fraction averaged over all stratiform grid cells within the 512 km model domain. These averages correspond to a space-timescale of 512 km by 1 hour, which is roughly comparable to the space-timescales of large-scale models. The lower two panels display the standard deviation and skewness of the total water mixing ratio r in the stratiform portion of the domain. These statistics are computed from PDFs that are constructed from the mean r in every 2 km stratiform grid cell. Although the cloud-resolving model has a sophisticated subgrid turbulence scheme that predicts the variance and skewness of r at scales less than 2 km, the data were not available for analysis. Thus the variance and skewness comes from scales of 2 km to 512 km that are explicitly resolved by the cloud-resolving model.

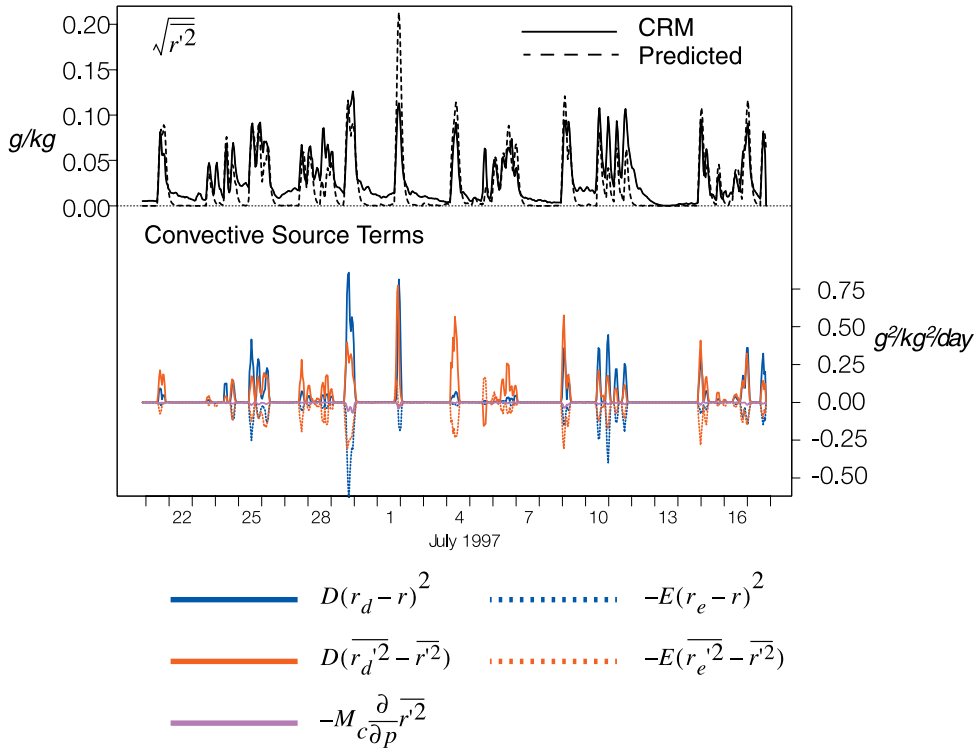

Figure 3. Time series of selected quantities at the cirrus detrainment level (~11 km) from the cloud-resolving model (CRM). The top portion of the figure illustrates the time series of the standard deviation of total water mixing ratio $\sqrt{\overline{r'^2}}$ from the CRM (solid) and that predicted from equation (12) using the convective source terms diagnosed from the CRM (dashed). The individual components of the convective source terms at this level are displayed in the lower portion of the figure. The peaks in the time series of $\sqrt{\overline{r'^2}}$ coincide very well with the peaks in the convective source terms diagnosed from the CRM, such that the convective source terms may be used to predict $\overline{r'^2}$.

[17] During this period, intermittent convection creates clouds in the upper troposphere while significantly impacting the variance and skewness of r. While the background vertical variation of $\overline{r'^2}$ is dominated by the exponential decrease with altitude of the saturation mixing ratio r_s, bursts of convection raise the variance of r, which subsequently decays to background levels when the convection ceases. Whenever convection occurs, the skewness of r increases markedly in the middle and upper troposphere. Although some regions have skewness exceeding +4 (and thus appear white in the lower panel of Figure 2), the skewness decays very rapidly. Also noteworthy is that when convection occurs, the skewness in the lower troposphere may become negative (i.e., the event near at the beginning of 2 July 1997). This probably corresponds to times when downdrafts bring air with low r into the boundary layer.

