J. Adv. Model. Earth Syst., **Vol. 3**, Art. M06004, 20 pp.

Parameterizing Convective Organization to Escape the Entrainment Dilemma

Brian Mapes[1] *and Richard Neale*[2]

[1] Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida, USA
[2] National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000

Revised, JAMES February 2011

Lateral mixing parameters in buoyancy-driven deep convection schemes are among the most sensitive and important unknowns in atmosphere models. Unfortunately, there is not a true optimum value for plume mixing rate, but rather a dilemma or tradeoff: Excessive dilution of updrafts leads to unstable stratification bias in the mean state, while inadequate dilution allows deep convection to occur too easily, causing poor space and time distributions and variability. In this too-small parameter space, compromises are made based on competing metrics of model performance. We attempt to escape this "entrainment dilemma" by making bulk plume parameters (chiefly entrainment rate) depend on a new prognostic variable ("organization," *org*) meant to reflect the rectified effects of subgrid-scale structure in meteorological fields. We test an *org* scheme in the Community Atmosphere Model (CAM5) with a new unified shallow-deep convection scheme (UW-ens, a 2-plume version of the University of Washington scheme). Since buoyant ascent involves natural selection, subgrid structure makes convection systematically deeper and stronger than the pure unorganized case: plumes of average (or randomly sampled) air rising in the average environment. To reflect this, *org* is nonnegative, but we leave it dimensionless. A time scale characterizes its behavior (here ~3 h for a 2° model). Currently its source is rain evaporation, but other sources can be added easily. We also let *org* be horizontally transported by advection, as a mass-weighted mean over the convecting layer. Linear coefficients link *org* to a plume ensemble, which it assists via: 1) plume base warmth above the mean temperature 2) plume radius enhancement (reduced mixing), and 3) increased probability of overlap in a multi-plume scheme, where interactions benefit later generations (this part has only been implemented in an offline toy column model). Since rain evaporation is a source for *org*, it functions as a time-lagged but positive feedback on deep convection development. This evades the entrainment dilemma, since fully developed *org*–enhanced convection is not overly dilute, avoiding stability bias, while the pioneering updrafts of new convection are suppressed by entrainment, encouraging more large-scale variability.

DOI:10.1029/2011MS000042

1. Introduction

Coarse-grid atmosphere models like General Circulation Models (GCMs) call for some parameterization of water phase changes and vertical fluxes that occur in unresolved (subgrid scale) convective drafts [Arakawa, 2004]. An appealingly physical basis for treating these effects, while respecting the relevant conservation laws, is to use a closed mathematical model of convective drafts and the associated condensate microphysics. In such mass-flux convection schemes, realizable mixtures of the air available within a grid cell are lifted to different pressures in a provisional manner to test their buoyancy. The buoyancy is assumed to drive vertical air motions, whose net vertical mass flux is used to compute a self-consistent set of tendencies in the large-scale governing equations. The steady-state plume model is a sturdy classic. Advanced versions used here (based on Bretherton et al. 2004) include both entrainment and detrainment of an assumed spectrum of mixtures, based on buoyancy sorting principles, and a vertical velocity equation for the plume core which can be used to drive microphysical activation processes.

Such an advanced plume model is sometimes called a "bulk" plume, meaning that it represents the bulk or aggregate effects of an implied ensemble of convective elements or

To whom correspondence should be addressed.
Brian Mapes, Rosenstiel School of Marine and Atmopsheric Sciences (RSMAS), University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149-1098
mapes@miami.edu

parcels or mixtures. However, this implicit ensemble is rather rigidly specified (or implied), so an explicit ensemble of a few different plume computations offers greater flexibility and a wider range of behaviors. Obviously, multiple plume computations incur additional computational costs, which must be weighed among the interests driving the coarse-grid modeling activity in the first place.

The steady state assumption is a disagreeable aspect of these schemes: it undercuts our ability to constrain scheme assumptions and parameters with time-dependent observations or cloud-resolving simulations. However, the bookkeeping complications and hidden structural parameters involved in running time-dependent cloud models inside time-evolving large-scale models are too daunting to undertake.[1] For present purposes, then, we seek to improve the results from a traditional steady-state plume based mass flux scheme, by correcting for some of its inherent biases.

Plume models differ in many ways from the myriad events and processes that occur in fields of real, 'organized' convective clouds and storms. The number of ways we could extend a coarse model's representation of convection is far larger than our appetite for coding. Model improvement work must therefore prioritize the leading large-scale *functional* differences between plumes and reality. Moist convection has two main functions in the general circulation: 1. Ventillating moisture from the subcloud layer, which would otherwise saturate; and 2. Heating (stabilizing) the troposphere in the face of radiative cooling (or more generally destabilization, including surface water vapor flux into the subcloud layer). These two functions (somewhat overlapping) can be mapped roughly onto shallow vs. deep convection, a distinction which has been hard-wired into many global models in the form of separate schemes, sometimes even with incompatible conceptual bases. An explicit plume ensemble is attractive as a unified treatment (e.g. Arakawa and Schubert 1974), but if it doesn't allow interactions among members, it is merely a buoyancy-driven competition among plumes for instability. Naturally this favors the least dilute, and becomes essentially a deep convection scheme, unless handicaps by plume height are devised and tuned. The systematic spatial alignments and interactions that occur in convecting skies among boundary-layer updrafts, shallow cumuli, and deep convective cells thus fall in the cracks between blocks of model physics code.

Convecting patches of sky the size of model grid cells exhibit many important forms of ''organization,'' defined as subgrid structures, relationships, and processes with functional roles (i.e., systematic or net impacts on larger scales). Since further definition is difficult, it is perhaps clearest scientifically to contrast real convection vs. a purely ''unorganized'' baseline of utterly random convection. Conveniently for our engineering plans, plume schemes are often close to this unorganized ideal: average air is lifted in plumes through the average environment, which is rather like randomly selected air rising through randomly selected environments. Randomness also means independence, as embodied in models' typical independence of shallow and deep ensemble members, or separate shallow and deep schemes.

The GCM modeling community has a large litany of 'functional' errors that are thought to be probably linked to traditional random convection treatments. These take the form of systematic large-scale model biases, and unwanted trade-offs [e.g. Dai and Trenberth 2004, Bretherton et al. 2007, Kim et al. 2011]. In our view, the way forward is to try and correct these errors by adding the missing *functionality* of organization. A new scheme must respect large-scale conservation laws, and should be parsimonious, and as realistic as possible, to the (limited) extent that steady-state plume schemes with multiple internal assumptions can be meaningfully constrained by temporal process empiricism (from convection observations and simulations).

A few qualitative truths or principles about convection organization include:

1. Natural selection is a powerful biasing force: the most buoyant air rises preferentially. For example, random (or average-air) plume buoyancy would be systematically low, even if the plume's mixing and other assumptions were 'right.'
2. Rain-driven downdrafts chill the grid cell average boundary layer, but the chilled air specifically doesn't enter updrafts (by principle 1). Furthermore, the sinking of chilled air actively helps trigger warmer air to rise and overcome inhibition.
3. Gravity waves emanating from convection (and other variability) produce above-PBL density fluctuations that make buoyancy and convective inhibition differ from their values as defined in a mean sounding. Again, principle (1) tells us that convection will exploit these fluctuations in a systematic way.
4. Humidity in partly cloudy skies is inhomogeneous, so if new updrafts are launched preferentially in the vicinity of prior clouds (for example, by effect 2), they will systematically benefit by entraining moister than average air.
5. Dilution of updrafts is reduced when cloud-base updrafts become wider or less transient. Effect (2) makes wider, steadier cloud-base updrafts – which in turn can produce more rain, a systematically positive feedback.

In short, organization (nonrandomness in meteorological fields in convecting regions) affects convection in many ways, but the net effect is surely a systematic *positive* influence relative to random or mean plumes. Furthermore, it is a positive *feedback* on the deepening and development of precipitating convection. While the positive *influence* might lie within the slop or uncertainty of plume parameter cal-

[1] with the notable exception of 'super-parameterization' or multiscale modeling framework (MMF) approaches (Grabowski 200x, Randall et al. 2004), using full-physics cloud resolving models.

ibration or tuning, an unaccounted positive feedback calls for a new parameterization.

In eddy - mean state bookkeeping for convection [Ooyama 1971, Yanai et al. 1973; Arakawa and Schubert 1974], it makes no difference how a given set of small-scale vertical drafts might be spatially arranged within a large-scale horizontal grid cell. Indeed, scale truncation is precisely the sacrifice of such information. The effects of organization that we seek to capture are therefore expressible entirely in terms of the area coverage, net vertical mass flux and detrainment profiles of plumes within a given large-scale model grid column.

There are two aspects of convective organization that we are *not* trying to parameterize here. One is the impacts of precipitating stratiform anvil cloudiness connected to deep convection, especially in mesoscale convective systems [Houze, 1997; Houze, 2004]. Some GCMs put anvil processes inside a deep convection scheme [Donner et al., 2001] while other GCM studies conflate stratiform rain (as defined in the texture of observed radar echo fields) with the rain produced by a model's large-scale condensation or fractional-cloud schemes [e.g. Scinocca and McFarlane 2004, Rasch et al., 2006]. To sidestep these ambiguities, we focus here on traditional cumulus parameterization: convective updrafts, treated as steady plumes. Another form of organization we are not addressing is the quasi-two-dimensionality, steadiness, and tilt of convective drafts that can lead to special momentum flux effects, as emphasized in 'organized' convection parameterization discussions such as Moncrieff [2004].

Section 2 describes our GCM implementation, then Section 3 shows its effects on global simulations. Discussion follows in section 4. Appendix 1 describes a more elaborate organization-dependent multi-plume convection scheme, which we have only implemented in an interpreted-language model driven in single-column runs with steady destabilization.

2. An organization scheme

Before describing our particular, partial implementations, we first lay out our *org* parameterization ideas in more general terms, grounded in the conceptual basis of convection as comprising one or a few updraft plumes. A *plume* is a solution to a set of ordinary differential equations in height, governing the evolution (but in a steady-state sense: merely the profile) of a current of air ascending from plume base (a lower boundary condition). A plume is defined by its profiles of vertical mass flux M_{up},[kg m^{-2} s^{-1}) lateral mass flux M_{in} and M_{out} (called *entrainment* and *detrainment*), along with profiles of the thermodynamic properties of all these airflows. A profile of vertical velocity w, driven by buoyancy, may also be computed. This further implies a fractional horizontal area coverage $a = M_{up}/\rho w$, where ρ is density.

Relaxing the steady state or diagnostic assumption in a multidimensional and multivariate way is too difficult, as mentioned above. Instead we settle for a separation of variables type approach, with a single prognostic scalar *org* that contains all the time dependence, as in Pan and Randall (1998). *Org* should ideally be able to govern the vertical structure as well as the strength of plumes. Figure 1a shows schematically four ways a central org(t) variable might impact plume computations (4 boxes at right, bottom to top), paralleling the principles enumerated in Section 1: (1) preferential rising of warmer air in updrafts; (2) dynamical fluctuations affecting plume-base vertical velocity, convective inhibition, and thus the 'closure': the probability (areal coverage) of plumes reaching their Level of Free Convection; (3) a shift of the size spectrum of motions, toward wider plumes with reduced lateral mixing rates; and (4) preferential growth of convective updrafts in pre-moistened areas near prior plumes, if they are being triggered in spatial clusters. The net effect of these preferential or biased subprocesses will tend to be more and deeper convection (thick

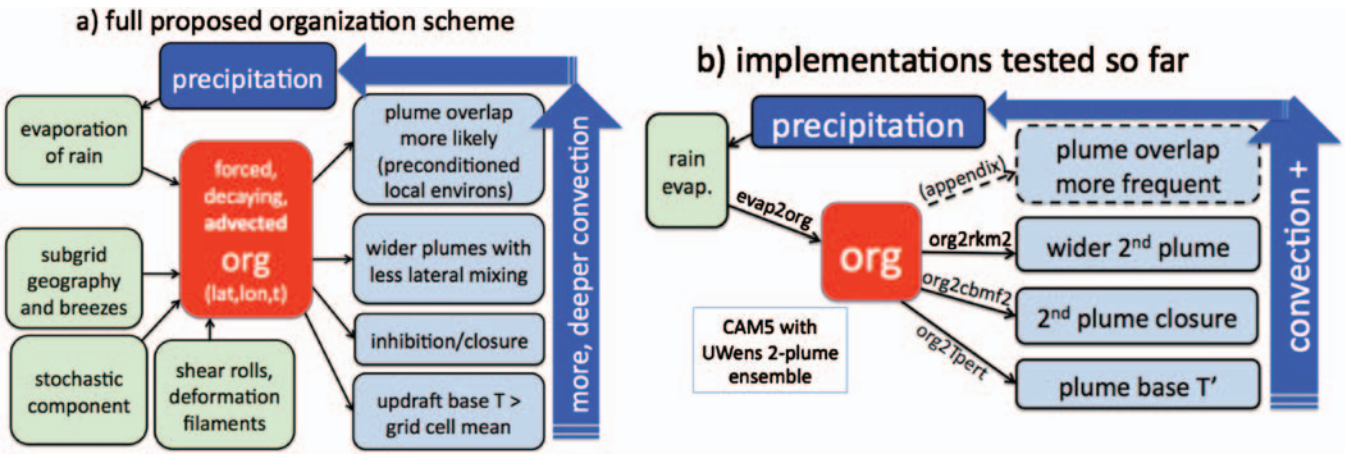

Figure 1. Flow charts of process interactions in (a) proposed and (b) implemented organization schemes for a coarse grid model with a plume treatment of convection.

blue arrows), leading to more precipitation and its re-evaporation in downdrafts, and thereby more organization.

