Quantifying the limits of convective parameterizations

Todd R. Jones[1] and David A. Randall[1]

Received 17 August 2010; revised 7 January 2011; accepted 3 February 2011; published 23 April 2011.

[1] Quasi-equilibrium (QE) closure is an approximation that is expected to apply to a large ensemble of clouds under slowly changing weather conditions. It breaks down under rapidly changing conditions or when the domain size is too small to provide an adequate sample of the cloud field. We explore fluctuations about an equilibrium state as simulated by a three-dimensional cloud-resolving model. An ensemble of simulations is used to determine how the response to prescribed periodic large-scale forcing changes with the period of the forcing and the size of the averaging domain. The vertical profile of the forcing is loosely based on GATE data. Results are compared with those from constant forcing simulations. In the constant forcing simulations, the noise-to-signal ratio is nearly independent of forcing magnitude. With time-varying forcing, a considerable range of responses is found. As expected, the more slowly the forcing varies, the better the response is approximated by QE. Errors become large when the period of the forcing is less than 30 h, suggesting that the diurnal cycle cannot be accurately simulated with a QE closure. Nondeterministic variability becomes more significant with smaller domain sizes. For the cases studied, a domain width of at least 180 km is needed to obtain an adequate sample of the cloud population.

Citation: Jones, T. R., and D. A. Randall (2011), Quantifying the limits of convective parameterizations, *J. Geophys. Res.*, *116*, D08210, doi:10.1029/2010JD014913.

1. Introduction

[2] As discussed by *Lin and Neelin* [2002], the global numerical models used for weather prediction and climate simulation purport to determine area-averaged precipitation rates for each grid cell; naturally, some places in a grid cell will get more precipitation than the average, and others will get less.

[3] The aim of a parameterization is to describe the statistics of unresolved processes in terms of resolved processes. In particular, the goal of cumulus parameterization is to determine changes in the simulated large-scale environment due to the collective influence of multiple cumulus clouds. Most cumulus parameterizations are deterministic, and are based on an assumed "quasi-equilibrium" (QE) balance between large-scale (resolved) forcing, such as upper level radiative cooling and low level moistening of the boundary layer, which act to destabilize the troposphere, and the net large-scale effects of unresolved, convective cloud processes, including upper level warming and boundary layer drying, which act to reduce the instability [*Arakawa and Schubert*, 1974; *Emanuel et al.*, 1994; *Randall et al.*, 1997].

[4] Departures from QE arise in part from the inherently unpredictable nature of small-scale motions. As is now well known, weather systems are subject to "sensitive dependence on initial conditions," in which fluctuations associated with dynamical instabilities on small scales grow and affect larger scales through nonlinear processes [e.g., *Lorenz*, 1969]. In cloud fields, important nonlinearities include not only the usual "eddy fluxes," but also phase changes, precipitation, and other microphysical processes [*Arakawa*, 2004]. Because of sensitive dependence on initial conditions, the grid cell–averaged precipitation rate and other statistics are *intrinsically* uncertain [e.g., *Hohenegger and Schär*, 2007]. This is a fundamental issue, unrelated to the details of any particular parameterization. Nonlinear interactions allow unpredictable small-scale fluctuations of the cloud field to influence larger scales. This can be particularly important in the presence of mesoscale organization [*Xu et al.*, 1992].

[5] Sample size is a second important issue. The number of convective clouds making up an "ensemble" within a model grid column is typically not large enough to yield robust statistics [*Molinari and Dudek*, 1992; *Palmer*, 1996]. Even under disturbed conditions, the number of large convective clouds within a typical $(100\ \mathrm{km})^2$ global circulation model (GCM) grid cell is typically on the order of ten to a hundred [e.g., *López*, 1977; *Johnson et al.*, 1999], not enough to give robust statistics. The problem is especially serious for large convective clouds; like large predators, they are widely spaced.

[6] Sensitive dependence on initial conditions and inadequate sample size conspire to introduce a stochastic component into the behavior of a cumulus ensemble, limiting the

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.

Copyright 2011 by the American Geophysical Union.
0148-0227/11/2010JD014913

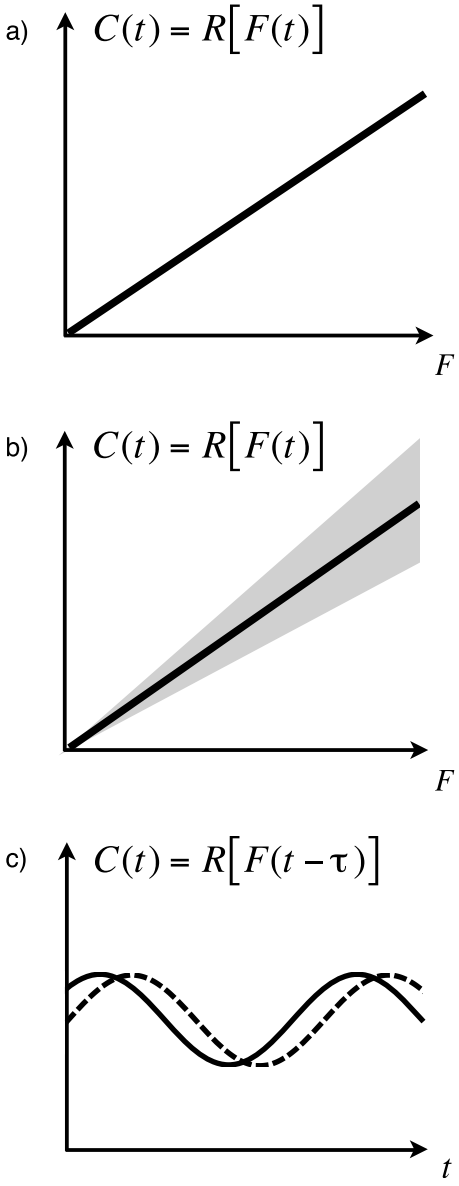


Figure 1. (a) In the quasi-equilibrium regime, the convective response (vertical axis) is strongly determined by the time-varying large-scale forcing. The black line represents a simple and strong relationship between forcing and response. (b) In the stochastic regime, the convective response is uncertain even when the convective adjustment time is short compared to the time scale for variation of the large-scale forcing. This occurs when the sample size is small, e.g., because the grid columns are small. The gray "fan" represents the uncertainty in the convective response. In this example, the uncertainty has been assumed to be proportional to the expected value of the convective response. (c) In the nonequilibrium regime, the convection (dashed curve) cannot keep up with the time-varying large-scale forcing (solid curve) because the convective adjustment time is too long. The forcing leads the response with a time shift of magnitude τ.

deterministic parameterizability of the convective statistics, such as heating and drying rates.

[7] A third issue is that classical QE closures break down when the time scale for changes in the resolved-scale weather is comparable to or shorter than the convective adjustment time [*Molinari and Dudek*, 1992; *Cohen and Craig*, 2006]. This problem is more severe in the emerging high-resolution global models [e.g., *Arribas*, 2004; *Bechtold et al.*, 2008], simply because convective systems with smaller spatial scales tend to have shorter time scales. This time scale issue does not prevent deterministic parameterization, but it does mean that parameterizations cannot be purely diagnostic. *Pan and Randall* [1998] proposed a simple prognostic closure for a cumulus parameterization.

[8] It is useful to distinguish several idealized regimes of statistical cumulus dynamics:

[9] 1. In the quasi-equilibrium regime, when the sample size is large and the weather regime changes slowly, the large-scale "forcing" and the sounding strongly determine the convective heating and drying rates, which adjust, without delay, to the evolving large-scale weather. The instantaneous forcing determines the instantaneous convective heating and drying rates. The past history of the convection does not matter. The expected value of the convective response closely approximates the true value. This idealized quasi-equilibrium regime is illustrated in Figure 1a.

[10] 2. In the nondeterministic regime, the convective heating and drying rates have a significant stochastic component; that is, they are only partially predictable. Figure 1b illustrates this nondeterministic regime. The uncertainty in the convective heating and drying rates creates an uncertainty in the larger scales. This can be true even when the convective adjustment time is small compared to the time scale for the variation of the large-scale circulation. This type of uncertainty is characteristic of real physical cloud systems [e.g., *Hohenegger and Schär*, 2007]; it cannot be eliminated by model refinements.

[11] 3. In the nonequilibrium regime, still deterministic, the sample size is assumed to be large enough to yield robust statistics. In this case, however, the large-scale weather evolves so rapidly that the convection cannot keep up. The past history of the convection therefore affects the instantaneous convective heating and drying rates. Although the convection is out of equilibrium, the convective heating and drying rates are still deterministic. The *sequence* of large-scale forcings strongly determines the *sequence* of convective heating and drying rates. This deterministic, nonequilibrium regime is illustrated in Figure 1c.

[12] In summary, two distinct problems with conventional deterministic, diagnostic parameterizations become more serious as model resolutions improve. The first is that deterministic parameterizations become less well justified because of the problem of sample size. The second is that diagnostic parameterizations become less well justified because of the nonseparation of time scales.

[13] It has been suggested that the lack of high-frequency, small-scale variability in the GCM-generated precipitation fields and convective heating rates may limit the ability of GCMs to simulate the large-scale, low-frequency aspects of climate variability [*Ricciardulli and Garcia*, 2000;

Horinouchi et al., 2003]. The failure of QE-based convective parameterizations to reproduce the full spectrum of convective variability has motivated the development of new approaches to cumulus parameterization. *Arakawa* [2004] discusses the current trend away from deterministic and diagnostic closures and even ensemble averaging methods toward prognostic and nondeterministic closures of varying types, including stochastic parameterizations, relaxation methods, triggered adjustments to equilibrium, and superparameterization, which includes the effects of small-scale variability through the use of an embedded CRM. A recognition of the need for stochastic convective parameterizations has developed from the realization that QE-based convective parameterizations employed in GCMs fail to reproduce the full spectrum of convective variability that is found in cloud resolving model (CRM) simulations and observational data. It has been suggested [*Neelin et al.*, 2008] that a realistic representation of stochastic convective variability is necessary for simulating large-scale weather systems, such as the Madden-Julian Oscillation. Stochastic parameterizations have been explored extensively in Europe [*Buizza et al.*, 1999; *Wilks*, 2005, 2008; *Williams*, 2005; *Palmer et al.*, 2005; *Palmer and Williams*, 2008; *Berner et al.*, 2008; *Ball and Plant*, 2008], while superparameterization has been a focus in the U.S. [*Khairoutdinov and Randall*, 2001; *Randall et al.*, 2003; *Räisänen et al.*, 2004; *Cole et al.*, 2005a, 2005b; *Khairoutdinov et al.*, 2005, 2008; *DeMott et al.*, 2007, 2010; C. A. DeMott et al., The Asian monsoon in the super-parameterized CCSM and its relation to tropical wave activity, submitted to *Journal of Climate*, 2011; *Tao et al.*, 2009; *Benedict and Randall*, 2009; J. J. Benedict and D. A. Randall, Impacts of idealized air-sea coupling on Madden-Julian Oscillation structure in the super-parameterized CAM, submitted to *Journal of Atmospheric Science*, 2011; *Thayer-Calder and Randall*, 2009; *Stan et al.*, 2010]. See *Jones* [2010] for further discussion.

[14] This study is aimed at achieving an improved understanding of nonequilibrium and nondeterministic cumulus convection. Our aim is to quantitatively explore the capabilities and limitations of cumulus parameterization as applied to models with grid spacing on the order of 20–50 km, or even finer. We analyze random fluctuations produced in simulations of convection using a CRM. Our goal is to determine the response of a variety of convection-related statistics to an imposed periodic large-scale forcing. A suite of CRM simulations is used to compare the response in two- and three-dimensional simulations, to explore how the response varies with changes in the length of the forcing period and the size of the averaging domain, and to see how the response to a time-varying forcing compares to the response to a steady forcing.

[15] Section 2 provides relevant scientific background. Section 3 gives a description of the CRM used and includes a discussion of the design of the numerical simulations. In section 4, we present the CRM results. Section 5 provides a discussion and conclusions.

2. Background

[16] *Xu and Arakawa* [1992] evaluated the parameterization of *Arakawa and Schubert* [1974, hereafter AS74] by performing semiprognostic tests against data from a two-dimensional cloud-resolving model (CRM). They found that although cumulus activity is strongly modulated by large-scale processes, the AS74 parameterization fails to capture some of the nonmodulated, high-frequency fluctuations as well as the systematic phase delays in modulation that are particularly noticeable in the presence of system-organizing wind shear. They showed that nondeterministic errors increase with smaller averaging distances. They concluded phase lags between the convection and the large-scale forcing are more significant for coarse-resolution models that do not resolve mesoscale features, while errors due to the nondeterministic (high-frequency variability) become more important with finer resolution.

[17] *Xu et al.* [1992, hereafter X92] attempted to determine the extent to which cumulus convection can be parameterized deterministically and diagnostically, and how the separation of scales that is required by some convective parameterizations is affected by mesoscale organization. Simulations using 2 km horizontal grid spacing over a 512 km domain and various cases of mesoscale organization (shear) and surface types showed that the magnitude of the fluctuations depends on forcing magnitude, wind shear, and domain size. X92 showed that the convective response to cyclic large-scale forcing is not identical at a given phase of the forcing, indicating nondeterministic fluctuations. The precipitation can have a standard deviation on the order of its mean; this was also noted by *Katzfey and Ryan* [2000]. X92 also showed that there are phase delays due to mesoscale organization when the wind shear is strong.

