The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations

AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE

10.1002/2015MS000511

The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations

C. E. Holloway[1] and S. J. Woolnough[2]

[1]Department of Meteorology, University of Reading, Reading, UK, [2]NCAS-Climate, Department of Meteorology, University of Reading, Reading, UK

Key Points:
- Convective self-aggregation in the Unified Model is driven mainly by direct radiative feedbacks
- Low-level circulations also aid aggregation, but are not primarily forced by radiation anomalies
- Convective aggregation occurs at SSTs as low as 290 K, mainly due to cloud-radiation feedbacks

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6

Correspondence to:
C. E. Holloway,
c.e.holloway@reading.ac.uk

Citation:
Holloway, C. E., and S. J. Woolnough (2016), The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations, *J. Adv. Model. Earth Syst.*, *8*, doi:10.1002/2015MS000511.

Received 7 JUL 2015
Accepted 11 JAN 2016
Accepted article online 14 JAN 2016

Abstract Idealized explicit convection simulations of the Met Office Unified Model exhibit spontaneous self-aggregation in radiative-convective equilibrium, as seen in other models in previous studies. This self-aggregation is linked to feedbacks between radiation, surface fluxes, and convection, and the organization is intimately related to the evolution of the column water vapor field. Analysis of the budget of the spatial variance of column-integrated frozen moist static energy (MSE), following Wing and Emanuel (2014), reveals that the direct radiative feedback (including significant cloud longwave effects) is dominant in both the initial development of self-aggregation and the maintenance of an aggregated state. A low-level circulation at intermediate stages of aggregation does appear to transport MSE from drier to moister regions, but this circulation is mostly balanced by other advective effects of opposite sign and is forced by horizontal anomalies of convective heating (not radiation). Sensitivity studies with either fixed prescribed radiative cooling, fixed prescribed surface fluxes, or both do not show full self-aggregation from homogeneous initial conditions, though fixed surface fluxes do not disaggregate an initialized aggregated state. A sensitivity study in which rain evaporation is turned off shows more rapid self-aggregation, while a run with this change plus fixed radiative cooling still shows strong self-aggregation, supporting a "moisture-memory" effect found in Muller and Bony (2015). Interestingly, self-aggregation occurs even in simulations with sea surface temperatures (SSTs) of 295 and 290 K, with direct radiative feedbacks dominating the budget of MSE variance, in contrast to results in some previous studies.

1. Introduction

Self-aggregation in idealized models of radiative-convective equilibrium (RCE) has been studied extensively since at least *Held et al*. [1993], who found that 2-D cloud-system resolving model (CRM) simulations of RCE organized into a single convective region surrounded by subsidence and suppressed conditions. Early studies of nonrotating RCE with 3-D CRMs found that homogeneous initial conditions evolved into band-like convective structures, which contained larger amounts of tropospheric water vapor than in the surrounding subsidence bands, and this organization was sensitive to interactions between convection, surface fluxes, and radiation [*Tompkins and Craig*, 1998; *Tompkins*, 2001]. *Su et al*. [2000] studied self-aggregation in runs using a model at 15 km grid spacing with parameterized convection and various profiles of prescribed domain-mean vertical ascent in a biperiodic, nonrotating, large domain and found some aggregation for all runs, but they found more aggregation for a top-heavy prescribed vertical ascent profile.

Bretherton et al. [2005] looked at large, biperiodic square domains with 3 km grid spacing (and no prescribed vertical motion) in the System for Atmospheric Modeling (SAM) and found that initially scattered convection clustered into a single moist, rainy patch after about 50 days. Not only did column water vapor (CWV) and rainfall become organized within the domain, but the domain-mean CWV decreased during this self-aggregation process (because there was less moistening of the subsidence regions by nearby convective detrainment) and the domain-mean outgoing longwave radiation (OLR) increased (because of less high cloud and water vapor). *Bretherton et al*. [2005] found that homogenizing either surface fluxes or radiative heating suppressed the self-aggregation process. They argued that a low-level circulation, forced by anomalous radiative cooling in the lower troposphere in drier regions, leads to an import of moist static energy into already moist regions, amplifying aggregation. *Stephens et al*. [2008] also found that interactive radiation was necessary for self-aggregation in large "bowling alley" CRM RCE runs, but they argued that interaction between upper-level cloud radiative warming and convection was the key feedback.

Muller and Held [2012, hereinafter MH12] studied self-aggregation in the SAM model at a variety of grid lengths (as small as 500 m) and domain sizes (square width $\approx$ 50–600 km) using both homogeneous and aggregated initial conditions. They also ran a number of mechanism denial experiments building on earlier work. They found that simulations started from homogeneous conditions were more likely to self-aggregate at grid spacing coarser than 2 km and domain length greater than 200 km, although with aggregated initial conditions there was no dependence on grid spacing. They found that longwave interaction with low-level liquid cloud (particularly below 1 km) was necessary for self-aggregation from homogeneous conditions, whereas clear-sky longwave interactive radiation is the necessary feedback to maintain already aggregated conditions. They argued, similarly to *Bretherton et al.* [2005], that the mechanism responsible for self-aggregation from homogeneous conditions is a low-level circulation that is forced by anomalous low-level radiative cooling in dry regions (which, in MH12, is mainly from low-level liquid cloud). This circulation then transports moist static energy (MSE) from dry to moist regions. The resolution dependence might then stem from different subgrid mixing which can affect low cloud amounts. MH12 argued that the clear-sky longwave feedback would also rely on anomalous low-level cooling in dry regions leading to an anomalous circulation with negative Gross Moist Stability (GMS).

Jeevanjee and Romps [2013] argued that the domain-size sensitivity seen in MH12 and other studies is due to a relatively constant negative feedback due to cold pools forced by downdrafts caused by the evaporation of rainfall in the lower troposphere. In their 2-D and 3-D CRM runs, this feedback appeared to be large enough to disaggregate convection (their runs are all initialized with a moist patch) in smaller domains (which have weaker large-scale circulations which oppose the cold pool effect). Preventing rain evaporation below 1 km resulted in the suppression of cold pools and led to maintenance of aggregation at all domain sizes (and stronger aggregation than in control runs of corresponding domain size which did have aggregation). *Wing* [2014] similarly found faster self-aggregation when rain evaporation was suppressed. *Muller and Bony* [2015] found that aggregation could be maintained when rain evaporation was suppressed even in the absence of radiation feedbacks (though wind-evaporation feedbacks were also suppressed in these simulations), suggesting that in this case, convection remains in the moist areas through a "moisture-memory aggregation" mechanism analogous to that discussed in *Tompkins* [2001]. *Craig and Mack* [2013] argued that self-aggregation was a coarsening process that results from a similar feedback between convection and free-tropospheric moisture.

Wing and Emanuel [2014, hereinafter WE14] introduced a novel analysis method to study self-aggregation based on the budget of the horizontal variance of vertically integrated MSE. They found that different feedbacks (averaged over the whole domain) were important for different stages of self-aggregation (starting from homogeneous conditions using the SAM), with surface fluxes and the direct diabatic radiation effect (particularly the longwave radiation) being most important at early stages, a "convergence" effect (from horizontal transport) dominating intermediate stages, and shortwave effects (and longwave cloud effects) maintaining the mature aggregated state. They also argued that clear-sky radiation effects seemed most important for the longwave feedback at early stages in their simulations, when their self-aggregation (like that in *Bretherton et al.* [2005]) developed as a single, slowly growing dry patch within a larger moist convective environment. In the mature stage, they found that cloud longwave effects in the cloudiest, moistest columns were the largest local positive feedbacks helping to maintain aggregation within the convective cluster, although shortwave feedbacks dominated the positive feedbacks in a domain-mean sense.

Khairoutdinov and Emanuel [2010] found that self-aggregation did not occur below a sea surface temperature (SST) of 298 K (although cold SSTs did not disaggregate an already aggregated initial state). They argued that this could explain why tropical SSTs in convective regions tend to be near this value. They hypothesized that SSTs above this threshold would lead to strongly aggregated conditions, which would then cause larger mean OLR and surface cooling. This cooling would continue until the SST was below the threshold, when convection would disaggregate and radiative fluxes would warm the surface again. In this way, tropical SSTs in convective regions would exhibit self-organized criticality, which could potentially reduce climate sensitivity in the tropics. WE14 also found no self-aggregation at colder SSTs, and *Emanuel et al.* [2014] argued that clear-sky longwave radiative feedbacks (and the dependence of the sign of these feedbacks on vertical differences in emissivity that ultimately depended on SST) were responsible for this SST threshold behavior. On the other hand, satellite-based studies of aggregated convection in nature [*Tobin et al.*, 2012, 2013] suggested that increased OLR in aggregated regions may not result in net surface

cooling because of compensating shortwave and turbulent surface fluxes. *Reed et al*. [2015] performed global-tropics RCE runs (at 25 and 100 km grid spacing using parameterized convection with no rotation) and found that global-mean SSTs as low as 295 K can result in large-scale aggregation within slab-ocean simulations, but these same colder SSTs result in much less organization when they are prescribed as fixed boundary conditions, suggesting that ocean coupling may also play a role in the sensitivity of aggregation to SST. *Wing and Cronin* [2016] found self-aggregation in the SAM for long-channel CRM simulations at SSTs ranging from 280 to 310 K and hypothesized that this was because of domain geometry; they also find strong positive cloud-longwave effects contributing to self-aggregation at early times for SSTs above 285 K, and they show negative "convergence" effects at all times, in contrast to WE14. *Abbot* [2014] also found self-aggregation in the SAM for snowball-Earth CRM simulations on small square domains (although this was prevented in runs with small imposed vertical wind shear).

There is disagreement in the literature about several points regarding aggregation in idealized models, including: the importance of the boundary layer [*Jeevanjee and Romps*, 2013] versus the importance of the free troposphere [*Craig and Mack*, 2013], the importance of clear-sky radiation feedbacks [*Wing and Emanuel*, 2014; *Emanuel et al*., 2014] or cloud-radiation feedbacks [*Stephens et al*., 2008; *Muller and Held*, 2012; *Wing and Cronin*, 2016] versus a convection-moisture feedback [*Held et al*., 1993; *Tompkins*, 2001; *Craig and Mack*, 2013; *Muller and Bony*, 2015], and the importance, especially at early times, of direct diabatic (and especially radiative) feedbacks (WE14) versus circulation feedbacks (MH12).

The present study is largely motivated by a desire to understand how important the mechanisms seen in studies of idealized convective aggregation are for the real world. This question can partly be addressed by looking at observations, and two recent studies [*Tobin et al*., 2012, 2013] have shown that some observed characteristics of aggregated convection, such as lower mean CWV and higher mean OLR, are similar to those seen in idealized models. Models can also be used to better understand mechanisms and to explore aggregation in more realistic conditions. The latter approach is used in C. E. Holloway (manuscript in preparation, 2016).

Several remaining questions about idealized aggregation are particularly relevant for its application to convective organization in the real world. For instance, how long does self-aggregation take to develop (or an already aggregated state to disaggregate)? Time scales are important because even if these processes act in nature, they will compete with other, potentially faster, processes in both directions. Which mechanisms are important for various stages (and spatial scales) of aggregation? The real world has a lot of existing organized convection, and mechanisms that are important for the very early stages of self-aggregation from homogeneous conditions, while scientifically important, may not be very relevant for most real-world situations. The present study is not a thorough investigation of all of these questions, but by comparing results from a different model with previous studies, we aim to make some progress along these lines.

This paper is organized as follows: section 2 describes the model used and the experimental setup. Section 3 describes the method of calculating budget terms from available model output, section 4 presents analyses of self-aggregation in the control simulation, section 5 presents results of mechanism denial experiments, section 6 describes sensitivity to initial conditions, section 7 presents results from simulations with colder SSTs, and section 8 presents discussion and conclusions. Appendix A provides a geometric argument for the appearance of bands versus circular patches in biperiodic RCE simulations.

2. Model Setup

We use the idealized RCE setup of version 7.5 of the Met Office Unified Model (UM) [*Davies et al*., 2005], which is semi-Lagrangian and nonhydrostatic. The horizontal spacing is 4 km, with 70 vertical levels. Vertical spacing between levels ranges from tens of meters in the boundary layer to around 250 m in the free troposphere. The model top is at 40 km, and there is a sponge layer in the upper few levels. The time step is 30 s.