The positive feedback between deep convection and *org* has a characteristic time scale τ, which we treat as a constant, ~3 hours for ~2 degree grid cells. This time scale choice is based loosely on visual impressions, time lapse movies, and model realizations such as Khairoutdinov and Randall (2006) of the diurnal organization and development process over land. It is a multicellular time scale, so if more physical elaboration were desired it could be expressed as a number (several to many) of parcel ascent time scales, which in turn could be linked back to the local instability and depth of the convecting layer. The common bias of too-early diurnal rain over land in climate models [Dai and Trenberth, 2004] is thus one obvious target for our organization scheme to ameliorate. Another is moisture sensitivity, which hinges on plume lateral mixing rate and is subject to the "dilemma" described in the Abstract.

Organization in convecting regions has other sources besides prior convection. Boxes in the lower-left corner of Fig. 1 represent three: geographically driven boundary layer structures like land–sea or mountain-valley breezes; internal atmospheric fine structures such as filaments due to deformation, or roll circulations driven by vertical shear; and stochastic sources or unpredictable artifacts of local mesoscale flow history.

2.1 *org's* budget and function in a CAM5 implementation

The arrows in Figure 1 represent relationships that we wish to capture or represent in model code. We chose to work with the Bretherton et al. (2004) plume model, as implemented in the Community Atmosphere Model (CAM5), where it is called the University of Washington (UW) "shallow" convection scheme (Park and Bretherton 2009, PB09). The UW scheme does produce precipitation, and also has ice thermodynamics, so its role as "shallow" scheme is not defined by its physics limitations, but rather mainly by its specified high lateral mixing rate (a z^{-1} dependence, equal to 700% of mass flux per km at 1 km altitude). We wish to work with a unified and self-consistent treatment of shallow and deep convection, allowing *org* to govern transitions and relationships between them, so we disabled CAM5's default deep convection scheme (Zhang and McFarlane 1995), and instead wrote explicit ensemble capability into the UW scheme. We call this modified scheme "UW-ens". Hohenegger and Bretherton (2011) describe another effort to use the scheme for both shallow and deep convection in an experimental version of CAM5, but with a single plume in each grid cell at each time step.

In our UW-ens scheme, an arbitrary number of plumes (hereafter, just two) are computed in sequence. Their fluxes of mass and conserved variables, and their area coverage, are combined to produce the total precipitation and all other convection-dependent fluxes and tendencies in the large-scale PDEs of the model. Different plumes in the ensemble may have different lateral mixing rates, cloud-base boundary conditions, or other internal parameters, and therefore reach different heights. In principle they can interact, although enforcing the constraints of conservation requires close bookkeeping in that case, so we only managed to implement this idea in an offline column model in a higher-level computer language (Appendix). We find that plume interaction effects are weaker than, and discouragingly dependent upon, very uncertain assumptions about lateral mixing and closure (the absolute amount of mass flux in each plume). For this reason, translating the plume interaction scheme into CAM's Fortran code basis was not deemed a priority, although the concept remains appealing because it could facilitate observational comparisons (Appendix).

In the code, each influence arrow linking the boxes in Fig. 1a is represented by a coefficient governing a simplest-possible assumed relationship. Arrows pointing out of the central "org" box are coefficients called org2xxx, to be read as "org to xxx", for example org2rkm where rkm is the code variable governing plume radius (lateral mixing rate). Arrows pointing in are called yyy2org, such as evap2org. Org is dimensionless, so org2xxx coefficients have units of the xxx variable (like m for plume radius). Coefficients named yyy2org are set up to yield an org production rate, with units of s^{-1}.

A prognostic, 2-dimensional *org* array was added to CAM5, accessed and used only within the UW-ens scheme's code file. A 3-dimensional passive tracer *org3D* was also introduced, for horizontal transport purposes, and had to be declared elsewhere in the model code and scripts. We allow org to be advected horizontally by a 'steering' flow that is the mass weighted mean over the convecting layer. Again, this assumption is based roughly on general observations and impressions of convective development and motion in midlatitude sheared flows. *Org* evolution occurs in the code by the following 3-step procedure at each time step: First, *org* is overwritten with the mass-weighted mean of *org3D* over the steering layer (defined as the surface to the highest plume top from the previous time step). After plumes are computed, *org* is subjected to local sources and linear decay, based on the forced-damped ordinary differential equation:

$$\frac{d(org)}{dt} = S - (org)/\tau \qquad (1)$$

where S is the sum of all sources (now just evap2org times rain evaporation rate), and τ is 10 ks (~3 hours). Our numerical scheme consists of relaxing *org* toward Sτ, the local instantaneous steady-state solution of (1), using the constant factor $\exp(-\Delta t/\tau)$ where Δt is the CAM time step. Finally, the *org3d* variable is overwritten at every altitude by this updated value of *org*. Since there is no vertical

gradient of *org3d* at this stage, advection and diffusion are purely horizontal as the model executes its next time step.

The value of *org* is dimensionless and arbitrary, so let us scale the coefficient evap2org to make *org's* typical (say, global mean) steady-state value Sτ be about 0.2. Since global mean rain evaporation is of order 1 mm/d, this is achieved by setting evap2org = 2 $m^2 kg^{-1}$ (units are clearer as s^{-1} org source per mm/s of rain evaporation rate). In grid cells with concentrated convective rain, org often reaches or exceeds 1. It is systematically larger over land, due to lower humidity and thus enhanced rain evaporation there.

Now that *org* has a budget, a timescale, a calculation method, and values of order unity, how shall we use it? Figure 1b shows the set of interactions in this study's CAM5 runs with the scheme. We focused on the simplest and most probably powerful effects: 1. temperature perturbations (Tpert) at plume base, representing self-selection of boundary-layer inhomogeneities (via coefficient org2Tpert), 2. cloud base mass flux in the second, wider plume (via coefficient org2cbmf2); and 3. the radius of plume 2 (rkm2, controlled via org2rkm2).

In our 2-plume UW-ens scheme with *org* effects, the first plume always has the default parameters of CAM5's UW shallow scheme (as described in PB09). The second plume has its lateral mixing rate divided by [1 + *org* x (org2rkm2)]. We will set org2rkm to values like 10, so that where org ~1 in significant rain the lateral mixing is reduced by about an order of magnitude. The second plume's cloud base area coverage (and thus mass flux, since w doesn't vary a lot at cloud base) is equal to the coverage by plume 1, multiplied by a factor [cbmf2 + *org* x (org2cbmf2)]. We sometimes set cbmf2 to 1 so that the second plume is always as active as the first, and/or set org2cbmf2 to values like 10 so that the second plume may have an order of magnitude more mass flux in rainy situations. The thermodynamic and w boundary conditions at the base of both plumes are taken to be the same.

The UW scheme closure is based on the ratio of convective inhibition (CIN) to boundary-layer turbulent kinetic energy (TKE). It embodies the assumption that the positive tail of a Gaussian distribution of w, i.e. the fractional area of the boundary layer where parcel w is large enough to overcome CIN, feeds into the base of moist plumes. It may seem that calling 2 plumes with this same area fraction is a "double counting" error. However, in time integrations, convective mass flux is closely governed by larger-scale budget constraints. Model CIN is a delicate, vertically subgrid quantity, with values of order 10 J/kg. Imperceptibly small adjustments in PBL state and/or low-level temperature are sufficient to adjust it and thereby enforce large-scale controls. Meanwhile, limiters in the code prevent nonsensical outcomes at individual locations and time steps. Thus double counting has no discernable impact, as we confirmed by experiment. If desired, one could back out an implied physical re-interpretation of our double use of the closure's area fraction, stating instead that we have adjusted one or more of the *ad hoc* physical assumptions underlying the scheme. For example, our approach could be expressed as a slight recasting (or *org* modulation of) the very debatable precise definition of CIN, or making mild adjustments to the specific assumptions that link bulk TKE to plume base w and area coverage (isotropic, Gaussian, etc.). All these physical deviations from default scheme assumptions surely occur in real convective weather, and their details and relative importance could perhaps be teased apart in cloud-resolving simulations or observations. However, in light of the large-scale governing process via immeasuarbly small mean-state adjustments (as discussed above), such efforts would have little impact on the model.

3. Tests in CAM5

We performed a set of experiments to test the effects of

1. disabling the Zhang-McFarlane deep convection scheme and introducing the two-plume (UW-ens) extension of the UW "shallow" convection scheme (experiment rkm0)
2. making the second plume in UW-ens undilute (rkminf)
3. introducing *org* dependence to the UW-ens scheme (exp. rkm10)

The results presented below culminate in showing how *org* effects ameliorate the entrainment dilemma. Advection-forced single column model (SCM) tests were a useful tool during code development, and suffice to show one or two points worth noting. All other results are from 5-year CAM5 runs, with a data ocean boundary condition consisting of climatological monthly sea surface temperature (SST) and sea ice conditions.

Table 1 shows parameter settings for the experiments described here. We deduced the effects of our changes by comparing a control run of CAM5 (with shallow and deep schemes) to experiments rkm0 (two identical UW "shallow" plumes only), rkminf (a UW shallow plus a second undilute plume with equal cloud base mass flux; strictly speaking, this experiment arose fortuitously from a code bug that made org values very large like 10^9 everywhere), and the org-dependent cases like rkm10.

3.1 SCM test of UW-ens scheme

To illustrate the division of labor that develops in UW-ens when two plumes of different radii (mixing rates) are computed, Figure 2 shows the time averaged mass fluxes from SCM experiments rkm0 and rkm10. Both were forced by identical advective terms, derived from GATE 1974 observations over the tropical Atlantic. With strong destabilization and abundant convection, this case illustrates how two independent plumes interact over time, via mean-state stability adjustments. Dashed blue curves show the identical mass flux profiles when the two independent "shallow"

Table 1. Experiment names and their parameter values. The second plume in UW-ens has its lateral mixing rate divided by [1 + *org* x (org2rkm2)], and its cloud base mass flux multiplied by [cbmf2 + *org* x (org2cbmf2)]. Plume base temperature in both plumes is augmented by adding [*org* x (org2Tpert)] K.

TABLE 1 *Experiment name*	org2rkm value	org2Tpert value	cbmf2, org2cbmf2 values
rkm0	0	0	1, 0
rkm10	10	0	1, 0
rkminf	very large	0	1, 0
rkm10+Tpert1	10	1	1, 0
rkm10+cbmf10	10	0	0, 10

plume calculations have identical parameters. This high-mixing-rate plume has to do all the jobs of both shallow and deep convection, so a severe unstable bias develops in this case (not shown). Red curves indicate that in the rkm10 eperiment, when the second plume has reduced lateral mixing (due to *org*), its enhanced buoyancy makes it behave more like deep convection: It reaches a higher top altitude, and sheds less mass flux through detrainment of negatively buoyant mixtures, making its mass flux profile more top-heavy on average (red curve in Fig. 1b). Meanwhile, the presence of this deeper, less-dilute plume 2 stabilizes the column, so that plume 1, with its large mixing rate, finds itself less buoyant and settles into the role of shallow convection (red curve in Fig. 1a).

3.2 GCM illustrations of the entrainment dilemma

To illustrate the entrainment dilemma, and how *org* parameterization evades it, we examine distributions of tropical precipitation (in space and time) and the mean stability of the tropical atmosphere, as the ZM scheme is switched off (experiment rkm0), an undilute second plume is added in UW-ens (rkminf), and finally an org-dependent second plume is allowed (rkm10).

Precipitation observations are shown in Fig. 3: a 22-year mean spatial pattern (from 1979–1998), and a 6-month time-longitude section at daily resolution (Jan-June 1997). Units are mm/d and color scales are held fixed in comparisons below. Figure 4 shows identically constructed plots for experiment rkm0 (top and lower left) and standard CAM5 (middle and lower right panels). In rkm0, the UW plume scheme with its high entrainment rate operates alone, so its ability to produce deep convection is limited to the moistest and most unstable conditions. As a result, precipitation is quite confined and concentrated in space and time (panels 4a, 4c). In fact, a substantial fraction of the rain in rkm0 is generated in the large-scale cloudiness scheme, not the convection scheme (not shown).

