THE PARAMETERIZATION OF DEEP CONVECTION: A REVIEW

Alan K. Betts
Atmospheric Research, R.D. 3, Box 3125, Pittsford, VT 05763
akbetts@aol.com

Abstract. Insights into the parameterization of convection from two decades of diagnostic studies are reviewed. The life cycle of a convective mesosystem mass flux is described using day 245 from the GARP Atlantic Tropical Experiment as example. The thermodynamic differences between non-precipitating convection and precipitating convection are discussed, as well as the importance of the mid-tropospheric freezing level in the tropics. Diagnostic studies which have identified three key vertical modes in the convective heating and drying structure are outlined. Two are related deep modes associated with precipitation and deep tropospheric ascent, but a variable upward equivalent potential temperature ($\theta_e$) flux. The other is a double mode structure with ascent in the upper troposphere over descent in the lower troposphere, coupled to inflow at the freezing level, with no net precipitation or transport of $\theta_e$; the mode associated with deep mesoscale anvils. We discuss the mass flux formulation of convective updrafts and downdrafts. We outline the concepts (but not the details) behind the Betts-Miller parameterization and suggest two extensions. One is a formulation of the adjustment time in terms of grid-scale and gravity-wave propagation speed for the two primary modes. The second is an explicit parameterization of the mesoscale anvil mode.

1. INTRODUCTION

There have been several reviews of convective parameterization in large-scale models including Betts [1974], Frank [1983], text-books such as Cotton and Anthes [1989] and Emanuel [1994] and particularly the recent American Meteorological Society monograph [Emanuel and Raymond, 1993]. This includes a recent discussion of the Betts-Miller scheme [Betts and Miller, 1993], which includes a formulation for unsaturated downdrafts driven by convection. Although I shall summarize the basis of the Betts-Miller scheme in section 4, and suggest a further extension, I will not discuss the details. For the most part, I shall review here some of the fundamental issues that have been uncovered by many researchers over the last twenty-five years; and try to reduce them to their simplest form. As a result this paper is part history, part review and part reinterpretation, together with a few suggestions for a way forward. This paper discusses only the energy and water transports by convection, not the issues of the momentum and vorticity transports. An extended version of this workshop paper will be published as Betts [1997].

2. DEEP CUMULUS DIAGNOSTIC RESULTS

2.1. Development of Cumulus Parameterizations

The development of cumulus parameterizations, which had started in the 1960's with the work of Kuo [1965] and Manabe and Strickler [1965] and two important conceptual papers by Ooyama [1969, 1971], received a stimulus from diagnostic models in the 1970's. The Arakawa and Schubert [1974] paper and the diagnostic study of Yanai et al. [1973], using data over the tropical Pacific, were closely interlinked. Their cloud model based on entraining updrafts has dominated this school of cumulus
parameterization up to this date. They showed how the $\theta_e$ flux could be calculated, independent of the precipitation, using this cloud mass flux model. Subsequently this model has been extended to include convective downdrafts [Arakawa and Cheng, 1993]. At the same time there were two parallel papers by Betts [1973a, 1973b]. The first [Betts, 1973a] introduced the mixed layer model for the dry boundary layer and the subcloud layer, which was included in the Arakawa-Schubert parameterization. This paper also extended the use of conserved variables (in particular, the liquid water potential temperature, dealing with the coupling of the enthalpy and liquid water fluxes in convection), and discussed the concept of lapse rate equilibrium for shallow convection. A second paper [Betts, 1973b] used data from the first Venezuela International Meteorological and Hydrological Experiment (VIMHEX) to formulate an updraft-downdraft budget model for mesoscale cumulonimbus systems, based on compositing rawinsonde observations in relation to radar echos. It showed how using $\theta_e$ conservation separately for the updraft and downdraft circulations permitted a separate evaluation of updraft condensation, evaporation into the downdraft and net precipitation. Figure 1a shows the mass fluxes associated with the high $\theta_e$ updraft circulation and the low $\theta_e$ downdraft circulation. The paper concluded that half the condensation in the updraft was evaporated into the downdrafts. The composite mesoscale system was also partitioned temporally into the growth and decay phase (Figure 1b). This showed how the net lifecycle upward mass flux was the resultant of an upward mass circulation in the growth phase, and a downward circulation in the decay phase. The paper also showed that the low level transformation by deep convection was dominated by cold dry downdraft outflows. The data used by Betts [1973b] was quite primitive, but the subsequent tropical field programs have supported the conclusions of the analysis. The second VIMHEX experiment explored in more detail the transports by unsaturated downdrafts [Betts 1976] and their modeling [Betts and Silva Dias, 1979]. Miller and Betts [1977] used a mesoscale model to study the distinction between convective and mesoscale downdrafts, reaching similar conclusions to pioneering papers by Zipser [1969, 1977] and the GATE squall-line study of Houze [1977].

2.2. GATE Diagnostic Studies

By the end of that decade it was clear from these tropical studies over land and from the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE) (see reviews by Betts [1978]; Houze and Betts [1981]) that parameterizing tropical convection needed a model for the lifecycle effect of mesoscale complexes, which evolve from lines of cumulonimbus to mesoscale anvil systems within 6-12 hours, on horizontal scales that are not well resolved in global models. Figure 2 is a schematic, typical of the GATE environment with strong shear in the low levels. Within the time-span of a typical mesoscale aircraft mission, the observer would see the development of cumulonimbus bands oriented along the low level shear, with inflows on one side and a developing anvil outflow to the rear. A typical spacing might be 60 km. As each line evolved, the strong convection would decay and a deep mesoscale anvil would persist often for many hours, with steady stratiform rain and a characteristic bright band at the freezing level [Leary and Houze, 1979a,b; 1980]. On some occasions, faster moving cross-shear lines with trailing anvils and some of the characteristics of squall-lines were observed [Barnes and Sieckman, 1984]. It was also clear from the detailed analyses of Ooyama in the late 1970's, that the freezing level, which in the tropics is in the mid-troposphere, is dynamically significant. In the mesoscale anvil stage, it is the level of maximum convergence, with ascent above in the region of condensation and freezing, and descent below driven
Figure 1. a) Vertical mass fluxes for high and low $\theta_e$ ranges showing simplified bulk updraft and downdraft (from Betts, 1973b). The peak mass flux of 150 (archaic) CGS units corresponds to 1.5 kg m$^{-2}$ s$^{-1}$ over a 25 km diameter echo for a time interval of 66 minutes. b) As Fig. 1a for vertical mass flux, partitioned into growth and decay phases, together with the life-cycle mean mass flux.