[18] Our study is an extension of X92. We use a three-dimensional model, and examine the effects of varying the period of the forcing and the size of the domain.

3. Numerical Simulations

3.1. Model Description

[19] We have analyzed simulations performed using a three-dimensional anelastic model based on the vorticity equation or vector vorticity model (VVM) developed by *Jung and Arakawa* [2008]. The model is included in an ongoing CRM intercomparison study [*Fridlind et al.*, 2010]; results indicate that the performance of the VVM is comparable to that of other CRMs.

[20] The simulations were performed on a doubly periodic *f* plane at 15°N, with a horizontal domain size of $(256\ \text{km})^2$ and a 2 km horizontal grid spacing. A stretched vertical grid [*Krueger*, 1988] is employed to enhance resolution near the surface. The depth of the domain is approximately 19 km, with 35 layers, giving a grid interval of approximately 100 m near the surface and 1 km near the model top. The model is stepped forward in 10 s intervals, and the time differencing scheme is second-order Adams-Bashforth.

[21] Our configuration of the VVM includes a three-phase microphysics parameterization [*Krueger et al.*, 1995; *Lin et al.*, 1983; *Lord et al.*, 1984], a surface flux parameterization using the flux-profile relationships given by *Deardorff* [1972], and a first-order turbulence closure [*Shutts and Gray*, 1994]. In place of the model's radiation parameterization, we used Newtonian cooling, relaxation toward the initial horizontally averaged profile with a time scale of 1 day above 15 km, linearly decreasing to zero at and below 10 km. The VVM also includes a Rayleigh-type

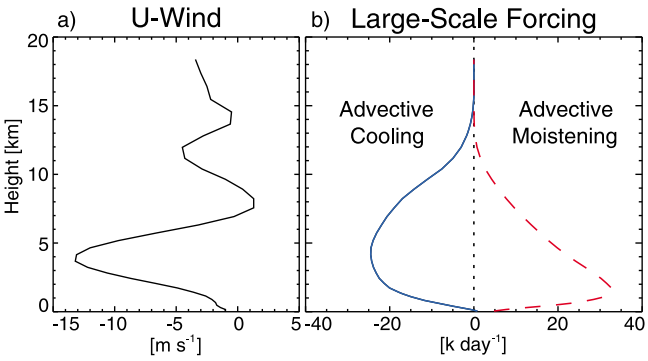


Figure 2. (a) The sheared, prescribed u wind profile with which each simulation is initialized and (b) the maximum value profiles of the prescribed large-scale advective cooling and moistening.

friction in the top five layers to damp upward propagating gravity waves.

[22] The surface is ocean with a prescribed sea surface temperature of 299.88 K and a surface roughness of 0.2 mm. The x component of the geostrophic wind is held constant with a form similar to the initial u wind profile, while the y component is assumed to be zero. Each simulation is initialized with horizontally uniform thermodynamic conditions based on the profiles in Figure 2, and clouds are initiated by the introduction of small (±0.25 K), random perturbations to the potential temperature in the lowest 100 m of the model for the first 30 simulated minutes. For all runs, we analyzed results after the first 300 simulated hours of integration.

3.2. Experimental Design

[23] The simulations are designed to simulate a field of cumulus clouds under a given set of large-scale conditions, as if the domain represented a single column of a GCM. Horizontally uniform large-scale advective cooling and moistening rates are prescribed.

[24] All of the model runs begin with the same initial conditions, which are based on soundings observed during GATE (Global Atmospheric Research Program's (GARP) Atlantic Tropical Experiment) Phase III [*Sui and Yanai*, 1986]. These initial conditions included moderate vertical wind shear (Figure 2a) and an enhanced profile of the GATE-III values for advective cooling and moistening (Figure 2b). The vertically integrated magnitudes of these maximum forcings are 1961.1 W m^{-2} and 60.8 mm d^{-1}. These latter two profiles show the maximum intensities of the time-varying forcings.

[25] The forcing profile varies in time, as in X92, with a periodic scaling factor of the form,

$$f(t) = \frac{1 - \cos\left(\frac{2\pi t}{T}\right)}{2}, \qquad (1)$$

where T is the period of the time variation and t is time. Thirteen periodically forced simulations were run for 15 cycles after the initial spin-up. We performed simulations with periods of 2, 4, 6, 8, 10, 12, 14, 16, 20, 24, 30, 60, and 120 h. Throughout the remainder of this paper, individual simulations are identified by the length of the period of the forcing preceded by an "F." For example, the simulation with a 6 h period is F06.

[26] To determine how the convective response to a periodic forcing compares to the convective response to a constant forcing, ten more 6.25 day simulations were run with steady or "constant" forcings chosen in increments of 10% of the maximum large-scale forcing profile (Figure 2b). These constant forcing runs were initialized in the same way as the variable forcing simulations, including spin-up. These simulations are referenced by "CF" (for "constant forcing") followed by a single digit representing the fraction of the maximum forcing. For example, CF8 refers to a constant forcing that is equal to 80% of the profiles in Figure 2b. The exception to this convention is the simulation run at 100% of the maximum; it is referred to as CFX, with "X" interpreted as the Roman numeral ten for some degree of consistency.

4. Results

4.1. Constant Forcing Simulations

[27] Here we briefly discuss the results of the constant forcing simulations. We use the surface precipitation rate as an indicator of the intensity of convection. Figure 3 shows the time evolution of the domain-averaged surface precipitation rate is shown for each of the constant forcing simulations. Figure 3 shows a progressive increase in the magnitude of the precipitation rate, represented by the black lines, as the forcing is increased. The red lines in each plot represent the time means. The increase in the mean precipitation rate is almost exactly linear with respect to the normalized forcing (Figure 4a). Quantitatively, there is an increase of approximately a quarter of a millimeter per hour for each ten percent increase in the large-scale forcing. Based on the prescribed forcing profiles (Figure 2), this corresponds to approximately −2.5 K d^{-1} of peak advective cooling and 3.5 K d^{-1} equivalent of peak advective moistening.

[28] The other very important feature of Figure 3 is that although the simulations are forced at constant rates, the domain average responses in precipitation rate are not constant in time. The blue lines in Figure 3 show plus and minus one standard deviation from the means. The domain averages are in statistical equilibrium, such that for each run the variance about the mean is small with respect to the mean. Figure 4b shows that the standard deviation of the precipitation rate increases almost linearly with increasing forcing, except for extreme forcing values.

[29] The ratio of the standard deviation to the mean (coefficient of variation, hereafter COV) for the precipitation rate is shown in Figure 4c. For most forcing values tested, the COV is nearly constant, meaning that variability scales with the mean, as many might expect. In fact, this is an assumption made in a number of convective parameterizations (e.g., in ECMWF models [*Buizza et al.*, 1999, 2000]). However, the results here indicate that the COV for weak forcing (CF1) is greater than twice that of the midrange forcings. For strong forcing (CFX), the COV decreases slightly, relative to CF2–CF9.

[30] A similar analysis has been performed for cloud fraction and vertically integrated nonprecipitating condensate (NPC), which is a measure of domain mean column

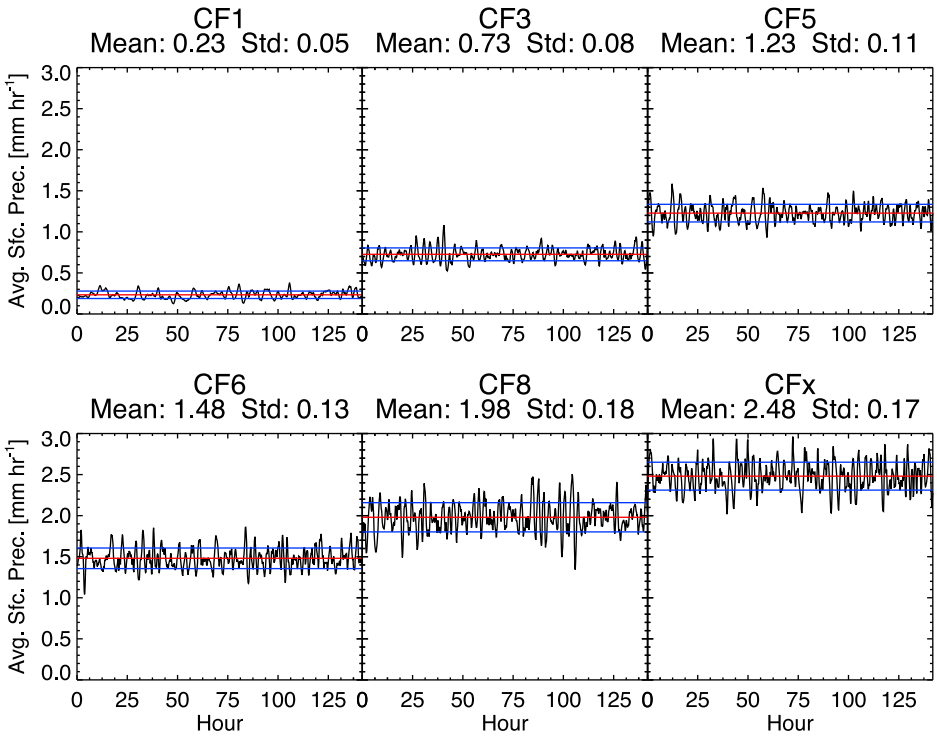


Figure 3. Time evolution of the domain-averaged surface precipitation rate (black line) for each of the constant forcing simulations. The time series mean (red line) and standard deviation bounds (blue lines) are overlaid; the actual numerical values are given above each plot.

cloud water and cloud ice. Figure 4 also shows the results for these variables. Cloud fraction decreases very slightly (~10%) as the forcing strengthens. One might expect more forcing to directly increase the amount of cloud in a given domain, but stronger convection can be more organized, with less stratiform cloud. The nonprecipitating condensate increases weakly with an increase in the large-scale forcing, indicating denser clouds. Trends in the standard deviations of these variables are weakly positive, but interactive radiation, not included here, might change the results.

[31] The COV increases with forcing strength for cloud fraction and shows no trend for NPC. In each of the three cases, the COV remains fairly constant for moderate forcing, such that the standard deviation is approximately 10% of the mean magnitude with notable departures at the forcing extremes. Parameterizations assuming constant COV would predict too little variability with weak forcing and too much variability with strong forcing.

4.2. Periodic Forcing Simulations

[32] The 15 cycles simulated in each run can be viewed as individual realizations of "the same" event, as predicted with different initial conditions. The cycles can be compared to analyze the predictability of the convective statistics for each period length.

4.2.1. Snapshot Views

[33] We begin with model output snapshots from the periodic forcing simulations. Figure 5 depicts surface precipitation rate (Figure 5, left) and cloud top temperatures (Figure 5, middle) during hour 402 of simulation of the F24 (24 h period) simulation. Figure 5 (right) shows the time series values of the mean surface precipitation rate (Figure 5, right top), domain cloud fraction (Figure 5, right middle), and corresponding normalized forcing (Figure 5, right bottom). Cloud top is defined as the height where the column integral of the cloud liquid and ice water exceeds 0.1 kg m^{-2} when integrating downward from the model top. The vertical red lines denote the time of the snapshot contours.

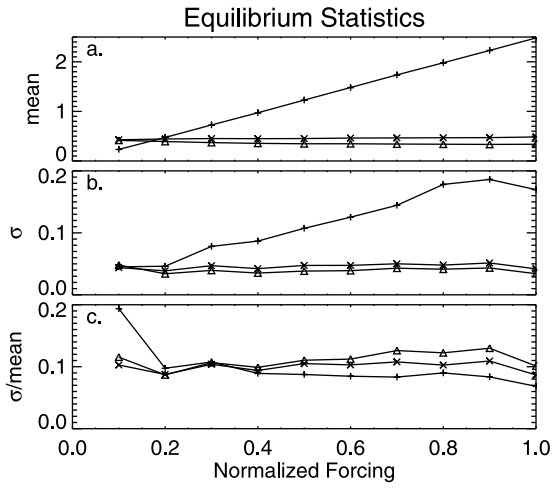


Figure 4. (a) The time mean, (b) the time-relative standard deviation, and (c) the coefficient of variation of the surface precipitation rate (pluses, mm h^{-1}), cloud fraction (triangles), and nonprecipitating condensate (crosses, kg m^{-2}) for constant forcing simulations CF1 through CFX.

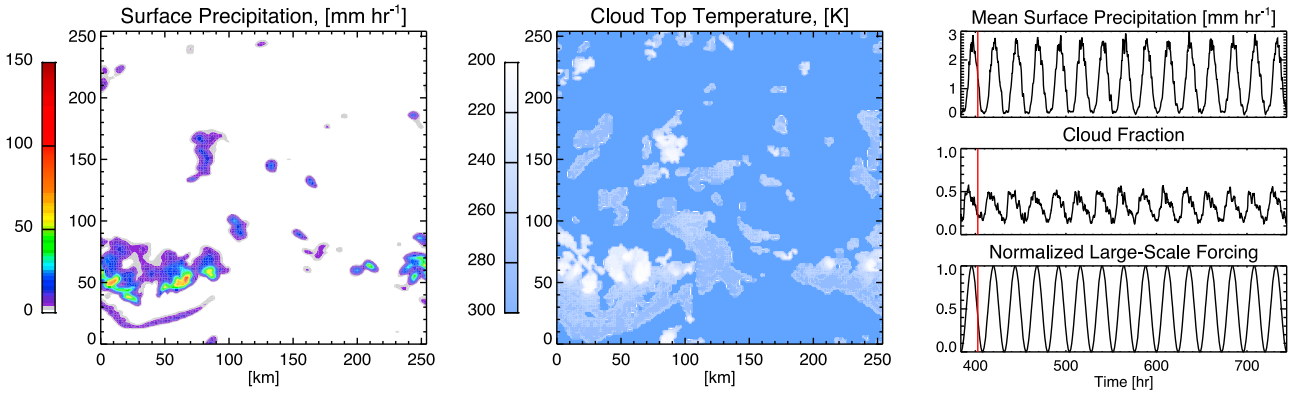


Figure 5. Snapshots and time series from F24.