[18] Figure 3 illustrates the time series of selected quantities from 11.2 km in the CRM, the cirrus detrainment level. This level is selected for examination both because it has the most clouds and because it has the largest value of $\overline{r'^2}/r_s^2$, the total water mixing ratio variance normalized by the square of the saturation mixing ratio, which is a key quantity for the prediction of clouds with a statistical cloud scheme approach. The solid line in the top panel shows that the CRM time series of $\overline{r'^2}$ has numerous "events" where the variance levels increase strongly. For display purposes, $\overline{r'^2}$ has been smoothed by application of a 3 hour running mean. Following each event, the variance initially decreases rapidly in the first 12 to 24 hours. After that, the variance decays more gradually on the order of days. A few exceptions exist where the variance increases slightly (26–27 June) or is constant (9–10 July) between convective events.

[19] The bottom panel of Figure 3 displays the magnitudes of the convective source terms as diagnosed from the CRM. (These convective source terms have also been smoothed.) The "events" in the $\overline{r'^2}$ time series correspond to instances where the convective source terms are large. Indeed, the inference that convection is a major source of variance is confirmed by the fact that 80% of the total variance at this level comes from the 30% of time when convective events occur, which are defined as when the convective source terms are greater than 0.01 $g^2/kg^2/day$. The average values of these source terms averaged over the duration of the simulation are recorded in Table 1. There are two interesting points to be made.

[20] First, it is initially surprising that the entrainment terms are significant at this level, the cirrus detrainment level. This is probably the result of the analysis method. Even if there is a net detrainment of mass from each convective cell, it is not necessarily the case that mass is flowing out of the convective cell at all points along the convective-stratiform interface. Mass might be leaving horizontally on one side of the convective cell but still be

Table 1. Convective Source Terms for Total Water Variance $\overline{r'^2}$ at 11.2 km Averaged Over the Last 27 Days of the Simulation

Convective Source Term	Average Value, 10^{-4} g^2/kg^2/day
$D(r_d - r)^2$	+371
$D(\overline{r_d'^2} - \overline{r'^2})$	+399
$-E(r_e - r)^2$	−193
$-E(\overline{r_e'^2} - \overline{r'^2})$	−217
$D(r_d - r)^2 - E(r_e - r)^2$	+177
$D(\overline{r_d'^2} - \overline{r'^2}) - E(\overline{r_e'^2} - \overline{r'^2})$	+182
$-M_c\frac{\partial \overline{r'^2}}{\partial p}$	−20

entering the same cell from the other side. With the analysis method used, this cell would have both detrainment and entrainment occurring. Note that each entrainment term is of opposite sign but about 50% of the magnitude of the corresponding detrainment term (Figure 3 and Table 1).

[21] Second, it appears that averaged over the simulation the magnitudes of the two detrainment terms are comparable although there is variability in their relative magnitudes between events (Table 1). This suggests that the impact of convection on the variance budget is not just through the detrainment of air parcels with different mean values of r, but also through the detrainment of air parcels with variance in excess of that present in the environment. Is this result to be believed? There are many limitations to this analysis. For example, only snapshots of model fields every 5 min were analyzed, the partitioning of the domain into convective and stratiform regions is arbitrary, and the variance of r at scales smaller than 2 km was not analyzed. In addition, because the model is not three-dimensional the scalar properties might be distorted and simulations with other CRMs may yield different results [*Khairoutdinov and Randall*, 2003]. In particular, as suggested by one reviewer, the very coarse horizontal resolution of the model implies that updrafts will fill only a few grid boxes. As a result, the variance of r in the detrained air comes from variability between different convective cores or variability over an hour of the mean r within a single core. A CRM with much higher horizontal resolution would allow one to quantify the variability inside the air detrained from a single updraft plume. Despite all these limitations, this result might be plausible because cumulus clouds are well-known to be turbulent. *Lin and Arakawa* [1997] showed in their Figure 2 that the prognosed subgrid turbulence was very large at the edges of cumulus clouds simulated with this same CRM. Thus one could envision that the variance of r in detrained air parcels is large, and a significant source of r variance for the stratiform environment.

[22] To illustrate the good correspondence between the convective source terms and the variance time series, the upper panel of Figure 3 displays the time series of $\overline{r'^2}$ predicted from the following equation:

$$\frac{\partial \overline{r'^2}}{\partial t} = S_{cv}\left(\overline{r'^2}\right) - \frac{\overline{r'^2}}{\tau}, \tag{12}$$