Figure 2. Schematic of GATE convective band development.
by the evaporation and melting of falling precipitation. However, although the analysis theory was published later in Ooyama [1987], only some of the fascinating lifecycle case studies have been fully published. (The Ooyama, Chu and Esbensen wind and thermodynamic analyses are available in an archive at National Center for Atmospheric Research: documentation is available in Esbensen and Ooyama [1993].) Tollerud and Esbensen [1985] and Esbensen et al. [1988] published some composites of mass, heat and moisture budgets of non-squall clusters based on that data. Other GATE composites were published [Frank, 1978, Johnson, 1980], based on larger scale analyses, but they do not show the mesoscale lifecycles well, because of the smoothing involved in the compositing. The interpretation of diagnostic budgets by Johnson [1984] in terms of convective and mesoscale components was however a significant step forward. The most detailed analysis of the Ooyama, Chu and Esbensen dataset was published by Cheng [1989a,b] and Cheng and Yanai [1989]. Using an Arakawa-Schubert type cloud model with a spectrum of entraining updrafts and downdrafts, together with a mesoscale circulation, Cheng [1989a] first explored a tilted updraft model and then systematically determined mass flux distributions for convective-scale updrafts and downdrafts. Cheng and Yanai [1989] used the results from the cloud model to extract a mesoscale mass flux and heating function, which explicitly shows the mesoscale couplet of heating and drying over cooling and moistening for several GATE cloud clusters. I will return to their analysis in section 2.6.

2.3. GATE Cloud Cluster Lifecycle On Day 245

As a brief review I will present an example here for the life cycle mass flux evolution of a major GATE cloud cluster on Julian Day 245. In the published literature, Betts [1978] gave a preliminary analyses of the mass flux evolution for this day; Mower et al. [1979] analyzed the synoptic, radar and aircraft data for this day, and Warner [1980] analyzed the cloud fields. This cluster evolution is an excellent example, as it occurred over the GATE ship array and the evolution can be seen in both time and space. Bands of convection developed during the day, evolved into a mesoscale complex and then decayed. The daytime evolution over the ship array was studied by a stack of five aircraft [Mower et al., 1979]; for which the present author was the airborne mission scientist. This cluster is case 31 in the Cheng and Yanai [1989] analysis.

Figure 3 shows a time-series (from the Ooyama analysis) of the mean vertical motion, $\omega$, from 0300 UTC to 2400 UTC on September 2, 1994 (Julian Day 245) at 3 hour intervals at 8.5°N, 22°W, just east of the center of the GATE ship array. The mass field shows upward vertical motion initially peaking in the lower troposphere at 0300 UTC. By noon there is a strong ascent through the whole troposphere with the peak vertical motion in the middle troposphere. Then the upward motion strengthens further in the upper troposphere (1500 and 1800 UTC), as it decays in the lower troposphere. This is the stage when strong mesoscale ascent develops in the extensive anvils in the upper troposphere, and underneath a mesoscale downdraft forms driven by the melting and evaporation of falling precipitation. The evolution continues until by 2400 UTC, the residual circulation is a dipole mass field with ascent above the freezing level and descent below. Figure 4 is a horizontal cross-section of the divergence field at the freezing level at 2100 UTC, showing a large peak in the convergence at this level, centered near the location of our time-series of $\omega$ in Figure 3. This freezing level convergence associated with the mesoscale couplet of ascent over descent is as large an inflow as the low level convergence at any time in the cloud cluster lifecycle. An important question is

169
Figure 3. Time-series of the mean vertical motion at 8.5°N, 22°W from 0300-2400 UTC on September 2 1974 (Day 245), from GATE analysis of Ooyama’s wind sounding data.

Figure 4. Horizontal cross-section of wind field and divergence at 565 hPa (near the freezing level) at 2100 UTC on day 245, showing peak convergence of $2.8 \times 10^5$ s$^{-1}$. The dotted lines are the GATE inner B-scale and outer A/B-scale ship arrays. Axes are latitude and longitude. Data is from Ooyama’s analysis of GATE wind sounding data.
whether this dipole circulation is well resolved by global models, or whether it should be parametrized as a heating-cooling couplet, which will then drive the dipole circulation in the large-scale model. We shall return to this later as well as to the Cheng and Yanai [1989] analysis of this case (in section 3.6), after discussing their diagnostic model.

3. THERMODYNAMIC ISSUES AND DIAGNOSTIC MODELS
The second topic I shall review is the thermodynamic issues which distinguish precipitating and non-precipitating convection. I will then discuss the key convective modes and finally outline the mass flux representation of deep convection used in diagnostic models, and some parametric models.

3.1. Non-precipitating Convection
As long as cloud droplets are small, they are carried with air parcels. Reversible thermodynamics apply provided there is no mixing; and ascent and descent of cloudy parcels follow saturated reversible adiabats [Betts, 1973a; Emanuel, 1994]. The enthalpy and water transports are well defined and well coupled to the mass flux, whether this is up or down. Indeed because the vertical motion is reversible, the net mass flux is enough to specify the net transports of enthalpy and water. However if non-precipitating parcels do not mix, their ascent and descent paths are identical and net mass transports would be very small. The shallow convection process we see in the atmosphere is characterized by an essential irreversibility caused by mixing between cloudy air parcels and their unsaturated and stably stratified environment. Cloudy parcels originating below cloud base, if they mix with air from above cloud base, which has a higher potential temperature, \( \theta \), as well as being unsaturated, cannot descend back as far as they ascended. Consequently the lifecycle of shallow clouds impose irreversible transports. Liquid water which is condensed in the lower part of the cloud layer is transported upwards and evaporated at higher levels [Betts, 1973a, 1975], resulting in a couplet of convective cooling at the inversion layer over warming below (with no integrated enthalpy source if the clouds do not precipitate). The mixing process, although irreversible, can however be treated to sufficient accuracy using conserved variables and mixing lines on a thermodynamic diagram [Betts, 1982a]. In oceanic boundary layers (BL), the entire convective BL reaches an equilibrium between the convective transports, the subsidence and the radiation field, which is relatively easy to understand and model (for example Betts, 1975; Betts and Ridgway, 1988, 1989). The cloud field transports moist air with subcloud layer properties upwards in the face of mean subsidence. The upward transport of water by the clouds balances the mean subsidence, and the low liquid-water potential temperature, \( \theta_L \), of cloud air (equal to the \( \theta \) of the subcloud layer), together with longwave radiation, cools the warm air sinking through the inversion at BL-top. Because the whole layer equilibrium has closely a mixing line structure, and liquid water is carried with air parcels, the convective transports can be well represented by a single mass flux [Siebesma and Cuijpers, 1995].