By contrast, most of the rain in standard CAM5 (Fig. 4b, 4d) comes from the Zhang-McFarlane (ZM) scheme, which has less mixing and greater buoyancy in its updrafts – but also less sensitivity to environmental conditions like dry air. As a result, CAM5's tropical rain field exhibits too little variability. This problem was worse in CAM3, but was ameliorated somewhat in CAM versions 3.5, 4 and 5 by increasing entrainment in the ZM closure, while compensating the resulting bias due to lost buoyancy somewhat by incorporating the latent heat of freezing (Neale et al. 2008).

Tropical instability may be gauged from Fig. 5, a mean profile of moist static energy (MSE) in a convectively active location (Truk Island in the west Pacific in January; other sites and months give generally similar results). In high-entrainment rkm0 (red dashed), the climatic balance between destabilization (cooling) and latent heating is only achieved when the upper troposphere is several degrees

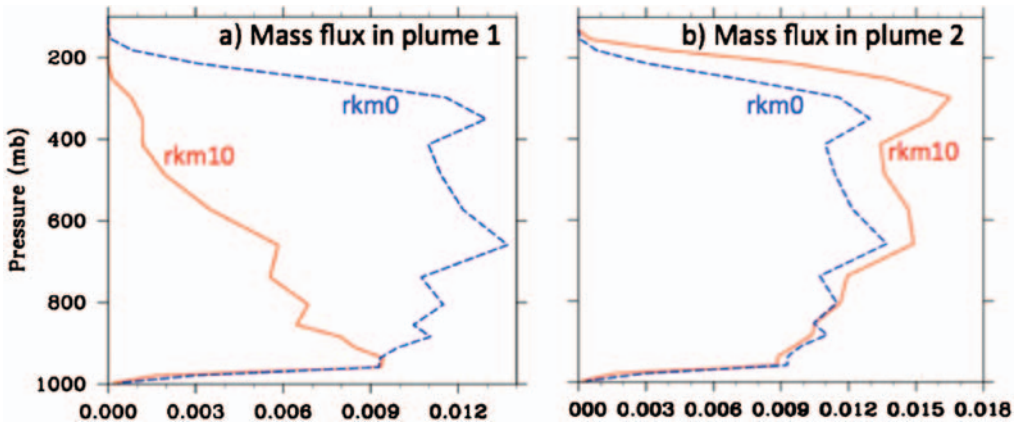

Figure 2. Time-mean mass flux profiles (kg m^{-2} s^{-1}) in SCM experiments rkm0 and rkm10, driven with advective forcings from GATE observations in the Tropical Atlantic.

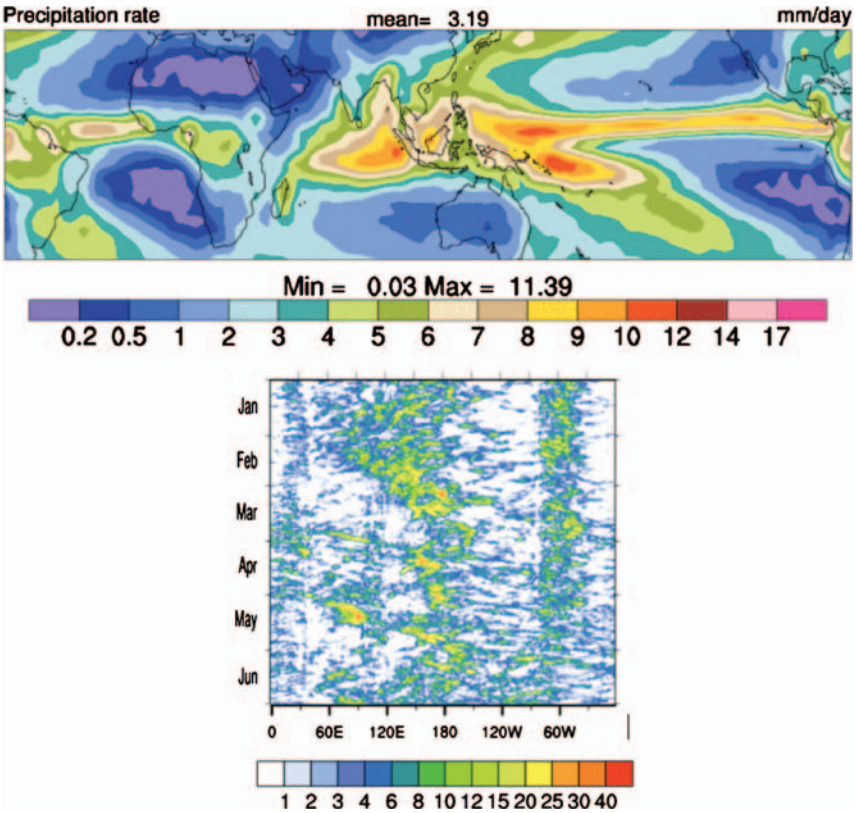

Figure 3. Observed precipitation. Top: Annual mean of CMAP data set (Xie and Arkin 1997). Bottom: Six-month daily tropical (15N-15S average) time-longitude section, (first 6 months of 1997 in Global Precipitation Climatology Project data).

cooler than observations (above about 400 mb, MSE differences are almost entirely due to T). Meanwhile the near-surface MSE is about right, over specified SST fields. We can already see that rain variability and mean state bias are linked by their dependence on mixing in updraft plumes. CAM5 has been extensively calibrated, and has upper troposphere T close to observations, since that is important for the realism of its jet streams and other mean climatic variables. However, its variability remains too weak as seen in Fig. 4d.

Allowing a second, undilute plume in our UW-ens scheme gives an extreme caricature of the problems arising from the ZM scheme (Fig. 6). The undilute plume rains throughout the conditionally unstable subtropics and tropics, essentially as soon as undilute parcel CAPE develops through surface latent heat flux. The rain field is thus more like the evaporation field: too flat in space and steady in time, and with a distinct minimum on the equator. Consistent with the reasoning above, the mean state has a severe stable bias in this case: upper-level MSE is about equal to near-surface values (although the latter exhibit a substantial low bias).

Figure 7 shows experiment rkm10, in which the second plume's entrainment rate is roughly an order of magnitude smaller wherever rainrate has been an order of magnitude greater than the global mean for more than a few hours, such that *org* has increased to be order 1. The mean state stability is very close to observations, yet the time-longitude diagram also exhibits variability whose magnitude is pretty similar to the observations of Fig. 3.

We hasten to admit that rkm10 does not reproduce the *character* of variability seen in observations. Detailed characterization is beyond the present scope, but some eyeball impressions can be mentioned. In high-entrainment rkm0, convection is tied strongly to humidity (along with large-scale condensation, not shown), so we speculate that the westward-moving variability seen in Fig. 4 is moist disturbances advected by easterly winds in the moist lower troposphere. In undilute-plume rkminf, there are weak eastward-moving rain waves in the Pacific, perhaps modulated by the low-level T field in Kelvin waves via the UW scheme's inhibition closure. Nature's broad spectrum of variability (Fig. 3), with sporadic, strong Madden-Julian Oscillation (MJO) events, remains very challenging for GCMs to simulate. Kim et al. (2011) discuss connections between MJO-like model variability and mean state biases in a suite of models, within a conceptual framework like ours that links these to convection-scheme restrictions and inhibitions.

Unfortunately, climatological spatial variance in mean rain maps like Fig. 7a is not a very good proxy for the space-time variance in time-longitude plots like Fig. 7b, due

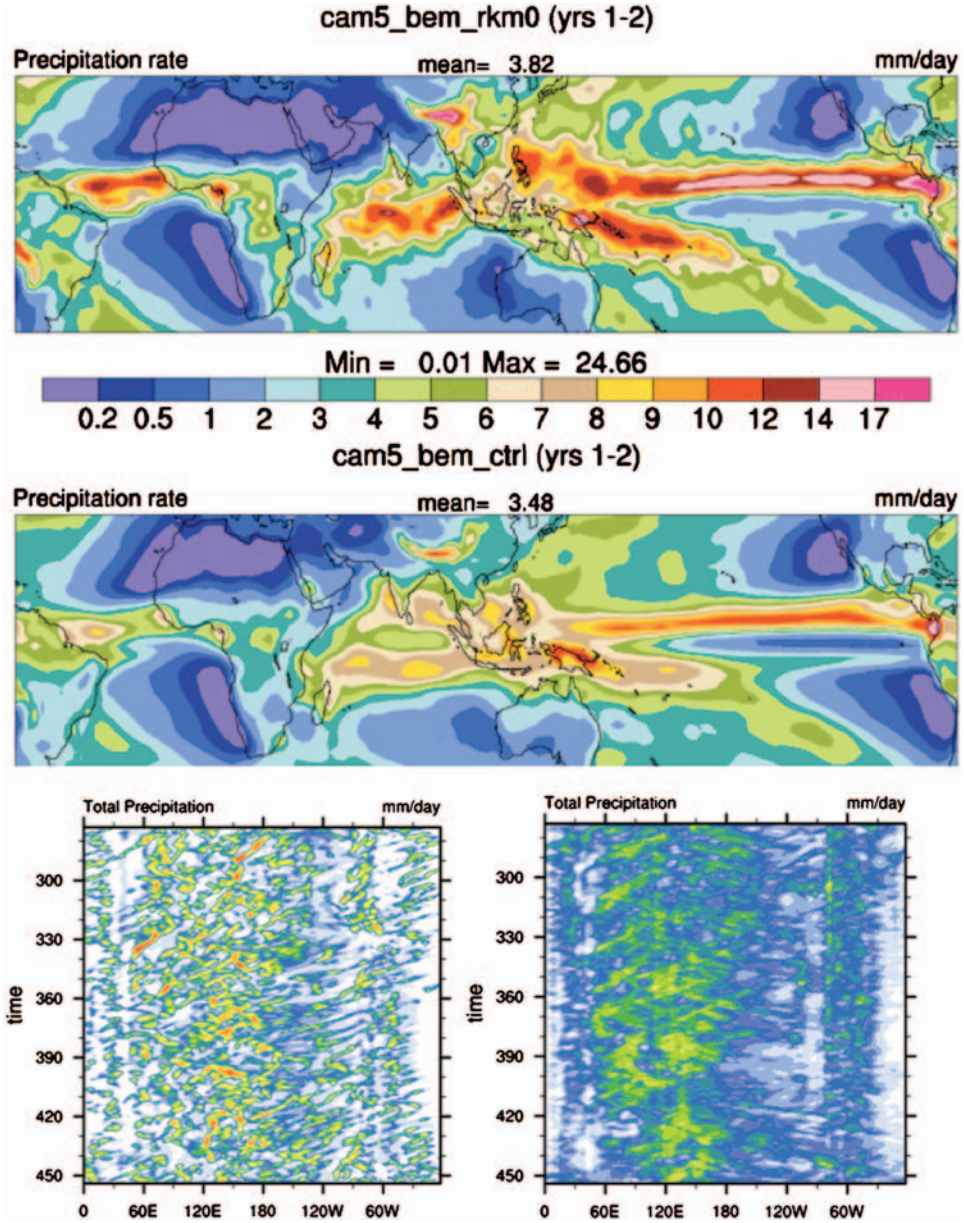

Figure 4. As in Fig. 3, from 2-year simulations of experiment rkm0 in which the Zhang-MacFarlane deep convection scheme is disabled (top and lower left panels), and standard CAM5 (middle and lower right).

to topographic features and different length climatologies. Our efforts to illustrate the ''entrainment dilemma'' across large libraries of runs of different CAM versions available on NCAR computers were not very conclusive (and are not shown), since we lack a clean way to measure variability from only monthly and seasonal data. A better joint suite of metrics for mean-state instability, transient variabiliy, and climatological rain-map structure could greatly extend studies like Kim et al. (2011) and this one, moving us toward a more robust understanding of convection-climate-weather interactions.