[34] In this example, large, organized mesoscale clusters/lines have produced precipitation via outflow lifting. The clusters persist for less than an hour. Many of the simulations produce similar linear mesoscale features.

[35] There is even larger-scale organization in a few instances. This is illustrated in Figure 6. In the case of F16 (Figure 6a) for a point where the large-scale forcing is locally maximized, the band of organized convection spans the entire width of the domain and moves as a cohesive unit from north to south. It is likely that the presence of this feature is the reason for the slightly broadened precipitation peak that immediately follows this point in the simulation. In the snapshot of theF120 simulation (Figure 6b), there is another large band of organized precipitation/convection in the northeast section of the domain. The feature is over 150 km from east to west, and is preceded by an outflow boundary as shown in F24 (Figure 5). Some simulations produce rotating convective systems (not shown).

4.2.2. Composite Analysis: Full Domain

[36] For the 15 realizations in each case, the results are averaged and the standard deviation across the realizations is determined. The standard deviation can be interpreted as a measure of the predictability of the simulated systems. To be clear, this is the standard deviation across realizations, rather than the spatial standard deviation at a point in time. A stochastic parameterization would generate the statistics corresponding to a particular realization.

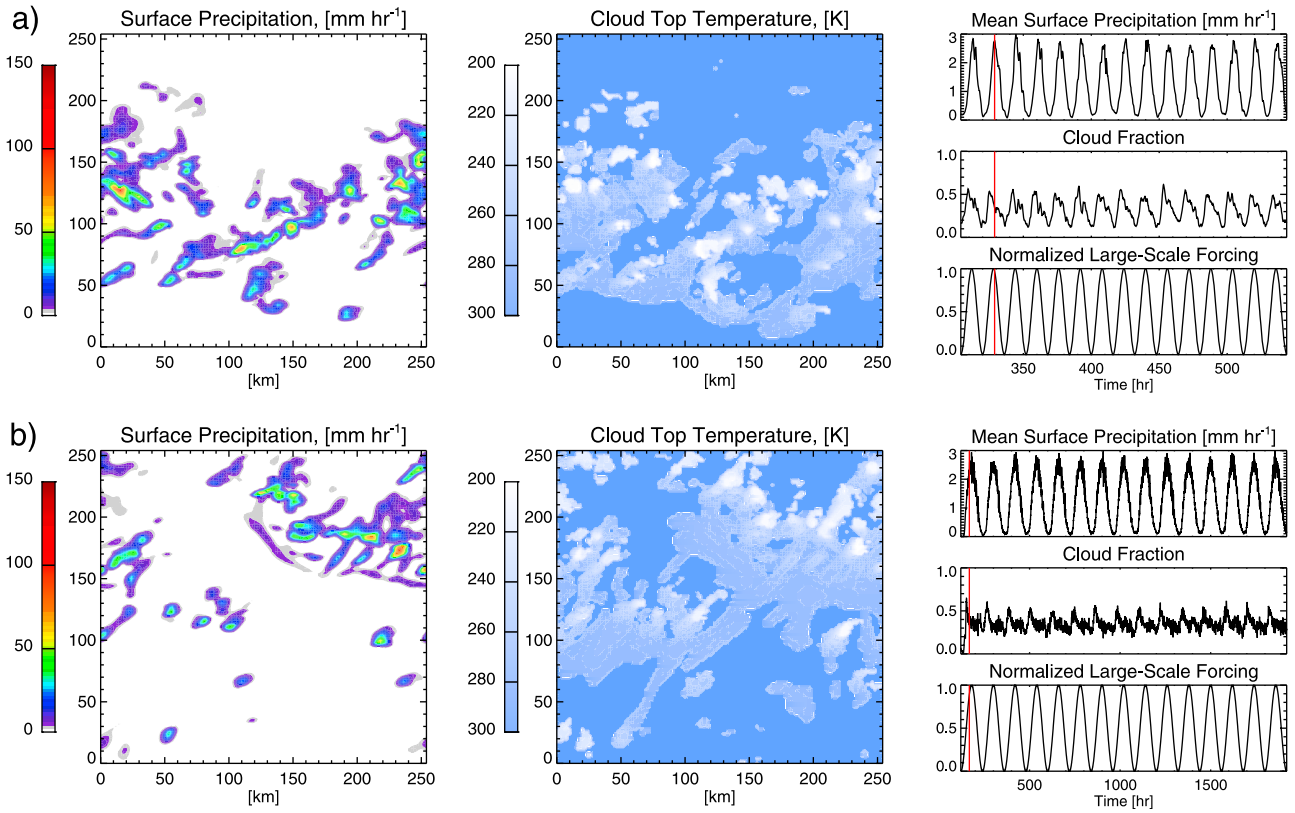


Figure 6. As in Figure 5. Snapshots from (a) F16 and (b) F120.

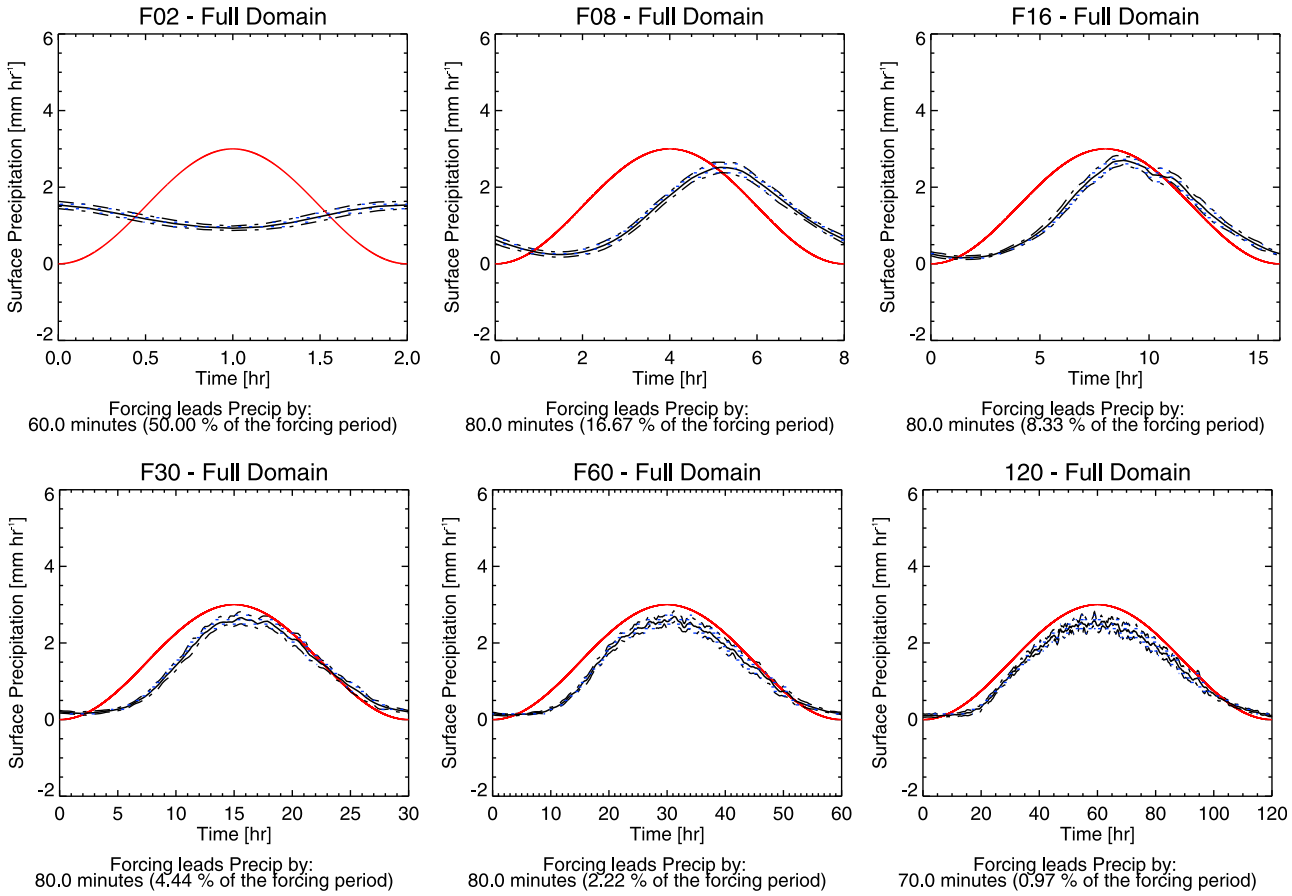


Figure 7. Whole domain composite of surface precipitation rate (mm h^{-1}) for the noted simulations. The black curve is the composite mean. The blue hash-filled region bounded by the dot-dashed lines denotes the ±1 standard deviation across 15 realizations. The red curve references the timing and relative magnitude (0%–100%) of the large-scale forcing; no specific values are implied.

[37] Returning briefly to Figure 6a (right top), we see that the precipitation appears to lag the forcing slightly and that there is considerable scatter. The variability in the time series appears to be much less than that obtained in the two-dimensional simulations of X92. This is expected, because the sample size is much larger in the 3D model. Also, the greater and more variable composite standard deviation shown by X92 likely has some contribution from the smaller ensemble size involved (their 9 realizations versus our 15).

[38] Composites of the realizations for the full domain of the F02, F08, F16, F30, F60, and F120 simulations are shown in Figure 7. Compared to the raw time series, the data in the composite plots are much more smooth, as the variability around the mean has been averaged out to some degree. This is shown by the blue stippled region (standard deviation) bounding the mean. Variability is larger for the long-period (F60, F120) simulations, perhaps because more persistent strong forcing permits the development of heavily precipitating outliers with self-perpetuating mesoscale organization. The standard deviation does not change much with changes in the period of the forcing, but it does change with the magnitude of the forcing.

[39] Phase lags between the precipitation and the forcing are both visually apparent and numerically documented in each plot of Figure 7. The phase lags were quantified by finding the lag that maximizes the correlation between the forcing and the response. Lags occur for two reasons. One is the delay associated with the presence of mesoscale organization (X92, Figure 9). This is evidenced by the rapid onset and more gradual decline of the precipitation rate, which are characteristic of the life cycles of mesoscale convective systems [X92; *Abdullaev et al.*, 2009]. Unfortunately, this mesoscale lag is not well captured by the lag detection method mentioned above. Second, lags occur when the period of the forcing is close to the convective adjustment time, which ranges from tens of minutes to a couple of hours depending on the cloud size [*Bechtold et al.*, 2008; *Yu and Neelin*, 1997]. With the shorter periods, the forcing changes too rapidly for the convection to keep up. Except for F02, the forcing leads the precipitation by ~80 min. This is a crude measure of the convective adjustment time [*Xu and Randall*, 1998]. As the period of the forcing increases, the relative lag decreases, while the absolute lag remains approximately constant.

[40] Again, the existence of a lag does not prevent parameterization. It simply points to the need for convective parameterizations to have a memory, or a finite adjustment time.