However once cloud droplets grow large enough to fall out of air parcels into unsaturated air, the cloud microphysics becomes important and the entire thermodynamic picture changes.

3.2. Precipitating Convection and the Limitation of Mass Flux Models
Firstly, determining the fraction of the precipitation that falls out is important. \( \theta_L \) is no longer conserved in the updrafts and the subsequent thermodynamics of updraft parcels is different, as their
cloud water is reduced considerably. Even in small Trade-cumulus clouds, once they become deep enough to precipitate, the fallout of precipitation produces layering in the atmospheric structure. The reason is that downdrafts formed by the evaporation of rain penetrate vertically more than those formed by mixing; simply because, for the same change of mixing ratio \( q \), virtual potential temperature, \( \theta_v \), changes faster along the moist adiabat (which is conserved in the evaporation of falling rain), than along typical shallow cumulus BL mixing lines, which constrain evaporation by mixing [Betts and Albrecht, 1987]. For evaporation driven downdrafts, the key thermodynamic issue is that only one air parcel parameter, \( \theta_v \), the equivalent potential temperature, is closely conserved in the evaporation process, and this does not determine the downdraft parcel subsaturation (unlike evaporation in a mixing process, when both \( \theta_v \) and \( \theta_e \) and saturation pressure mix conservatively). By definition precipitation droplets are larger, they have a smaller surface to mass ratio, and they do not evaporate fast enough to keep the air into which they fall saturated, since this air, once cooled by evaporation, sinks in downdrafts seeking a new level of buoyancy equilibrium. In this process, downdraft subsaturation is a result of an internal balance, which can be formulated conceptually in terms of a pressure scale for evaporation [Betts and SilvaDias, 1979]

\[
\Pi_E = \rho g W_d \tau_E
\]

where \( W_d \) is a characteristic downdraft speed and \( \tau_E \), a characteristic evaporation time-scale for the water flux of the falling droplets, dependent in turn on their microphysical properties; mean size and number density. For the simple case of a uniform population of \( N \) drops of size \( r \), one can show [Kamburova and Ludlam, 1966; Betts and SilvaDias 1979]

\[
\tau_E = 4\pi DC_v Nr
\]

where \( D \) is a coefficient of diffusion of water vapor in air, \( C_v \) is a ventilation coefficient for the evaporation of falling drops. The rainfall rate, downdraft speed and negative buoyancy and the stratification are all interactive, but typically we observe in cumulonimbus downdraft outflows, values of \( \Pi_E \) from 30-120 mb corresponding to low level relative humidities from 85-55%. In unstable atmospheres, where downdraft speeds are larger, we tend to see higher values of \( \Pi_E \), corresponding to lower relative humidity.

The key consequence of (1) is that, unlike moist updrafts, which remain close to saturation, the thermodynamics of moist downdrafts are not well determined: their subsaturation depends on small-scale dynamical and microphysical parameters, which must be formulated in terms of large-scale model variables. This is in sharp contrast to the updraft circulation, which is very close to saturation (because the droplets are small and supersaturations are small, see for example Ludlam, [1980]). Diagnostically if we measure downdraft inflow and outflow we can infer evaporation into them [e.g. Betts, 1973b, 1976], but only a few parametric schemes [e.g. Emanuel, 1991; Betts and Miller, 1993] have attempted to treat unsaturated downdrafts even in a simple manner. Unfortunately in many convective parameterizations, moist downdrafts, if included at all, are formulated simply as a fraction of the updraft mass flux, and often treated as saturated. This successfully avoids the real complexity of the deep convective process by reducing the parametric problem to determining a single updraft.

172
mass flux, but it is an unsatisfactory simplification. Unlike shallow convection, deep convection is not a simple mass flux problem, because the precipitation is falling freely and interacting with the atmosphere. For both the updraft and downdraft circulations, only one variable $\theta_E$ is (approximately) conserved, and the submodels which handle the microphysics of precipitation fallout and evaporation into downdrafts are critical. The widespread use of mass flux models for cumulus parameterization has perhaps partially obscured this important issue.

3.3. The Freezing Level
The freezing level is important in the tropics: it is in the middle troposphere near 550 mb. Usually it is the level where the profile changes from unstable in, saturation equivalent potential temperature, $\theta_{ES}$, to stable [Betts, 1982a]. Often there is a visible kink in the thermal $\theta_{ES}$ structure, and typically the level of minimum $\theta_E$ is near the freezing level. This has been known for many years, but it was not explicitly incorporated as a feature in a convective parameterization scheme until Betts [1986], Betts and Miller [1986]. Undoubtedly the stratiform precipitation phase change plays a role in the maintenance of this characteristic structure. We have already mentioned in section 2 that in the decay phase of convective mesosystems, the inflow peaks at the freezing level (Figure 4) with ascending motion above, and descending motion below. We shall discuss this convective mode further in section 3.4 and the relationship of the mid-tropospheric $\theta_E$ minimum to the upward $\theta_E$ flux in section 3.5.