The *org* evolution timescale τ is short compared to most GCM rainfall variability, so quasi-equilibrium prevails and the org field broadly resembles its source, the rain field (with some evaporation enhancement over the drier continents). However, the small delay time τ may still be important in the development phase of convective events, for example in the timing of diurnal rainfall over warm land areas. Unfortunately, a subtle competition of mechanisms complicates diurnal timing in our simulations: Org peaks a few hours after rain evaporation, and facilitates deep convection, so it may indeed act to delay the daily rainfall peak. On the other hand, simulations with little or no org effect (like rkm5 or rkm0) have plumes with large entrainment rates, so diurnal rain may be delayed in those runs because greater instability must build up before such plumes can become

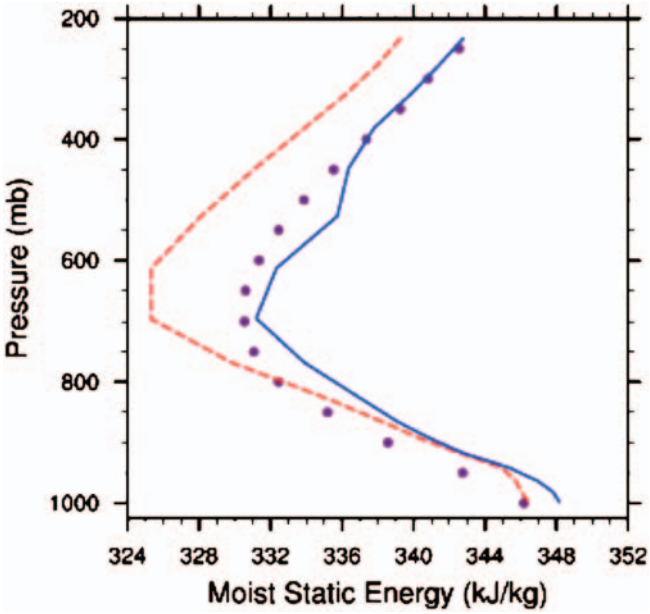

Figure 5. Mean moist static energy profiles in CAM5 (blue) and experiment rkm0 (red dashed), compared to observations (purple circles). Values shown are for Truk Island in January.

deep. Comparisons must therefore be devised carefully, and can never be fully clean since org has time mean as well as diurnally varying effects in the different experiments. After much effort, we found no figure compelling enough to emphasize by inclusion here.

4. Summary and Conclusions

We have introduced a new scheme to represent the systematic, functional (i.e., large-scale or rectified) effects of subgrid structures and relationships in convecting skies that are not reflected in typical GCM parameterizations. Our approach is parsimonious (minimal), and *corrective*: we are trying to advance the goals of large-scale modeling at its margin or frontier, not enumerate details of subgrid phenomena. For this reason, we tried to capture the net effect of many phenomena in one variable *org*, rather than undertaking physically elaborated treatments (e.g. Qian et al. 1998, Rio et al. 2010). As a result, our org related parameters, few as they may be, are admittedly difficult to address observationally. However, that is already the case for the other parameters in model convection schemes.

Most profoundly, what is the "right" effective lateral mixing rate for a steady-state bulk plume? Certainly not the large true tracer mixing rates of convective elements as defined in diagnostic studies with cumulus resolving models (Romps and Kuang 2010). In the face of such ambiguities, the heuristic nature of *org* is arguably the most honest approach. The "best" values of parameters for GCMs are those which make them perform best *at the large scales they are designed to simulate*. Optimization exercises at large scale, especially via data assimilation, are the way forward in this view. When the casting of a subgrid scheme produces a dilemma that pits model virtues against each other (like mean state bias vs variability), true optimization becomes impossible and metric dependent, leaving only compromise. If the problem can be recast to avoid the dilemma, at the cost of a couple more parameteris to optimize, that may constitute progress (time will tell). On the other hand, we are doing physical as opposed to statisitcal modeling, so schemes and parameters should represent something real.

We suspect that there may be further dilemmas inherent in a bulk plume with the same mixing rate governing both the mass flux profile and a single buoyancy profile assumed to drive the ascending motion. It is likely that a deeper recasting of convection treatments could be fruitful. The enduring generality of Ooyama 1971's "dispatcher function" for rising bubbles of air of various sizes has yet to be realized as a convection scheme. Casting convection in terms of a set of statistics of a much more numerous ensemble of time dependent entities, rather than a bulk plume, might open the way for process observations to contribute more. Organization parameters would probably still have a role in such statistical descriptions, again because organization is real: convective elements are systematically different from a set of random, independent samples of the air within a box of convecting sky.

This work has many precedents and influences in the literature. *Org*'s strategic purpose of making convection development more sensitive to dry air (see Derbyshire 2004 and others in that special issue), without overly diluting updrafts in active areas, has also been attempted by other means. For example, a one-line formula making entrainment rate proportional to subsaturation [Bechtold et al., 2008] has produced great improvements in European Center forecast model performance, albiet without physical justifications of even a qualitative nature. The prognostic convection scheme of Pan and Randall (1998) envisioned a "cumulus kinetic energy" (CKE) governed by an equation like (1), for each cloud type (plume top height), and with similar time constant (hours). Since kinetic energy budgets are not really closed even approximately within a local column, by cloud type, or in the vertical component, CKE is arguably just as heuristic as org, so the difference may be largely aesthetics of labeling. Another effort that informed and predated ours is Piriou et al. (2007), where a prognostic organization variable was used to reduce entrainment rate, corresponding to larger plume radii after rain develops.

In short, our work is a modest, 'corrective' extension of existing cumulus parameterization treatments, but we hope it opens fruitful avenues of thought, debate, and refinement. We do hope to genuinely improve a widely used global atmosphere model, although much work remains to optimize a version and demonstrate skill across multiple metrics of performance. Mixing at small scales is the hard problem at the heart of many large-scale geophysical flows, and

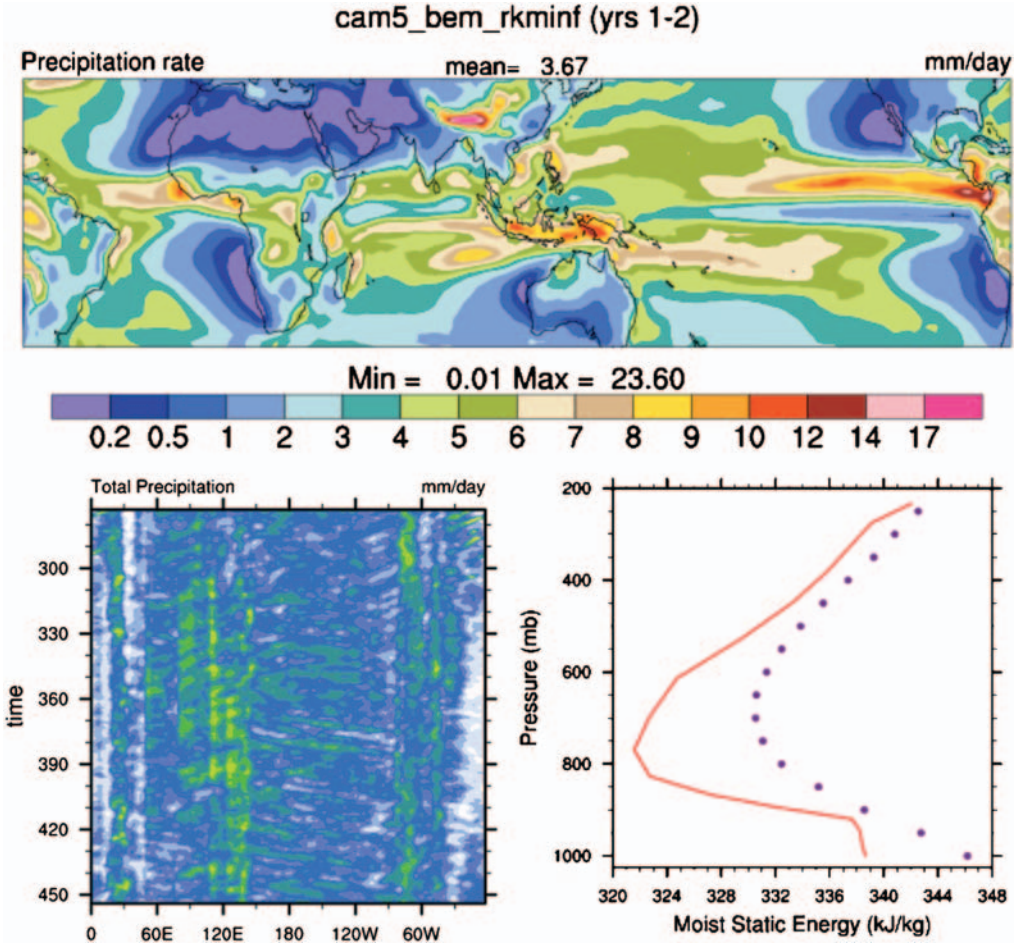

Figure 6. Panels as in Figs. 3–5, for the rkminf experiment in which the second plume has no entrainment and equal mass flux to the first.

remains a core difficulty here, but better appreciation of its influence on larger-scale weather and climate seems a worthy scientific goal, and many model applications studies suggest that there is still plenty of room for unambiguous improvements over current models.

Appendix 1: an organizable multi-plume ensemble

This section describes a more ambitious implementation of org dependence in a plume ensemble scheme: allowing for plume interactions. In particular, shallower initial plumes are allowed to assist the development of deeper plumes by detraining warm moist air into them, within a closed and conservative steady-state framework. Our plume code follows B04 and PB09, but is implemented in the Interactive Data Language (IDL), an interpreted language with implicit loops and integrated graphics. This IDL coding was thus much easier than Fortran coding within CAM. Our basic plume code is based on conservation of moist entropy s and total water mixing ratio q, on a 5 hPa pressure grid.

The B04 plume both entrains and detrains mass. At each altitude, a total mixing mass flux M_{mix} is specified as a fraction of updraft mass flux M_{up}. The mixed air is assumed to consist of a spectrum of sub-parcels with a broad (here, uniform) distribution of dilution (mixing fraction χ), from 0 (pure plume air) to 1 (pure environment). The more buoyant of these mixtures may be entrained, bringing some environmental air properties into the well-mixed plume updraft, while the less buoyant tend to be detrained, carrying some diluted plume air into the environment. The detrained or outflowing mass flux M_{out} has humidity q_{out} and entropy s_{out} that are in general greater than those of environmental air. In a traditional independent-plume calculation, such detrained air would merely have the effect of moistening the environment slightly for the next large-scale time step. The novelty here is to allow that detrained air to be entrained by another plume before it is mixed away. A plume that entrains preconditioned air will tend to rise higher. In fact, we terminate the multi-plume loop when the latest member doesn't achieve a higher top than its predecessor.

The question of organization almost asks itself: how much of the mass flux detrained by one generation of plume is mixed into the next generation? Time-lapse animations of

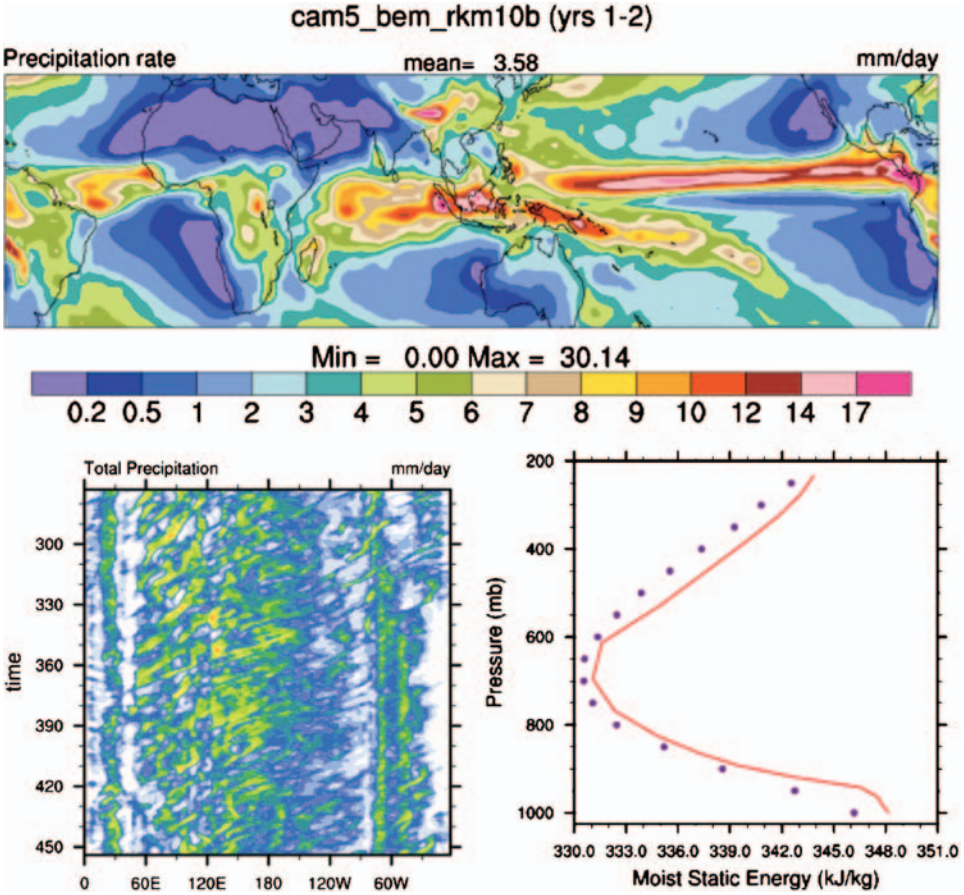

Figure 7. As in Fig. 6, for the rkm10 experiment in which the second plume has an org-dependent entrainment rate (see Table 1).

observed cumulus development suggest to us that such mass passing from one convective 'element' to another is an important effect, but a quantification using conserved quantities and accurately computed buoyancy is necessary to make more testable claims.