where τ is a fixed dissipation timescale of 1 hour and S_{cv} ($\overline{r'^2}$) is taken directly from that diagnosed from the CRM (and for which the components of S_{cv} ($\overline{r'^2}$) are shown in the lower panel of Figure 3). The dissipation term is meant to represent all of the non-convective source/sink terms in the variance equation not expressly calculated. The arbitrary choice of a very short dissipation timescale may correspond to that associated with the dissipation of variance by precipitation (3); *Khairoutdinov and Randall* [2002] found that the dissipation of variance by precipitation is a major sink in the presence of precipitating convection. The correspondence between the predicted and actual CRM time series, while not perfect, is very strong; the linear correlation coefficient between these time series is 0.84. During convective events, the mean values of $\sqrt{\overline{r'^2}}$ from the CRM and predicted by (12) are 0.060 and 0.062 g/kg, respectively, indicating little bias in the prediction of variance during times of convection. However, the use of an hour timescale is probably not appropriate at periods in between convective events, as the predicted variance is considerably smaller than that resolved by the CRM. At these times, it may be that variance is being dissipated slowly by subgrid-scale mixing between CRM grid boxes, a slow process with a timescale considerably greater than one hour. Given that non-convective times occur 70% percent of the time and that clouds still persist during these times, the selection of an appropriate decay timescale due to subgrid-scale mixing would be important. However, as a whole, this analysis suggests that the source terms as parameterized in (10) are reasonable.

[23] The analysis for the third moment of total water $\overline{r'^3}$ is presented in Figure 4 and Table 2. The faster decay of $\overline{r'^3}$ relative to that of the variance is noticeable, with the skewness of the distribution being essentially zero between convective events, whereas the variance retains a finite value. Of the convective source terms, the cross terms ($3D(r_d - r)(\overline{r_d'^2} - \overline{r'^2}) - 3E(r_e - r)(\overline{r_e'^2} - \overline{r'^2})$) have the largest magnitude, being roughly equal to the sum of the other two types of terms ($D(r_d - r)^3 - E(r_e - r)^3$ and $D(\overline{r_d'^3} - \overline{r'^3}) - E(\overline{r_e'^3} - \overline{r'^3})$, Table 2). If these diagnosed source terms are used to predict $\overline{r'^3}$ according to:

$$\frac{\partial \overline{r'^3}}{\partial t} = S_{cv}\left(\overline{r'^3}\right) - \frac{\overline{r'^3}}{\tau}, \tag{13}$$

then the correspondence is striking. Although some events are not predicted well, the linear correlation coefficient is 0.69.

4. Practical Considerations for Inclusion of Convective Source Terms in Large-Scale Models

[24] Terms representing the impact of convection on the variance and third moment of total water mixing ratio r have been presented. These source terms could be used in a prognostic equation for these moments in a model that also has a mass-flux convection scheme. Comparison with the variance and third moment budget at the cirrus detrainment level as resolved by a CRM suggests that the approach has promise.

[25] Can these terms be profitably incorporated into a large-scale model? While the physical basis of these source terms presents a theoretical improvement over the ad-hoc convective source terms in the work of *Tompkins* [2002], it may not be possible to implement them in a large-scale model because the large-scale model lacks the information