The model physics settings are similar to those described for the "4-km 3Dsmag" model version in *Holloway et al*. [2013]. This model version includes Smagorinsky-type subgrid mixing in the horizontal and vertical dimensions (with no local or nonlocal boundary layer scheme). There is a single-moment mixed-phase microphysics scheme with three components: ice/snow, cloud liquid water, and rain [*Wilson and Ballard*, 1999]. The model uses a CAPE-limited version of the convective parameterization that asymptotes to a

30 min CAPE time scale at zero CAPE but has a CAPE time scale that rapidly increases with increasing CAPE such that for typical tropical values and in all the simulations in this study, virtually all rainfall is generated explicitly [*Roberts*, 2003; *Lean et al*., 2008]. As mentioned in *Holloway et al*. [2013], the parameterization is on mainly to slightly increase numerical stability and generate shallow convection in conditions without mean ascent. Removing the parameterization entirely did not make a noticeable difference in self-aggregation when tested against a control in simulations run previously with a slightly different model configuration.

There is no diurnal cycle, with the interactive radiation simulations having constant incoming solar radiation of 434.12 W m^{-2} to match conditions at equatorial equinox; the effective zenith angle is 50.5°. The simulations all have a fixed constant SST and a Coriolis parameter of zero. We enforce moisture conservation with a Priestly scheme; we find fairly accurate moisture conservation based on the domain-mean moisture budget, although there is still about 0.05 mm d^{-1} of excess precipitation judging by all other terms, or about 1% of the total mean precipitation. A Charnock relationship is used to calculate surface roughness lengths interactively, and we set a minimum surface wind speed of 1 m s^{-1} within the bulk surface exchange equations.

All runs have a domain size of 576 km × 576 km with doubly periodic lateral boundaries. The runs which start with homogeneous initial conditions are given a constant profile of temperature, moisture, and pressure, and a random temperature perturbation (up to ±0.1 K) is applied at the first time step within the lowest 200 m. The initial profile for these runs comes from the mean of the final 10 days of 70 day RCE simulations, each with the same SST as the corresponding larger run, on a 96 km × 96 km domain that is too small to exhibit significant self-aggregation. These small RCE runs also provide the constant surface flux values and radiation temperature increments (again, averaged over the final 10 days) which are used for fixed forcing for mechanism denial experiments. The wind is set to zero everywhere initially in all runs that start from homogeneous conditions, and there is no nudging of wind during any of the runs.

3. Methods

Most of the analysis methods are explained in the main body of the paper, but here we discuss the model output frequency and spatial covariance term calculations. We do not output the spatial covariance terms between frozen moist static energy and surface and radiation fluxes at each time step, as in WE14, but instead calculate these terms from instantaneous hourly thermodynamic fields and hourly mean fluxes. However, the latter method, when tested for the simulation in WE14, produced very similar results to the former method except for terms representing the separate wind speed and air-sea disequilibrium components of the surface fluxes (A. Wing, personal communication, 2014). We do not decompose the surface flux term into these separate components in this study, so our results should not be affected by approximating the covariance terms in this way.

4. Self-Aggregation in Control Run

We first analyze a 40 day control simulation with a setup similar to that of *Bretherton et al*. [2005]. The SST is fixed at 300 K, and there is interactive radiation and interactive surface sensible and latent heat fluxes. The run is initialized from homogeneous thermodynamic profiles (see section 2 for details) with small perturbations in boundary layer temperature and zero initial wind fields. More details of the model are described above in section 2.

Figures 1 and 2 show daily mean precipitation, total turbulent surface heat flux and surface wind vectors, column water vapor (CWV), and outgoing longwave radiation (OLR) for day 1.5 and day 39.5 (where days are labeled by the mean time in days, and the run starts at time 0). Near the beginning of the run at day 1.5, Figure 1 shows scattered clusters of rainfall, and these tend to correspond to slightly larger surface fluxes (which are dominated by latent heat flux) and surface winds, slightly larger CWV, and slightly lower OLR. By the final day of the run (Figure 2), the convection is clustered in one general location, with high CWV and low OLR there and much lower CWV and higher OLR in the surrounding environment. High surface fluxes and relatively high wind speeds surround the precipitating region as low-level winds converge toward the convection; there are also high surface fluxes in the regions with lowest CWV, where thermodynamic near-surface disequilibria are largest. These findings are consistent with simulations using the SAM model [*Bretherton et al*., 2005;

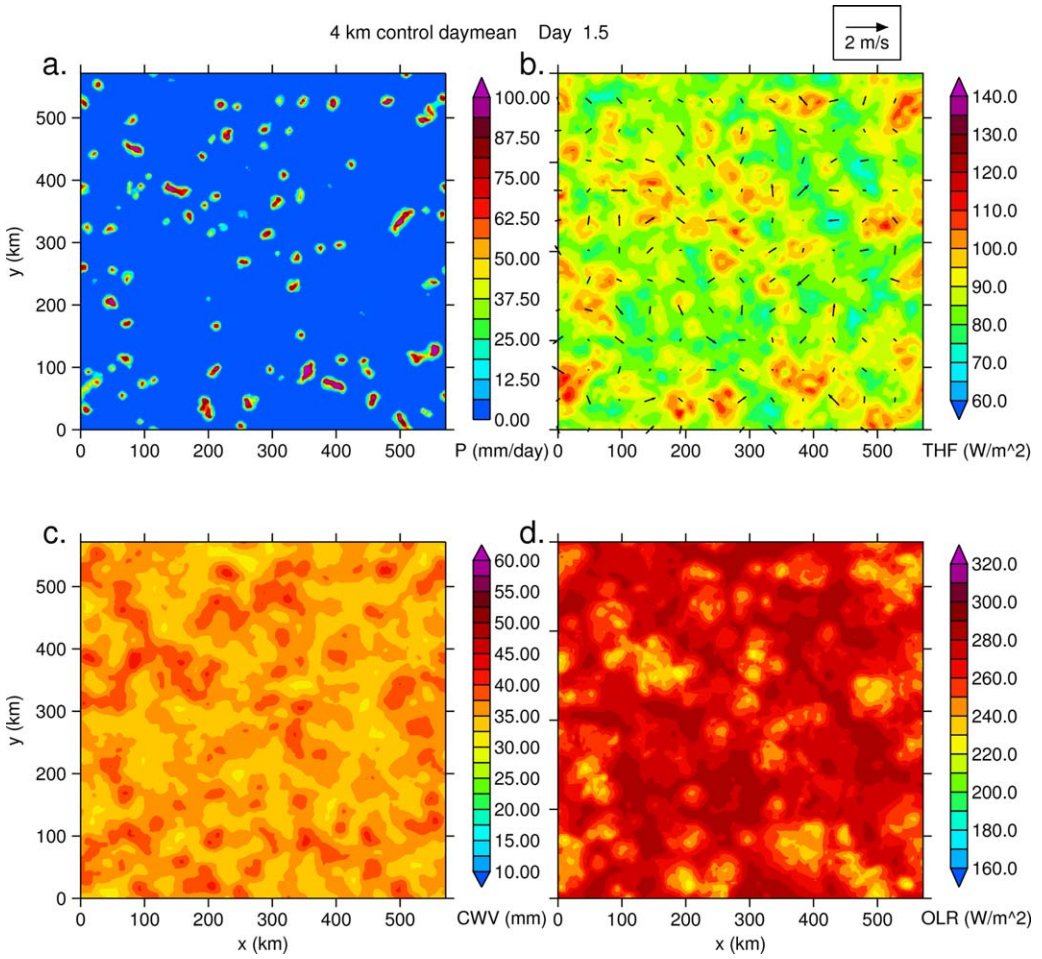

Figure 1. Day 1.5 mean values of (a) precipitation (mm d^{-1}), (b) sensible plus latent heat flux (colors, W m^{-2}) and surface wind (vectors), (c) CWV (mm), and (d) OLR (W m^{-2}) for control run with 300 K SST started from homogeneous conditions.

Muller and Held, 2012; *Wing and Emanuel*, 2014], although the region of high precipitation and CWV is less circular and exhibits more horizontal precipitation variance (even in daily means) in our simulations.

The aggregation in the control simulations evolves as follows. Some of the small moist patches of high CWV grow larger and increase in mean CWV, and the precipitation clusters mostly within these, while other smaller moist patches stop having precipitation and they shrink and their CWV is reduced to that of the environment, which is distributed rather evenly around the moist patches (not in distinct dry patches as in MH12 or WE14) and in which the CWV steadily decreases. This process continues, with moist patches growing mostly by a process resembling coarsening as described by *Craig and Mack* [2013]. During intermediate stages of aggregation, a single large moist patch sometimes forms a rectangular "band" shape, stretching completely across one horizontal dimension and a fraction of the domain in the other horizontal dimension, rather than a circular shape, and this has been seen in other studies such as WE14 and *Jeevanjee and Romps* [2013]. This is likely to be a consequence of moist patches seeking to minimize their perimeter-to-area ratio in a biperiodic domain, and patches with an area larger than about a third of a square domain do this by forming a band rather than a circle, as described below in Appendix A.

Figure S1 in the supporting information, which shows daily mean CWV plots for 6 days spanning the simulation, reveals that our control run self-aggregation consists of simultaneous growth and merging of both moist and dry patches. This can also be seen in Animation S1 showing hourly fields (including CWV) during the first 7 days of the simulation. WE14 link the early stages of self-aggregation in their simulations to clear-sky radiative feedbacks occurring within a single, expanding dry patch; however, we do not see this asymmetry in the growth of dry versus moist patches in our simulations.

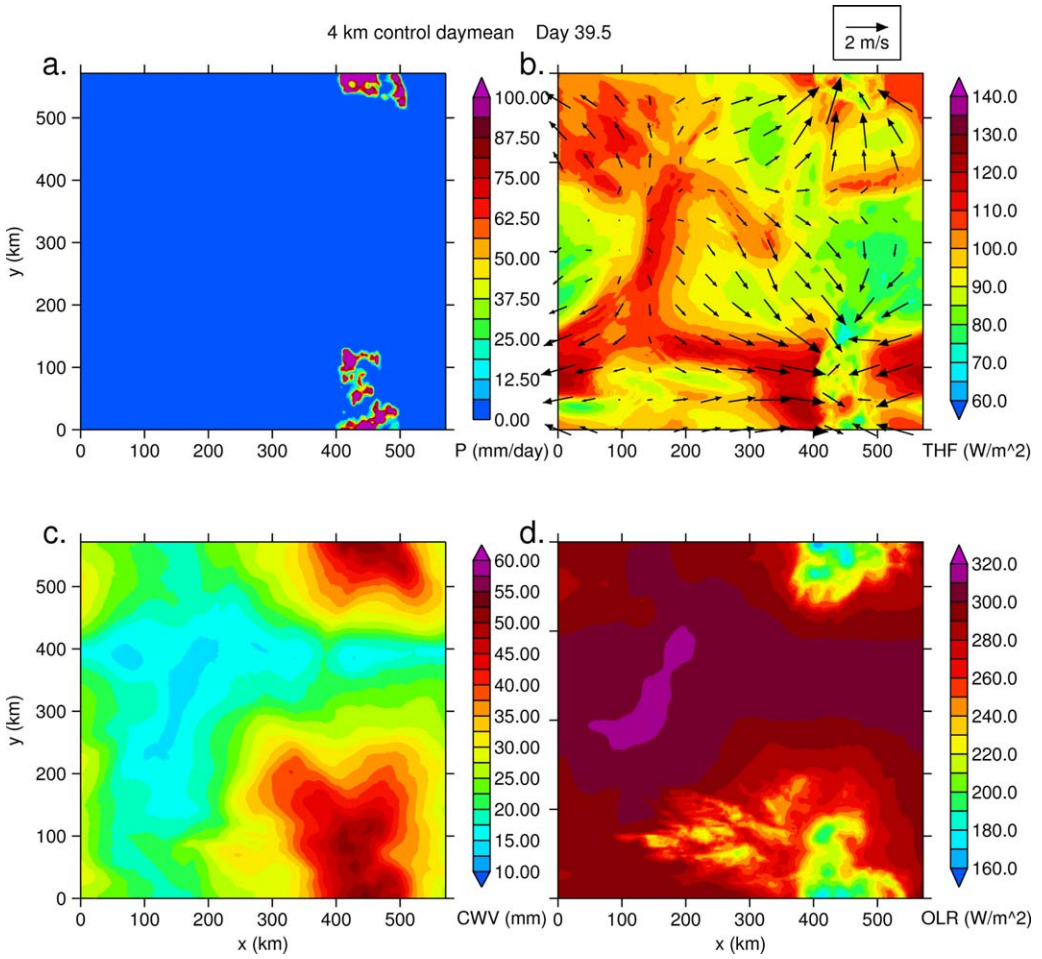

Figure 2. Day 39.5 mean values of (a) precipitation (mm d^{-1}), (b) sensible plus latent heat flux (colors, W m^{-2}) and surface wind (vectors), (c) CWV (mm), and (d) OLR (W m^{-2}) for control run with 300 K SST started from homogeneous conditions.