[41] Figure 8 shows a similar analysis for cloud fraction. The cloud fraction is not modulated in the same fashion as

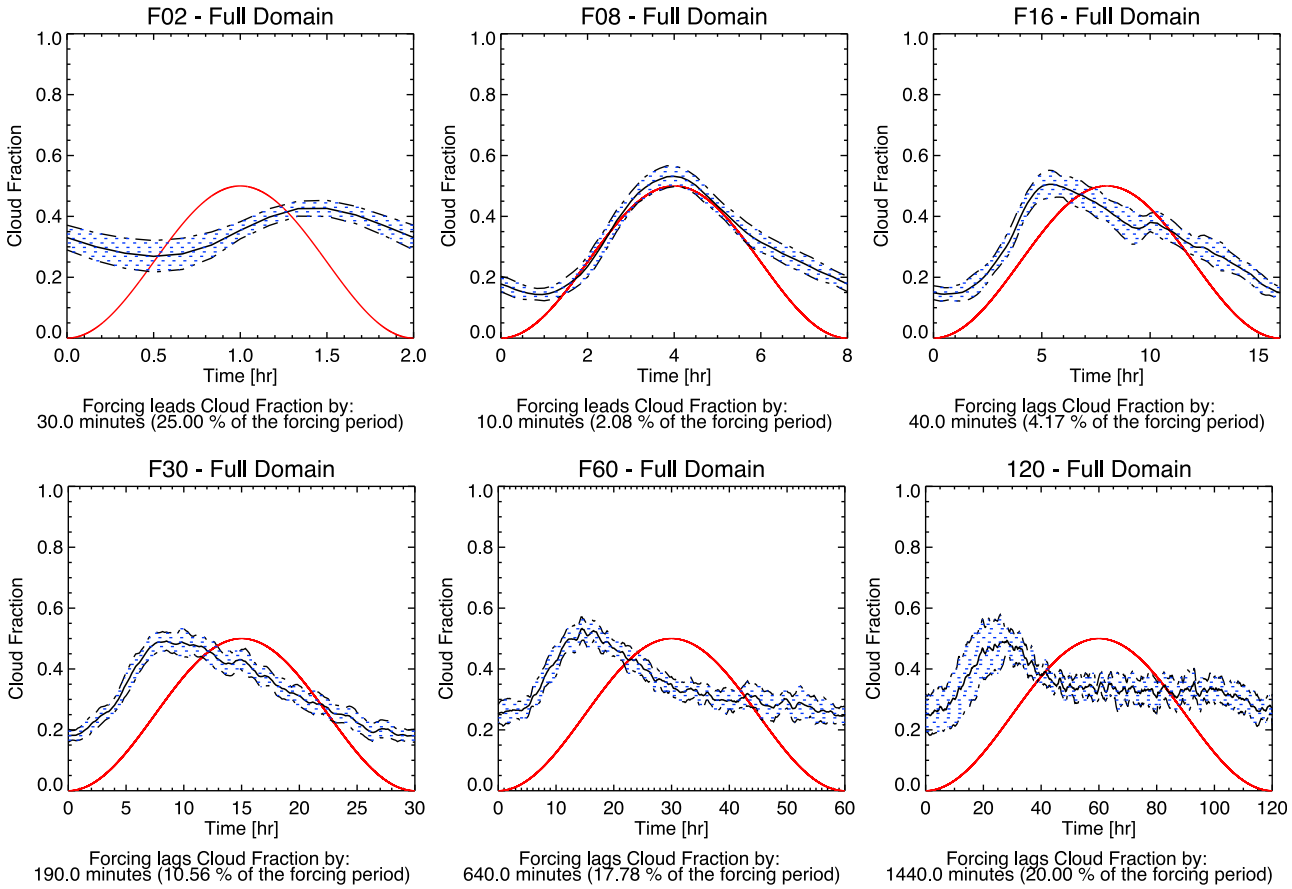


Figure 8. As in Figure 7 but for cloud fraction.

the precipitation. In F02, the cloud fraction is 90 degrees out of phase with the large-scale forcing. The two are almost exactly in phase for the F08 simulation, and the cloud fraction actually leads the forcing in all longer-period simulations. The explanation seems to involve remnants of nonprecipitating (most often cirrus or stratus) clouds left over from the previous cycle. For example, when the F120 forcing is at its minimum (not shown), there is nearly zero precipitation, but a significant amount of cloud cover remains in the domain. The following peak in cloud cover corresponds to the increasing forcing of the following cycle, and the following rapid cloud cover decline and plateau corresponds to a transition toward organized convection that has the effect of clustering clouds into more densely packed regions of taller clouds. In this way, areal cloud cover does not immediately correspond to more intense convection.

[42] We define the NPC as the vertically integrated sum of cloud water and cloud ice. The NPC is somewhat related to cloud fraction, and behaves similarly to it (not shown). The standard deviations about the mean remain fairly constant for all period lengths, though they tend to be slightly larger when the mean NPC is at its maximum, as was the case for precipitation and cloud fraction. This follows from the standard deviations scaling with the mean in the equilibrium simulations in section 4.1. The NPC and cloud fraction both decrease with the rapid increase in precipitation that accompanies mesoscale development, but the NPC lags a bit more than cloud fraction because it is associated with taller, denser clouds that develop later. Our confidence in these cloud-related results is diminished by the lack of interactive radiation and the periodic domain.

[43] For the purposes of this analysis, we define the net cloud mass flux as the domain mean vertical mass flux through the 2.5 km level, including only grid cells where cloud liquid water or ice are present. The results for the net cloud mass flux are very similar to those for surface precipitation in terms of mean and standard deviation, as well as lag times (not shown). The standard deviation about the mean is small, indicating a deterministic convective response. The response transitions from 90° out of phase with very short period forcings to nearly in phase for longer forcings.

[44] Different results are seen when the cloud mass flux is partitioned into upward and downward components (Figure 9). To focus on the more vigorous elements, only velocity magnitudes greater than 0.1 m s^{-1} are considered. The absolute lags are greater for both updrafts and downdrafts than for the net mass flux, and are greatest for the downdrafts in the F30 case. Relative lags decrease with increasing forcing period as with the net cloud mass flux. The maximum updraft mass flux, averaged over realizations, more than doubles going from F02 to F60, suggesting that short forcing periods do not allow full development of the convective activity. An increase in intensity is also seen for the downdrafts, if an increased air speed threshold is used.

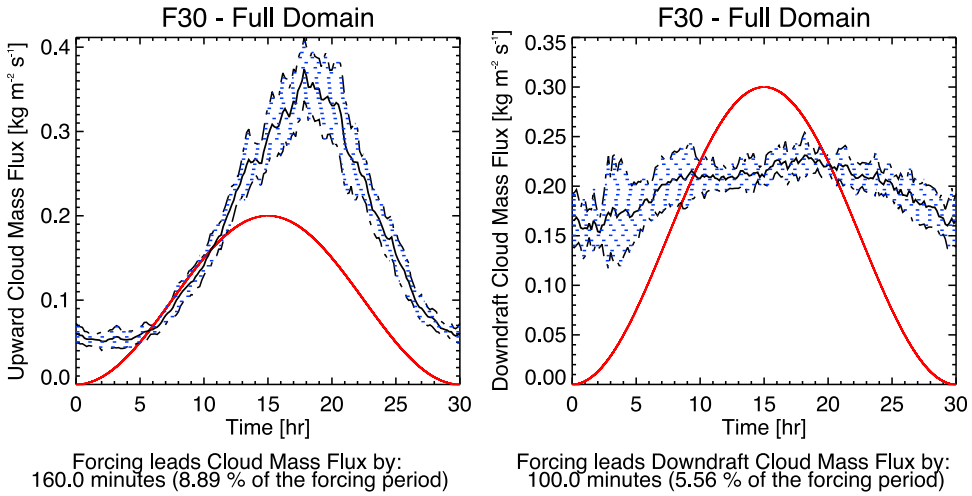


Figure 9. As in Figure 7 but for (left) upward and (right) downward cloud mass flux.

[45] The intensity of convection is very strongly linked to the vertically integrated buoyancy flux, B, which is defined as

$$B = \int_{z_S}^{z_T} \rho c_p \overline{w' T_v'} dz. \tag{2}$$

Here T_v is the virtual temperature, and the integral runs from the surface to the top of the model. The composite analysis results for the buoyancy flux are shown in Figure 10. Many of the features seen in other variables in the composite analyses are present here as well. Variability scales roughly with the mean, and the relative lag decreases with increasing period. The absolute lag relationship, though, is less well

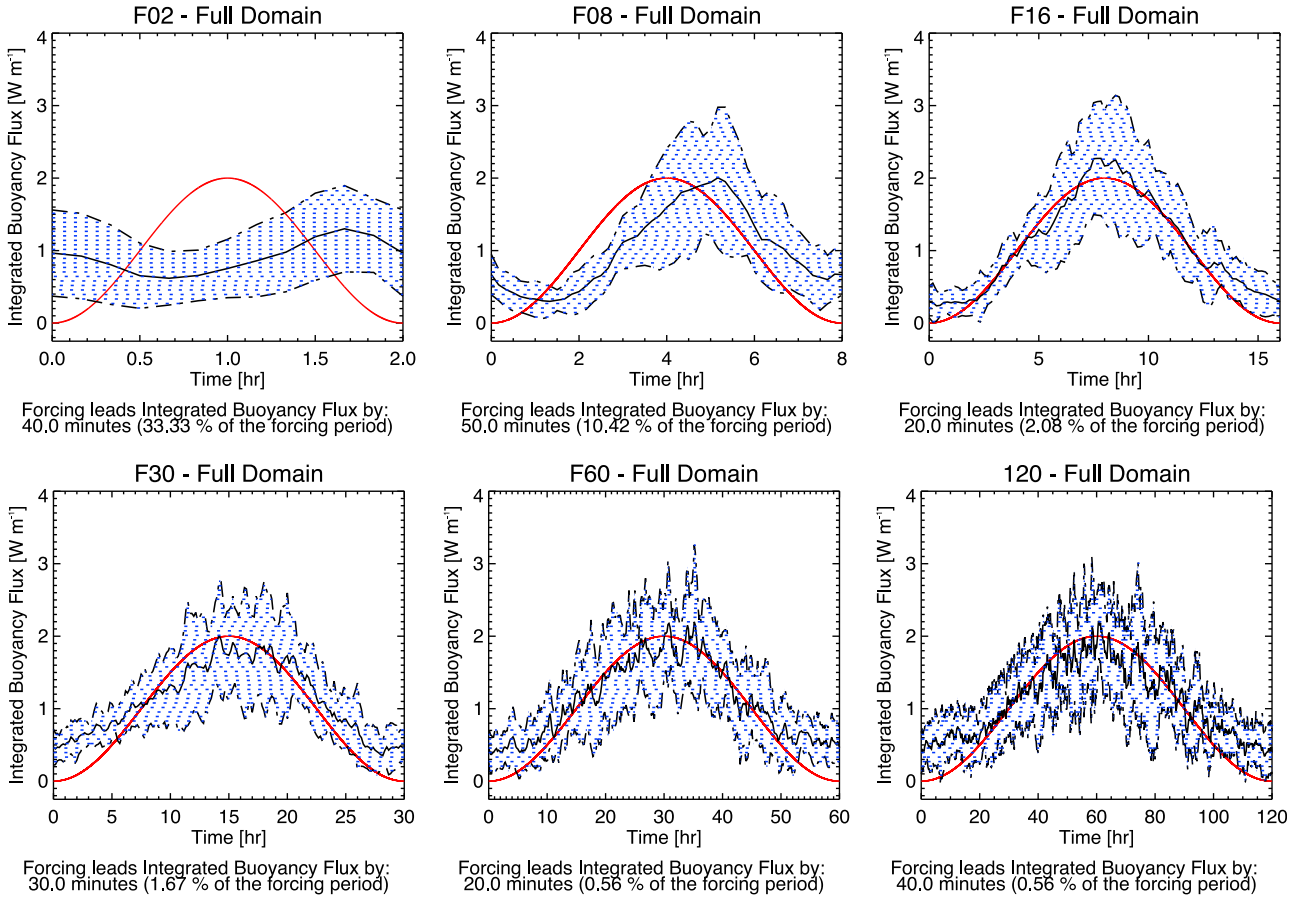


Figure 10. As in Figure 7 but for the integrated buoyancy flux.

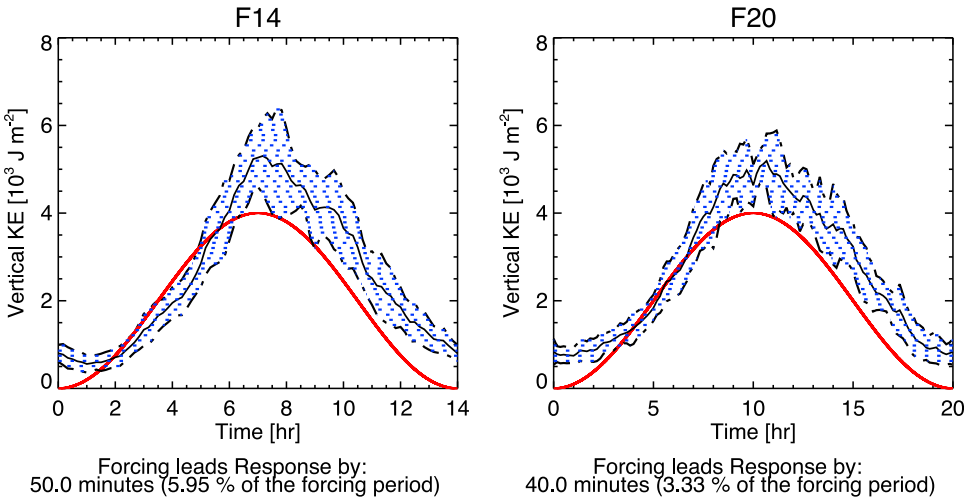


Figure 11. As in Figure 7 but for the domain mean vertical eddy flux.

defined than with the other variables and lacks a coherent trend. Of all of the variables shown thus far, the buoyancy flux is the most closely synchronized with the large-scale forcing, especially for longer periods.

[46] An additional measure that can be used as an indicator of convective activity is the vertical eddy kinetic energy (EKE). This is an effective indicator because it is mainly a measure of updraft and downdraft intensity. In order to obtain the EKE, the procedure outlined in X92 was followed. A given quantity, A, can be separated into large-scale and eddy-scale components, such that

$$A \equiv [A] + A', \tag{3}$$

where $[A]$ is the large-scale component and A' is the eddy-scale component. Here, $[A]$ is defined as

$$[A] \equiv \frac{1}{\tau}\int_{t-\tau/2}^{t+\tau/2}\frac{dt}{L_x L_y}\int_0^{L_x}\int_0^{L_y} A dx dy. \tag{4}$$

In (4), L_x, L_y, and τ are the horizontal extents and the time interval for defining the large-scale component, respectively. The horizontal extent used in the averaging is the full domain size of 256 km, and the time interval was selected based on the period of the forcing and ranged from forty minutes to ~4 h.