The code described here is in 3 modules. First, there is a plume subroutine, whose inputs are an environmental profile and cloud base conditions. Outputs are mass flux profiles in the well-mixed plume updraft, inflow, and outflow, along with the thermodynamic properties of those 3 airflows. Second, there is a multi-plume subroutine, which loops over 'generations' (the individual plume calls) and aggregates the ensemble's net fluxes. Third, there is a time-dependent single column model (SCM), which calls the multiplume ensemble every time step, after specified column cooling, surface fluxes, and a mixed-layer PBL scheme have acted in that order. The SCM integrates the large-scale average entropy s(p) and water mixing ratio q(p) profiles.

a. Module 1: the plume

The vertical coordinate is pressure at 5 mb spacing, with B04's grid staggering. Our codes for computing, mixing, and inverting entropy came from Raymond and Blyth (1992). Condensed water in the plume is capped at a fixed value (0.2 or 1 g/kg in this paper); any excess is tracked as precipitation. Precipitation is a sink of plume q, and of s due to condensate's (small) heat capacity. Freezing and melting effects, assumed to occur at equilibrium, are optionally included as entropy source/sink terms. Buoyancy is computed with water loading effects and is expressed as density temperature excess.

Plumes are solved by upward integration from a lowest or "launch" level. A lateral environmental mixing mass flux in each thermodynamic layer i is assumed, proportional to M_{up} on the interface at the base of the layer (also indexed as i):

$$M^i_{mix} = \varepsilon M^i_{up} dp \tag{2}$$

The mixtures produced add up to $2M_{mix}$ since an equal amount of plume and environmental mass are involved. These mixtures are partitioned into entrainment M_{in} and detrainment M_{out}. In B04, all mixtures that are buoyant, or rising fast enough to ascend more than a critical stopping distance, are entrained. A problem with that formulation for present purposes is that often nothing is detrained in layers where the plume is buoyant. To combat this, we force all mixtures composed of more than some tunable percentage (75%) of environmental air to detrain.

Entrained air is instantly mixed within the plume, and the resulting net M_{up} and averaged s_{up} and q_{up} are carried to the next interface (i+1), where the computation continues. In discrete form, the entrainment and detrainment satisfy:

$$M_{up}^{i+1} = M_{up}^{i} + M_{in}^{i} - M_{out}^{i} \tag{3a}$$

$$(sM)_{up}^{i+1} = (sM)_{up}^{i} + (sM)_{in}^{i} - (sM)_{out}^{i} + \Delta(sM)_{precip,ice} \tag{3b}$$

$$(qM)_{up}^{i+1} = (qM)_{up}^{i} + (qM)_{in}^{i} - (qM)_{out}^{i} + \Delta(qM)_{precip} \tag{3c}$$

In addition to its mass flux, the plume has a vertical velocity at interface levels satisfying a discrete version of:

$$\frac{Dw}{Dt} = w\frac{dw}{dz} = \frac{1}{2}\frac{dw^2}{dz} = B - b\varepsilon w^2 \tag{4}$$

The drag term in (4) has w^2 rather than w to avoid the mathematical and coding complications detailed in B04's Appendix. B in the thermodynamic layer midpoint (i) is used to increment this equation upward from the lower (i) to upper (i+1) interface levels. We follow B04 with b=2 as a crude representation of pressure drag (i.e., drag is treated as a doubled entrainment of environmental air with w=0).

Frequently, w^2 goes to zero at a 'crisis' level in an overshooting layer where B<0, even while M_{up} remains positive since detrainment mass loss cannot exceed the specified M_{mix}. In that case, a 'forced detrainment' term of required magnitude is computed, equal to M_{up} at the level of neutral buoyancy (LNB) divided by the pressure distance from the LNB to the crisis level. The code's loop index is then reset to the LNB, and the plume is re-integrated upward with this new forced detrainment term added to (3) above, overwriting the solution in this overshooting layer.

b. Module 2: The multiplume ensemble and use of *org*

The plume code above is called repeatedly to construct ensemble-total mass flux and detrainment for the large-scale tendency terms. Because plumes interact, yet must conserve entropy and water, the closure (which decides the *absolute* values of mass fluxes) is integral to the multiplume calling module. We can no longer rescale individual ensemble member mass fluxes separately, after the fact. The B04 and PB09 closure is well suited, since it is based on low-level controls of plume base mass flux and doesn't require results from the whole plume computation.

The first plume computation represents a 'pioneer' generation of updrafts which mix with the large-scale mean or environmental s(p) and q(p) in the traditional manner. It produces a detrained mass flux $M_{out}(p)$ and detrained entropy and water fluxes $(Ms)_{out}$ and $(Mq)_{out}$ in (3) above. To best justify computational expense, additional plumes should be as different as possible, but they should also represent an important amount of mass flux. These goals are in conflict: if the limited amount of preconditioned air $(Ms)_{out}$ and $(Mq)_{out}$ is mixed into a large next-generation plume mass flux, it will have little effect; yet if the next generation carries little mass, its fate is of little interest. After much consideration, the following closure strategy involving org was chosen.

The closure has two cases: the first generation, and subsequent generations. The cloud base mass flux of the first-generation plume is $M_b = \rho a_b w_b$, where ρ is density, a_b is the area coverage of plume bases, and w_b is plume base vertical velocity. Given w_b, we estimate the area fraction where boundary-layer w exceeds w_b by integrating over the upper tail of an assumed distribution of w. The distribution is characterized by a proto-plume kinetic energy (PKE), which we equate with a turbulent energy (TKE) in the subcloud layer following B04 and PB09. If we further assume isotropic turbulence, then PKE = TKE = 3 $\{w^2/2\}$, where brackets {} denote a PBL average. The standard deviation of w is thus $w_{ref} = (2TKE/3)^{1/2}$, and if we assume a Normal distribution then integrating over its positive tail yields ½ erfc(w_b/w_{ref}), which lies in [0, 0.5]. In summary, our closure is a_b = ½ erfc(w_b/w_{ref}) where $w_{ref} = (2TKE/3)^{1/2}$.

Setting plume base w_b is the next difficulty. How much of the PBL w distribution tail is able to leave the subcloud layer and enter the moist plumes? This should depend on a Convective Inhibition (CIN) energy, which needs to be computed from a lifted-parcel ascent, or ideally from the plume computation itself for consistency (although this might require iteration). For present purposes, we just fix w_b = 1 m/s. The calculations here all have TKE of order 1 m^2s^{-2}, so the erfc() function is safely in its useful ramp range, not in one of its flat tail regimes giving tiny or large (near 0.5) values of a_b.

PKE, the energy of "proto-plumes," should really involve only an upper tail of some *size* distribution of boundary layer puffs, not the full TKE. Isotropy and Gaussian w are already poor assumptions for the roots of convective clouds. The assumptions here are very crude, but suffice to demonstrate the main point: plume interactions. Plumes are also given a 1.5 g/kg moisture increment over the mean environment at their base level, following the findings of Fletcher and Bretherton (2010).

Higher generation plumes are a second steady-state plume computation giving a different result from the first, not simply a temporal 'crop' of additional updrafts that might have redundant fates with the pioneer plume computation. The closure for higher generations therefore depends on cloud field organization: If plumes are sparse and isolated, they will essentially all be the same, and can be represented by one generation. Only if they have very high fractional coverage, or enhanced probability of occurring in the same vicinity (organization) will additional generations with different outcomes be common enough to be worth computing.

For higher-generation plumes, we assume the same w_b = 1 m/s, which amounts to assuming the same CIN and cloud-roots PKE: these plumes not different in their cloud base characteristics, they are simply the lucky ones that happen to touch the previous-generation plume and entrain the air it detrained. The a_b closure for higher generation plumes then yields to simple spatial overlap probability reasoning: In a *random overlap* cloud field, the coverage of "lucky" or overlapping plumes of generation g+1 is proportional to the square of the area fraction of generation g. *Maximum overlap* would make the two area fractions equal. Org is here invoked as a clipped variable restricted to [0,1] that linearly slides between the random and maximum overlap limits:

$$a_{g+1} = a_g^2 + org \times \left(a_g - a_g^2\right) \quad (5)$$

The effect of *org* plays out as follows. When high *org* creates a large cloud base mass flux of generation g+1, and thus large demand for lateral mass inflow to it M_{mix}, the detrained mass flux M_{out} of generation g may not be enough to satisfy the demand, at least at some levels. The rest of the generation g+1 plume's M_{mix} demand must be filled with ambient environment air. Often the result will be a plume g+1 that is not very different from generation g. On the other hand, when *org* is very small, high-generation plumes have little coverage and mass flux, and can obtain all their specified M_{mix} in the form of preconditioned air from lower-generation plumes. However, their rarity gives them little large-scale impact. Somewhere in the middle, org has its biggest large-scale influence. We loop over plume generations until an additional plume call fails to produce a higher plume top height. In practice, this is usually just a few generations, even with 5 hPa vertical resolution where fairly small gains count.

The time has come to specify the most important parameter in convection schemes [Rougier et al., 2009]: lateral mixing rate (ε from Eq. 2). Classical plume reasoning, with $\varepsilon \sim 0.2/R$, might begin from plume base, with initial radius R chosen as plume base height and evolving upward based on M_{up} and w profiles. But for simplicity here, we chose the simpler fixed form $\varepsilon = c_0/(p_{sfc} - p)$, with $c_0 = 1$, similar to B04 and PB09. This decreasing ε with height is consistent with dilution inferred in cloud model studies [Lin and Arakawa, 1997; Kuang and Bretherton, 2006]. For a purely entraining plume, this form would make M_{up} linear in p, and seamless across the plume launch level p_{launch} if the convergence feeding the plume's base mass flux is uniform through the PBL. These are appealing properties, since horizontal divergence is subject to dynamical constraints that keep its profile somewhat smooth. The plumes do detrain some mass on their way up, making the plume mass convergence profile bottom-heavy rather than constant, but precipitation evaporation contributes a low-level cooling whose dynamical impact will add top-heaviness in the GCM setting. Constant convergence with height through the lower troposphere is a decent approximation to observed profiles in deep convection [Mapes and Lin, 2005; Mapes and Houze, 1995].

An example plume ensemble with org = 0.5 is shown in Fig. 8, for a mean tropical western Pacific sounding with its lowest 50 mb made into a well mixed layer. Here the plumes, launched at 950 hPa, have ambient T and 1.5 g/kg q excess. We have also assumed TKE =1 J/kg and w_b = 1 m/s. In this case, the ensemble has 8 plumes. The first (purple, thickest curves) travels only 3 grid levels to 935 mb before negative buoyancy terminates it, but it has the largest mass flux (panel g). The other plumes are a set of successively deeper versions of deep convection. Figure 7a shows the plume entropy profiles (colors) against the backdrop of environmental dry, moist, and saturated entropy (black curves), while panel (b) shows the mean entropy of outflowing mixtures at each level. Outflowing values are in between the plume and ambient values, just as they are in panel (c) showing outflow q with environmental and saturation values as the background black curves. Buoyancy (panel d) has a small negative or inhibition layer, before turning robustly positive for the deep plumes (although far short of an undiluted reference buoyancy, the dotted curve). Vertical velocity (e) lags its buoyancy forcing term in the vertical, allowing plumes to overshoot. The critical mixing fraction (panel f) for mixtures to either entrain into the plume or become outflow tends to swing between 0 and the maximum allowed (0.75), with a few noisy spikes in the second generation to intermediate values ($\sim$0.4) in the marginally buoyant layer around 750 hPa.

The mass flux of the first generation plume is greatest (Fig. 7g), while others have less according to the dictates of equation (5) with org=0.5. Inflows and outflows, expressed in units of horizontal flow divergence (panels h,i), illustrate the outer or environment-centric view of the plume activity. The increasingly energy-rich (panels b,c) detrained air is available in ever-smaller quantities (panels h,i), as each generation has about half the area coverage and mass flux of the previous.

The top panel of Fig. 8 displays data from Fig. 7 in a geometric depiction of the implied cloud field. Because both M_{up} and w are computed, plume width (area coverage) at each height can be derived, and depicted as a shape. Furthermore, since generation g plumes owe their existence to generation g-1, they must be displayed as touching. However, not all the mass flux of generation g-1 is required to support generation g, so the area coverage of generation g-1 can be broken into 2 parts: the part connected to generation g and the part unconnected to it. From these principles, one can work down from the highest generation and build an objective depiction of the convective cloud *systems* implied by the plume generations. Since the systems are independent, we display them spaced as far apart as possible within the grid cell. There are just two systems in Fig. 8a: a deep multi-generational system (depicted at 0.66), and some unused mass flux in the first generation shallow

plume (depicted at 0.33). In addition to the plume area coverage (solid fill, quadrupled for clarity, depicted to the left of each system's center line), we can compute how much area the plume's condensate outflow could saturate in a given time period (here 1 ks, about 20 minutes), based on the condensate loading of detrained air (here capped at 0.2 g/kg). That 'anvil' area is filled with hatching in Fig. 8, to the right of each system's centerline.