3.4. The Key Convective Modes
An important paper by Johnson [1984] showed one method of partitioning diagnostic heat and moisture budgets into cumulus and mesoscale components. He assumed that the condensation in the mesoscale anvils was a certain fraction (=0.2) of the total precipitation, and then derived the characteristic mesoscale couplet signature of the warming and drying above freezing level and cooling and moistening below.

This paper was followed by a review by Arakawa and Chen [1987] which contained some significant diagnostic insights. They distinguished different types of closure assumptions in parameterization schemes. In particular, they defined a Type II closure as one that constrained the coupling of the convective heat source $(Q_1, -Q_R)$ and moisture sink $(Q_2)$, using the notation of Yanai et al. [1973], in which $Q_1$ is the total diabatic source term and $Q_R$ is the radiative contribution to this term. They used canonical correlation analyses on the GATE Phase III data [of Esbensen and Ooyama, 1983] and an Asian data set [from He et al., 1987] to show there were three principal modes of coupling of $(Q_1, -Q_R)$ and $Q_2$. We show them schematically in Figures 5a, and 5b. Mode 1 is the principal deep convection mode associated with cumulonimbus updrafts and downdrafts through the deep troposphere. (We will associate their third mode with a modulation of Mode 1: see below). Since there is heating throughout the atmosphere and net precipitation, this mode is associated with a single cell of mean upward vertical motion in the troposphere, although within that there are moist updrafts and downdrafts. There is a net upward flux of $\theta_E$, peaking in the mid-troposphere where the $(Q_1, -Q_R)$ and $Q_2$ curves cross. Mathematically, integrating over the troposphere [Yanai et al., 1973; Betts, 1978] and neglecting the surface sensible and latent heat fluxes, the surface precipitation flux $F_{pr}(0)$ is given by integrals of the convective source terms. For notational brevity, we use height rather than pressure co-ordinates and do not include density, which varies with height.
\[ LF_{PR}^{\infty}(0) = \int_0^{z_T}(Q_1 - Q_R)dz = \int_0^{z_T}Q_2dz \]

where \( z_T \) is a level high in the atmosphere, where the convective source terms are small. There is a net upward transport of \( \theta_e \) and moist static energy \( h \) associated with this mode since \( Q_1 - Q_R < Q_2 \) at low levels, given by [Yanai et al., 1973]

\[ (C_p T/\theta_E) F_{\theta_E}(z) = F(h(z)) = -\int_0^{z_T}(Q_1 - Q_2 - Q_R)dz \]

Arakawa and Chen, [1987] describe their third mode, as one which increases the separation of the \( Q_2, Q_1 - Q_R \) peaks in Figure 5a. This Mode 3 has been drawn on Figure 5a as a modulation of Mode 1, which increases the mid-tropospheric \( \theta_e \) flux, while having little impact on net precipitation. Thus this key diagnostic study shows that the upward \( \theta_e \) flux is not uniquely coupled to the precipitation. Conceptually one might perhaps associate a larger upward \( \theta_e \) flux with a dynamical structure feeding more low \( \theta_e \) air into the system in mid-levels.

![Figure 5 Modes of interaction of convective heating (Q - Q) and drying (Q), freely adapted from Arakawa and Chen [1987]. Panel (a) shows modes 1 and 3; panel (b) is mode 2.](image)

In terms of the feedback to the larger scale, the precipitation is important because it is associated with heating and a deep tropospheric ascent mode. The importance of the upward \( \theta_e \) flux is that this lowers boundary layer \( \theta_e \) [Betts, 1978], and increases the ocean surface \( \theta_e \) flux (primarily the moisture fluxes are involved). The importance of this process in regulating convection is discussed in detail by Raymond [1995, 1996].

The mode 2 in Figure 5b is described by Arakawa and Chen as the component representing deviations of "large-scale" condensation and evaporation, since \( (Q_1 - Q_R) = -Q_2 \). Note that, as we have drawn it in Figure 5b, it represents a condensation over evaporation couplet with no net precipitation and zero \( \theta_e \) flux (from Equations (3) and (4)). This diagnostically derived mode can be thought of as the signal coming from the variable presence of mesoscale anvils. The key consequence of this heating over cooling couplet (which in reality involves condensation and freezing over evaporation and melting)
is to force a 2-cell vertical structure with ascent over descent, and a larger-scale convergence in mid-
levels near the freezing level as seen in Figure 4.

In the light of this diagnostic study, which is consistent with our GATE September 2 GATE example
and those shown in Johnson [1984], I propose that the minimum requirements for a convective
parameterization scheme is whether it can represent these 3 modes correctly. (a) A deep convective
precipitating mode with an upward $\theta_e$ flux, not uniquely coupled to the precipitation (Modes 1 and
3), and (b) a heating/cooling couplet with no net precipitation and no $\theta_e$ flux (Mode 2).

A scheme would then need sufficient closures to be able to determine the magnitude of the net
precipitation, $\theta_e$ flux and the heating over cooling couplet; and perhaps their time evolution for an
evolving unresolved mesoscale convective system.

One immediate question is does the couplet Mode 2 have to be parameterized at all? If it is "large-
scale" precipitation, why can't the grid-scale processes handle it (provided there is an adequate
prognostic cloud-scheme being fed liquid and solid precipitation from the convective scale)? We are
approaching the heart of the so-called scale-interaction problem. What scales are well represented by
the large-scale model? It is clearly unreasonable to expect a hydrostatic global climate model with
a horizontal grid of 250 km to represent the mesoscale at all, but can a hydrostatic model with a 50
km grid develop a crude representation of a mesoscale anvil? The key test I would propose is whether
the mid-level convergence shown in Figure 4 develops in tropical convective systems in the model.
If not, I would argue it should be forced by parametrically representing the Mode 2 couplet. In section
4.3, a simple formulation will be proposed.

From an observational perspective, what we see in the GATE data is that while convective bands
initially develop in favorable regions of large-scale waves, and therefore might be regarded as a
response to large-scale destabilization, the subsequent convective and mesoscale developments control
the evolution of the mass field. In nature all the scales interact dynamically and can evolve together;
but the convective and mesoscales have shorter time scales than the 'large-scale'. Because we only
simulate the dynamics of the large-scale in our global models, the faster processes must be
parameterized. Just as a convection scheme, by introducing precipitation before saturation on the grid-
scale is reached, can change the phasing of large-scale dynamical development, so if we introduce a
parameterized mesoscale couplet forcing, this too will feed back on the large-scale model dynamics
sooner, than if we wait for grid-scale processes to reach saturation. Since we know that this inflow
at the freezing level is dynamically important in the tropics, it is likely that the impact of this Mode
2 parameterization will be significant.