[47] Using the eddy-scale component of the vertical wind, w', the vertical EKE can be computed as

$$K_Z = \frac{1}{2}\int_0^{\infty} w'^2 \rho_0 dz. \tag{5}$$

Here, $\tilde{n}_0$ is the model reference state density. Figure 11 shows examples of results (F14 and F20) from the analysis of K_Z. The kinked nature of the early response and slow decline in magnitude are evident in some instances this variable, though these features become less evident with increasing period length (not shown). The relative lag with respect to the forcing tends to decrease as we have seen with other variables, but the absolute lag declines with increasing period length. This suggests that the mesoscale influence is less for longer forcing periods. In no case does the forcing lag the mean response; even at the greatest forcing lengths (60 and 120 h), the forcing leads by at least 10 min.

[48] The eddy kinetic energy data were further analyzed to obtain the component that is due only to mesoscale features. Here, as in X92, the mesoscale is defined as having scales larger than 32 to 64 km. To separate out the mesoscale, a low-pass filter [*Ormsby*, 1961] is applied in physical space as a two-dimensional array of weights derived by rotating and scaling the following 1-D formula:

$$W(\delta) = \frac{\cos(2\pi f_C \delta) - \cos(2\pi f_T \delta)}{2\pi^2 (f_T - f_C)\delta^2}. \tag{6}$$

Here, $\ddot{a}$ is the distance from the point being filtered. The filter is designed to allow full retention of features greater than the "cutoff frequency," f_C, and excludes features smaller than the "roll-off termination frequency," f_T, with a gradual transition between the two. The cutoff and roll-off termination frequencies correspond to wavelengths of 64 km and 32 km, respectively. See *Jones* [2010] for additional details on the application and performance of this filter. In order to estimate the contribution from only the mesoscale spectral analysis was performed on each of the EKE fields at all points in time for each simulation. The percentage of power contained in wave numbers 1–8 (wavelengths greater than 32 km) was used to estimate the fraction of the domain mean EKE that is contributed by the mesoscale.

[49] Figure 12 shows examples of the filtered domain mean vertical EKE, corresponding to those in Figure 11. Comparing these, we can observe the signature of the mesoscale, namely, the delayed, slow ramp-up of convection followed by an even slower, shallower decline. Note the reduced scale on the ordinate axis. The lag increases with increasing period, e.g., 60 to 70 min from F14 to F20, as seen in Figure 12, and on to 80, 90, and 130 min for F30, F60, and F120, respectively, suggesting an increasing mesoscale influence on the lag. These results are very similar to those found by X92. Because we did not vary the shear profiles, we cannot discuss how the statistics vary with different types of mesoscale organization. In fact, without the results

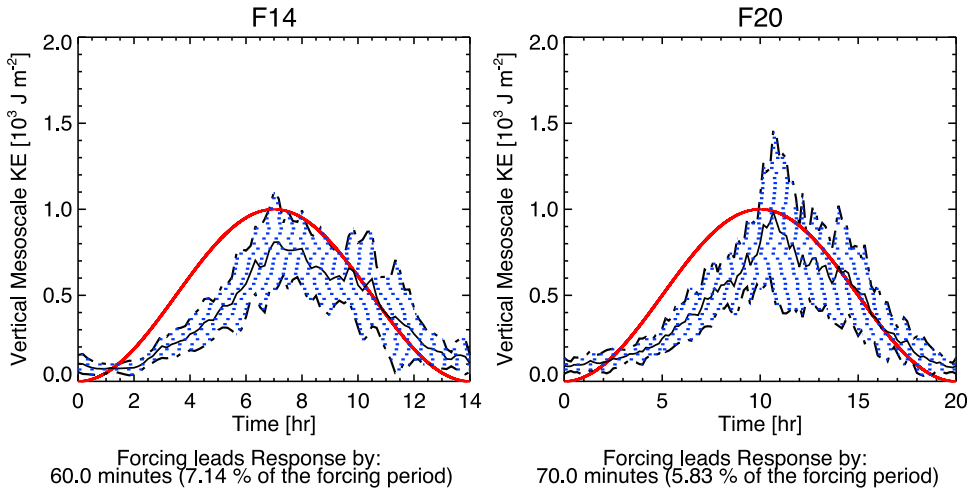


Figure 12. As in Figure 7 but for the domain mean mesoscale vertical eddy flux.

from X92 for comparison, it would be difficult to make the case for the influence of the mesoscale on the modulation of the convection.

4.2.3. Composite Analysis: Subdomains

[50] To determine how the statistics vary with domain size, a subsampling technique was employed. The full domain size of $(256\ \text{km})^2$ was broken down into fractions of one half (rectangular, not square), one quarter, one sixteenth, one 64th, and one 256th. This allows for statistics corresponding to a total of six GCM horizontal grid spacings, namely 256, ~181, 128, 64, 32, and 16 km. The goal is to see the effects of changing a GCM's grid spacing from O(100 km) to O(10 km). We want to determine at which resolution the assumptions of QE tend to break down and to characterize how this occurs.

[51] In the composite analysis, the mean of a quantity is calculated for each individual subdomain at each time. In the case of the 256th subdomain size $(16\ \text{km})^2$, there are 256 time series of 15 cycles, similar to the full domain composite analysis. The standard deviation over realizations

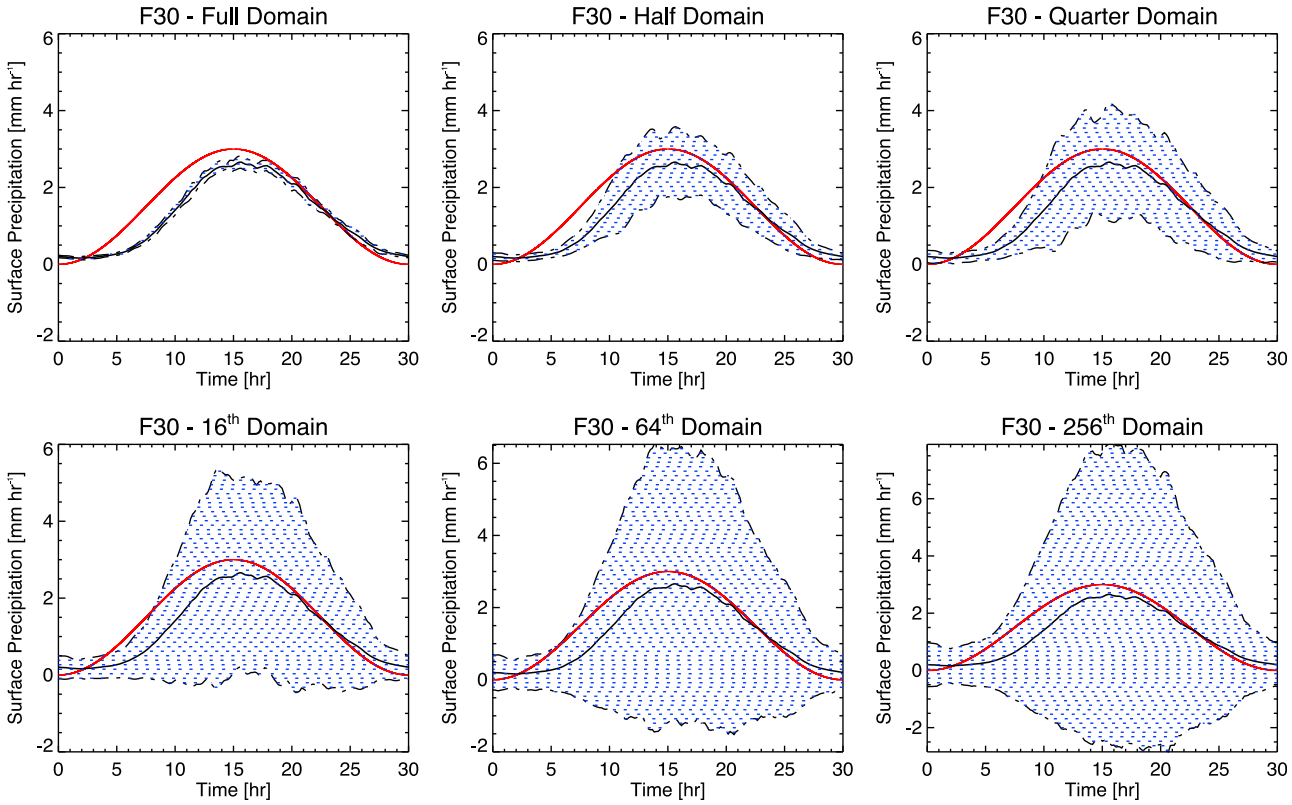


Figure 13. As in Figure 7 but for noted subdomains of the F30 simulation. The mean curve shown here is the same in each plot and corresponds to Figure 7 (bottom left).

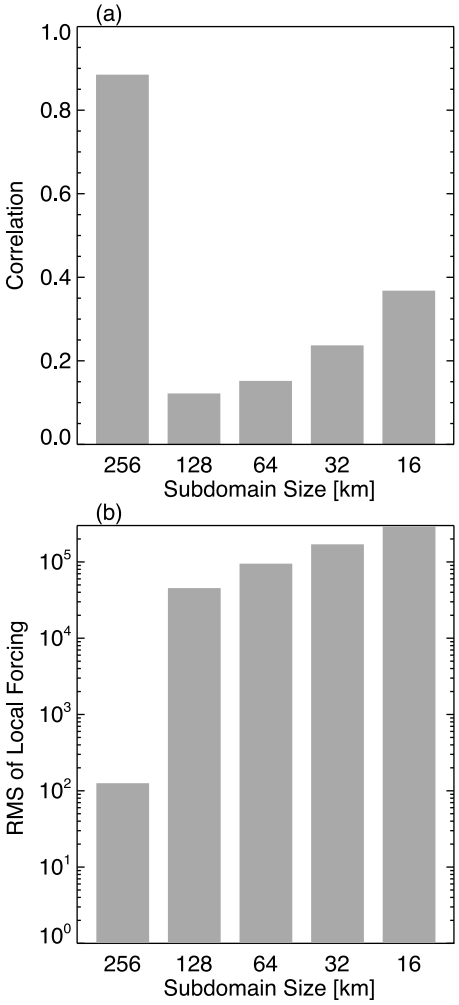


Figure 14. (a) The correlation between the local forcing and the response and (b) the root-mean-square of the local forcing for the listed subdomain sizes.

is calculated for individual subdomains, and the data is averaged across each of the subdomains. After such averaging of the subdomain statistics, the average of the means is of course the same as the average over the full domain, but the standard deviations will differ from the standard deviations over the full domain.

[52] Figure 13 shows the surface precipitation rate for the F30 simulation. The red curve in each plot is the large-scale forcing, just as in Figures 7–12. The black curve is the mean for each subdomain, averaged over all subdomains, so it is identical to the large-scale mean, as already shown in Figure 7. The point of Figure 13 is that, as the averaging domain is made smaller, the variability in the response, as measured by the standard deviation over realizations, increases greatly. Near the peak precipitation response for the F30 simulation, the standard deviation of the surface precipitation rate grows by about a factor of 50, from roughly one tenth of a millimeter per hour for the full domain to nearly 5 mm per hour in the 256th. Obviously, the negative values that appear in the standard deviation curves are simply statistical artifacts and do not imply negative precipitation.

[53] The applied large-scale forcing is not the effective forcing for a subdomain. In order to investigate the trend of lag times as the subdomain area decreases, we need to define the local forcing. It was calculated based on the following relations:

$$\frac{\partial \overline{\theta}}{\partial t} = F_0 - \nabla \cdot \left(\overline{\mathbf{V}}\,\overline{\theta}\right) - \frac{\partial}{\partial z}\left(\overline{w}\,\overline{\theta}\right) + \frac{\theta}{\rho_0 c_p T} LC - \frac{\partial}{\partial z}\left(\overline{w'\theta'}\right) \quad (7)$$

$$\frac{\partial \overline{q_v}}{\partial t} = F_1 - \nabla \cdot \left(\overline{\mathbf{V}}\,\overline{q_v}\right) - \frac{\partial}{\partial z}\left(\overline{w}\,\overline{q_v}\right) - C - \frac{\partial}{\partial z}\left(\overline{w'q_v'}\right) \quad (8)$$

Here, F_0 and F_1 refer to the imposed large-scale advective tendencies of potential temperature and water vapor, respectively. C is the condensation rate, and L is the latent heat of condensation. Overbars denote horizontal means over a subdomain, and primes are deviations from that mean. In equations (7) and (8), the first three terms on the right-hand sides denote the "local" forcing and the last two denote the response. It is the vertically integrated effect of both equations that is of interest. Vertical integration of the forcing eliminates the vertical advection terms, leaving only the imposed large-scale forcing, the CRM-resolved horizontal flux divergence, and the surface fluxes. If we further integrated over the entire domain, the horizontal flux divergence would drop out, leaving just the imposed large-scale forcing and the surface fluxes.

[54] In QE, the forcing and response are highly correlated. Figure 14a shows the correlation between the local forcing (as defined by combining (7) and (8) to obtain the local vertically integrated moist static energy forcing) and the response, for the full, 1/4, 1/16, and 1/64, and 1/256 subdomains, and for a forcing period of 30 h. For the 265 km domain, the correlation is about 0.9, consistent with QE. For the 128 km domain, the correlation drops drastically, to about 0.1. As the domain size decreases further, the correlation gradually increases, reaching about 0.5 for the 16 km domain. The correlation increases for smaller domain sizes because the forcing and response both become very large, because the vertical velocities become large, while the local time rate of change remains fairly small. The dramatic increase of the root-mean-square forcing for the smaller domains is shown in Figure 14b, in which the ordinate is logarithmically scaled. In the limit of "cloud resolving" resolution (not shown), the distinction between the forcing and the response disappears.