The effect of org, with the same sounding and all other parameters held constant, is illustrated in Fig. 8b. Here org = 0.1, so higher-generation plumes have much less likelihood (area coverage), although they do exist. In this case 4 cloud systems are implied in the 4 plume generations that emerge (a 5th generation rises no higher than generation 4 and thus terminates the succession). Because the higher-generation plumes have so little mass flux, none exhausts the availability of the detrained outflow from the previous generation. As a result, each generation has some mass flux that is not used by a higher generation, and thus there are 4 independent cloud system types implied by the 4 generations (or types) of plumes.

The case of Fig. 8a (org = 0.5) happens to be an optimum for this observed sounding, with 8 generations. The most obvious effect of org is the direct one – how rapidly cloud base mass flux decreases with generation, by (5) – but subtler relationships are also implied. When org is small, high-generation plumes have little mass flux, and so demand little mass for their lateral mixing M_{mix} via Eq. (2). They tend to reach higher altitudes quickly, but with little mass flux. When org is large (like >0.5 in this COARE sounding, not shown), generation g's demand for M_{mix} is large, so the limited amount of pre-moistened M_{out} from generation g-1

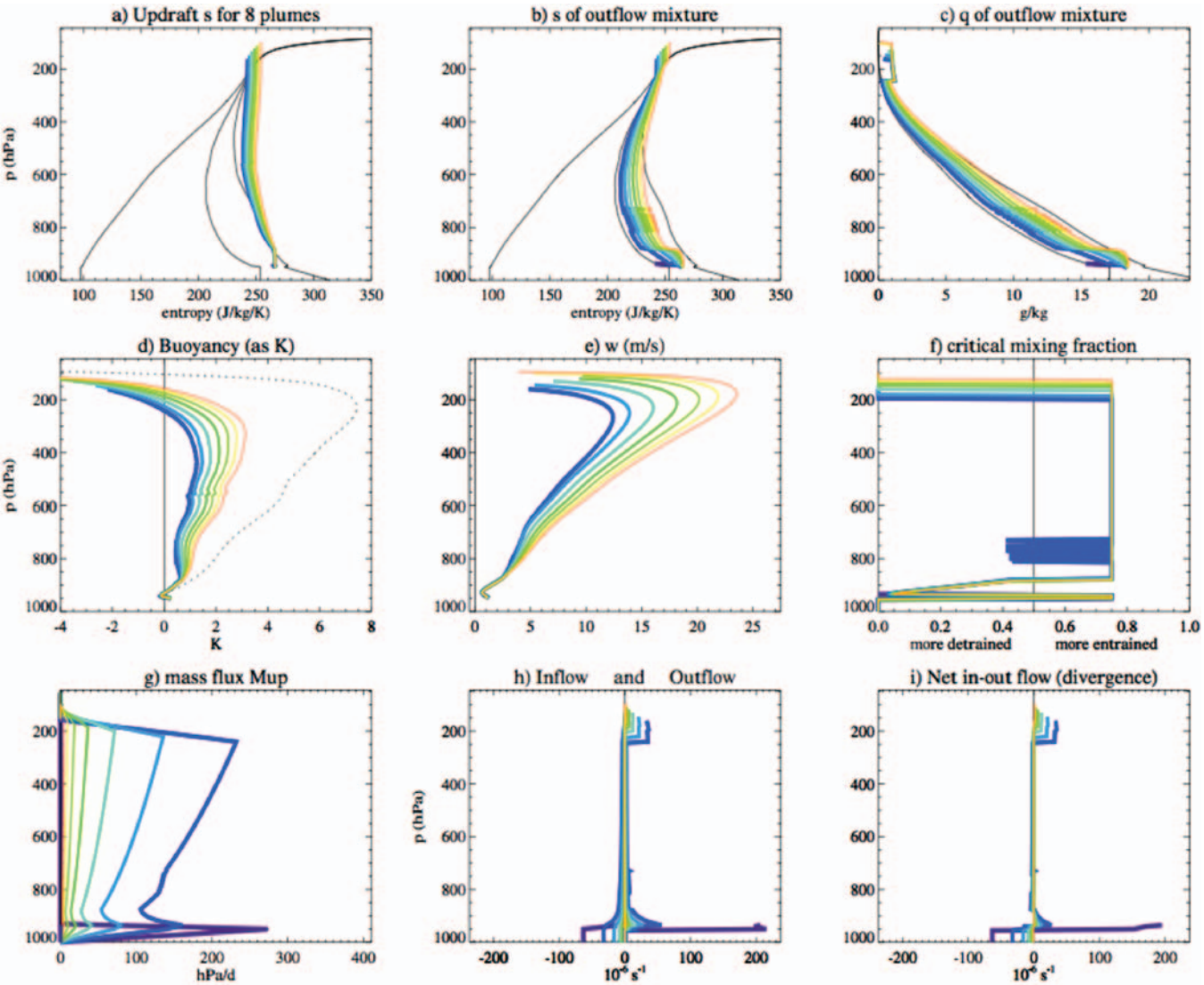

Figure 8. Plume characteristics in the offline multiplume scheme described in the Appendix, acting on a mean tropical western Pacific sounding. Each successive generation is depicted by a thinner line and a warmer color (purple to red).

has relatively little effect on g. As a result, plume deepening with generation doesn't occur effectively, and the succession is terminated fairly early.

Although the closure of each generation is decided during computation, and affects future generations, the whole set of mass fluxes can still be rescaled at the end of the process without violating conservation laws for s and q.

c. The SCM code calling multiple plumes

This section describes a minimal 1-dimensional time-dependent column model for testing the convection scheme above in a tropics-like setting. Destabilization is specified as a constant dT/dt cooling of 4 K/d up to about 250 mb, then smoothly tapered to 0. Compensating this cooling are surface fluxes of sensible and latent heat computed with bulk aerodynamic formulae

$$SHF = C_d \rho |V| C_p \times (SST - T_0) \tag{6a}$$

$$LHF = C_d \rho |V| L_c \times (q_{sat}(SST, p_0) - q_0) \tag{6b}$$

where the subscript 0 denotes surface air values. Here p_0 = 1000 hPa, C_d = 0.001, wind speed |V| = 10 m/s, and SST=28C.

These fluxes are deposited initially in the lowest thermodynamic level (a 5 hPa layer), where they cause large changes to s and q. The next process in a time step is surface-based dry mixing. A virtual potential temperature profile is computed, and its cumulative average up to all different heights is computed to represent the buoyancy at each level of a thoroughly mixed ascending unsaturated parcel rising off the surface. A dry convective energy profile is computed as the indefinite integral of this buoyancy over height. Mixed-layer top is defined as the highest level where the integral is positive. Below mixed layer top, s and q are then homogenized, and the dry convective energy is taken as a target TKE toward which the model TKE is relaxed with a 3 hour time constant.

With a cooled troposphere, a flux-driven mixed layer, an altitude from which to launch the plumes (mixed layer top), and a value for TKE, the multiplume convection subroutine described above can then be called. Some value of org is needed (to be discussed below). The scheme returns M_{up}, $(Ms)_{out}$ and $(Mq)_{out}$ which are sufficient to compute tendencies of s and q in the usual manner. The compensating subsidence term is computed by replacing each gridpoint's value with an upstream value from a location dp = (M_{up}/g)dt above the gridpoint in question. For numerical stability, if this distance exceeds the grid spacing (5 hPa), the entire set of plume fluxes is rescaled to safe levels. If this rescaling is required frequently, the time step dt should be shortened (10 or 5 minutes tend to be short enough).

Precipitation's source S_P is both in the plumes and (sometimes) in a post-convection large-scale process: when convective tendencies drive q above its saturation value, like via the detrained condensate seen at the top of Fig. 7c, the excess is treated as large-scale precipitation. Precipitation flux P(p) is determined by a downward integration of

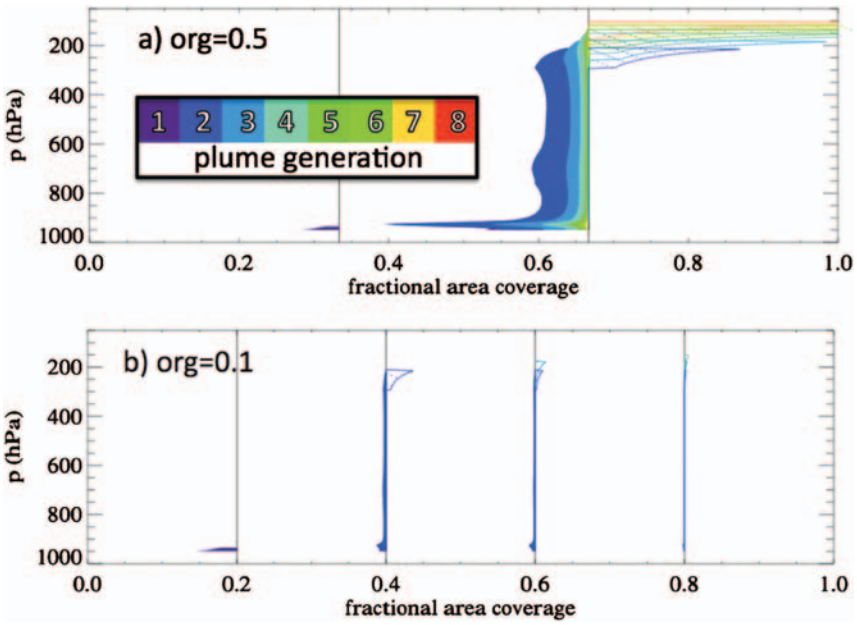

Figure 9. Area coverage by plumes (solid fill, x4 for clarity) and anvils (hatched fill: the fractional area that detraining condensate could just saturate in 1000 s). The two panels are for (a) org=0.5 and (b) org=0.1, with all other input profiles and parameters the same. The dependence of higher generation plume types on physical contact with the previous generation logically implies this smaller set of independent plume systems: two in panel a, four in panel b. These are depicted at equally spaced intervals within the 'grid cell' or domain [0,1].

$$\frac{dP}{dp} = S_P - f_{evap} P \qquad (7)$$

Evaporated precipitation is treated as a source of q at constant s. The evaporation fraction per hPa, f_{evap} in (7), is constructed as the product of 3 factors: (1) a T dependence of droplet evaporation from Rogers and Yau (1992, eq. 7.17), which steeply enhances evaporation in the lower troposphere; (2) a particle residence time (1/fallspeed, with fallspeed crudely set to differ for snow (1 m/s) and rain (6 m/s); and (3) subsaturation (1-RH). The factors (1-RH) and the precipitation production profile $S_P(p)$ are more variable, and undergo mutual adjustments that shape the mean state in climatic balance. The absolute value of f_{evap} was scaled to give about 50% evaporation for a reference calculation in the mean tropical sounding of Fig. 1. Large-scale precipitation encounters ambient RH, while convective precipitation encounters subsaturation (1-RH) that is a tunable fraction (here 0.5) of the ambient value. Equation (7) is integrated downward for each of the 2 precipitation types. No attempt has been made to account for convergence of eddy fluxes by evaporation-driven downdrafts. Rain evaporation drives *org* via Equation (1).