As mentioned by many previous studies, the self-aggregation process in idealized RCE models results not in a conservative redistribution of CWV and OLR but in changes in the domain means of these fields. Figure 3 shows the evolution of several daily domain-mean quantities, including the decrease in CWV (solid line, Figure 3a). This change, as mentioned in previous studies, is likely due to increased separation between regions of subsidence drying and regions of convection where water vapor and cloud condensate detrain and moisten nearby surroundings. The domain-mean OLR also increases (shown as a negative and offset dotted line in Figure 3d) as the subsidence areas dry and the total area of upper-tropospheric cloud cover diminishes, again in agreement with previous studies. Figures 3b and 3d show that the domain-mean precipitation and net atmospheric radiation stay relatively constant throughout the simulation, although there are some fluctuations of domain-mean precipitation on time scales of a few to 10 days, as seen also in other studies such as *Bretherton et al*. [2005]. The latent and total heat fluxes (Figure 3c) increase during the period of increasing aggregation due to increased surface winds and mean thermodynamic disequilibrium.

The simulation is largely in radiative-convective equilibrium, with precipitation balancing latent heat flux and net atmospheric radiative heating mostly balancing total surface heat flux. There is a slight imbalance between radiative heating and surface fluxes, with the radiative cooling averaging around 8 W m^{-2} larger in magnitude than the surface fluxes during the last 10 days of the simulation. This is mainly accounted for by a slight disequilibrium in the upper model levels (above the tropopause), and the lower stratosphere cools about 5 K during the entire simulation. The net downward surface energy flux including turbulent and radiative fluxes (dashed line, Figure 3c) shows a large positive implied warming tendency for the surface, but this decreases with aggregation (since latent heat flux increases while net downward surface

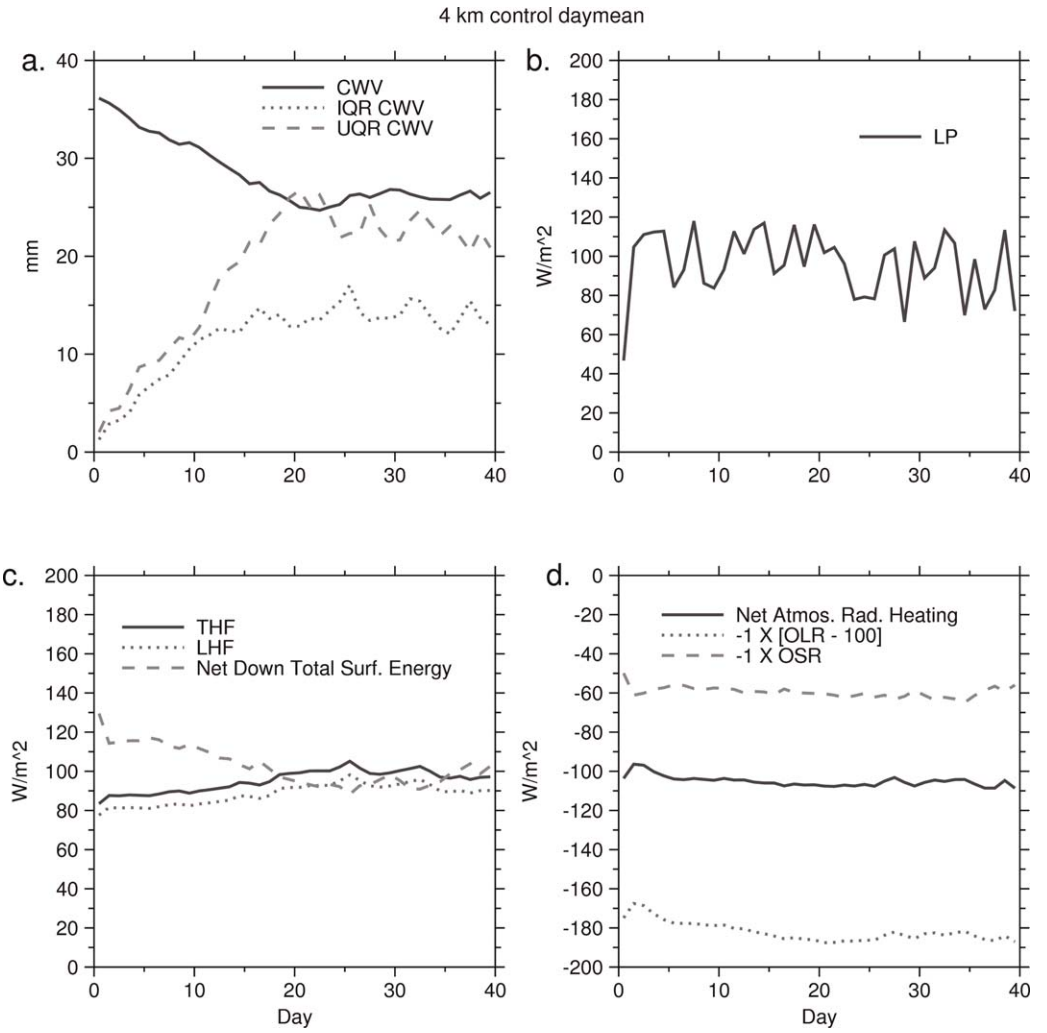

Figure 3. Daily mean values of (a) domain-mean CWV, interquartile range, and "upper-quantile" range (95th percentile minus median) from hourly CWV (b) precipitation rate (W m^{-2}), (c) surface latent heat flux, surface total (sensible plus latent) heat flux, and net downward surface energy including radiation (W m^{-2}), and (d) net atmospheric column radiative heating, −[OLR − 100], and −[OSR] (W m^{-2}) for control run with 300 K SST started from homogeneous conditions.

longwave decreases, not shown). Net downward surface shortwave stays fairly constant (not shown), similar to WE14 but unlike observations in *Tobin et al*. [2012]. Outgoing shortwave at the top of the atmosphere (reflected shortwave) also stays fairly constant (shown as a negative dashed line in Figure 3d), implying that low-level clouds increase in environmental regions as upper-level clouds decrease. Note that OLR increases even though net atmospheric radiative heating stays fairly constant because of the corresponding decrease in net downward surface longwave radiation.

Two simple metrics for aggregation state, shown in Figure 3a, are the Interquartile Range (IQR) of CWV, which is the 75th percentile minus the 25th percentile of CWV (taken from instantaneous model output at the 4 km grid for each hour and then averaged here for each day), and the here-defined "Upper-quantile Range" (UQR) of CWV, which is the 95th percentile minus the median. The IQR of CWV was found to be a good proxy for self-aggregation (as compared with the evolution of domain-mean fields and subjective determination of convective clustering) by MH12, and a similar quantity, the ratio of mean CWV in the driest quartile to mean CWV in the moistest quartile, was used by *Bretherton et al*. [2005]. We can compare our control IQR evolution to that of MH12 (their Figure 7), which show simulations that have 300 K SST like ours. Our maximum of ≈15 mm is reached after about 16 days, which is 5–10 days earlier than theirs, and their maximum IQR values range from 15 to 25 mm (for simulations with 2–2.8 km grid spacing and 128 × 128 grid point domains). The control run in WE14 had a maximum IQR of almost 40 mm for a 301 K SST run with

3 km grid spacing and a 768 km × 768 km domain (A. Wing, personal communication, 2015), although that simulation took roughly 60–80 days to fully aggregate. Both the IQR and UQR serve to measure the spread of the CWV distribution, specifically the separation between the moist convective regions and the dry environment. Although the two metrics are fairly similar overall, we favor the UQR slightly because the 25th percentile and the median are always located in the environment in our simulations, and the 95th percentile is always associated with convective regions, whereas in highly aggregated states, the 75th percentile can begin to be more similar to environmental regions. The domain-mean CWV and UQR in Figure 3a all show steady self-aggregation between day 1 and around day 20 (while the IQR stops its steady rise several days earlier), and all three metrics show an aggregated state with some fluctuations in the strength of that aggregation thereafter.

4.1. Moist Static Energy Variance Budget Analysis

The frozen moist static energy (MSE, or h) should be approximately conserved for moist adiabatic processes and is defined as in WE14

$$h = c_p T + gz + L_v q_v - L_f q_i, \tag{1}$$

where c_p is the specific heat of dry air, T is the air temperature, g is gravitational acceleration, z is height above the surface, L_v is the latent heat of vaporization, q_v is the specific humidity with respect to water vapor, L_f is the latent heat of fusion, and q_i is the specific humidity with respect to ice condensates (including frozen precipitation). Following WE14, we analyze the budget of the spatial variance of the vertically integrated frozen moist static energy. The budget equation for the mass-weighted vertical integral of h, $\hat{h}$, is

$$\frac{\partial \hat{h}}{\partial t} = \mathrm{THF} + \mathrm{S\hat{W}} + \mathrm{L\hat{W}} - \nabla_{\mathbf{h}} \cdot \hat{\mathbf{v}h}, \tag{2}$$

where THF is the total surface heat flux (latent plus sensible), SW and LW are the net shortwave and longwave heating, respectively, and $\nabla_{\mathbf{h}} \cdot \hat{\mathbf{v}h}$ is the horizontal divergence of the vertically integrated flux of h. For quantity x, $\hat{x} = \int_0^{z_{\mathrm{top}}} x \rho dz$, where ρ is the density.

We also define $\{x\}$ as the horizontal domain mean of x, and we define the anomaly from this mean as x'. Then, again following WE14 to get their equation (9), we subtract the horizontal mean of equation (2) from the original equation and multiply the result by $\hat{h}'$ to get the budget for the horizontal spatial variance of $\hat{h}$

$$\frac{1}{2}\frac{\partial \hat{h}'^2}{\partial t} = \hat{h}'\mathrm{THF}' + \hat{h}'\mathrm{S\hat{W}}' + \hat{h}'\mathrm{L\hat{W}}' - \hat{h}'\nabla_{\mathbf{h}} \cdot \hat{\mathbf{v}h}. \tag{3}$$

The four terms on the right-hand side (RHS) of equation (3) are the spatial covariances between $\hat{h}$ and the four budget terms on the RHS of equation (2). As in WE14, we calculate the last term, due to advective processes or "convergence" of MSE, as a residual from the remaining budget; the remaining terms are calculated from instantaneous hourly thermodynamic fields and hourly mean fluxes as described in section 3. We can study the covariance terms to determine how much each process in equation (2) contributes to changes in the variance of $\hat{h}$ and thus to changes in aggregation (although, as with the IQR and UQR, there is no mathematically required relationship between the variance of $\hat{h}$ and the spatial scales of convective organization). In the control run, the shape of the evolution of $\{\hat{h}'^2\}$ (discussed in section 5) closely resembles the shape of the evolution of the UQR, underlining the link between $\{\hat{h}'^2\}$ and aggregation.

The evolution of daily mean, domain-mean values of four terms on the RHS of equation (3), as well as the sum of the three diabatic terms (longwave, shortwave, and surface heat flux), all normalized by $\{\hat{h}'^2\}$, is shown in Figure 4a. Analogous terms calculated using only clear-sky longwave and shortwave radiation, in which cloud effects are ignored in the radiative flux calculations, are also shown. In the first 6–8 days, the total diabatic forcing is large and positive, while the convergence is smaller but also mainly positive, which means that both diabatic and advective forcing leads to increased $\hat{h}$ variance and increased aggregation. The diabatic forcing is composed of a large positive contribution from longwave radiation and significant but smaller contributions from surface fluxes and shortwave radiation. However, from days 5 to 10, the surface fluxes become much less important, and the longwave decreases closer toward the level of the shortwave radiation, which stays fairly constant (in a normalized sense) over the entire simulation. Between days 5 and 15, meanwhile, the convergence term decreases to become slightly negative on average. After day 15, when the simulation has reached