3.5. Mass Flux Representation Of Deep Convective Updrafts And Downdrafts
There have been many discussions of the mass flux representation of cumulus transports following
[1987], Cheng [1989a,b] and Cheng and Yanai [1989] and many others. For this review only a
condensed summary will be given to illustrate the key issues. The subgrid-scale heating and drying
by convective updrafts and downdrafts can be written in bulk form as
We define the terms as follows: The updraft and downdraft mass fluxes are $M_u$, $M_d$, respectively, and they satisfy mass conservation equations

\[ \frac{\partial M_u}{\partial z} = \epsilon_u - \delta_u \]  

(7)

\[ -\frac{\partial M_d}{\partial z} = \epsilon_d - \delta_d \]  

(8)

where $\epsilon$, $\delta$ represent entrainment and detrainment rates. The bulk properties of the updraft are its liquid water static energy $s_{Lu}$ and total water $q_{Tu}$; the downdraft is assumed to have no cloud water, so its properties are $s_d$, $q_d$. The environmental mean with properties $\bar{s}$, $\bar{q}$ is also assumed unsaturated. $F_{Pr}(z)$ is the flux of precipitation, which is related to 3 terms

\[ -\frac{\partial F_{Pr}}{\partial z} = C_u - E_d + Q_M \]  

(9)

where $C_u$ is the fallout of precipitation from the updraft and $E_d$ is the evaporation of falling precipitation into the downdraft. Following Arakawa and Chen [1987] and Cheng and Yanai [1989], and the schematic in Figure 5b, we include, in addition to the convective terms, a mesoscale condensation/evaporation couplet term, $Q_m$, which is not linked to a mass circulation, and which has both zero moist static energy ($h$) flux and no net precipitation flux. Thus $Q_M$ satisfies

\[ Q_M = Q_{1M} = Q_{2M} \]  

(10a)

and

\[ \int_0^{z_T} Q_M dz = 0 \]  

(10b)

In (5) and (6) we have also not included for brevity the surface sensible and latent heat fluxes or any representation of ‘turbulent’ boundary layer fluxes.

From energy conservation in the updraft, which entrains and detains and condenses precipitation as it ascends, one can write the updraft budget equation

\[ \frac{\partial}{\partial z} (M_u s_u) = C_u + \epsilon_u \bar{s} - \delta_u s_{Lu} \]  

(11)

Similarly for the downdraft

\[ -\frac{\partial}{\partial z} (M_d s_d) = -E_d + \epsilon_d \bar{s} - \delta_d s_d \]  

(12)
where in both (11) and (12) it is assumed that air entrained into both updraft and downdraft has the properties of the mean environment. Substituting (11) and (12) in (5) and (6), using (7), (8), and (9), and rearranging, gives

\[ Q_1 - Q_R = M_c \frac{\partial \bar{s}}{\partial z} + \delta_u \left( s_{u} - \bar{s} \right) + \delta_d \left( s_d - \bar{s} \right) + Q_M \]  

(13)

\[ -Q_2 = M_c \frac{\partial \bar{q}}{\partial z} + \delta_u \left( q_{Tu} - \bar{q} \right) + \delta_d \left( q_d - \bar{q} \right) - Q_M \]  

(14)

where the net convective mass flux is (defining both \( M_u \) and \( M_d \) as positive)

\[ M_c = M_u - M_d \]  

(15)

The equations (13) and (14) are the heart of many diagnostic studies of convection, and the parametric mass flux representation. Many of the authors cited above have used a spectral representation of the convective transports, but the bulk representation here is sufficient to illustrate the key issues, since the magnitudes (and the mechanisms) of the entrainment and detrainment terms are poorly known. Note that the convective updraft and downdraft mass fluxes can be formally combined in the leading terms, which have often been described in the literature as "compensating" subsidence terms. They are the terms which represent the bulk heating and drying by deep convection. The physical reality is that the mean grid scale ascent, \( \bar{M} \), is simply carried upward on small scales as a net transport within convective towers [Riehl and Malkus, 1958], that is \( M_c = \bar{M} \), and the environmental motion between deep convective towers is small.

Cho [1977] recognized that if the convective outflows are considered to be at buoyancy equilibrium after evaporating any remaining cloud water (note that this depends on the precipitation parameterization), then one can formally drop the detrainment terms in (13) (if virtual temperature effects are neglected), and calculate an \( M_c \) from (also dropping the mesoscale term)

\[ Q_1 - Q_R = M_c \frac{\partial \bar{s}}{\partial z} \]  

(16)

It could be argued that this is a satisfactory treatment, except near the surface, where cold downdraft outflows cannot sink to buoyancy equilibrium. However Cho noted that even if \( M_c \) is calculated from (16), there is no equivalent condition to buoyancy equilibrium in the moisture budget. The moisture content of convective outflows must be determined in (14) to solve the parameterization problem, since it cannot be assumed that \( q_{Tu} = q_d = \bar{q} \). Downdraft outflows are typically unsaturated and must be modeled. The precipitation fallout determines the water content of updraft outflows and they too will not be saturated, after sinking to buoyancy equilibrium [Betts, 1982a].

The net vertical mass transport of the deep convective mode is directly related to the net precipitation. Integrating (16) gives

\[ \int_{0}^{z_T} (Q_1 - Q_R) dz = \int_{0}^{z_T} (M_c \frac{\partial \bar{s}}{\partial z}) dz \]  

(17)

Substituting (9) in (5) and integrating through the troposphere gives

177
since the convective fluxes disappear at the integration limits and the mesoscale term disappears using (10b). (If we had not dropped the surface sensible heat flux it would also be included here). \( F_{pr}(0) \) is again the surface precipitation flux. Thus (17) and (18) show that the deep mode with net precipitation is formally related to a net deep convective mass flux.