[55] The relationship between subdomain area and lag time between the local forcing and response as defined by equations (7) and (8) is irregular and there are only six points to consider; as a result, we were not able to find a simple relationship. When calculating the lag between the this local forcing and the surface precipitation rate for the F30 case, there is also a general, but irregular, trend such that lag time tends to decrease as subdomain area decreases, though the converse is true in a few cases. This trend is in general agreement with the findings of *Xu and Arakawa* [1992] and *Xu and Randall* [1998].

[56] Similar results are seen in other simulations. For example, in simulations F08 and F120, the standard deviation increases greatly with decreasing averaging area (not shown), and the data do not meet the requirements for QE

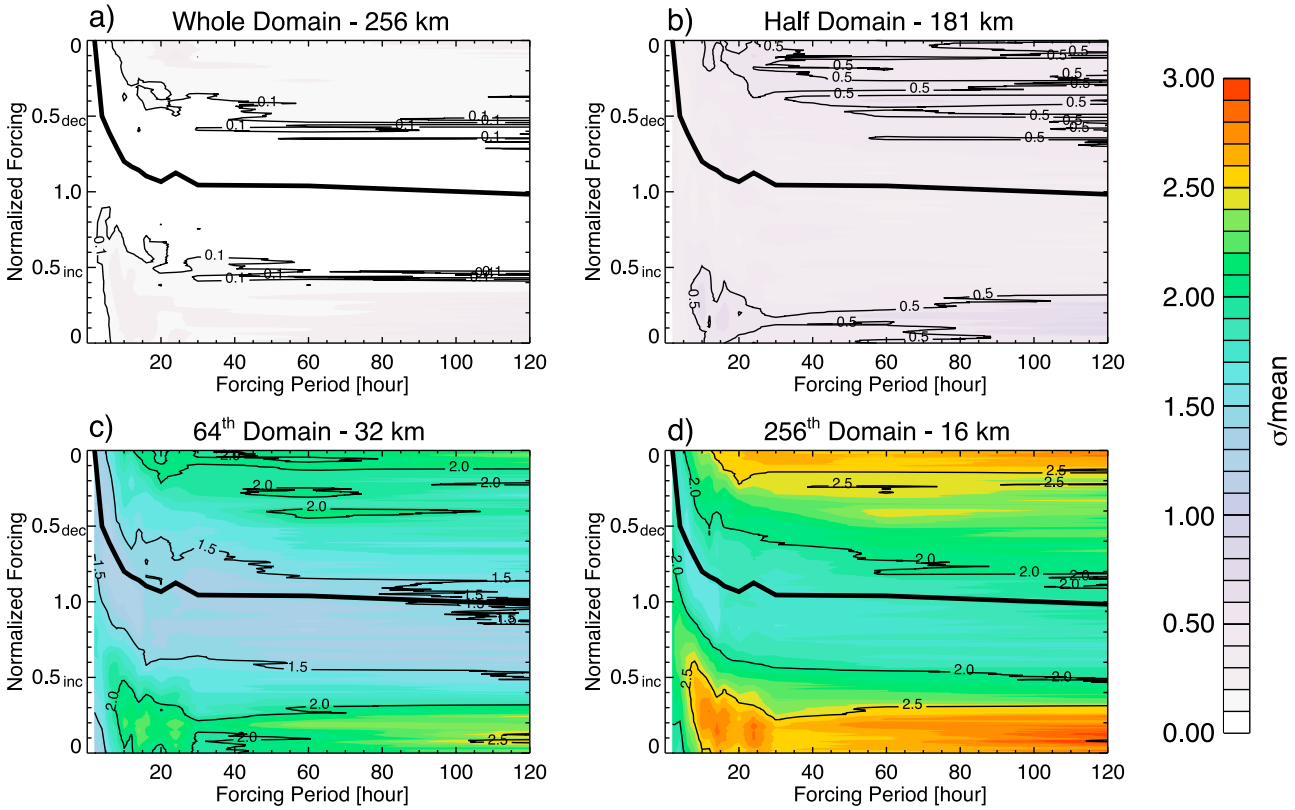


Figure 15. The coefficient of variation of surface precipitation rate for the indicated subdomains. The abscissa is the forcing period, and the ordinate is the normalized forcing, which is a proxy for the phase of the forcing, or for time. The forcing increases upward along the ordinate and reaches its maximum halfway up. The thick line is a simple graph of the average time (in the sense of forcing phase) of the maximum mean response. For periods longer than about 20 h, the response is a maximum when the forcing is a maximum.

for subdomains smaller than half of the full domain size. We also examined the F30 composite analysis results, for the six domain sizes, for cloud fraction, NPC, and net cloud mass flux through the ~2.5 km level (not shown). The changes in variability with domain size are similar to those shown in Figure 13. The standard deviation about the mean tends to increase by at least an order of magnitude when the averaging area decreases from the full domain to the 1/256th domain. This same increase in the variability over realizations is also seen in the other simulations (not shown). Further analysis of the relationship between lag and subdomain size for these variables did not yield a coherent trend.

[57] For what grid spacing does nondeterministic convection become important? In section 4.1 and Figures 3 and 4, we showed for the full domain, the COV is about 0.1. This means that, with 256 km grid spacing, a deterministic parameterization would be expected to have errors on the order of 10% of the mean, for this weather regime. The standard deviation might decrease if the domain was made larger with the same imposed large-scale forcing, but on the other hand a grid spacing coarser than 256 km is not good for simulating the large-scale circulation. As discussed below, with a smaller domain size, the COV increases. We conclude that, for the weather regime considered here, and with a grid spacing of 256 km, a "perfect" deterministic parameterization will have errors on the order of 10%. Evidently stochastic parameterization can be useful even with a 256 km grid spacing.

4.2.4. Extended Composite Analysis

[58] This section contains an extension of the information contained in sections 4.2.2 and 4.2.3 covering all tested periods and subdomains, and focusing solely on the surface

Table 1. The Coefficient of Variation as a Function of Large-Scale Forcing Period Length and Subdomain Size, Computed for Each Subdomain and Averaged Over the Length of 15 Realizations

Period (hours)	Subdomain Side Length (km)					
	256	181	128	64	32	16
2	0.077	0.412	0.612	1.058	1.515	2.020
4	0.075	0.301	0.486	0.963	1.474	1.991
6	0.094	0.345	0.547	1.035	1.548	2.056
8	0.125	0.394	0.608	1.095	1.629	2.147
10	0.121	0.442	0.688	1.185	1.719	2.205
12	0.123	0.431	0.662	1.179	1.706	2.202
14	0.117	0.407	0.628	1.156	1.698	2.209
16	0.125	0.452	0.698	1.205	1.745	2.215
20	0.113	0.433	0.686	1.185	1.715	2.200
24	0.124	0.450	0.698	1.222	1.760	2.226
30	0.113	0.419	0.656	1.177	1.693	2.185
60	0.116	0.418	0.664	1.222	1.760	2.227
120	0.147	0.454	0.707	1.282	1.815	2.257

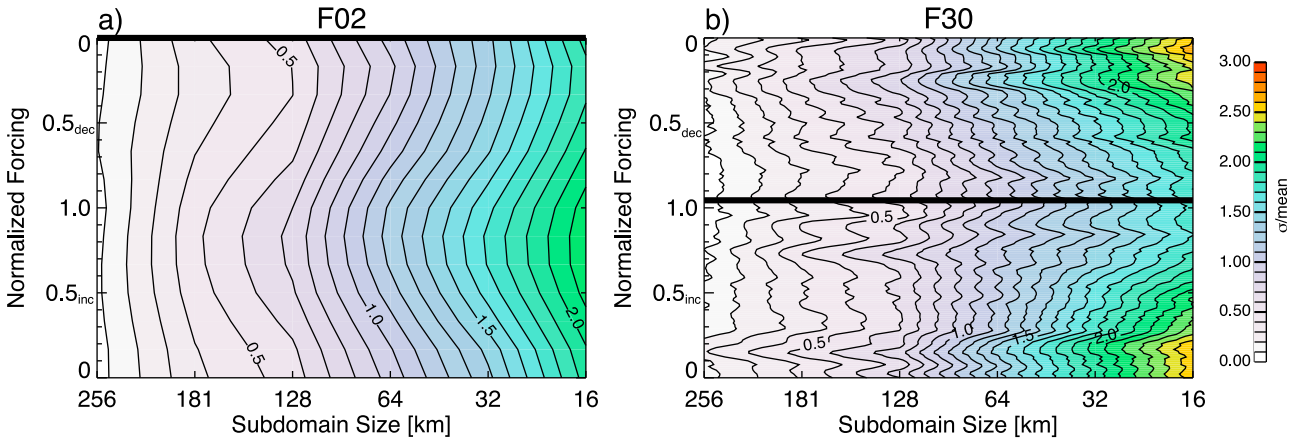


Figure 16. As in Figure 15 but for the coefficient of variation of surface precipitation rate for the noted large-scale forcing periods at all subdomains and large-scale forcings. The thick black line in each plot is the time of the peak response.

precipitation rate. In the case of the full-domain standard deviation (across realizations) of the surface precipitation rate, there is little discernible dependence on forcing strength or period length, consistent with Figure 7. The COV does have such dependencies, however.

[59] Figure 15 shows how the COV of the surface precipitation rate varies with the period of the forcing, and with time during a forcing cycle, for four different subdomain sizes. The abscissa of Figure 15 is the forcing period, and the ordinate is the normalized forcing, which is a proxy for the phase of the forcing, or for time. The forcing increases upward along the ordinate, and reaches its maximum halfway up. The thick line is a simple graph of the average time (in the sense of forcing phase) of the maximum mean response. For periods longer than about 20 h, the response is a maximum when the forcing is a maximum. The contours show the COV. Comparison of Figures 15a–15d shows that the COV increases as the domain size decreases. In addition, the COV varies over a cycle (vertically in the contour plots), by up to a factor of two. The COV tends to be largest when the large-scale forcing is weak, a time when any convection is likely to be dominated by self-sustaining mesoscale convective systems. There is a weak tendency for longer periods to have larger COVs, indicating that the predictability of the response does not depend strongly on the rate of change of the large-scale forcing. Similar results were obtained for the net cloud mass flux, the cloud fraction, and the NPC (not shown).

[60] Table 1 shows the same results in a different way. The entries in Table 1 are the COV for the precipitation rate, averaged over each full simulation. This averaged COV increases strongly as the subdomain size decreases, but it depends only weakly on the period of the forcing.

[61] Figure 16 shows how the COV varies with subdomain size and through a cycle, for two very different forcing period lengths. The thick black line in each plot is the time of the peak response. Adjusting for the location of the peak mean response, the COVs are very similar even though the forcing periods are very different. For both periods, the COV tends to increase with decreasing grid spacing. It is the largest away from the time of the peak response. This second result is enhanced in the presence of greater mesoscale contributions to convection. The greatest difference between the two period lengths is the variability of the COV with forcing strength (vertically in the contour plots), which is greater for longer periods, because strong forcing is applied for a longer time.

[62] The COV increases nonlinearly with decreasing domain size along the heavy black maximum mean response line in Figure 16, roughly according to

$$COV = c - b\ln(A), \tag{9}$$

where b and c are constants that depend on the variable in question, and A is the subdomain area. The constant b controls the shape of the curve, while c most strongly controls the domain size for which the standard deviation becomes small (greater c shifts the zero to a larger area). One might expect that there would be a $1/\sqrt{A}$ dependence instead, but r^2 was greater than 0.99 for equation (9) versus 0.7 for a $1/\sqrt{A}$ dependence.

[63] Table 2 gives the best fit values of b and c for surface precipitation, cloud fraction, net cloud mass flux, and NPC. All of them have high correlation coefficients showing that the fits are quite good. The values of b are similar for the precipitation rate mass flux, and they are similar for the cloud fraction and NPC. This supports the idea, mentioned earlier, that there are two distinct variable types, namely the "storage" variables, NPC and cloud fraction, and the "flux" variables, surface precipitation rate and net cloud mass flux. The variation with subdomain size is stronger for the flux variables, consistent with the idea that the storage variables are advected between subdomains and have significant inertia of "Lagrangian memory."