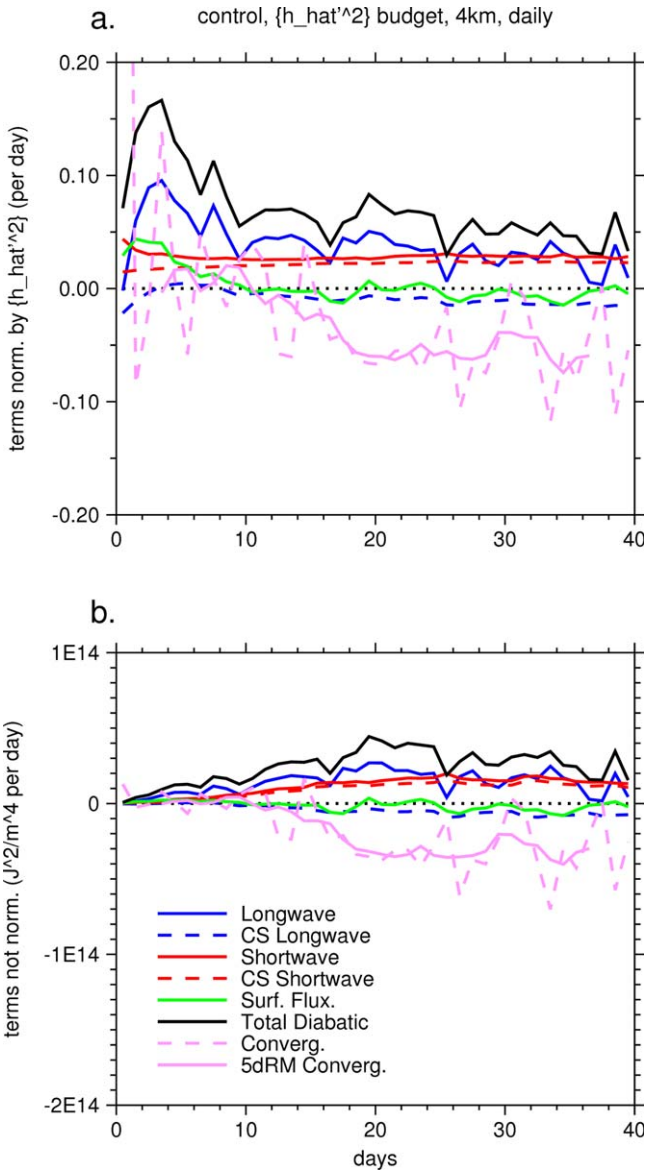

Figure 4. Daily mean values of domain-mean terms in the $\hat{h}'^2$ budget, with terms in (a) each normalized by $\{\hat{h}'^2\}$, in units day^{-1}, and terms in (b) not normalized, in units J^2 m^{-4} d^{-1}. The lines show the sum of all diabatic terms (black) as well as contributions from longwave radiation (blue solid), clear-sky-only longwave (blue dashed), shortwave radiation (red solid), clear-sky-only shortwave (red dashed), surface heat flux (green), and horizontal convergence of h flux (pink dashed), with a 5 day running mean of the latter shown as pink solid; for the control run with 300 K SST started from homogeneous conditions.

its aggregated state, there is a balance between diabatic forcing acting to increase aggregation (with roughly equal contributions from longwave and shortwave radiation) and convergence acting to decrease aggregation.

Figure 4a can be compared directly to WE14 (their Figure 5) and to *Wing and Cronin* [2016, Figures 4 and 5]. WE14's aggregation takes several times longer to reach a relatively stable level of aggregation, but there are similarities in the evolution of the different covariance terms. For instance, the normalized total diabatic term is large and positive in the early part of their simulation, with surface fluxes being important then and with longwave radiation also particularly important. The shortwave is also similar, contributing a relatively constant normalized amount throughout the simulation of about 0.03 day^{-1}. The radiative effect is mainly due to cloud-longwave and all-sky-shortwave contributions. The all-sky and clear-sky shortwave are almost the same after the first few days, but much of the positive shortwave effects come from the moistest, cloudiest columns, so clouds may change the vertical distribution of the shortwave heating from what it would be in clear-sky conditions (see discussion of vertical radiation distribution in section 4.2 below). The strong cloud effects for longwave (even at early times) is similar to *Wing and Cronin* [2016] for their runs with SSTs larger than 290 K but contrasts with arguments in WE14 and *Emanuel et al*. [2014].

The biggest difference between WE14 and our Figure 4a is the convergence term. Our convergence term is fairly small during self-aggregation (days 0–16) and then becomes strongly negative during the mature phase, while WE14's convergence term is negative in the early part of the simulation (while the surface flux contribution is significantly positive) and then changes sign and becomes the dominant term during the later stages of self-aggregation (which is similar to the mature phase in *Wing and Cronin* [2016], especially for cooler SSTs). WE14 describe this later stage of aggregation, dominated by a positive convergence term, as the stage in which a self-aggregation mechanism described by *Bretherton et al*. [2005] and MH12 likely takes place. This mechanism involves circulations driven by horizontal gradients in low-level radiative cooling which transfer near-surface high-entropy air from drier to moister columns and lower-entropy air in the lower free troposphere from moister to drier columns. While this mechanism (or other mechanisms contributing to the convergence term) may play a role in our control simulation as well, it is clearly not as significant as in WE14. We discuss this more in section 4.2 below.

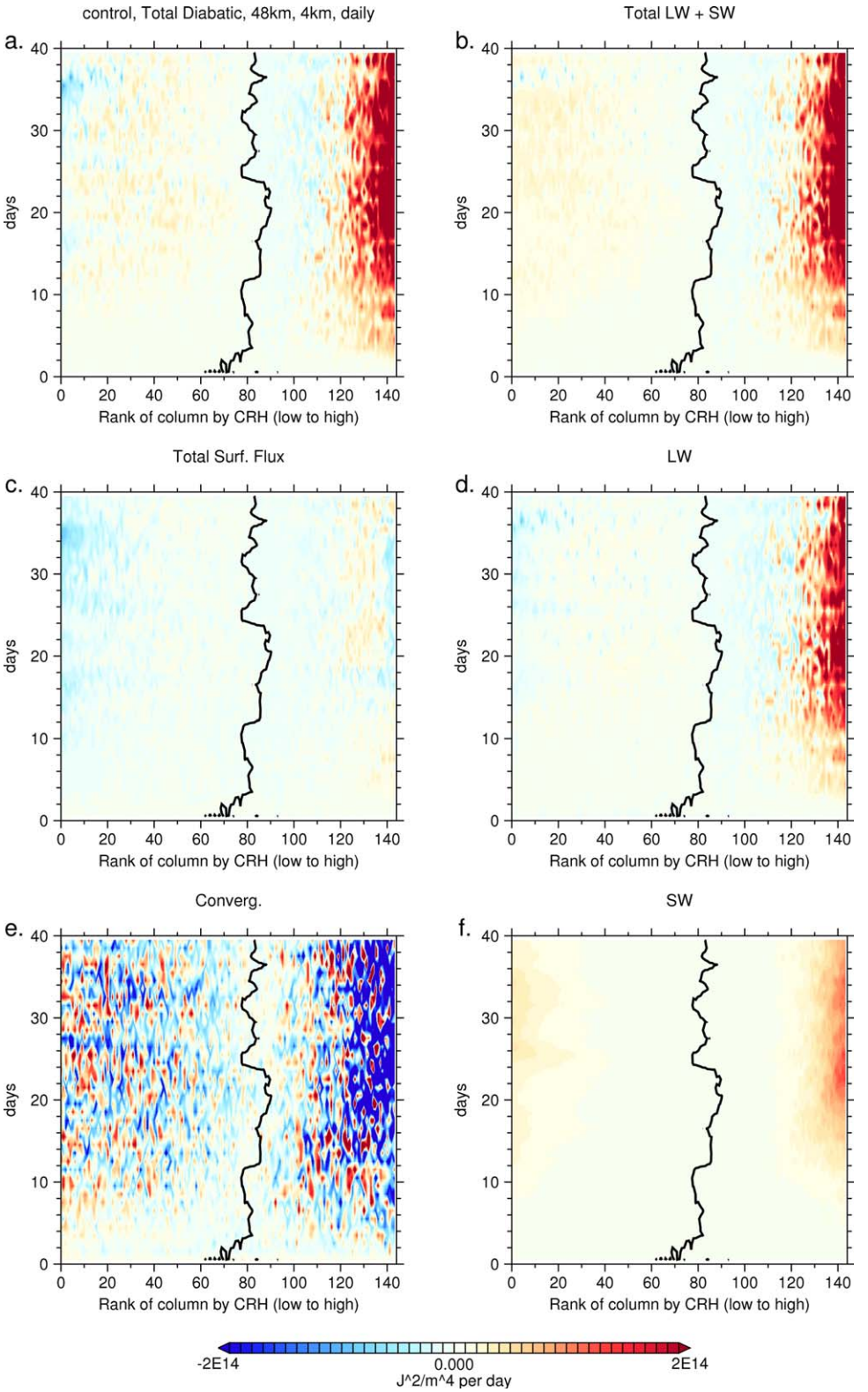

Figure 5. Daily mean values of 48 km × 48 km blocks for terms in the $\hat{h}'^2$ budget (not normalized), in units $J^2\ m^{-4}\ d^{-1}$. The plots show (a) the sum of all diabatic terms, (b) the sum of contributions from longwave and shortwave radiation, (c) surface heat flux, (d) longwave radiation, (e) horizontal convergence of h flux, and (f) shortwave radiation, for the control run with 300 K SST started from homogeneous conditions. The x axis shows the rank of the boxes sorted by column relative humidity (CRH), driest at the left; the sorting is done for each day. The black line is the $\hat{h}' = 0$ contour.

MH12 found that interactive shortwave radiation actually opposes self-aggregation, preventing self-aggregation starting from homogeneous conditions for a slightly larger number of runs in their grid-spacing/domain-size parameter space (compared with sensitivity tests with homogenized shortwave radiation), although it did not cause an already aggregated state to disaggregate. They hypothesized that this negative feedback is a direct radiative effect (shortwave radiation warming dry regions more than moist, cloudy regions), conflicting with our study and with WE14, which both find a positive direct diabatic effect from shortwave radiation. It may be that the results in MH12 relating to shortwave effects were actually the results of weak overturning circulations forced by horizontal shortwave anomalies; our estimates of radiatively forced circulations using WTG assumptions, discussed in section 5 below, suggest that shortwave-forced circulations slightly oppose aggregation, especially at upper levels (not shown).

The evolution of the nonnormalized covariance terms shown in Figure 4b illustrates the main points above. The end of the main self-aggregation phase, starting around day 15, is clearly marked by the rapid decrease in the convergence term. After this point, there are further large swings in the total diabatic term but these tend to be balanced on longer time scales by opposite changes in the convergence term (although the aggregation state does fluctuate somewhat as mentioned above).

Again following WE14, the evolution of these nonnormalized terms can be further investigated by averaging each term in 48 km $\times$ 48 km blocks for each day and ranking them by mean column relative humidity (CRH), which is the CWV divided by the column-integrated saturation specific humidity (ranking by $\hat{h}$ instead gave very similar results, not shown). Figure 5 shows the four covariance terms, along with the sum of the three diabatic terms and the sum of the radiation components, in color for the 144 ranked blocks and each day, with the $\hat{h}'=0$ contour plotted in black as a reference. (Note that blocks can change rank over time, so the same block may be in a different horizontal position in Figure 5 on different days.) The total diabatic term shows that the largest positive contributions come from the highest-ranked 10–15% of the blocks, with most of this contributed by longwave radiation along with significant contributions from shortwave radiation in the period after day 10. The shortwave term also contributes significant positive values in the lowest-ranked 15% of the blocks after day 10, which tends to balance some negative contributions from surface fluxes and longwave radiation in those blocks. These positive shortwave contributions (anomalous shortwave *cooling*) in low-CRH columns agree with WE14, who link them to reduced shortwave absorption by water vapor in dry regions. The convergence term is more noisy, but it is interesting that before day 10, the largest positive convergence values occur in the lowest-ranking 20% of the blocks, whereas in days 10–14, while these continue, there are also large positive contributions (and a lack of large negative contributions) in the highest-ranking 20% of blocks.

The diabatic panels in Figure 5 can be compared with the (normalized) diabatic panels in WE14 (their Figure 6). The radiative components seem largely similar, especially at later times and higher-ranking blocks, whereas our surface flux contributions are much smaller in magnitude. The black $\hat{h}'=0$ contour in our figure stays closer to the median CRH ranking than in WE14, where this line dips into much lower CRH rank during early stages of self-aggregation. This is likely due to the nature of self-aggregation in WE14, in which a dry patch develops and causes a small fraction of columns to dry significantly at early times while the rest of the domain stays moist. In our simulation, as mentioned above, both moist and dry patches grow and merge at similar rates.

Another way of visualizing the covariance terms is to bin them by CRH for each day and then plot the mean of each term in each CRH bin multiplied by the fraction of total blocks in that bin for each day. This visualization technique is inspired by a similar figure [*Wing*, 2014, Figure B-1], although the covariance terms used in that figure are normalized by $\{\hat{h}'^2\}$ and are not multiplied by the fraction of blocks in each bin. Our pdf-weighted contour plots are shown in Figure 6, whereas Figure S12 in the supporting information shows the PDF in these bins and Figure S13 shows the equivalent of Figure 6 without multiplying by the PDF. Some of the points made above with respect to Figure 5, such as the times and CRH ranks of significant positive contributions from the convergence term, can be seen more clearly here alongside the spreading of the CRH distribution. Figure 6b shows how important the positive radiative feedbacks from very moist regions are for developing and, especially, maintaining aggregation. These are mainly longwave effects, and they happen in regions with lots of cloud, supporting results above that clear-sky effects do not dominate even early stages of aggregation in our control run.