If we add (13) and (14), the mesoscale term, which has no \( h \) transport because of (10a), again disappears and we get

\[
Q_1 - Q_2 - Q_R = M \frac{\partial \bar{h}}{\partial z} + \delta_u h_u + \delta_d h_d
\]

(19)

where \( h_u = h_1 - \bar{h} \), and \( h_d = h_d - \bar{h} \). In Betts [1973b], the inflows and outflows were measured directly and the terms in (19), including separate updraft and downdraft mass fluxes were evaluated. As mentioned earlier, at low levels, where \( M_c \) is small, the convective source term is dominated by the outflows of unsaturated downdraft air (since it was assumed that the inflow to the updraft had environmental mean properties, only the outflows appear in (19)). In other diagnostic studies using sondes networks to derive \( Q_1 - Q_2 \) (\( Q_s \) is usually calculated), (19) cannot be inverted by itself as it contains 2 unknowns, the updraft and downdraft mass fluxes. Only If downdrafts are neglected, can (19) be immediately inverted, given a cloud model for \( h \), to derive a net convective mass flux \( M_c \) [Yanai et al. 1973]. In some other diagnostic studies, the ratio of downdraft to updraft mass fluxes has been prescribed [e.g. Johnson, 1976] to solve (19). Any two of (13), (14) and (19) can however be regarded as independent. Nitta [1977] showed that, if a precipitation parameterization is introduced to determine the remaining cloud water in \( s_{Lu} \) in (13), and the downdrafts are assumed to be saturated, then (13) and (19) can be solved simultaneously for both an updraft and downdraft mass flux.

The key conclusion is that the mass flux representation of convection is only adequate to the extent that the detrainment terms in (13) and (14) can be either neglected or calculated. In the middle troposphere, the vertical mass flux terms do dominate, but as pointed out in Betts [1973b], near the surface, the low level cooling and drying by convection depends essentially on downdraft outflows. Thus a satisfactory mass flux parameterization for deep convection must calculate both an updraft and a downdraft mass circulation and the properties of the outflows of both the updrafts and downdrafts. All these depend on microphysical and cloud-scale dynamical processes, so a satisfactory general solution has not yet been found.

Convective precipitation is linked to a convective mass flux (see (17 and (18)) but this alone does not determine the upward \( \theta_e \) flux, which is also linked to the environmental \( \theta_e \) structure, and the convective-scale dynamics. Combining (4) and (19), consider the integral to the freezing level \( z_f \) in the mid-troposphere

\[
(C/T^0)_{\theta_e}(z_f) = F_h(z_f) = \int_0^{z_f} M_c \frac{\partial \bar{h}}{\partial z} dz - \int_0^{z_f} \delta_u h_u dz - \int_0^{z_f} \delta_d h_d dz
\]

(20)
At the freezing level, $\overline{h}$ is typically a minimum, so below $z_p$, $\partial \overline{h}/\partial z$ is negative, and the first term is positive. The downdraft outflows have typically $h_d$ negative, so the third term is also positive. If we neglect updraft outflows in the lower troposphere (which means ignoring shallow clouds), it is clear that the strength of mid-tropospheric upward deep convective flux of $h$ (and $\theta_e$) is related to the net convective mass flux, the value of the $\overline{h}$ minimum, and the strength and properties of the downdrafts. Since downdrafts bring down mid-tropospheric low $\overline{h}$ air, a low value of mid-tropospheric $\overline{h}$ contributes in both terms to a larger upward $h$ flux. It is this upward flux of $h$ or $\theta_e$ which plays a key role in the convective interaction with the surface fluxes as discussed earlier (see also Raymond [1996]).

### 3.6. Diagnostic Retrieval of Mesoscale Source Terms

Arakawa and Chen [1987] and Cheng and Yanai [1989] discussed a method of extracting the mesoscale flux information from (13) and (14). They defined a parameter $H$, as follows, so as to eliminate the convective mass flux $M_c$.

$$
\frac{\partial \overline{h}}{\partial z} = \frac{Q_i - Q_R}{(-L \partial \overline{q}/\partial z)} - \frac{(Q_i - Q_R)}{\overline{\delta s}/\partial z} \\
= -\left( \frac{\delta u q_{Tu} + \delta_d q_d}{-L \partial \overline{q}/\partial z} \right) \cdot \overline{\delta s}/\partial z \\
\cdot Q_M \left( \frac{1}{-L \partial \overline{q}/\partial z} - \frac{1}{\overline{\delta s}/\partial z} \right)
$$

(21)

Cheng and Yanai argue that the detrainment terms are dominated by the detrainment of water from the updrafts ($\delta u q_{Tu}$); this term makes a negative contribution to $H$. In contrast the mesoscale term can be rearranged as

$$
\frac{Q_M \partial \overline{h}/\partial z}{(\overline{\delta s}/\partial z)(-L \partial \overline{q}/\partial z)}
$$

(22)

Since the typical mesoscale couplet has $Q_M$ positive in the upper troposphere where $\partial \overline{h}/\partial z$ is positive, and negative in the lower troposphere where $\partial \overline{h}/\partial z$ is negative, this mesoscale term is positive at all levels. Cheng and Yanai (1989) noticed that $H$ was positive during GATE convective cluster episodes, so the mesoscale term must dominate over the negative convective scale detrainment term in these cases. They used a cloud model to calculate the detrainment terms, and successfully estimated the mesoscale heating/cooling couplet. Figures 6a and 6b from Cheng and Yanai [1989] show the results of their diagnostic model at 1800 UTC for the GATE Case Study on Day 245, which we discussed earlier in section 2.3. Figure 6a shows a cross-section of what they called the mass flux in the cumulus environment.

$$
\overline{M} = \overline{M} - M_c
$$

(23)

where $\overline{M}$ is the observed mean vertical mass flux and $M_c$ is the net cumulus mass (updrafts and downdrafts) diagnosed by their model. In the presence of cloud clusters, the authors regarded $\overline{M}$ as a measure of the mesoscale mass flux. Figure 6b is the corresponding cross-section of the mesoscale convective heating term, $Q_M$, found by subtracting the convective contribution (diagnosed by their model) from the total $(Q_i - Q_R)$, diagnosed from the sonde budget analysis. The characteristic heating
over cooling couplet can be seen. Their longitude cross-section at 1800 UTC is along 8.5°N through the center of the GATE array. Note that Figure 6a shows mesoscale ascent over descent at 1800 UTC, while the total mass flux at that time in Figure 3 (at 22°W) does not show descent in the lower troposphere. The diagnostic analysis leading to Figure 6 has however removed the convective mass flux. Figure 3 only shows the ascent over descent pattern later in the lifecycle of the convective system, presumably after the convective circulations have decayed further. It would be useful if the Cheng and Yanai analysis could be repeated using other cloud models, since it is clear that their general conclusion is not cloud model dependent.