Starting from the COARE mean sounding of Figs. 7 and 8, this time dependent single-column model (SCM) was integrated with 300 s time steps for 20 days to produce Fig. 10. In the first few days, the system drifts: mixed-layer bursts of increasing depths occur as the troposphere cools and dries. Eventually the surface layer adjusts to give the large required

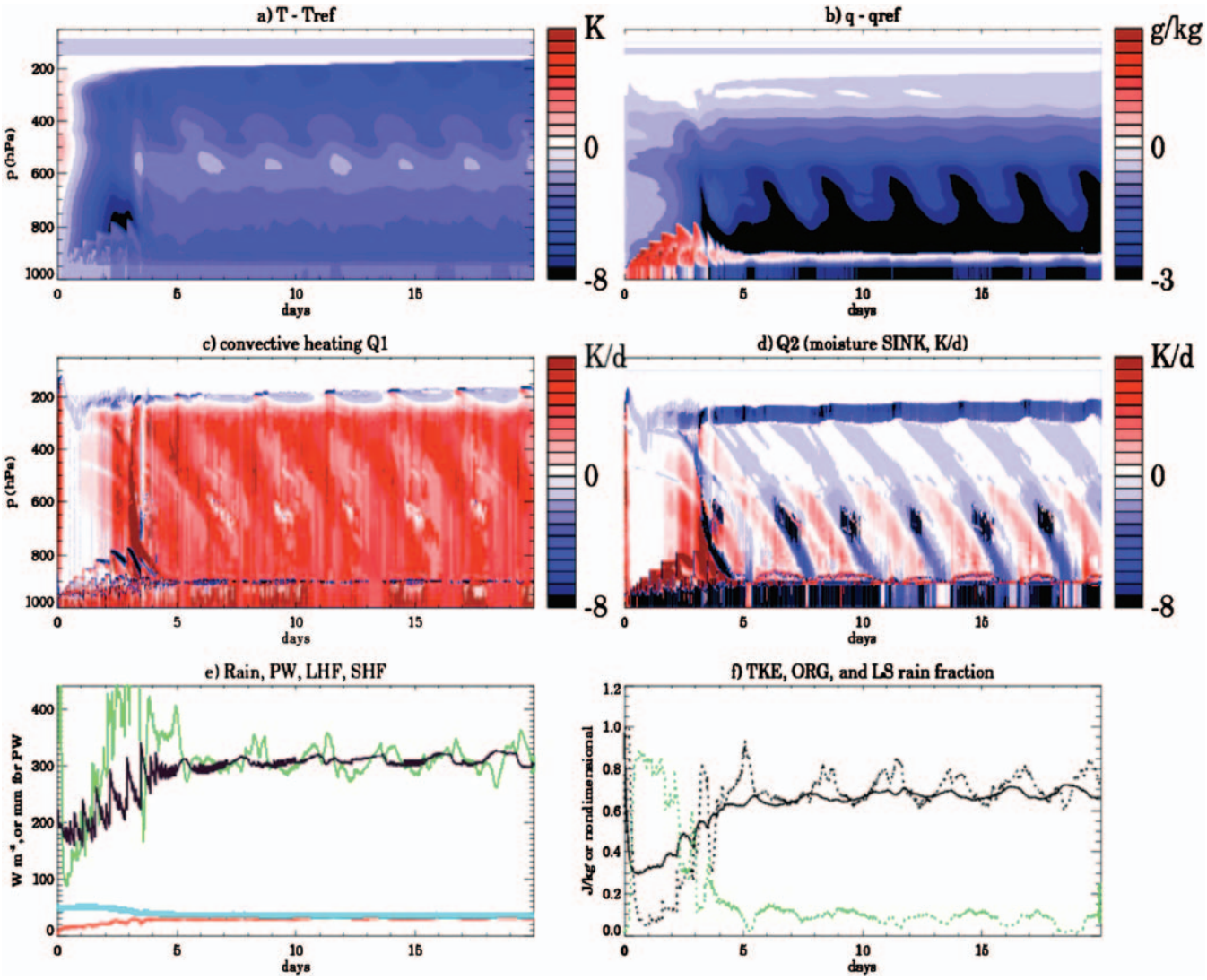

Figure 10. 20-day SCM series, starting from the observed sounding used for Fig. 1. (a) T minus its initial profile, (b) q minus initial, (c) convective heating rate, (d) convective drying Q2 (in K/d units), (e) Rain flux in energy units (green), surface evaporation in energy units (black), sensible flux (red), and precipitable water in mm (thick light blue), (e) TKE (J/kg, solid), org(dotted), and large scale rain fraction (green). All rain series were smoothed with a 5 h running mean except at the ends.

surface fluxes (> 300 W m-2) with the specified 10 m/s wind speed. Plume tendencies evolve to equilibrate with the tropospheric state, and then the system is quasi-steady. A mild ~3-day oscillation is seen, involving downward advection of anomalies through the troposphere by the compensating subsidence between plumes, along with some fast intermittent noise.

The SCM's quasi-steady tropospheric state is cooler (panel 10a) and drier (10b) than the initial state. A middle level near 500 hPa is the least cool-biased, presumably because coolness in the initial (observed) sounding (see saturated entropy in Fig. 8a) already makes plume buoyancy robustly positive there. This midlevel feature of observed soundings actually motivated this work 10 years ago: how can a middle-depth or congestus population of convective clouds (Johnson et al. 1999) be explained in light of such robust midlevel lifted-parcel buoyancy? Perhaps by protection from dry-air entrainment by neighboring clouds up to middle levels, then exposure to environmental mixing that can quench the buoyancy?

Convective heating Q1c (from the mixed layer plus multiplume convection scheme, Fig. 10c) and drying Q2 (Fig. 10d) show the 3-day downward-propagating features, amidst some noisy intermittency. Surface rainfall (green curves) is about 90% from convection and 10% from large-scale rain, the latter being fed by the 1 g/kg lofted condensate in plumes, detrained into the upper troposphere (200–300 hPa) where the specified cooling profile weakens with height. Surface fluxes are mostly latent (Fig. 10e), and TKE (in J/kg) and org both hover near 0.7 (Fig. 10f).

The 3-day oscillations occur in org along with other variables, but fixing org's value doesn't stop the oscillation (Fig. 11). Figure 11′s shorter time axis allows easier scrutiny of the oscillation, and colors are more balanced as the reference values for T and q are the mean sounding of Fig. 10′s last 16 days. One effect of fixed org in Fig. 11 is a cold 175 hPa level: its relative warmth in the control run is maintained by occasional overshooting convection facilitated by org fluctuations. Reducing org more drastically, dividing its production rate coefficient Eq. 1 by 3 (Fig. 12), also causes the upper troposphere to cool: more instability is needed for less-organized convection to do its climatic job. This effect is comparable in magnitude to disabling the latent heating of freezing in updrafts (not shown).

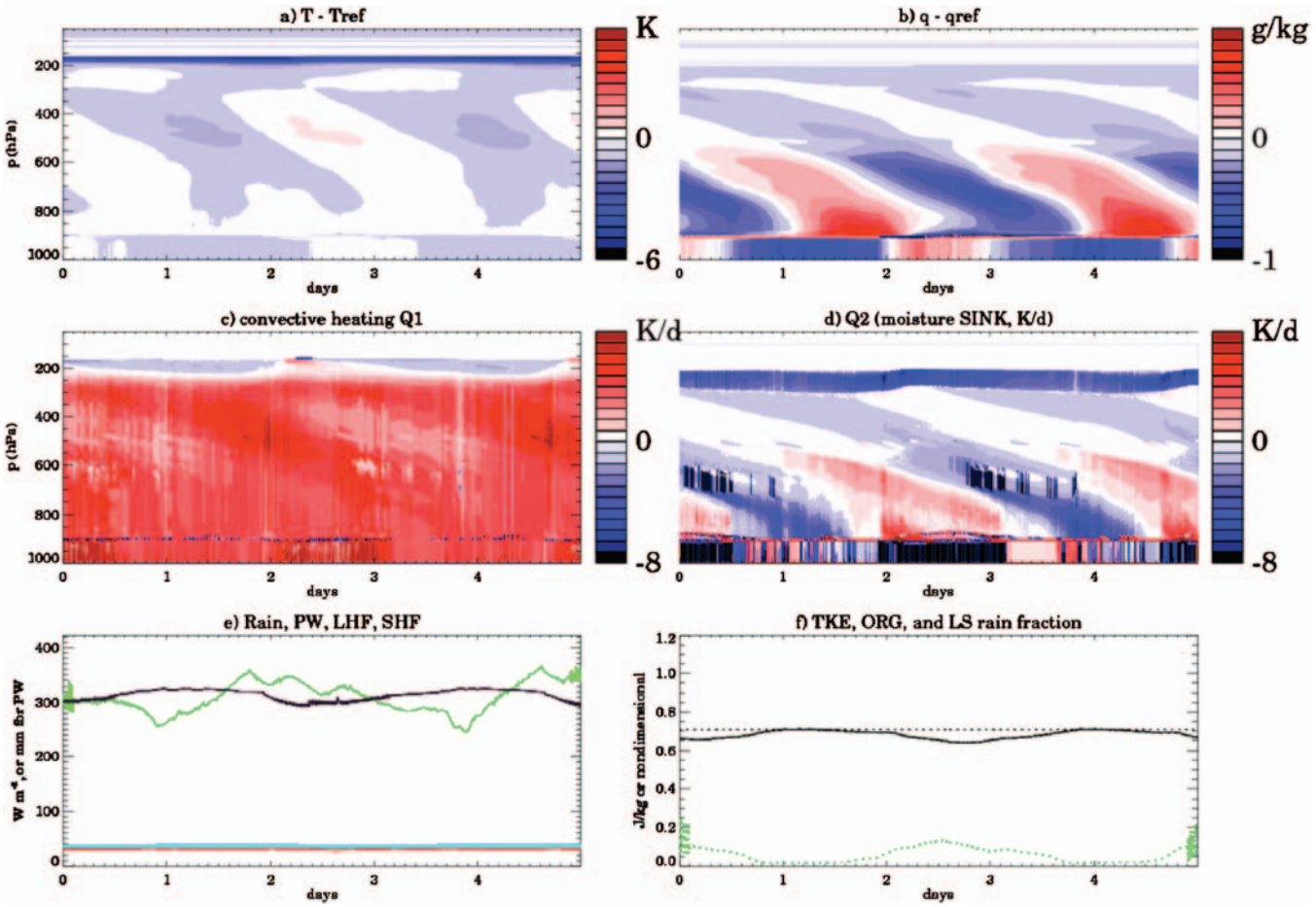

Figure 11. As in Fig. 10 but for the subsequent 5 days, with org fixed at 0.71 (its value at the end of Fig. 4). The reference profile for the top panels is mean T and q for the last 16 days of Fig. 10.

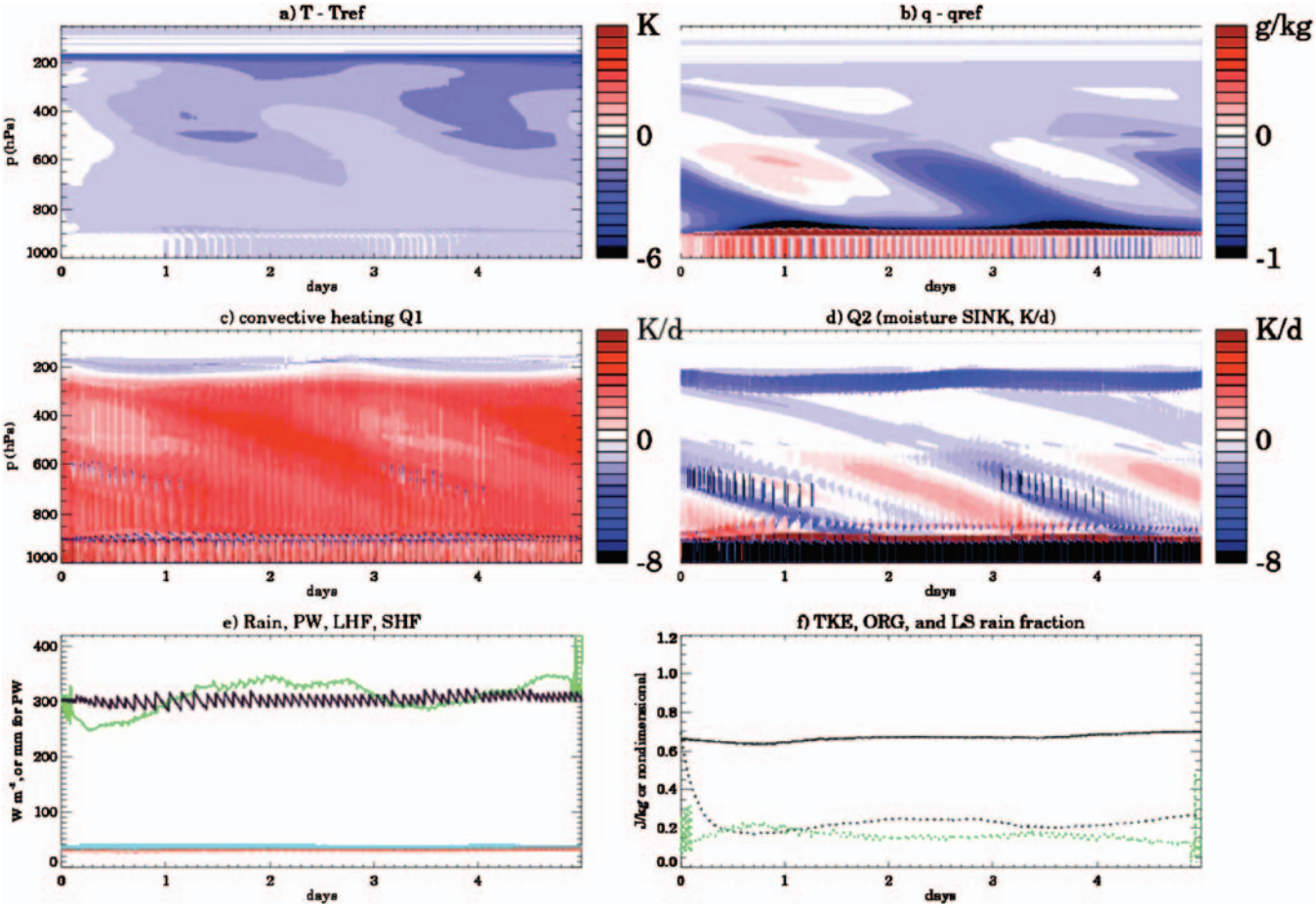

Figure 12. As in Fig. 11 but with variable org, with its source coefficient a reduced to 1/3 the control value used in Fig. 10.