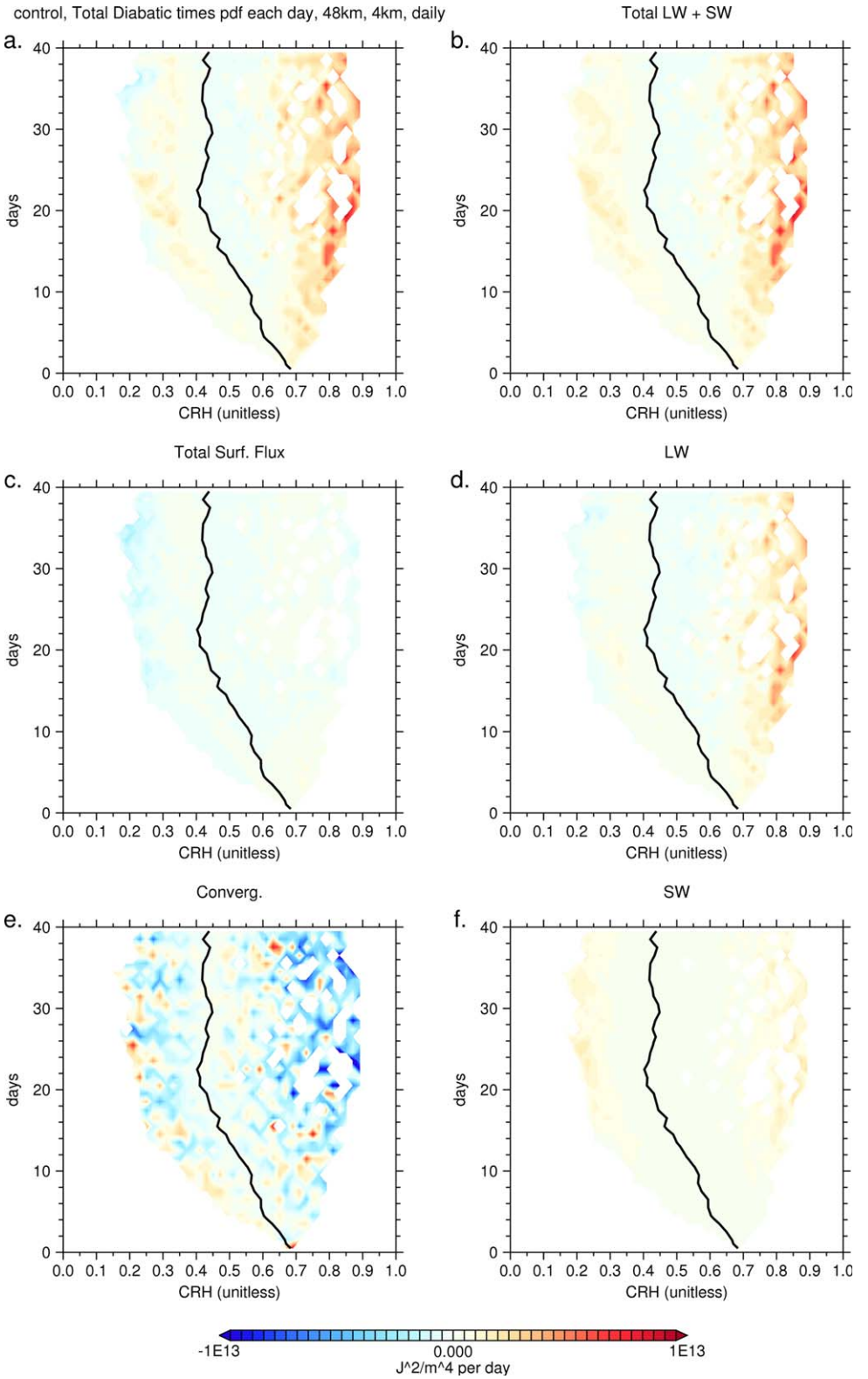

Figure 6. Daily mean values of 48 km × 48 km blocks for terms in the $\hat{h}'^2$ budget (not normalized) conditionally averaged by column relative humidity (CRH) for each day and multiplied by the probability density for that CRH and day, in units $J^2\ m^{-4}\ d^{-1}$. The plots show (a) the sum of all diabatic terms, (b) the sum of contributions from longwave and shortwave radiation, (c) surface heat flux, (d) longwave radiation, (e) horizontal convergence of h flux, and (f) shortwave radiation, for the control run with 300 K SST started from homogeneous conditions. The black line is the $\hat{h}'=0$ contour.

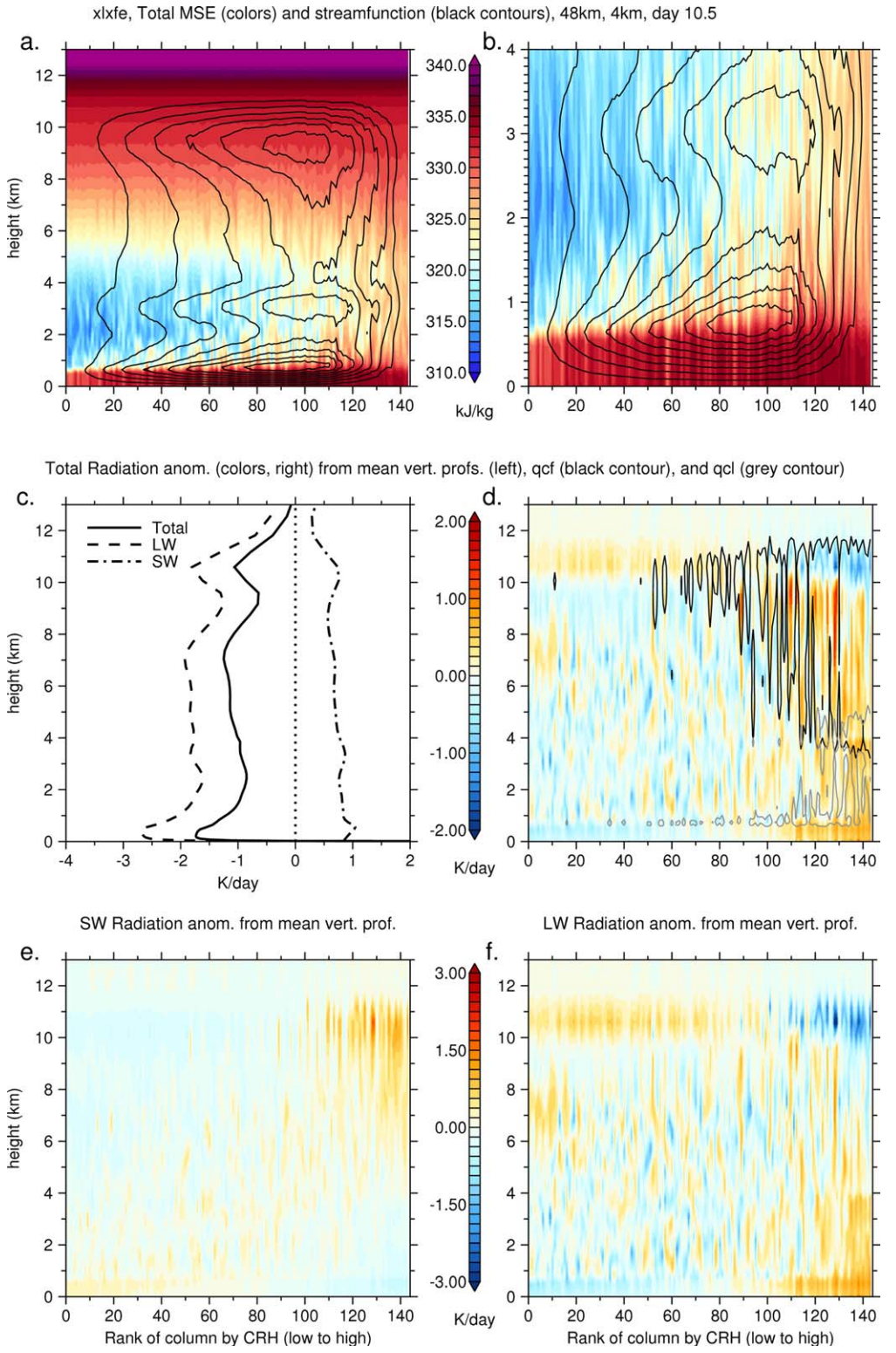

Figure 7. Day 10.5 mean values of 48 km × 48 km blocks, sorted in the *x* dimension by CRH, for (a) *h* (kJ kg^{-1}, colors) and Ψ (black lines, negative lines solid, contour interval is 0.06 kg m^{-2} s^{-1} starting from ±0.06), (b) lowest 4 km heights zoomed from Figure 7a, (c) profiles of domain-day-mean total, longwave, and shortwave radiative heating (K d^{-1}), (d) horizontal anomaly (from domain-day-mean profile) of total radiation (K d^{-1}) (colors) with 5% cloud fraction contours of ≥1 mg kg^{-1} cloud ice/snow (black line) and ≥1 mg kg^{-1} cloud liquid (grey line), (e) horizontal anomaly of shortwave radiation (K d^{-1}), and (f) horizontal anomaly of longwave radiation (K d^{-1}). The *x* axis shows the rank of the boxes sorted by column relative humidity (CRH), driest at the left; the sorting is done for each day.

4.2. Vertical Profiles of Mass Flux, Radiation, and MSE

Bretherton et al. [2005] and MH12 find evidence for a low-level circulation between dry and moist columns that transports MSE up-gradient, leading to self-aggregation at early-to-intermediate stages of the process. To characterize this upgradient transport, *Bretherton et al*. [2005] define an effective stream function Ψ. At each time (day mean), the N coarse-grained blocks are ranked from low to high CRH and assigned an index value $i-1/2$, $i=1,2,\ldots,N$. Setting $\Psi_0=0$, Ψ_i is then a cumulative sum of the upward mass flux over the i driest blocks

$$\Psi_i(z)=\Psi_{i-1}(z)+\{\rho(z)\}w_{i-1/2}(z), \tag{4}$$

where $\{\rho(z)\}$ is taken here to be the domain-mean density profile for each day and $w_{i-1/2}(z)$ is the block-average daily mean vertical velocity profile. Although this stream function does not represent a true horizontal circulation, it should represent the main resolved flow between moist and dry blocks. While our total domain-mean convergence term is not a significant positive effect on self-aggregation, we are still interested in the evolution of circulations and their implied effects on MSE transport between blocks.

Figures 7 and 8 show this stream function (top panels, black contours) along with block-averaged h, radiation anomalies from domain-day-mean profiles, and contours for 5% cloud fraction of q_i and q_l (the cloud liquid specific humidity) for day 10.5 and day 20.5, respectively. The cloud fractions are calculated by counting all 4 km instantaneous hourly q_i and q_l values above 1 mg kg^{-1} in each block for each day and dividing by the total number of values. Day 10.5 has a significant positive convergence covariance term (in the $\hat{h}^{\prime 2}$ budget, see pink dashed line in Figure 4b) whereas day 20.5 has a strong negative contribution. At day 10.5 (Figure 7), there is strong flow in the lowest 600 m from dry to moist blocks, and h in this subcloud layer is slightly higher than h in the upper-tropospheric return flow from 11 to 12 km height, suggesting slight import of h into moist regions by this deep tropospheric circulation. At the same time, a low-level circulation transports air at roughly 0.4–0.7 km height from moderately dry blocks to moist blocks with slightly higher h than air in a return flow around 1–2 km height, again importing h into moist regions. There is more net radiative cooling in the lower troposphere for drier blocks, notably in the lowest 600 m but also more generally between 0–2 km height, but, as discussed in section 5 below, this is not the main factor driving subsidence and forcing this low-level circulation, in contrast to hypotheses in *Bretherton et al*. [2005] and MH12. The thick ice and liquid cloud occurs in the moister blocks, and these tend to have less radiative cooling both near the surface and throughout the troposphere. There are some thin liquid clouds near the top of the boundary layer even in the dry columns, and these may contribute to the larger radiative cooling there, similar to MH12.

At day 20.5 (Figure 8), the upper troposphere has warmed during the aggregation process as seen in previous studies [e.g., *Bretherton et al*., 2005], which attribute this to a less-dilute ascent within updrafts. The upper-level outflow from moist to dry blocks is now of a similar h value to the inflow below 600 m, and there is also a larger gradient of h below 600 m from moist to dry blocks. Both of these imply no import, and perhaps a slight export, of h from moist to dry blocks by this deep circulation. The low-level circulation is still of a similar magnitude to that at day 10.5, but it is now embedded further within the moist blocks and its lower and upper branches have h values that are of more similar magnitudes to each other, suggesting little net transport of h by this circulation as well. Finally, there is an additional, somewhat weaker, mid-level circulation at day 20.5, apparently forced in large part by strong radiative cooling near the freezing level at moderately dry and moderately moist blocks (see also section 5), which exports h from the moistest blocks. The combination of these circulations would result in a net transport of h out of the moistest blocks at day 20.5, tending to work against aggregation, as seen in Figure 4b.