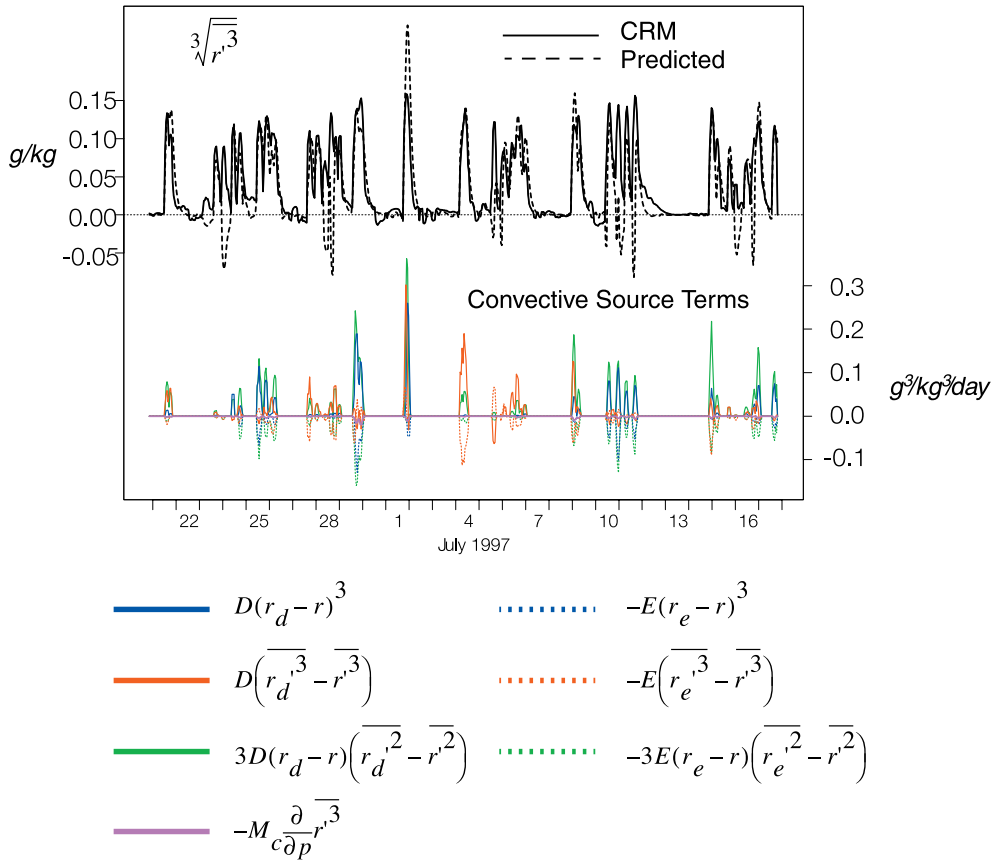

Figure 4. As in Figure 3 but for quantities related to the third moment of total water $\overline{r'^3}$. The top portion of the figure illustrates the time series of the cube root of the third moment of total water $\sqrt[3]{\overline{r'^3}}$ from the CRM (solid) and that predicted from equation (13) using the convective source terms diagnosed from the CRM (dashed). The individual components of the convective source terms at this level are displayed in the lower portion of the figure. As is true for the variance budget, the peaks in the time series of $\sqrt[3]{\overline{r'^3}}$ coincide very well with the peaks in the convective source terms diagnosed from the CRM.

necessary for these source terms. In particular, while most of the variables in these terms are available from current mass-flux schemes, the variance and third moment of r for the air parcels being exchanged between the stratiform and convective parts of the grid box are not available from current convection schemes. If the contributions of the source terms that contain these variables were small, one might ignore them. However, the evidence presented here suggests that they are important for the higher-order moments. How might one predict these terms? One might considering solving the variance budget for a convective updraft (Appendix C), if one has enough confidence that one can model this term well. Note that the variance and third moment of the air detrained from convection is not necessarily the same as the variance and third moment in the convective region, which one might diagnose from a mass-flux scheme that contains a spectrum of updrafts. Another approach might be to use simple scaling relationships; for example, one might assume that the dispersion of total water in the detrained air, $\sqrt{\overline{r_d'^2}}/r_d$, is fixed. Some support for this is provided by Figure 5, which shows a relatively good correlation between $\sqrt{\overline{r_d'^2}}$ and r_d; the linear correlation coefficient is 0.81. Some caution with respect to this result should be exercised as a CRM with much finer spatial resolution may yield a different result.

[26] In implementing these terms into a large-scale model, further simplifications might be needed. For example, the skewness present in the CRM while important is very transitory, suggesting that skewness might be best handled as a diagnostic rather than prognostic variable of a large-

Table 2. Convective Source Terms for the Third Moment of Total Water $\overline{r'^3}$ at 11.2 km Averaged Over the Last 27 Days of the Simulation

Convective Source Term	Average Value, 10^{-4} g^3/kg^3/day
$D(r_d - r)^3$	+78
$D(\overline{r_d'^3} - \overline{r'^3})$	+69
$3D(r_d - r)\,(\overline{r_d'^2} - \overline{r'^2})$	+153
$-E(r_e - r)^3$	−34
$-E(\overline{r_e'^3} - \overline{r'^3})$	−37
$-3E(r_e - r)\,(\overline{r_e'^2} - \overline{r'^2})$	−75
$D(r_d - r)^3 - E(r_e - r)^3$	+44
$D(\overline{r_d'^3} - \overline{r'^3}) - E(\overline{r_e'^3} - \overline{r'^3})$	+32
$3D(r_d - r)\,(\overline{r_d'^2} - \overline{r'^2}) - 3E(r_e - r)\,(\overline{r_e'^2} - \overline{r'^2})$	+78
$-M_c \frac{\partial \overline{r'^3}}{\partial p}$	−6