4. THE BETTS-MILLER SCHEME

4.1. Brief Review
I will not outline the details of the Betts-Miller scheme here, as they are adequately covered in the recent review by Betts and Miller, [1993], but the underlying concepts will be mentioned. It was clear in the decade after the GATE experiment, one of whose key objectives was to resolve the parameterization and “scale-interaction” problem [Betts, 1974], that we had not found a simple solution. Some advocated more detailed cloud models [Frank, 1983] with hierarchies of convective and mesoscale updrafts and downdrafts, but it was clear that the key issue of closure, the linking of all the submodels to each other and to large-scale parameters, was unresolved. The Betts-Miller scheme [Betts, 1986; Betts and Miller, 1986] was one response to this. It is an attempt to formulate the convective forcing in a very simple mathematical way, so that perhaps the coupling can be explored in some detail. The idea was lagged convective adjustment towards convective equilibrium profiles of T and q. Since we see convection in the tropics adjust the atmosphere towards quasi-equilibrium structures, can we not directly model this process, perhaps more easily than trying to get it as an outcome of complex convection sub-models.
I introduced three concepts:

a) The moist virtual adiabat (the reversible adiabat), rather than the pseudoadiabat, as a reference adiabat up to the freezing level. This was an inference from observations. The scheme adjusts towards a thermal reference profile, which has a specified instability in the lower troposphere defined with reference to this moist virtual adiabat.

b) The freezing level was built into the parameterization in calculating the quasi-equilibrium reference profiles, because observationally it appeared to be significant. This was a recognition that the freezing-melting process also plays a role in determining the characteristic thermal structure with a $\theta_{es}$ minimum.

c) The adjustment was lagged, to represent the response time of the convective and mesoscales to changes on the large-scale. This gives a smoothed convective feedback which seems physically more realistic than the on-off behavior of “instantaneous” convection schemes. In addition the mathematical structure allows the possibility of simplified analytical solutions. Indeed versions of this scheme have since proved useful in simplified tropical climate models [Seager and Zebiak, 1995, 1996].

The key idea here is that, while convection is occurring, the atmosphere is never allowed to get too far from the type of thermodynamic structures we observe. Even if we cannot adequately model the convective terms in detail, if we constrain a model in this way in the face of large-scale forcing, we are imposing the convective sources of heat and moisture that we would derive by diagnostic methods.

In our second paper [Betts and Miller, 1993], we introduced a unsaturated model downdraft circulation with its own adjustment time, based on a simple coupling of the evaporation into the downdraft to the net precipitation. This was an attempt to directly model unsaturated downdraft outflows into the boundary layer, and some improvements in the tropical climate resulted [Betts and Miller, 1993; Slingo et al. 1994]. In Betts and Miller, [1986], the adjustment near the surface was not well constrained.

Even a decade later, this lagged adjustment approach still has validity, as we have yet to answer key questions of closure, and find ways of linking the dynamical and thermodynamic transports on all the unresolved scales (from individual cells to say the 50 km scale) to the scales resolved by global models. Two further extensions will now be proposed.

4.2. Adjustment Time-scales for the Betts-Miller Scheme

Betts and Miller [1993] present no theory for determining their convective adjustment time ($\tau$), other than the empirical approach of setting $\tau$ short enough to make saturation on the grid scale infrequent. They recognized that this required smaller $\tau$ in higher resolution models as maximum values of grid-scale $\omega$ increase. However, it is now possible to propose a simple dynamical basis for $\tau$. Elsewhere in this volume, Mapes [1996] discusses the critical gravity-wave modes for the convection feedback to the large-scale. The first, a vertical wave-mode 1, spanning the deep troposphere has the fastest gravity wave propagation speed of 50 ms$^{-1}$ (Table 1), while the second wave-mode 2, the dipole with a node at the freezing level has a slower gravity wave speed $\sim$25 ms$^{-1}$. These are the two convective modes which interact rapidly with the larger scale flow. These are also the Modes 1 and 2 that we discussed in section 3.4, for which the diagnostic study of Arakawa and Chen [1987] showed
characteristic structure and transports. Consequently using the phase speeds for these modes, one can derive corresponding adjustment times for the Betts-Miller scheme. Table 1 shows values for three model grid lengths: 60 km, characteristic of the ECMWF forecast model (with spectral truncation of T-213); 120 km and 400 km, characteristic of a relatively low resolution climate model. This would give a formal basis for changing \( \tau \) with horizontal resolution, although again some empirical adjustment may be necessary.

<table>
<thead>
<tr>
<th>Wave-speed (m/s)</th>
<th>Adjustment time ( \tau )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(60 km)</td>
</tr>
<tr>
<td>Wave-Mode 1 (Deep Troposphere)</td>
<td>50</td>
</tr>
<tr>
<td>Wave-Mode 2 (Inflow at Freezing Level)</td>
<td>25</td>
</tr>
</tbody>
</table>

The adjustment time scales suggested by Betts and Miller [1993] (1 hr at T-106, corresponding roughly to 120 km horizontal resolution) lie in between the values given for the fast Mode 1 and the slower Mode 2, suggesting that their empirical approach had some merit. However this dynamical model for \( \tau \) does raise clearly a new issue. Gravity wave propagation can adjust quickly the thermal (buoyancy) structure on larger scales with these adjustment times. However it is not obvious that the moisture structure is adjusted on the same time scale, as is assumed in Betts and Miller [1993]. This relates back to the discussion in section 3.5. Vertical displacements by gravity waves can simulate the thermal mass transport model represented by equation (16), but in the equivalent moisture equation the “detrainment terms” cannot be neglected. This needs further study.