Both of these days have significant positive longwave and shortwave covariance terms as well. For shortwave, the larger absorption at low levels for dry blocks is more than offset by larger absorption at mid-to-upper levels for moist blocks, particularly by day 20.5 when the moistest blocks have very large shortwave absorption near the top of the upper-level ice cloud layer. Note that the clear-sky shortwave term has strong contributions from moist blocks as well (not shown), suggesting that, in the absence of cloud, there would still be large positive shortwave anomalies in these moist blocks due to water vapor absorption, though this would likely be concentrated in the lower troposphere. Longwave cooling is stronger for moderately dry blocks for most levels except this upper level ice-cloud layer (although opposite tendencies a few km below this tend to offset this effect). There is some anomalous warming at middle levels in the

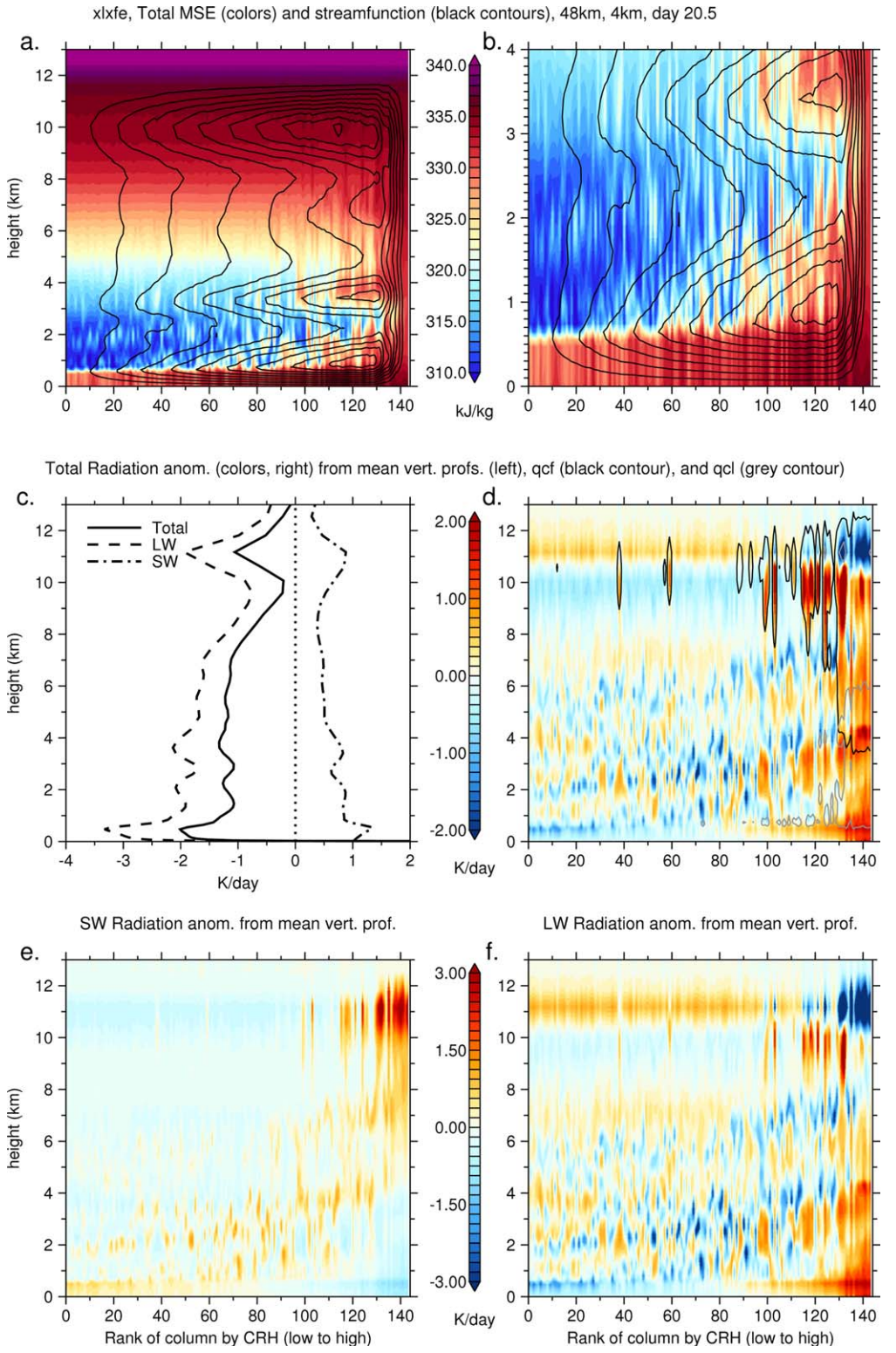

Figure 8. Day 20.5 mean values of 48 km × 48 km blocks, sorted in the *x* dimension by CRH, for (a) *h* (kJ kg^{-1}, colors) and Ψ (black lines, negative lines solid, contour interval is 0.06 kg m^{-2} s^{-1} starting from ±0.06), (b) lowest 4 km heights zoomed from Figure 8a, (c) profiles of domain-day-mean total, longwave, and shortwave radiative heating (K d^{-1}), (d) horizontal anomaly (from domain-day-mean profile) of total radiation (K d^{-1}) (colors) with 5% cloud fraction contours of ≥1 mg kg^{-1} cloud ice/snow (black line) and ≥1 mg kg^{-1} cloud liquid (grey line), (e) horizontal anomaly of shortwave radiation (K d^{-1}), and (f) horizontal anomaly of longwave radiation (K d^{-1}). The *x* axis shows the rank of the boxes sorted by column relative humidity (CRH), driest at the left; the sorting is done for each day.

driest blocks (likely because they are too dry to emit much longwave) but this is moderated by especially large cooling in the lowest 600 m of these blocks. In both days, the main positive longwave feedback comes from low and middle levels in the moistest blocks.

5. Sensitivity to Fixed Surface Fluxes and Radiative Cooling, and to Nonevaporation of Rain

To test the effects of interactive radiation and surface fluxes, and the effects of rain evaporation and subsequent cold pools, we compare the control run with five sensitivity tests. Each simulation was started from the same homogeneous initial state (with small perturbations to boundary layer temperature) that was used in the control run, and all of these simulations were run for 40 days. Three of the simulations are like the control run except that they have either fixed prescribed surface fluxes, fixed prescribed radiative cooling, or both, where "fixed" means held constant in horizontal space and time. In these simulations, the prescribed forcing is taken from the average of the forcing diagnosed in the final 10 days of the 70 day small-domain RCE runs. The other two simulations are like the control run except that liquid hydrometeors are not allowed to evaporate at any model level, and one of these additionally has fixed prescribed radiative cooling as described above.

The mechanism denial experiments starting from homogeneous initial conditions show that self-aggregation is generally favored by interactive radiation and interactive surface fluxes, while suppressing rain evaporation speeds up self-aggregation and can allow strong self-aggregation even in the absence of interactive radiation. Figures 9a, 9b, and 9d–9f show that runs with fixed prescribed surface fluxes (long-dashed lines) have the least amount of self-aggregation, with higher mean CWV and lower IQR, UQR, $\{\hat{h}'^2\}$, and UQR normalized by daily mean saturation CWV, which is equivalent to a UQR based on CRH assuming weak-temperature gradients. They also have significantly less horizontal structure in the CWV field as measured by average autocorrelation length scales, which are calculated by first finding the largest spatial lag at which the spatial autocorrelation drops below e^{-1} for each individual row and column at a single hour and then averaging these together (and then taking a daily mean), similar to the method of *Craig and Mack* [2013] (Figure 9c). However, noting that we had seen strong self-aggregation for fixed surface fluxes in a previous run based on initial conditions and constant flux values from a small domain run with artificially high surface roughness lengths (and thus higher surface fluxes and a warmer, moister state), we ran several additional sensitivity tests to determine why these runs differed. The key factor was the value of the constant prescribed surface fluxes. For the forcing used in the runs shown in Figure 9, the values are {5.9, 79.1} for the sensitive and latent heat fluxes, respectively (in W m^{-2}). We also saw little self-aggregation for a run in which we used {6.9, 91.6} for these fluxes, which was taken from the mean of the last 10 days of our control run in the present study, which develops stronger surface fluxes at later times. However, we did get significant self-aggregation when we used {3.7, 98.8}, which was taken from the small domain run with very high prescribed roughness lengths mentioned above. These results were not sensitive to the initial profiles of temperature and humidity. One possible explanation is that cold pools are much stronger and more active, with much larger extents, in the run with fixed surface fluxes than in the control run, since the control has very strong (over twice as large) surface fluxes in cold pools that damp their temperature and moisture perturbations and, therefore, their propagation speeds (compare Animations S1 and S2 in the supporting information). This effect will be stronger for lower mean surface flux forcing (although only the latent heat flux varies monotonically in the tests described above). It is also possible that the larger mean precipitation and atmospheric radiation required by larger surface fluxes, or larger boundary layer buoyancy fluxes, are important for self-aggregation via radiation or circulation feedbacks or by setting the lower-tropospheric saturation deficit which also affects downdraft and cold pool intensity. Note that we still find that constant surface fluxes do not disaggregate an already aggregated state, even when we use the lowest values mentioned above (as shown in section 6).

Another important factor differentiating the evolution of these simulations is interactive radiation, as seen by comparing the red lines (simulations with interactive radiation) with the blue lines (simulations with fixed prescribed radiation) in Figure 9. The run with fixed prescribed radiation (solid blue line) shows fairly little aggregation by most measures, although it does allow for significant horizontal structure (large CWV autocorrelation length scales), and even the runs with fixed surface fluxes show fairly large values of this metric,

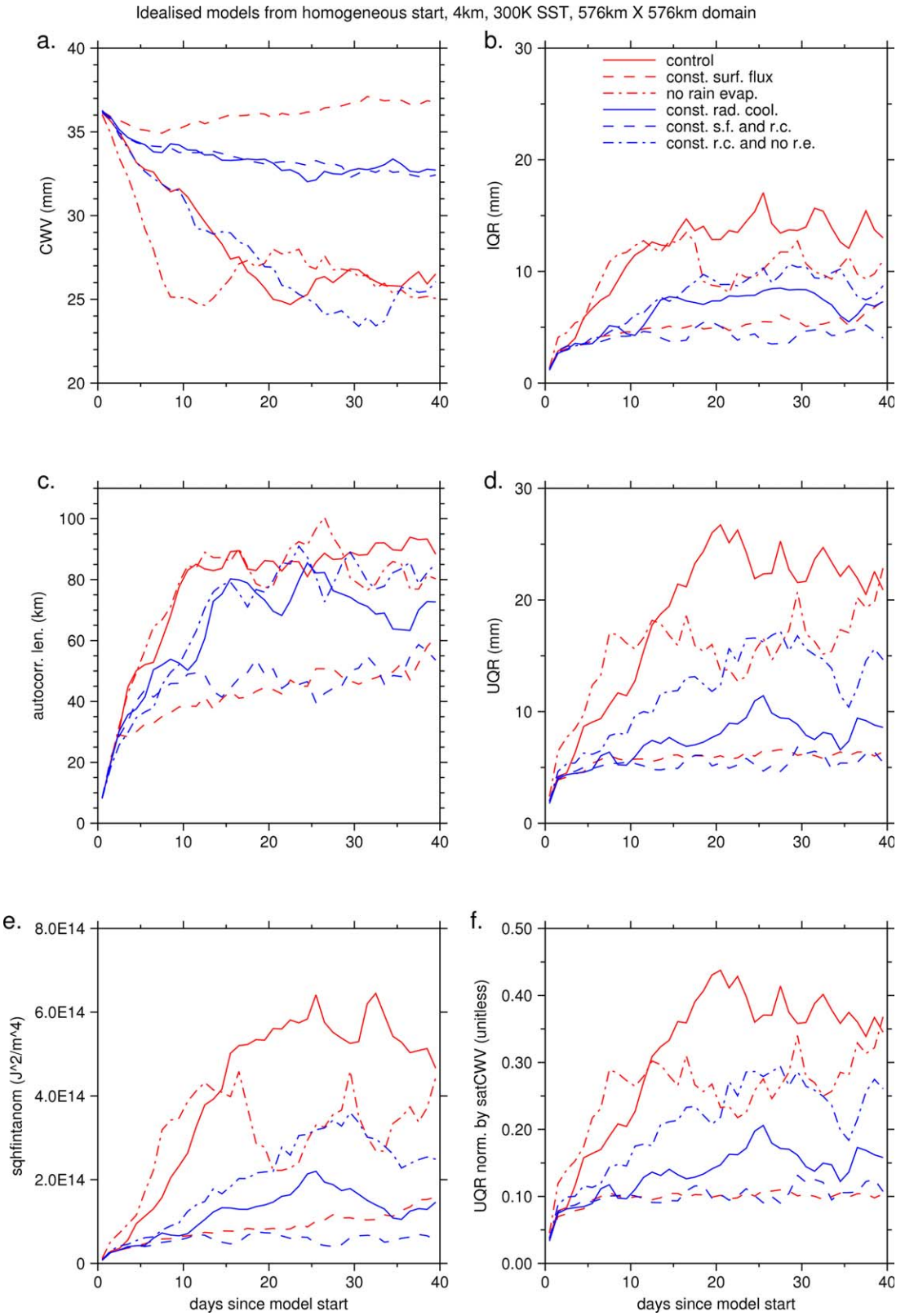

Figure 9. Daily mean values of (a) domain-mean CWV (mm), (b) interquartile range of CWV over domain at each hour (mm), (c) mean of autocorrelation length scale for all rows and columns at each hour (km), (d) "upper-quantile" range (UQR, 95th percentile minus median) of CWV over domain at each hour (mm), (e) $\{\hat{h}'^2\}$ ($J^2\ m^{-4}$), and (f) UQR divided by daily mean domain-mean saturation CWV (unitless), for six idealized simulations which all start from homogeneous conditions.