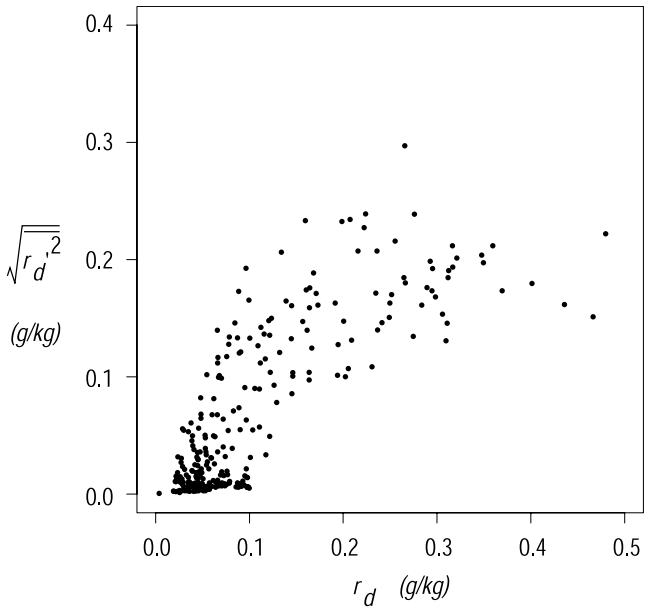

Figure 5. Scatterplot of the standard deviation $\sqrt{\overline{r_d'^2}}$ versus the mean r_d of total water mixing ratio in air detrained at 11.2 km. The data are plotted only for hours with mass detrainment rates greater than 5 day^{-1}. The good correlation suggests that $\sqrt{\overline{r_d'^2}}$, a quantity generally not available from convection parameterizations, may be parameterized in terms of r_d.

scale model. This would be advantageous because the source terms for the third moment from other processes such as turbulence are complicated enough that one might not be able to model them accurately. One practical suggestion, inspired by *Bony and Emanuel* [2001], is that one could adjust the skewness of the distribution at each time step so that the tail of the assumed PDF encompasses the high value of total water being detrained from the convection scheme. This would be advantageous because the PDF diagnosed from a prognosed variance and/or third moment of total water may not encompass the value of the total water being detrained from convection. This would help to ensure greater consistency between the convection scheme and the large-scale clouds.

Appendix A

[27] In this appendix, the derivation of (4) is outlined. The definition of variance of total water mixing ratio r is:

$$\overline{r'^2} = \overline{(r-\bar{r})^2}. \tag{A1}$$

To derive (4), (A1) is applied to the region consisting of the detrained mass and the stratiform environment before detrainment (Figure 1). For example, the mean total water mixing ratio of this region, which is the mean total water mixing ratio at the next time step r^{n+1}, is given by:

$$r^{n+1} = \Delta a r_d + (1-\Delta a) r^n, \tag{A2}$$

where r_d is the mean total water mixing ratio in the detrained air and r^n is the mean total water mixing ratio in the stratiform environment before detrainment. Expanding (A1) into averages over the detrained air and the stratiform environment before detrainment one has:

$$\overline{r'^2}^{n+1} = \Delta a \overline{(r-r^{n+1})^2}^{\Delta a} + (1-\Delta a)\overline{(r-r^{n+1})^2}^{1-\Delta a}. \tag{A3}$$

Substitution of (A2) into the first term on the right-hand side of (A3) gives:

$$\Delta a \overline{(r-r^{n+1})^2}^{\Delta a} = \Delta a \, \overline{((r-r_d)+(1-\Delta a)(r_d-r^n))^2}^{\Delta a} \tag{A4}$$

which after expansion of the quadratic yields:

$$\Delta a \, \overline{(r-r^{n+1})^2}^{\Delta a} = \Delta a \, \overline{(r-r_d)^2}^{\Delta a} + 2\Delta a(1-\Delta a) \cdot (r_d - r^n)\overline{(r-r_d)}^{\Delta a} + \Delta a(1-\Delta a)^2 (r_d - r^n)^2. \tag{A5}$$

The second term on the right-hand side of (A5) contains an average over the detrained air of the deviation of total water mixing ratio from the mean total water mixing ratio in detrained air, an average which equals zero. Thus:

$$\Delta a \, \overline{(r-r^{n+1})^2}^{\Delta a} = \Delta a \, \overline{r_d'^2} + \Delta a(1-\Delta a)^2 (r_d - r^n)^2, \tag{A6}$$

where the average contained in the first-term on the right-hand side of (A5) has been identified by inspection as $\overline{r_d'^2}$, the variance of total water mixing ratio in the detrained air. Substitution of (A2) into the second term on the right-hand side of (A3) yields after a similar procedure:

$$(1-\Delta a)\overline{(r-r^{n+1})^2}^{1-\Delta a} = (1-\Delta a)\overline{r'^2}^n + (1-\Delta a)\Delta a^2 (r_d - r^n)^2. \tag{A7}$$

After substituting (A6) and (A7) into (A3) and combining the like terms, one arrives with (4).

Appendix B

[28] To calculate the convective source terms, a means of calculating mass-fluxes, detrainment/entrainment rates, and the properties of air undergoing detrainment/entrainment is needed. This appendix describes the details of these calculations.

[29] The first step is to partition the CRM domain into convective and stratiform regions. For each 5 min snapshot, every 2 km grid box in the 512-km domain is declared convective or stratiform according to the algorithm of *Xu* [1995]. Convective regions are identified by searching for columns with high values of either surface precipitation rate or vertical velocity beneath the melting level. It is important to note that the same decomposition applies to every vertical level of the CRM; that is, no 2 km column may contain a mixture of convective and stratiform grid cells.