The model desceibed in this Appendix shows a few main points: 1. Multiple-plume interactions can be formulated in a closed and rigorous manner, with org having a satisfying geometric interpretation as the probability of interaction (overlap). 2. Doing so is tantamount to parameterizing convection as an org-dependent set of plume *systems* (Fig. 8), which might facilitate observational comparisons, although the steady state assumption is still an obstacle. 3. Organization and its peak values in fluctuations are seen to play a rectified role in upper-tropospheric temperature budget, as in the main text. Unfortunately, porting these mechanisms to CAM Fortran code reliably proved too frustrating, due to the awkward nature of coding and diagnosis in that setting. Since these interaction effects are smaller than, and dependent upon, the uncertainties inherent in steady state plumes with specified mixing coefficients, the effort is not being pursued actively.

Acknowledgments: This material is based upon work supported by the National Science Foundation under grant ATM-0555796, and the Department of Energy Office of Science through grant DE-SC0000823. Visitor funds from NCAR's Climate and Global Dynamics division facilitate our ongoing collaboration. Thanks to Akio Arakawa, Chris Bretherton, Zhiming Kuang, and Sungsu Park for inspiration over time and on this project specifically.

References

Addis, R. P., M. Garstang, and G. D. Emmitt, 1984: Downdrafts from tropical oceanic cumuli. *Bound.-Layer Meteorol.*, **28**, 23–49, doi: 10.1007/BF00119455.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *J. Clim.*, **17**, 2493–2525, doi: 10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2.

Arakawa, A., W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part 1. *J. Atmos.Sci.*, **31**, 674–701, doi: 10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2.

Balaji V., and T. L. Clark, 1988: Scale selection in locally forced convective fields and the initiation of deep cumulus. *J. Atmos. Sci.*, **45**, 3188–3211, doi: 10.1175/1520-0469(1988)045<3188:SSILFC>2.0.CO;2.

Bechtold, P., M. Koehler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. R. Meteorol. Soc.* **134**, 1337–1351, doi: 10.1002/qj.289.

Bjerknes, J., 1938: Saturated-adiabatic ascent of air through a dry-adiabatically descending environment. *Quarterly Journal of the Royal Meteorological Society,* **64**, 325–330.

Bretherton, C. S., J. R. McCaa, and H. Grenier, 2004: A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: Description and 1D results. *Mon.Weather Rev.,* **132**, 864–882, doi: 10.1175/1520-0493(2004)132<0864:ANPFSC>2.0.CO;2.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785, doi: 10.1175/1520-0493(1996)124<1767:SOMCFB>2.0.CO;2.

Dai, A., K. E. Trenberth, 2004: The diurnal cycle and its depiction in the Community Climate System Model. *J. Clim.,* **17**, 930–951, doi: 10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2.

Derbyshire, S. H., I. Beau, P. Bechtold, J. Y. Grandpeix, J. M. Piriou, J. L. Redelsperger, and P. M. M. Soares, 2004: Sensitivity of moist convection to environmental humidity. *Q. J. R. Meteorol. Soc.* **130**, 3055–3079, doi: 10.1256/qj.03.130.

Donner, L. J., C. J. Seman, R. S. Hemler, and S. M. Fan, 2001: A cumulus parameterization including mass fluxes, convective vertical velocities, and mesoscale effects: Thermodynamic and hydrological aspects in a general circulation model. *J. Clim.,* **14**, 3444–3463, doi: 10.1175/1520-0442(2001)014<3444:ACPIMF>2.0.CO;2.

Emanuel, K. A., and M. Zivkovic-Rothman, 1999: Development and evaluation of a convection scheme for use in climate models. *J. Atmos. Sci.,* **56**, 1766–1782, doi: 10.1175/1520-0469(1999)056<1766:DAEOAC>2.0.CO;2.

Fletcher, J., C. S. Bretherton, 2010: Evaluating boundary-layer based mass flux closures using cloud-resolving model simulations of deep convection. *J. Atmos. Sci.*, **67**, 2212–2225, doi: 10.1175/2010JAS3328.1.

Fovell, R. G., G. L. Mullendore, and S. Kim, 2006: Discrete propagation in numerically simulated nocturnal squall lines. *Mon.Weather Rev.,* **134**, 3735–3752, doi: 10.1175/MWR3268.1.

Grabowski, W. W., M. W. Moncrieff, 2001: Large-scale organization of tropical convection in two-dimensional explicit numerical simulations. *Q. J. R. Meteorol. Soc.,* **127**, 445–468, doi: 10.1002/qj.49712757211.

Grabowski, W. W., M. W. Moncrieff, 2004: Moisture-convection feedback in the tropics. *Q. J. R. Meteorol. Soc.* **130**, 3081–3104, doi: 10.1256/qj.03.135.

Hohenegger, C., and C. S. Bretherton, 2011: Simulating deep convection with a shallow convection scheme. *Atmos. Chem. Phys. Discuss.,* **11**, 8385–8430.

Houze, R. A., 1997: Stratiform precipitation in regions of convection: A meteorological paradox? *Bull. Am. Meteorol. Soc.,* **78**, 2179–2196, doi: 10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2.

Houze, R. A., 2004: Mesoscale convective systems. *Rev. Geophys.* **42**, RG4003. doi: 10.1029/2004RG000150.

Johnson, R. H., T. M. Rickenbach, S. A. Rutledge, P. E. Ciesielski, and W.H. Schubert, 1999: Trimodal characteristics of tropical convection. *J. Climate*, **12**, 2397–2418, doi: 10.1175/1520-0442(1999)012<2397:TCOTC>2.0.CO;2.

Khairoutdinov, M., D. Randall, 2006: High-resolution simulation of shallow-to-deep convection transition over land. *J. Atmos. Sci.,* **63**, 3421–3436, doi: 10.1175/JAS3810.1.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev.Geophys.* **47**, RG2003, doi: 10.1029/2008RG000266.

Kuang, Z., C. S. Bretherton, 2006: A mass-flux scheme view of a high-resolution simulation of a transition from shallow to deep cumulus convection. *J. Atmos. Sci.,* **63**, 1895–1909, doi: 10.1175/JAS3723.1.

Kuang, Z., 2009: Linear response functions of a cumulus ensemble to temperature and moisture perturbations and implication to the dynamics of convectively coupled waves. *J. Atmos. Sci.* in press.

Lac, C., J. P. Lafore, and J. L. Redelsperger, 2002: Role of gravity waves in triggering deep convection during TOGA COARE. *J. Atmos. Sci.*, **59**, 1293–1316, doi: 10.1175/1520-0469(2002)059<1293:ROGWIT>2.0.CO;2.

Lilly, D. K., 1959: On the theory of disturbances in a conditionally unstable atmosphere. *Journal of Geophysical Research,* **64**, 1114–1114.

Lin, C. C., A. Arakawa, 1997: The macroscopic entrainment processes of simulated cumulus ensemble .2. Testing the entraining-plume model. *J. Atmos. Sci.,* **54**, 1044–1053, doi: 10.1175/1520-0469(1997)054<1044:TMEPOS>2.0.CO;2.

Mapes, B., S. Tulich, J. Lin, and P. Zuidema, 2006: The mesoscale convection life cycle: Building block or prototype for large-scale tropical waves? *Dyn. Atmos. Oceans,* **42**, 3–29, doi: 10.1016/j.dynatmoce.2006.03.003.

Mapes, B. E., 2000: Convective inhibition, subgrid-scale triggering energy, and stratiform instability in a toy tropical wave model. *J. Atmos. Sci.,* **57**, 1515–1535, doi: 10.1175/1520-0469(2000)057<1515:CISSTE>2.0.CO;2.

Mapes, B. E., R. A. Houze, 1995: Diabatic divergence profiles in western pacific mesoscale convective systems. *J. Atmos. Sci.,* **52**, 1807–1828, doi: 10.1175/1520-0469(1995)052<1807:DDPIWP>2.0.CO;2.

Mapes, B. E., J. L. Lin, 2005: Doppler radar observations of mesoscale wind divergence in regions of tropical convection. *Mon. Weather Rev.,* **133**, 1808–1824, doi: 10.1175/MWR2941.1.

Moncrieff, M. W., 2004: Analytic representation of the large-scale organization of tropical convection. *J. Atmos. Sci.,* **61**, 1521–1538, doi: 10.1175/1520-0469(2004)061<1521:AROTLO>2.0.CO;2.

Neale, R. B., J. H. Richter and M. Jochum. 2008: The impact of convection on ENSO: From a delayed oscillator to a

series of events. *Journal of Climate*, **21**, 5904–5924, doi: 10.1175/2008JCLI2244.1.

Ooyama, K. V., 1971: A theory of parameterization of cumulus convection. *J. Meteor. Soc. Japan*, **9** (special issue), 744–756.

Pan, D. M., D. A. Randall, 1998: A cumulus parametrization with a prognostic closure. *Q. J. R. Meteorol. Soc.,* **124**, 949–981.

Park, S., C. S. Bretherton, 2009: The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the Community Atmosphere Model. *J. Clim.,* **22**, 3449–3469, doi: 10.1175/2008JCLI2557.1.

Peng, M. S., J. A. Ridout, and T. F. Hogan, 2004: Recent modifications of the emanuel convective scheme in the navy operational global atmospheric prediction system. *Mon. Wea. Rev.*, **132**, 1254–1268, doi: 10.1175/1520-0493(2004)132<1254:RMOTEC>2.0.CO;2.

Piriou, J., J. Redelsperger, J. Geleyn, J. Lafore, and F. Guichard, 2007: An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations. *J. Atmos. Sci.,* **64**, 4127–4139, doi: 10.1175/2007JAS2144.1.

Qian, L. Y., G. S. Young, and W. M. Frank, 1998: A convective wake parameterization scheme for use in general circulation models. *Mon.Weather Rev.,* **126**, 456–469, doi: 10.1175/1520-0493(1998)126<0456:ACWPSF>2.0.CO;2.

Ramirez, J. A., and R. L. Bras, 1990: Clustered or regular cumulus cloud fields: The statistical character of observed and simulated cloud fields, *J. Geophys. Res.*, **95**(D3), 2035–2045, doi: 10.1029/JD095iD03p02035.

Ramirez, J. A., R. L. Bras, and K. A. Emanuel, 1990: Stabilization functions of unforced cumulus clouds - their nature and components. *J. Geophys. Res.-Atmos.,* **95**, 2047–2059, doi: 10.1029/JD095iD03p02047.

Rasch, P. J., M. J. Stevens, L. Ricciardulli, A. Dai, A. Negri, R. Wood, B. A. Boville, B. Eaton, and J. J. Hack, 2006: A characterization of tropical transient activity in the CAM3 atmospheric hydrologic cycle. *J. Clim.,* **19**, 2222–2242, doi: 10.1175/JCLI3752.1.

Raymond, D. J., A. M. Blyth, 1992: Extension of the stochastic mixing model to cumulonimbus clouds. *J. Atmos. Sci.,* **49**, 1968–1983, doi: 10.1175/1520-0469(1992)049<1968:EOTSMM>2.0.CO;2.

Rogers, R. R. and M. K. Yau, 1989: A short course in cloud physics. Pergamon, New York, 293 pp.

Rougier, J., D. M. H. Sexton, J. M. Murphy, and D. Stainforth, 2009: Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *J. Clim.,* **22**, 3540–3557, doi: 10.1175/2008JCLI2533.1.

Scinocca, J. F., N. A. McFarlane, 2004: The variability of modeled tropical precipitation. *J. Atmos. Sci.,* **61**, 1993–2015, doi: 10.1175/1520-0469(2004)061<1993:TVOMTP>2.0.CO;2.

Tulich, S. N., B. E. Mapes, 2008: Multiscale convective wave disturbances in the tropics: Insights from a two-dimensional cloud-resolving model. *J. Atmos. Sci.,* **65**, 140–155, doi: 10.1175/2007JAS2353.1.

Tulich, S. and B. E. Mapes, 2009: Transient environmental sensitivities of explicitly simulated tropical convection. *J. Atmos. Sci.,* **67**, 923–9401, doi: 10.1175/2009JAS3277.1.

Weger, R. C., J. Lee, T. Zhu, and R. M. Welch, 1992: Clustering, randomness and regularity in cloud fields: 1. Theoretical considerations, *J. Geophys. Res.*, **97**(D18), 20,519–20,536.

Yanai, M., S. Esbensen, and J. H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.,* **30**, 611–627, doi: 10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2.

Zhang, G. J., N. A. Mcfarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate center general-circulation model. *Atmosphere-Ocean,* **33**, 407–446.

Zhu, T., J. Lee, R. C. Weger, and R. M. Welch (1992), Clustering, randomness, and regularity in cloud fields: 2. Cumulus cloud fields, *J. Geophys. Res.*, **97**(D18)20, 537.20,558.