especially in the final days of the simulations. This organization takes the form of fairly large, slightly above-average moist regions within which precipitation occurs across wide areas (and sometimes clusters in sub-regions, especially in the run which still has interactive surface fluxes) but is less confined to specific regions than in the control run (see Figures S4 and S5 and Animations S4 and S5 in the supporting information). In the run without interactive surface fluxes or interactive radiation, this organization must be due only to circulation effects, including the "moisture-memory" moisture-convection feedback [*Muller and Bony*, 2015; *Tompkins*, 2001].

Similar to *Muller and Bony* [2015], we find that a run with fixed radiative cooling and no rain evaporation (blue dash-dotted line) shows strong self-aggregation, which they also attribute to the "moisture-memory" effect, hypothesizing that this effect is stronger in the absence of cold pools fed by rain evaporation. A run similar to the control except with no rain evaporation and no cold pools (red dash-dotted line) shows self-aggregation that occurs several days earlier than the control (though the degree of aggregation does not peak as high), agreeing with evidence from *Jeevanjee and Romps* [2013] that cold pools actually slow down the aggregation process by increasing low-level mixing between moist and dry regions. *Wing* [2014] and *Muller and Bony* [2015] also found that removing rain evaporation (only below 1 km, as in *Jeevanjee and Romps* [2013]) led to faster self-aggregation; as in our simulation, these studies also found that the aggregation process develops similarly to the coarsening process in *Craig and Mack* [2013], with moist areas growing in place and merging or dying out with little horizontal drift of moist and dry regions. This can be seen, both with and without interactive radiation, in Figures S3 and S6 and Animations S3 and S6 in the supporting information.

Most of the mechanism denial simulations discussed above result in a reduced number of nonzero terms in equation (3), and while this can lead to challenges when comparing their budgets with the control budget, it is still useful to investigate these simulations using the $\hat{h}'^2$ budget framework. Figure 10a shows the terms for the $\hat{h}'^2$ budget (normalized by $\{\hat{h}'^2\}$) for the run with fixed prescribed surface fluxes. Comparing these to Figure 4a, the fixed surface flux run has a negative convergence term almost throughout, consistent with the hypothesis discussed above that increased cold pools, which are not damped by strong interactive surface fluxes, cause increased low-level h mixing. This run also has a much smaller longwave radiation effect for all skies, and a significantly negative clear-sky longwave term. It is not clear why this is true, but it suggests possible limitations to using mechanism denial experiments to study processes that are not independent of each other.

The run with no rain evaporation, and therefore no cold pools (Figure 10b), has a relatively large contribution from the longwave effect (though the clear-sky longwave term is slightly more negative due to significant negative contributions from the driest blocks) compared to the control. The positive longwave effect is concentrated in the moistest 5–10% of blocks from around day 2 throughout the rest of the run (not shown), indicating the importance of the cloudiest convective areas for this run. There is a significant contribution from surface fluxes at later times (but not during the self-aggregation period) which comes from the moistest 10% of blocks. There is a slightly larger convergence term during the rapid self-aggregation phase (the first 6 days) compared to the control, and there is an absence of large negative contributions from this term (the pink dashed line) during this period, giving some support to the hypothesis of *Jeevanjee and Romps* [2013] that cold pools disrupt aggregation at early times by transporting high-h air out of moist regions at low levels (this slightly counterintuitive process may involve increased h at the edges of cold pools along with recovery of h in the interior of cold pools due to increased surface fluxes). However, the convergence term in the run with no rain evaporation does not dominate the self-aggregation period and becomes even more negative (in a normalized sense) than in the control at later times, suggesting that the "moisture-memory" effect or other advective processes may not contribute much to self-aggregation in this run.

On the other hand, Figures 10c and 10d show that for two runs with fixed radiative cooling (in which the total diabatic heating, shown by the black line, represents only surface fluxes), it is mainly the convergence term that is significantly nonzero and contributes to self-aggregation. This is especially true for the run which no rain evaporation in addition to fixed radiative cooling (Figure 10d), for which the convergence term is the only positive term, and which exhibits strong self-aggregation as discussed above. This suggests either "moisture-memory" or some other advection-related process.

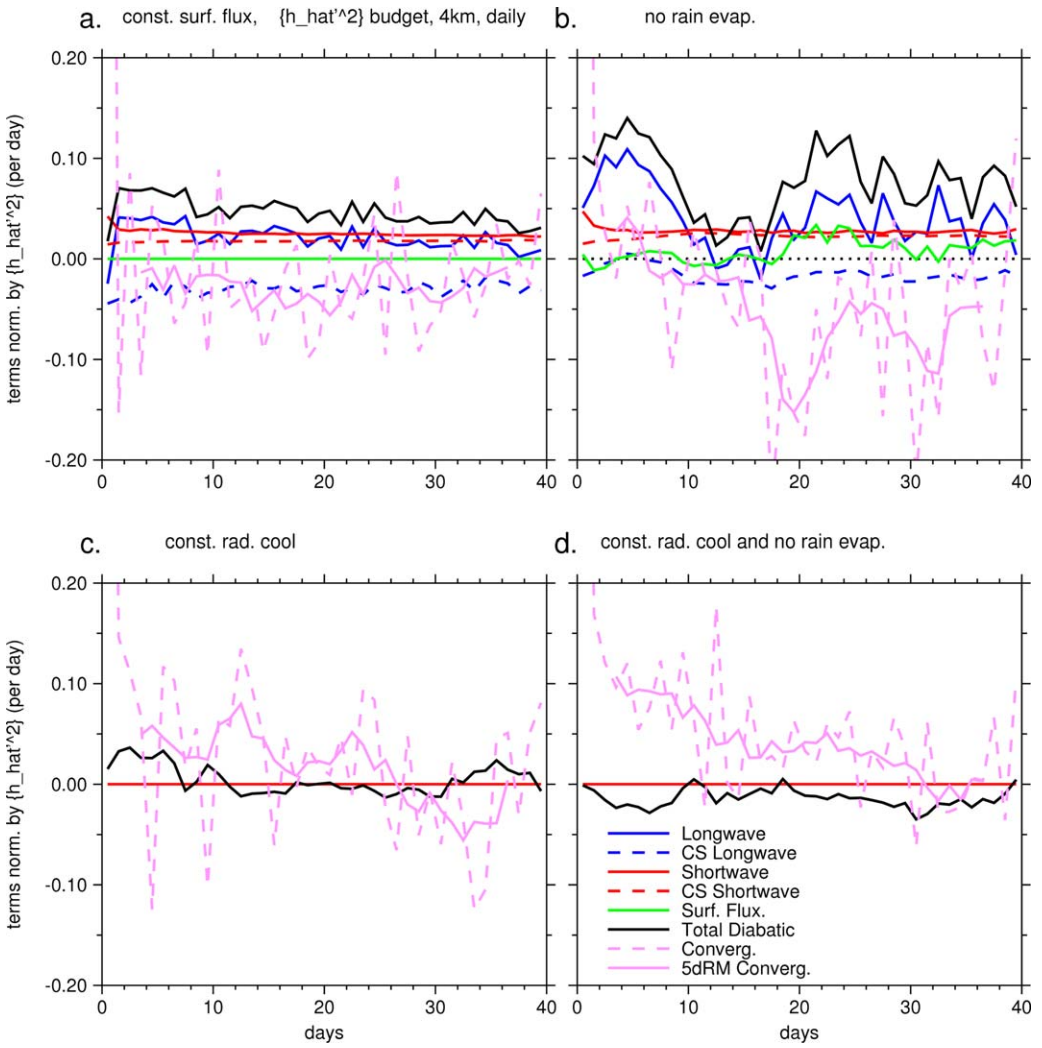

Figure 10. Daily mean values of domain-mean terms in the $\hat{h}'^2$ budget (each normalized by $\{\hat{h}'^2\}$), in unit day^{-1}. The lines show the sum of all diabatic terms (black) as well as contributions from longwave radiation (blue solid), clear-sky-only longwave (blue dashed), shortwave radiation (red solid), clear-sky-only shortwave (red dashed), surface heat flux (green), and horizontal convergence of h flux (pink dashed), with a 5 day running mean of the latter shown as pink solid; for runs with 300 K SST started from homogeneous conditions and (a) fixed prescribed surface fluxes, (b) no rain evaporation, (c) fixed prescribed radiative cooling, and (d) fixed prescribed radiative cooling and no rain evaporation.

To investigate the effects of radiation on circulation, we calculate a radiative stream function based on a vertical velocity calculated from horizontal radiation anomalies and weak-temperature gradient (WTG) ideas [*Sobel and Bretherton*, 2000; *Raymond and Zeng*, 2005]. We define the radiative stream function, Ψ_{rad}, to be the same as Ψ except that w in equation (4) is replaced with w_{rad} which is defined for levels above 600 m height as

$$w_{\mathrm{rad}}(z) = \frac{Q'_{\mathrm{rad}}}{\{\Pi\}\frac{\partial\{\theta\}}{\partial z}}, \tag{5}$$

where Q'_{rad} is the horizontal anomaly of total radiative temperature forcing, θ is the potential temperature, and $\Pi = T/\theta$ is the Exner function. For this calculation, we use daily mean block-averaged values of each variable. Note that w_{rad} and Ψ_{rad} are only defined above the boundary layer top (here approximated as 600 m) because $\partial\theta/\partial z$ is close to zero in the boundary layer; this means we neglect the effects of significant anomalous radiation in the boundary layer, which presumably has indirect effects on circulation via convective instability and convective mass flux along with its direct diabatic effects on $\hat{h}$.