[30] To calculate the rates at which mass is entrained into or detrained from the convective portion of the model, the method of *Siebesma* [1998] is used (Figure B1). The convective detrainment rate D, with units of s^{-1}, is calculated from a line integral of the outward flow of mass across

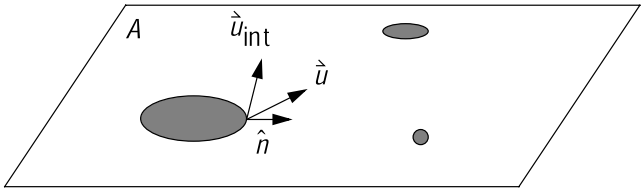

Figure B1. Schematic diagram illustrating the calculation of the detrainment rate. Grey shaded areas indicate convective regions in the horizontal plane with a total model area A. $\vec{u}$ is the horizontal wind vector, $\hat{n}$ is a unit vector normal to the convective-stratiform interface, and $\vec{u}_{\rm int}$ is the horizontal velocity of the interface. The detrainment rate is the rate at which mass flows from the convective region to the stratiform region across the moving convective-stratiform interface.

the interfaces between the convective and stratiform regions:

$$D = \frac{\oint \max(0, \hat{n} \bullet (\vec{u} - \vec{u}_{\rm int}))dl}{A}. \tag{B1}$$

In (B1), $\hat{n}$ is a unit vector normal to the convective-stratiform interface and pointed in the direction of the stratiform region, A is the total horizontal area of the model including both stratiform and convective areas, dl is the differential of length along the convective-stratiform interface, $\vec{u}$ is the horizontal wind vector, and $\vec{u}_{\rm int}$ is the vector for the horizontal velocity of the convective stratiform interface. Note that one needs to know the horizontal motion of the convective-stratiform interface, as only flow relative to the interface contributes to detrainment. The maximum operator present in (B1) selects only the portions of the interfaces where the flow is from the convective region to the stratiform region.

[31] To calculate r_d, the mean total water mixing ratio in the detrained air, one weights the total water mixing ratio at the convective-stratiform interface with the rate of flow relative to the interface:

$$D \times r_d = \frac{\oint r \times \max(0, \hat{n} \bullet (\vec{u} - \vec{u}_{\rm int}))dl}{A}; \tag{B2}$$

$\overline{r_d'^2}$ and $\overline{r_d'^3}$, the variance and third moment of total water mixing ratio in the detrained air, are defined likewise:

$$D \times \overline{r_d'^2} = \frac{\oint (r - r_d)^2 \times \max(0, \hat{n} \bullet (\vec{u} - \vec{u}_{\rm int}))dl}{A} \tag{B3}$$

$$D \times \overline{r_d'^3} = \frac{\oint (r - r_d)^3 \times \max(0, \hat{n} \bullet (\vec{u} - \vec{u}_{\rm int}))dl}{A}. \tag{B4}$$

When calculating these quantities in a two-dimensional model, some simplifications arise. Equation (B1) reduces to a summation over all convective-stratiform interfaces:

$$D = \frac{\sum \max(0, n \times (u - u_{\rm int}))}{L}, \tag{B5}$$

where L is the total horizontal length of the model (512 km), and n is equal to +1 if the interface has a stratiform region to the right of a convective region, that is, at a larger value of the horizontal coordinate in whose units the horizontal wind velocity is measured. If the interface has a stratiform region to the left of a convective region, then n is equal to -1. Likewise r_d is computed from:

$$D \times r_d = \frac{\sum r \times \max(0, n \times (u - u_{\rm int}))}{L}, \tag{B6}$$

where the value of r at the interface is taken from the grid cell on the convective side of the convective-stratiform interface. $\overline{r_d'^2}$ and $\overline{r_d'^3}$ are calculated similarly. However, to eliminate statistical noise in their calculation, $\overline{r_d'^2}$ and $\overline{r_d'^3}$ are set to missing when there are less than 10 convective-stratiform interfaces in an hour.

[32] The only complicated matter is the calculation of $u_{\rm int}$, the horizontal motion of the convective-stratiform interface, from the 5 min model snapshots. For each interface a search is made in both the previous and next snapshot for an interface within the adjacent 3 grid cells. If such an interface is found in both snapshots, the difference in horizontal position of the interfaces between the next and previous snapshot is used to compute the velocity of the interface. If an interface is found in only one of these snapshots, then the difference in position between the interface in the snapshot with the interface and the current snapshot is used to define the interface velocity. If an interface cannot be found in either the previous or next snapshots, as occurs if an interface appears and disappears within the time between the previous and next snapshots, then the interface velocity is set to zero.