One unanswered question, which was raised in section 3.4, is whether convection parameterization schemes should have an explicit formulation of this wave-mode 2 forcing of the large-scale. One could argue that a global model with 60 km resolution and an explicit cloud scheme might possibly generate a marginal but adequate representation of the stratiform precipitation and this mesoscale couplet on the grid-scale. However without the non-hydrostatic mesoscale dynamics, it may not. This could be explored numerically since the key feature, the development of the strong mid-level inflow in the mature and decaying phase of a mesoscale system, should be visible. Here we will argue the reverse: namely that only by improving the representation of a mesoscale wave-mode 2 couplet, linked to the convective-scale, will the convective forcing of the large-scale be adequately represented in most if not all hydrostatic global models. In any event, this is a key scale separation issue, which needs exploring by a variety of techniques.

4.3. Explicit Parameterization Of Mesoscale Condensation-evaporation Couplet

We will illustrate this suggestion by proposing a simple extension to the Betts-Miller scheme (or any other convection scheme) to include a mesoscale precipitation-evaporation couplet. This is a further extension to Betts and Miller [1993], who proposed a formulation for an unsaturated downdraft
circulation, with its own time scale linked to the precipitation. This Mode 2 couplet has no \( \theta_E \) transport, and no net precipitation, but does redistribute enthalpy by means of condensation/freezing aloft above evaporation/melting below. In the tropics the freezing level is almost half the pressure depth of the deep convective layer, so we can then represent the convective mesoscale forcings shown schematically in Figure 5b, as a simple sine function in pressure coordinates

\[
Q_M = \frac{C_p}{M} \frac{\partial T_M}{\partial t} = \frac{-L \delta q_M}{\delta t} = \frac{\pi LP_M}{(p_o - p_T)} \sin[2\pi(p - p_T)/(p_o - p_T)]
\]  

(24)

where \( p_o \) is the surface pressure, and \( p_T \) is the top of the deep convective layer. This satisfies the constraints (10a) and (10b) (in pressure coordinates). We need to link this exchange to the net convective heating, so the simple closure is proposed

\[
P_M = \beta F_{PR}
\]

(25)

where \( F_{PR} \) is the surface precipitation rate, and the tunable parameter \( \beta \) is perhaps \( \approx 0.2 \) [Johnson, 1984]. The magnitude of \( \beta \) effectively couples the adjustment time and amplitude of this mode to that of the deep convective mode. This mesoscale mode parameterization would be imposed in addition to the convective parameterization. It does not affect the net precipitation or vertical \( \theta_E \) transport, but it does change the vertical redistribution of enthalpy and water. The issue of the lifecycle of mesoscale systems has still not been addressed. In climate models in which the resolved horizontal scale is larger, one might plausibly argue that within one model grid cell, large enough to include several convective mesosystems, that the simultaneous representation of convective and mesoscale components is adequate. In higher resolution forecast models, this may not be satisfactory and some means of representing convective system lifecycles may still be needed. Indeed since the evolution of many of the GATE mesoscale systems also appeared to be linked to the diurnal cycle, further work on this is needed.

5. SUMMARY

This paper summarizes some important concepts in the parameterization of shallow and deep convection in large-scale numerical models. Starting in the early 1970’s, diagnostic models have influenced the development of cumulus parameterizations, and field programs in the tropics, such as VIMHEX and GATE, showed the complexity of the life-cycle evolution of mesoscale systems with convective and mesoscale updrafts and downdrafts coupled to the condensation and evaporation process. It is important to appreciate the thermodynamic differences between nonprecipitating and precipitating convection. Shallow, non-precipitating convection can be modeled using a mass transport model, and any two independent conserved variables (or alternatively using convective adjustment to quasi-equilibrium convective structures). Ironically, despite its simplicity and its key role in controlling the surface fluxes both over the ocean and over land, we have yet to parameterize shallow convection satisfactorily in numerical models. In contrast, precipitating convection is much more complex and difficult to parameterize. Precipitation falls from updrafts and evaporate driving downdrafts, so \( \theta_E \) is not conserved. Only the \( \theta_E \) flux depends on the updraft and downdraft mass fluxes. Cloud microphysics control precipitation and in addition the downdraft thermodynamics is
poorly known as downdrafts remain unsaturated, with a subsaturation that depends on small-scale dynamical and microphysical balances. Consequently the $\theta_e$ flux is not tightly coupled to the enthalpy and water fluxes, which depend greatly on the precipitation flux, and its change with height through condensation and evaporation of falling precipitation. Diagnostic studies have shown that while the principal deep convective mode is tied to precipitation, and a deep tropospheric upward mass circulation, the upward $\theta_e$ flux which interacts strongly with the subcloud layer over the oceans is not uniquely coupled to the precipitation. In addition, diagnostic studies show there is a key second convective mode associated with the mesoscale anvil couplet of ascent over descent, which might be parameterized in the same way as large-scale condensation as a condensation-evaporation couplet without any $\theta_e$ transport or surface precipitation.

The concepts behind the Betts-Miller scheme are briefly reviewed and two extensions are proposed. One is a theoretical basis for the adjustment time scale, based on the gravity wave propagation speed of the two primary modes. The second is a suggestion that the mesoscale couplet mode be explicitly parameterized, and coupled to the surface precipitation. Although this does not address the cloud cluster life cycle issue, which may matter in high resolution global models or because of coupling to the diurnal cycle, it might provide new insight into how these two convective modes interact with the large-scale flow.

Acknowledgments
I was stimulated to write this work by the faculty and students of the NATO Advanced Study Institute at Kloster Seeon in August 1996. My research has been supported for many years by the National Science Foundation, currently under Grant ATM95-05018. I am grateful to Vic Ooyama for discussions in 1980, and to Steve Esbensen for retrieving the data in Figure 3.

References


BETTS, A.K.: THE PARAMETERIZATION OF DEEP CONVECTION: A REVIEW