Table 2. List of Parameters b and c Corresponding to the Best Fit Curve Given by Equation (9) and the Correlation Coefficient for Each Best Fit Curve

Variable	b	c	r^2
Surface precipitation rate	0.283	0.091	0.9976
Cloud fraction	0.106	0.075	0.9974
Cloud mass flux	0.253	0.054	0.9988
Nonprecipitating condensate	0.132	0.084	0.9962

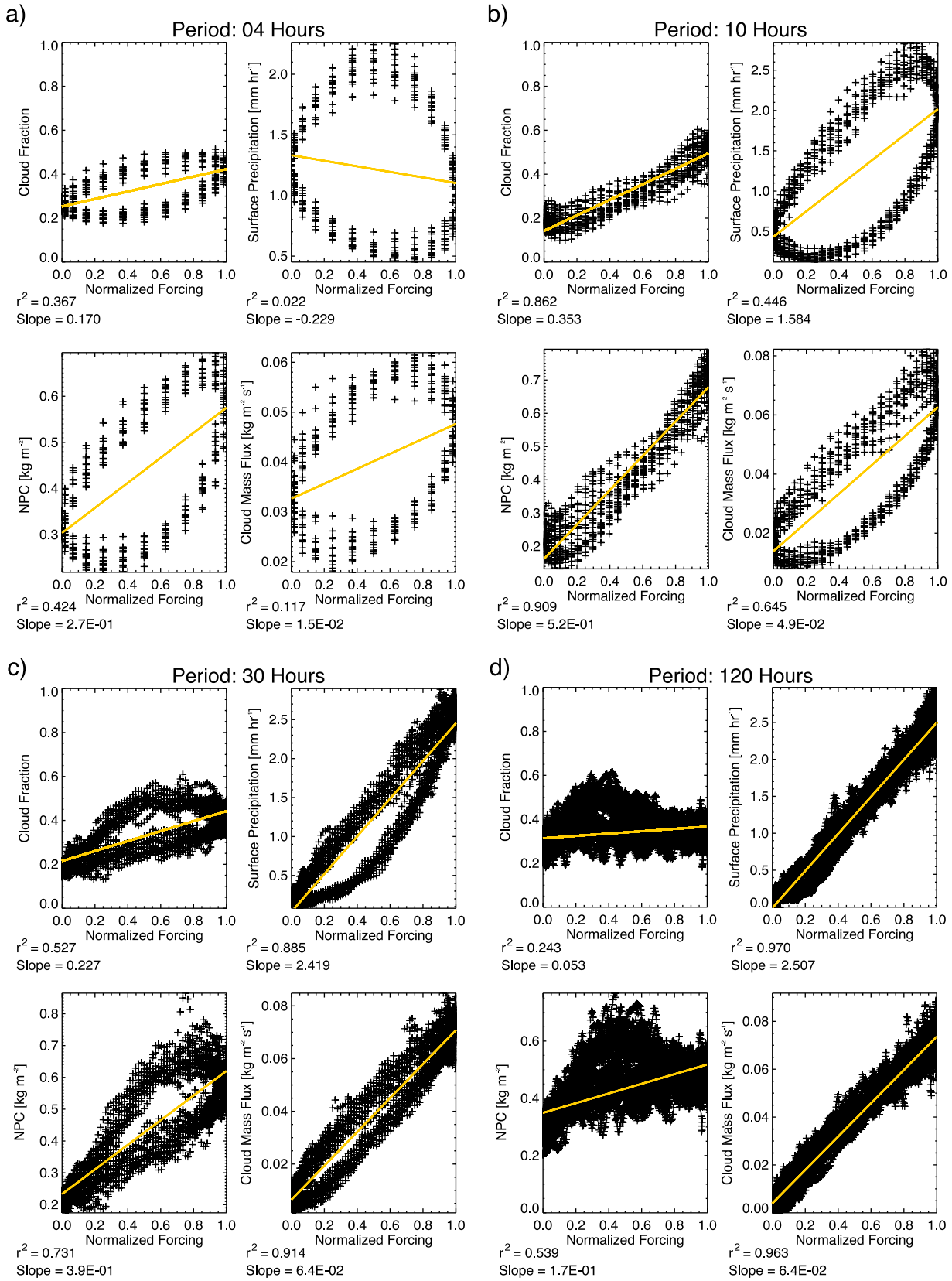


Figure 17. Scatterplots for the full 15 cycles of the noted simulations of the full domain average cloud fraction, surface precipitation rate, nonprecipitating condensate, and cloud mass flux versus the normalized forcing. The golden line is the linear regression trend line. The slope of this line and the correlation coefficient (r^2) are given to the bottom left of each plot.

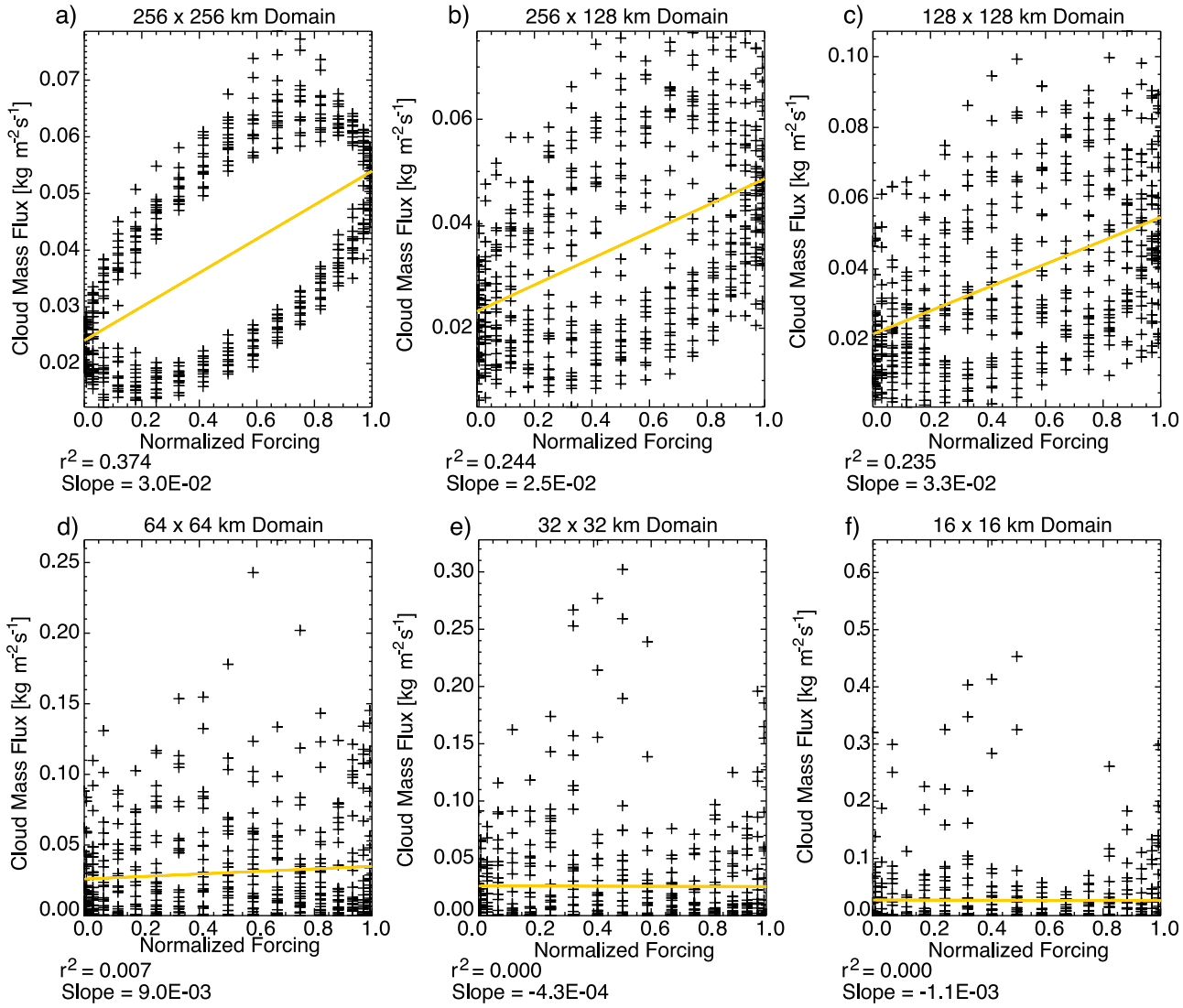


Figure 18. Scatterplots for the full 15 cycles of the F06 simulation of the (a) full domain, (b) half domain, (c) quarter domain, (d) 16th domain, (e) 64th domain, and (f) 256th domain average values of cloud mass flux versus the normalized forcing. The golden line is the linear regression trend line. The slope of this line and the correlation coefficient (r^2) are given to the bottom left of each plot.

4.2.5. Correlation Analysis

[64] If the convective ensemble is close to QE, the forcing should closely track but slightly lead the response. Figure 17 shows scatterplots of the response as a function of the forcing for simulations F04, F10, F30, and F120. For the shorter periods, the scatterplots take the shape of hysteresis loops. The upper and lower portions of the loop are associated with increasing or decreasing forcing, depending on the period and variable in question. For instance, the lower portion of the scatterplot in Figure 17a, for cloud fraction in the F04 simulation, corresponds to increasing forcing. When the forcing period is changed to 30 h, as in Figure 17c, it is the upper portion of the loop that corresponds to the increasing forcing. Recall that the cloud fraction maximum leads the forcing maximum for a longer period. There is a tendency for the data points to "hug" the regression line more tightly as the period length is increased. This is consistent with QE for the longer periods.

[65] The convergence toward a single regression line is very clear for the "flux variables," surface precipitation rate and net cloud mass flux. The "storage variables" show a more complicated behavior. The slopes of the flux variables monotonically increase from short to long periods, with a slowing of the rate of change toward longer periods. More interesting, perhaps, are the values to which the slopes converge. Let us focus on the case of the surface precipitation rate. Referring once again to Figure 4a, the expected slope for a simulation run at an approximation to statistical equilibrium is ~2.5 mm h^{-1} per a full increase of the large-scale forcing, and interestingly enough, that also the slope for simulations F60 (2.522 mm h^{-1}, not shown) and F120. This supports the idea that while statistical equilibrium is not well approximated for rapidly varying large-scale forcings, it is well approximated for slowly varying forcings. For the surface precipitation rate, the correlation of the convective statistic with the phase of the forcing increases

with increasing period length. The correlation coefficient for F30 is 0.885, and it is 0.970 for the F120 simulation. QE holds for long periods on the full domain but not for the short ones.

[66] For the F06 simulation, Figure 18 shows the scatterplots, regression lines, and correlation coefficients for the whole, half, quarter, 16th, 64th, and 256th domain as applied to the net cloud mass flux. In these plots, we can see the range of variation of the net cloud mass flux for each subdomain size. Note that the ordinate axis values differ among the plots. The correlation coefficients fall off quickly with decreasing domain size. This shows that convection sampled over small subdomains is not close to statistical equilibrium. There is no clear trend for the regression slopes. There may be an increase in regression slope with decreasing domain size, but this is generally due to very large average cloud mass flux values that are the result of averaging over such a small area and it does not hold over all averaging areas. When the subdomain size is very small, there are cases in which the slope becomes negative due to the sheer number of values that are much smaller than the maxima. Extending this analysis to other variables (not shown), it is clear that that correlations become small for subdomains smaller than the full domain.

[67] Considering the same data in terms of the local forcing as defined in equations (7) and (8), there are some subtle differences that would arise in a plot like Figure 18. Local forcing tends to relate to surface precipitation and cloud mass flux in a more variable fashion than the large-scale forcing. This is especially true for the largest local forcings, which appear in the smallest subdomains. As an example, over the range of the forcing seen in the whole domain average, which is only about 2% of the range seen at the 16 km subdomain, precipitation rates deviate from expected QE rates (as determined by linear regression of the forcing and precipitation) by maximum of 1.1 mm h^{-1} for the 256 km domain and by 11.4, 20.5, 45.6, and 119.9 mm h^{-1} for the 128, 64, 32, and 16 km subdomains. These values were obtained by an additional F30 simulation with output recorded every minute.

5. Concluding Discussion

[68] Cloud parameterization will be needed for the foreseeable future, because cloud resolving models will continue to be too computationally expensive with large domain sizes. The results of our experiments highlight need for new approaches to convection parameterization. Deterministic and diagnostic QE assumptions, while useful, break down at resolutions that are now commonly used in weather prediction models and are starting to be used in climate models.

[69] We have investigated the departures from convective QE, with a view to motivating and guiding the development of nonequilibrium and nondeterministic parameterizations. A CRM was driven by periodic large-scale forcing. For comparison, the model was also subjected to a series of constant forcing simulations for a variety of large-scale forcing magnitudes. We analyzed the results to see how the nonequilibrium and nondeterministic response of the convection vary with domain size and forcing period.

[70] The more slowly the forcing varies, the better the response is approximated by the equilibrium solution. QE approximations fail when the time scale for changes in the resolved-scale weather is close to or less than the convective adjustment time. This is a problem is more severe in newer high-resolution models [e.g., *Arribas*, 2004; *Bechtold et al.*, 2008] simply because shorter time scales are inherent in convective systems with smaller spatial scales.

[71] For deep convection, QE is viable when the time scale for the variation of the large-scale forcing is greater than about 30 h. For some fields, variability increased with forcing period length, suggesting a role for mesoscale organization. Lags arise from a number of factors, including mesoscale organization, which allows the convection to become self-sustaining. Lags do not make parameterization impossible. They can be dealt with by endowing the parameterization with a memory.

[72] The breakdown of deterministic prediction was anticipated by *Lorenz* [1975], who wrote "I believe that the ultimate climatic models…will be stochastic, i.e., random numbers will appear somewhere in the time derivatives." For a given weather regime, nondeterministic variability is much stronger with smaller domain sizes; this is the problem of insufficient sample size, which grows in importance as grid spacing decreases. Our results show that the coefficient of variation is nearly independent of the magnitude of the forcing, lending support to the assumption that variability scales with the mean. The COV was found to follow a logarithmic function of the domain area.

[73] Our results suggest that QE applies for domain sizes of 256 km or greater, and for forcing periods of 30 h or greater. Even then, a deterministic parameterization will have "expected errors" on the order of 10%.