[33] Admittedly these calculations are somewhat crude and have their limitations. However, they may be sufficient because the values of detrainment and entrainment diagnosed from this method reasonably satisfy the mass-flux budget equation (7) (not shown). This represents a somewhat independent test, as the mass-flux, whose vertical divergence is the left-hand side of (7), is diagnosed from the CRM vertical velocity directly.

Appendix C

[34] The convective source terms present in this paper include quantities that are not generally available from convection schemes, such as the variance $\overline{r_d'^2}$ and third moment $\overline{r_d'^3}$ of the total water mixing ratio in the air detrained from convection. In this appendix, the budget equation for the variance of total water mixing ratio in a convective updraft is provided as an illustration of a possible way to close $\overline{r_d'^2}$.

[35] The budget equation in pressure coordinates for total water mixing ratio in a convective updraft r_c is:

$$\frac{\partial r_c}{\partial p} = -\frac{E}{M_c}(r_e - r_c) + \frac{D}{M_c}(r_d - r_c) + G_{pr,c}, \tag{C1}$$

where r_e is the total water mixing ratio of air entrained into the updraft, r_d is the total water mixing ratio of air detrained from the updraft, and $G_{pr,c}$ is the sink of total water due to precipitation with units kg water (kg air)$^{-1}$ Pa^{-1}. Following

arguments similar to the derivation in section 2, the budget equation for $\overline{r_c'^2}$, the variance in total water mixing ratio in a convective updraft would be:

$$\frac{\partial \overline{r_c'^2}}{\partial p} = -\frac{E}{M_c}(r_e - r_c)^2 - \frac{E}{M_c}\left(\overline{r_e'^2} - \overline{r_c'^2}\right) + \frac{D}{M_c}(r_d - r_c)^2 + \frac{D}{M_c}\left(\overline{r_d'^2} - \overline{r_c'^2}\right) + 2\,\overline{G_{pr,c}' r_c'} + \varepsilon, \quad \text{(C2)}$$

where $\overline{r_e'^2}$ and $\overline{r_d'^2}$ are respectively the variance of total water mixing ratio of air entrained into and detrained from the convective updraft, $\overline{G_{pr,c}' r_c'}$ is the covariance between total water and the precipitation sink in the convective updraft, and ε is the dissipation of variance per unit pressure. The physical meaning of each of these terms is as follows. The variance of total water in the updraft increases with altitude (or decreasing pressure) if the air entrained into the updraft has a different mean value of total water than present in the updraft, or if the variance within the entrained air is greater than that in the updraft. Likewise, the variance will decrease with altitude (or decreasing pressure) if the air detrained from the updraft has a different mean value of total water than present in the updraft, or if the variance in the detrained air is greater than the variance present in the updraft. The next to last term in (C2) indicates that if the precipitation sink is greater in the portion of the updraft that contains greater than average total water, then the variance will decrease with altitude (or decreasing pressure). The last term is the dissipation of variance due to mixing within the updraft.

[36] If one assumes that the properties of air detrained from the convective updraft do not differ from the mean properties of the updraft and that the properties of air entrained into the updraft do not differ from those of the stratiform environment, then (C2) reduces to:

$$\frac{\partial \overline{r_c'^2}}{\partial p} = -\frac{E}{M_c}(r - r_c)^2 - \frac{E}{M_c}\left(\overline{r'^2} - \overline{r_c'^2}\right) + 2\,\overline{G_{pr,c}' r_c'} + \varepsilon. \quad \text{(C3)}$$

The value of $\overline{r_d'^2}$ needed for the convective source terms in (10) and (11) is equal to the value of $\overline{r_c'^2}$ at the detrainment level that is determined from the solution of (C3). The boundary condition on (C3) is that the value of $\overline{r_c'^2}$ at the base of the updraft equals the value of $\overline{r'^2}$ in the stratiform environment.

[37] **Acknowledgments.** Thoughtful comments provided on the manuscript by Anning Cheng, Christian Jakob, Ben Johnson, and the three anonymous reviewers are appreciated. This research is funded in part by the Office of Science (BER), U.S. Department of Energy, Interagency Agreements DE-AI02-00ER62934, DE-AI02-03ER63562, and DE-AI02-02ER63318. This work was performed under the auspices of the U.S. Department of Energy at the University of California Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

References

André, J. C., G. De Moor, P. Lacarrère, and R. du Vachat (1978), Modeling the 24-hour evolution of the mean and turbulent structures of the planetary boundary layer, *J. Atmos. Sci.*, *35*, 1861–1883.

Bony, S., and K. A. Emanuel (2001), A parameterization of the cloudiness associated with cumulus convection; evaluation using TOGA-COARE data, *J. Atmos. Sci.*, *58*, 3158–3183.

Bougeault, P. (1982), Cloud-ensemble relations based on the gamma probability distribution for the higher-order models of the planetary boundary layer, *J. Atmos. Sci.*, *39*, 2691–2700.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider (1994), The albedo of fractal stratocumulus clouds, *J. Atmos. Sci.*, *51*, 2434–2460.

Golaz, J.-C., V. E. Larson, and W. R. Cotton (2002), A PDF-based model for boundary layer clouds. part I: Method and model description, *J. Atmos. Sci.*, *59*, 3540–3551.

Khairoutdinov, M. F., and D. A. Randall (2002), Similarity of deep continental cumulus convection as revealed by a three-dimensional cloud-resolving, *J. Atmos. Sci.*, *59*, 2550–2566.

Khairoutdinov, M. F., and D. A. Randall (2003), Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, *J. Atmos. Sci.*, *60*, 607–625.

Lappen, C.-L., and D. A. Randall (2001), Toward a unified parameterization of the boundary layer and moist convection. part I: A new type of mass-flux model, *J. Atmos. Sci.*, *59*, 2021–2036.

Larson, V. E., R. Wood, P. R. Field, J.-C. Golaz, T. H. Vonder Haar, and W. R. Cotton (2001), Systematic biases in the microphysics and thermodynamics of numerical models that ignore subgrid-scale variability, *J. Atmos. Sci.*, *58*, 1117–1128.

Larson, V. E., J.-C. Golaz, and W. R. Cotton (2002), Small-scale and mesoscale variability in cloudy boundary layers: Point probability density functions, *J. Atmos. Sci.*, *59*, 3519–3539.

Lin, C., and A. Arakawa (1997), The macroscopic entrainment processes of simulated cumulus ensemble. part I: Entrainment sources, *J. Atmos. Sci.*, *54*, 1027–1043.

Luo, Y., S. K. Krueger, G. G. Mace, and K.-M. Xu (2003), Cirrus cloud properties from a cloud-resolving model simulation compared to cloud radar observations, *J. Atmos. Sci.*, *60*, 510–525.

Mellor, G. L. (1977), The Gaussian cloud model relations, *J. Atmos. Sci.*, *34*, 356–358.

Pincus, R., and S. A. Klein (2000), Unresolved spatial variability and microphysical process rates in large-scale models, *J. Geophys. Res.*, *105*, 27,059–27,065.

Siebesma, A. P. (1998), Shallow cumulus convection, in *Buoyant Convection in Geophysical Flows*, edited by E. J. Plate et al., pp. 441–486, Springer, New York.

Smith, R. N. B. (1990), A scheme for predicting layer clouds and their water-content in a general-circulation model, *Q. J. R. Meteorol. Soc.*, *116*, 435–460.

Sommeria, G., and J. W. Deardorff (1977), Subgrid-scale condensation in models of nonprecipitating clouds, *J. Atmos. Sci.*, *34*, 344–355.

Tompkins, A. M. (2002), A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover, *J. Atmos. Sci.*, *59*, 1917–1942.

Xu, K.-M. (1995), Partitioning mass, heat, and moisture budgets of explicitly simulated cumulus ensembles into convective and stratiform components, *J. Atmos. Sci.*, *52*, 551–573.

Xu, K.-M., and D. A. Randall (1996), Evaluation of statistically based cloudiness parameterizations used in climate models, *J. Atmos. Sci.*, *53*, 3103–3119.

Xu, K.-M., et al. (2002), An intercomparison of cloud-resolving models with the Atmospheric Radiation Measurement summer 1997 intensive observation period data, *Q. J. R. Meteorol. Soc.*, *128*, 593–624.

Yanai, M., and R. H. Johnson (1993), Impacts of cumulus convection on thermodynamic fields, in *The Representation of Cumulus Convection in Numerical Models*, edited by K. A. Emanuel and D. J. Raymond, pp. 39–62, Am. Meteorol. Soc., Boston, Mass.

Zhang, M.-H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S.-C. Xie (2001), Objective analysis of ARM IOP data: Method and sensitivity, *Mon. Weather Rev.*, *129*, 295–311.

C. Hannay, National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, CO 80305, USA.

S. A. Klein, Atmospheric Science Division (L-103), Lawrence Livermore National Laboratory, Livermore, CA 94551, USA. (klein21@llnl.gov)

R. Pincus, Climate Diagnostics Center, University of Colorado, 216 UCB, Boulder, CO 80309-0216, USA.

K.-M. Xu, NASA Langley Research Center, Mail Stop 420, Hampton, VA 23681-2199, USA.