[74] **Acknowledgments.** This work was made possible through support from the Atmospheric Radiation Measurement (ARM) program, award DE-0000000SC01904, sponsored by the U.S. Department of Energy's Office of Science. We would like to acknowledge computing support provided by the National Energy Research Scientific Computing Center (NERSC), at the Lawrence Berkeley Laboratory of the U.S. Department of Energy. The authors would like to thank Joon-Hee Jung for providing the model code and guidance and reviewers Zhiming Kuang, Kuan-Man Xu, and Kerry Emanuel for their very useful comments and discussion.

References

Abdullaev, S. M., A. A. Zhelnin, and O. Y. Lenskaya (2009), Life cycle of mesoscale convective systems, *Russ. Meteorol. Hydrol.*, *34*, 285–292, doi:10.3103/S1068373909050033.

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

Arribas, A. (2004), Results of an initial stochastic physics scheme for the Met Office Unified Model, *Forecast. Res. Tech. Rep.*, *452*, Met Off., Exeter, U. K.

Ball, M. A., and R. S. Plant (2008), Comparison of stochastic parameterization approaches in a single-column model, *Philos. Trans. R. Soc. A*, *366*, 2603–2621, doi:10.1098/rsta.2008.0050.

Bechtold, P., M. Köhler, 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.

Benedict, J. J., and D. A. Randall (2009), MJO structure in the superparameterized CAM, *J. Clim.*, *66*, 3277–3296.

Berner, J., F. J. Doblas-Reyes, T. N. Palmer, G. Shutts, and A. Weisheimer (2008), Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model, *Philos. Trans. R. Soc. A*, *366*, 2559–2577, doi:10.1098/rsta.2008.0033.

Buizza, R., M. Miller, and T. N. Palmer (1999), Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *125*, 2887–2908, doi:10.1002/qj.49712556006.

Buizza, R., J. Barkmeijer, T. Palmer, and D. Richardson (2000), Current status and future developments of the ECMWF ensemble prediction system, *Meteorol. Appl.*, *7*, 163–175, doi:10.1017/S1350482700001456.

Cohen, B. G., and G. C. Craig (2006), Fluctuations in an equilibrium convective ensemble. Part II: Numerical experiments, *J. Atmos. Sci.*, *63*, 2005–2015, doi:10.1175/JAS3710.1.

Cole, J. N. S., H. W. Barker, W. O'Hirok, E. E. Clothiaux, M. F. Khairoutdinov, and D. A. Randall (2005a), Atmospheric radiative transfer through global arrays of 2D clouds, *Geophys. Res. Lett.*, *32*, L19817, doi:10.1029/2005GL023329.

Cole, J. N., H. W. Barker, D. A. Randall, M. F. Khairoutdinov, and E. Clothiaux (2005b), Interactions between clouds and radiation at scales unresolved by global climate models, *Geophys. Res. Lett.*, *32*, L06703, doi:10.1029/2004GL020945.

Deardorff, J. W. (1972), Parameterization of the planetary boundary layer for use in general circulation models, *Mon. Weather Rev.*, *100*, 93–106.

DeMott, C. A., D. A. Randall, and M. Khairoutdinov (2007), Convective precipitation variability as a tool for general circulation model analysis, *J. Clim.*, *20*, 91–112, doi:10.1175/JCLI3991.1.

DeMott, C. A., D. A. Randall, and M. Khairoutdinov (2010), Implied ocean heat transports in the standard and super-parameterized Community Atmosphere Models, *J. Clim.*, *23*, 1908–1928, doi:10.1175/2009JCLI2987.1.

Emanuel, K. A., J. D. Neelin, and C. S. Bretherton (1994), On large-scale circulations in convecting atmospheres, *Q. J. R. Meteorol. Soc.*, *120*, 1111–1143, doi:10.1002/qj.49712051902.

Fridlind, A., A. Ackerman, J. Petsch, P. Field, A. Hill, G. McFarquhar, S. Xie, and M. Zhang (2010), ARM/GCSS/SPARC TWP-ICE CRM intercomparison study, *NASA Tech. Memo., NASA-TM-2010-215858*.

Hohenegger, C., and C. Schär (2007), Atmospheric predictability at synoptic versus cloud- resolving scales, *Bull. Am. Meteorol. Soc.*, *88*, 1783–1793, doi:10.1175/BAMS-88-11-1783.

Horinouchi, T., et al. (2003), Tropical cumulus convection and upward-propagating waves in middle-atmospheric GCMs, *J. Atmos. Sci.*, *60*, 2765–2782, doi:10.1175/1520-0469(2003)060<2765:TCCAUW>2.0.CO;2.

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

Jones, T. R. (2010), Quantifying the limits of convective parameterizations: A statistical characterization of simulated cumulus convection, M.S. thesis, Dep. of Atmos. Sci., Colo. State Univ., Fort Collins.

Jung, J.-H., and A. Arakawa (2008), A three-dimensional anelastic model based on the vorticity equation, *Mon. Weather Rev.*, *136*, 276–294, doi:10.1175/2007MWR2095.1.

Katzfey, J. J., and B. F. Ryan (2000), Midlatitude frontal clouds: GCM-scale modeling implications, *J. Clim.*, *13*, 2729–2745, doi:10.1175/1520-0442(2000)013<2729:MFCGSM>2.0.CO;2.

Khairoutdinov, M. F., and D. A. Randall (2001), A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model: Preliminary results, *Geophys. Res. Lett.*, *28*, 3617–3620, doi:10.1029/2001GL013552.

Khairoutdinov, M., D. A. Randall, and C. A. DeMott (2005), Simulation of the atmospheric general circulation using a cloud-resolving model as a super-parameterization of physical processes, *J. Atmos. Sci.*, *62*, 2136–2154, doi:10.1175/JAS3453.1.

Khairoutdinov, M., C. A. DeMott, and D. A. Randall (2008), Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU Multiscale Modeling Framework, *J. Clim.*, *21*, 413–431, doi:10.1175/2007JCLI1630.1.

Krueger, S. K. (1988), Numerical simulation of tropical cumulus clouds and their interaction with the subcloud layer, *J. Atmos. Sci.*, *45*, 2221–2250, doi:10.1175/1520-0469(1988)045<2221:NSOTCC>2.0.CO;2.

Krueger, S. K., Q. Fu, K. N. Liou, and H.-N. Chin (1995), Improvements of an ice-phase microphysics parameterization for use in numerical simulations of tropical convection, *J. Appl. Meteorol.*, *34*, 281–287.

Lin, J. W.-B., and J. D. Neelin (2002), Considerations for stochastic convective parameterization, *J. Atmos. Sci.*, *59*, 959–975, doi:10.1175/1520-0469(2002)059<0959:CFSCP>2.0.CO;2.

Lin, Y.-L., R. D. Farley, and H. D. Orville (1983), Bulk parameterization of the snow field in a cloud model, *J. Clim. Appl. Meteorol.*, *22*, 1065–1092.

López, R. E. (1977), The lognormal distribution and cumulus cloud populations, *Mon. Weather Rev.*, *105*, 865–872, doi:10.1175/1520-0493(1977)105<0865:TLDACC>2.0.CO;2.

Lord, S. J., H. E. Willoughby, and J. M. Piotrowicz (1984), Role of a parameterized ice-phase microphysics in an axisymmetric, nonhydrostatic tropical cyclone model, *J. Atmos. Sci.*, *41*, 2836–2848, doi:10.1175/1520-0469(1984)041<2836:ROAPIP>2.0.CO;2.

Lorenz, E. N. (1969), The predictability of a flow which possesses many scales of motion, *Tellus*, *21*, 289–307, doi:10.1111/j.2153-3490.1969.tb00444.x.

Lorenz, E. N. (1975), Climatic predictability, in *The Physical Basis of Climate and Climate Modelling*, *WMO GARP Publ. Ser.*, vol. 16, pp. 132–36, World Meteorol. Organ., Geneva.

Molinari, J., and M. Dudek (1992), Parameterization of convective precipitation in mesoscale numerical models: A critical review, *Mon. Weather Rev.*, *120*, 326–344, doi:10.1175/1520-0493(1992)120<0326:POCPIM>2.0.CO;2.

Neelin, J. D., O. Peters, J. W.-B. Lin, K. Hales, and C. E. Holloway (2008), Rethinking convective quasi-equilibrium: Observational constraints for stochastic convective schemes in climate models, *Philos. Trans. R. Soc. A.*, *366*, 2579–2602, doi:10.1098/rsta.2008.0056.

Ormsby, J. F. A. (1961), Design of numerical filters with application to missile data processing, *J. ACM*, *8*, 440–466, doi:10.1145/321075.321087.

Palmer, T. N. (1996), On parametrizing scales that are only somewhat smaller than the smallest re- solved scales with application to convection and orography, paper presented at Workshop on New Insights and Approaches to Convective Parametrization, Eur. Cent. for Med.-Range Weather Forecasts, Reading, U. K., 4–7 Nov.

Palmer, T. N., and P. D. Williams (2008), Introduction. Stochastic physics and climate modeling, *Philos. Trans. R. Soc. A*, *366*, 2419–2425, doi:10.1098/rsta.2008.0059.

Palmer, T. N., G. J. Shutts, R. Hagerdorn, F. J. Doblas-Reyes, T. Jung, and M. Leutbecher (2005), Representing model uncertainty in weather and climate prediction, *Annu. Rev. Earth Planet. Sci.*, *33*, 163–193, doi:10.1146/annurev.earth.33.092203.122552.

Pan, D.-M., and D. A. Randall (1998), A cumulus parameterization with a prognostic closure, *Q. J. R. Meteorol. Soc.*, *124*, 949–981.

Räisänen, P., H. W. Barker, M. F. Khairoutdinov, J. Li, and D. A. Randall (2004), Stochastic generation of subgrid-scale cloudy columns for large-scale models, *Q. J. R. Meteorol. Soc.*, *130*, 2047–2067, doi:10.1256/qj.03.99.

Randall, D. A., P. Ding, D.-M. Pan, and G. G. Cripe (1997), Quasi-equilibrium, in *The Physics and Parameterization of Moist Atmospheric Convection*, edited by R. K. Smith, pp. 359–385, Kluwer Acad., Dordrecht, Netherlands.

Randall, D. A., M. Khairoutdinov, A. Arakawa, and W. Grabowski (2003), Breaking the cloud-parameterization deadlock, *Bull. Am. Meteorol. Soc.*, *84*, 1547–1564, doi:10.1175/BAMS-84-11-1547.

Ricciardulli, L., and R. R. Garcia (2000), The excitation of equatorial waves by deep convection in the NCAR Community Climate Model (CCM3), *J. Atmos. Sci.*, *57*, 3461–3487, doi:10.1175/1520-0469(2000)057<3461:TEOEWB>2.0.CO;2.

Shutts, G. J., and M. E. B. Gray (1994), A numerical modeling study of the geostrophic adjustment process following deep convection, *Q. J. R. Meteorol. Soc.*, *120*, 1145–1178.

Stan, C., M. Khairoutdinov, C. A. DeMott, V. Krishnamurthy, D. M. Straus, D. A. Randall, J. L. Kinter III, and J. Shukla (2010), An ocean-atmosphere climate simulation with an embedded cloud resolving model, *Geophys. Res. Lett.*, *37*, L01702, doi:10.1029/2009GL040822.

Sui, C.-H., and M. Yanai (1986), Cumulus ensemble effects on the large-scale vorticity and momentum fields of GATE, part I: Observational evidence, *J. Atmos. Sci.*, *43*, 1618–1642.

Tao, W.-K., et al. (2009), A multi-scale modeling system: Developments, applications, and critical issues, *Bull. Am. Meteorol. Soc.*, *90*, 515–534, doi:10.1175/2008BAMS2542.1.

Thayer-Calder, K., and D. A. Randall (2009), The role of convective moistening in the formation and progression of the MJO, *J. Clim.*, *66*, 3297–3312.

Wilks, D. S. (2005), Effects of stochastic parameterizations in the Lorenz '96 system, *Q. J. R. Meteorol. Soc.*, *131*, 389–407, doi:10.1256/qj.04.03.

Wilks, D. S. (2008), Effects of stochastic parameterization on conceptual climate models, *Philos. Trans. R. Soc. A*, *366*, 2475–2488, doi:10.1098/rsta.2008.0005.

Williams, P. D. (2005), Modelling climate change: The role of unresolved processes, *Philos. Trans. R. Soc. A*, *363*, 2931–2946, doi:10.1098/rsta.2005.1676.

Xu, K.-M., and A. Arakawa (1992), Semiprognostic tests of the Arakawa-Schubert parameterization using simulated data, *J. Atmos. Sci.*, *49*, 2421–2436, doi:10.1175/1520-0469(1992)049<2421:STOTAS>2.0.CO;2.

Xu, K.-M., and D. A. Randall (1998), Influence of large scale advective cooling and moistening effects on the quasi-equilibrium behavior of explicitly simulated cumulus ensembles, *J. Atmos. Sci.*, *55*, 896–909, doi:10.1175/1520-0469(1998)055<0896:IOLSAC>2.0.CO;2.

Xu, K.-M., A. Arakawa, and S. K. Krueger (1992), The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model, *J. Atmos. Sci.*, *49*, 2402–2420, doi:10.1175/1520-0469(1992)049<2402:TMBOCE>2.0.CO;2.

Yu, J.-Y., and J. D. Neelin (1997), Analytic approximations for moist convectively adjusted regions, *J. Atmos. Sci.*, *54*, 1054–1063, doi:10.1175/1520-0469(1997)054<1054:AAFMCA>2.0.CO;2.

T. R. Jones and D. A. Randall, Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523-1371, USA. (toddj@atmos.colostate.edu)