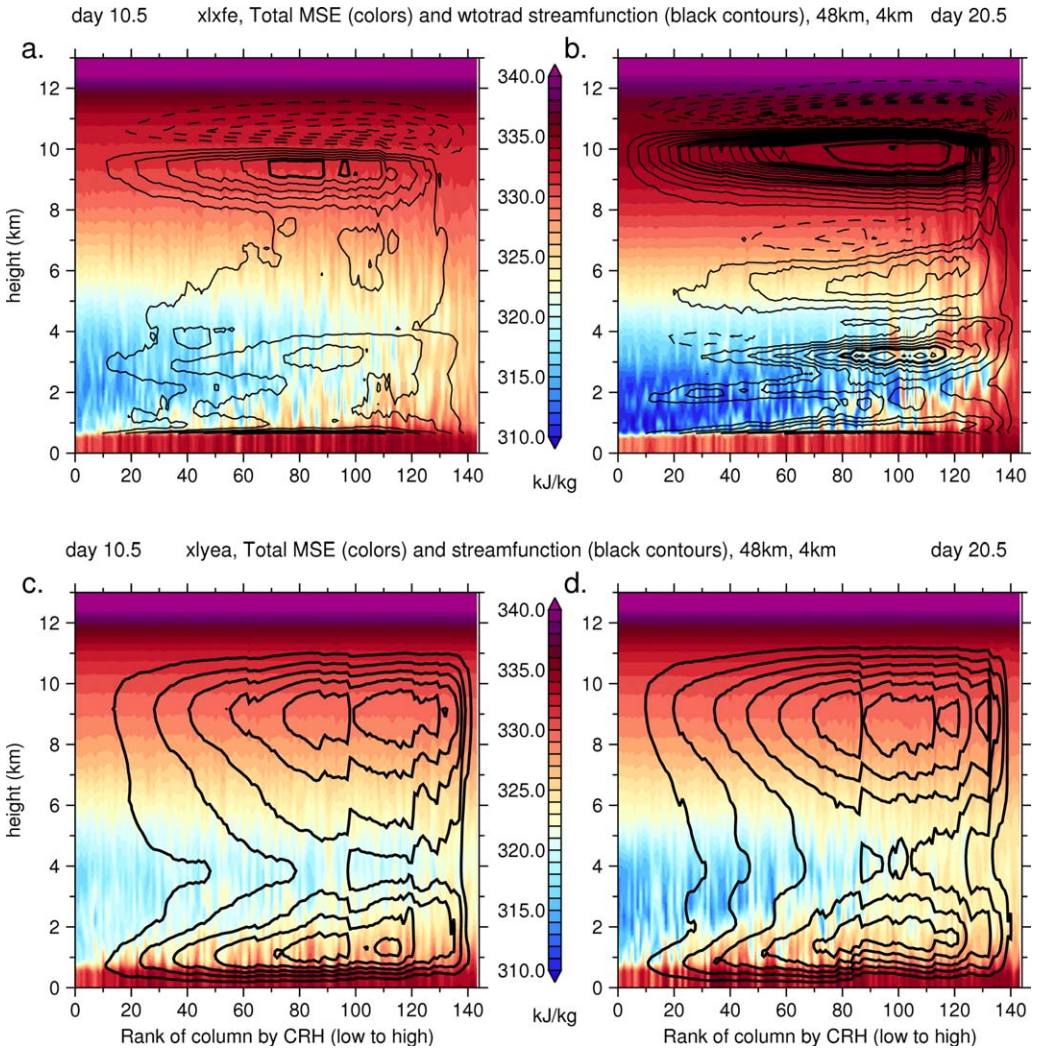

Figure 11. Daily mean values of 48 km × 48 km blocks for the control run, sorted in the *x* dimension by CRH, of *h* (kJ kg^{-1}, colors) and Ψ_{rad} (black lines, negative lines solid, contour interval is 0.01 kg m^{-2} s^{-1} starting from ±0.01, with thick black lines at 0.06 kg m^{-2} s^{-1} starting from ±0.06), for (a) day 10.5 and (b) day 20.5. Lower plots show corresponding *h* for the constant radiation run along with normal Ψ from that run (thick black lines, negative lines solid, contour interval is 0.06 kg m^{-2} s^{-1} starting from ±0.06) for (c) day 10.5 and (d) day 20.5. The *x* axis shows the rank of the boxes sorted by column relative humidity (CRH), driest at the left; the sorting is done for each day.

Ψ_{rad} describes the circulation that would develop if free-tropospheric horizontal anomalous radiative heating was exactly balanced by vertical advection of the mean dry static energy profile. Figures 11a and 11b show block-averaged daily mean *h* and Ψ_{rad} for the control run at days 10.5 and 20.5. Note that the contour intervals for the thin black contours in Figure 11 are only one sixth the magnitude of the thick black contours, which have the same interval as the contours of Ψ in Figures 7 and 8. Comparing Figures 11a and 11b with Figures 7a and 8a, we see that there is little or no lower-tropospheric circulation described by Ψ_{rad} below 2 km on day 10.5, whereas the Ψ_{rad} circulation below 2 km is roughly one tenth the magnitude of the corresponding circulation described by Ψ for day 20.5, suggesting that anomalous radiative forcing is at most a minor component of this circulation. A separate analysis with an analogous $\Psi_{convective}$ calculated using diabatic heating from convective and turbulence processes (approximated by subtracting the radiative and advective model temperature increments from the total temperature increment) confirms that the full Ψ is very well approximated by this, suggesting that horizontal convective heating anomalies drive the main overturning circulations including those below 2 km (not shown). On the other hand, the mid-level circulation around 3–4 km height described by Ψ_{rad} on day 20.5 is about half the size of the equivalent circulation in Ψ on that day, so this circulation (which is a negative feedback on aggregation) does appear to have

a significant component due to anomalous radiative forcing. There are also significant radiative contributions to the upper-tropospheric circulation, which would also oppose aggregation, especially on day 20.5.

Although the simulation with fixed prescribed radiative cooling does not have strong aggregation, it still develops low-level circulations that are somewhat analogous to those in the control run. Figures 11c and 11d show Ψ for this simulation at days 10.5 and 20.5, respectively, with all black contours in bold to show that they have the same interval as those in Figures 7 and 8. Both days have a small positive convergence term in this run (Figure 10b). While the horizontal h gradients are relatively small due to weak aggregation, the low-level inflow is slightly higher up, and the low-level circulations are somewhat less embedded in the moistest columns compared with the control, the presence of these kinds of circulations in this run is more evidence that interactive radiation is not a necessary driver for them. The run with fixed radiative cooling and no rain evaporation also has low-level circulations which would favor aggregation at early times, although they also have a deep circulation that would favor aggregation during the self-aggregation period as well (not shown).

6. Sensitivity to Initial State

To test the sensitivity to the initial state, we have performed three 40 day simulations that are initialized with the final, aggregated state of the control run at the end of its fortieth day. The three simulations have either fixed prescribed surface fluxes, fixed prescribed radiative cooling, or both, and the prescribed forcing is the same as that used for some of the simulations described in section 5 above. The main finding from these tests is that the run with interactive radiation but fixed surface fluxes stays aggregated (Figure 12), unlike the corresponding run starting from homogeneous conditions. This, along with our findings that a similar version of the latter run did show self-aggregation with higher constant surface flux values, provides more evidence that surface flux feedbacks are not a fundamental process causing self-aggregation in our model and certainly are not necessary to maintain an already aggregated state. The latter point is supported by WE14 and MH12.

The two runs with fixed prescribed radiative cooling (but with rain evaporation included) have a final state that does not depend on the initial aggregation state, as shown by the blue lines in Figure 12. These runs return to values of domain-mean CWV, IQR, UQR, CWV autocorrelation length scale, $\hat{h}'^2$, and saturation-normalized UQR that are very similar to those that were reached by the corresponding runs shown in Figure 9 that were initialized with homogeneous conditions (see also Figures S7–S9 in the supporting information). Furthermore, they reach these equilibrium levels in only about 10 days (for mean CWV and UQR), roughly half the time that the control simulation starting from homogeneous conditions took to self-aggregate (but also only changing about half as much in these metrics). Note that MH12 found sometimes quite long disaggregation times, over 80 days in some cases as seen in their Figure 8, and that these times were longer for larger domains and/or finer grid scale: however, these were all for runs with fully interactive radiation and surface fluxes. These simulations disaggregated (within 100 days or fewer) for domains of about 200 km $\times$ 200 km or smaller, and circulation effects leading to disaggregation would have been competing with strong diabatic effects maintaining aggregation.

The simulation with constant surface fluxes (the red dashed line in Figure 12) exhibits significant variability relative to the final state of the control run, with higher mean CWV, IQR, and $\{\hat{h}'^2\}$ but lower UQR and autocorrelation length scale. The mean CWV is still significantly lower than that of the corresponding run from homogeneous initial conditions which did not self-aggregate.

7. Effects of Colder SST

Several studies have mentioned a sensitivity of self-aggregation to SST in idealized simulations of RCE [*Khairoutdinov and Emanuel*, 2010; *Wing and Emanuel*, 2014; *Emanuel et al*., 2014]. For instance, WE14 found that simulations with their model setup did not produce self-aggregation with SST at or below 300 K; also, they found no self-aggregation for simulations with SST at or above 310 K. For the high-SST cases, WE14 speculate that they do not self-aggregate because they require a larger area of subsidence and thus a larger domain to support the aggregation process, since the larger dry static stability results in weaker subsidence

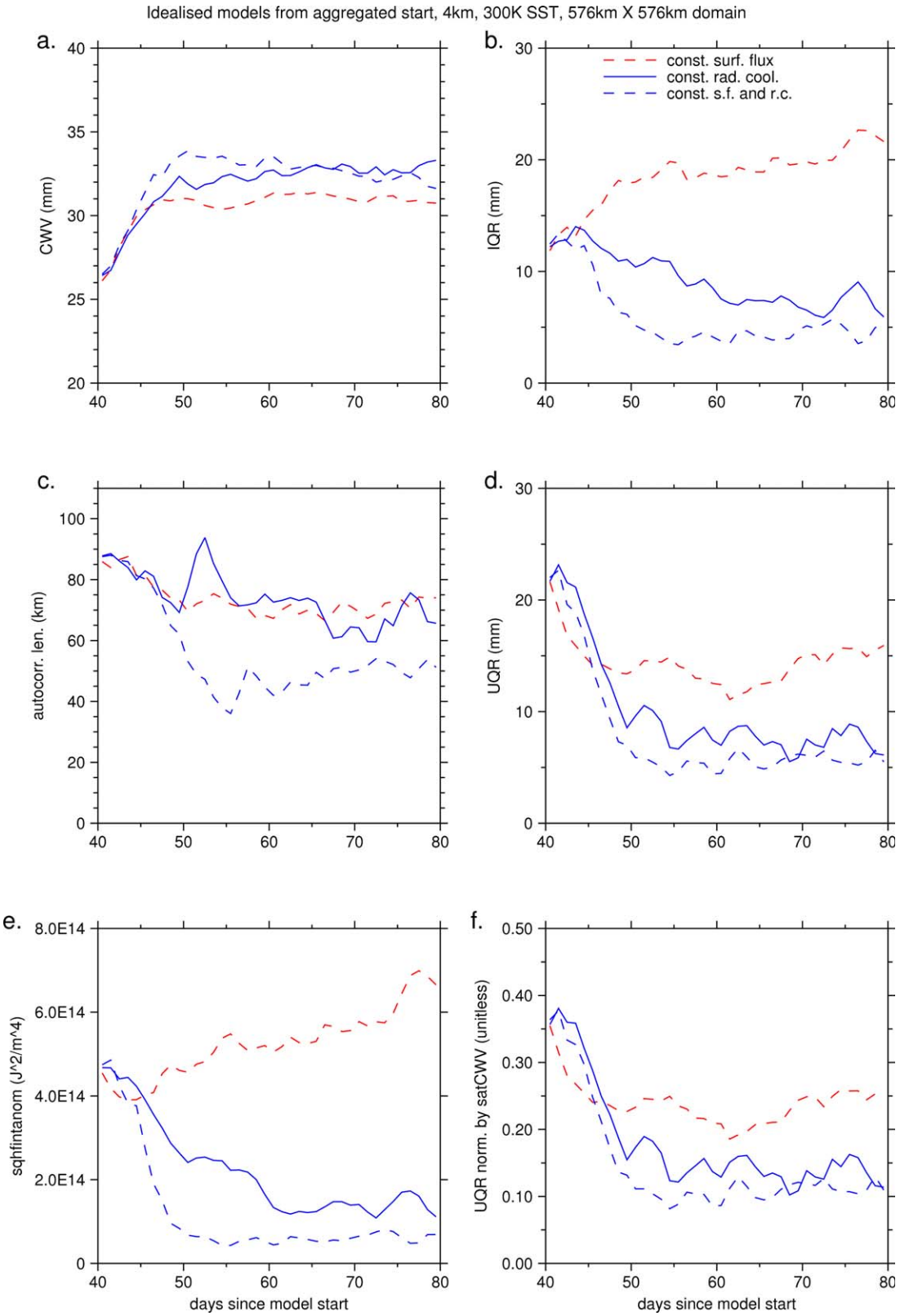

Figure 12. Daily mean values of (a) domain-mean CWV (mm), (b) interquartile range of CWV over domain at each hour (mm), (c) mean of autocorrelation length scale for all rows and columns at each hour (km), (d) "upper-quantile" range (UQR, 95th percentile minus median) of CWV over domain at each hour (mm), (e) $\{\hat{h}'^2\}$ ($J^2\ m^{-4}$), and (f) UQR divided by daily mean domain-mean saturation CWV (unitless), for three idealized simulations which all start from the final state of the 40 day control run. Note that day "40" is the start time for these runs.

for a given unit of area. In fact, they do see aggregation at 310 K SST when they increase the domain size to be 1536 km × 1536 km, four times their original domain.

To investigate the sensitivity of self-aggregation to colder SSTs, we have run two simulations, similar to the control discussed above, except that the SST is fixed at 295 K in one and 290 K in the other. The initial temperature and humidity profiles for each of these two simulations are taken from a mean of the final 10 days of a smaller 70 day, 96 km × 96 km domain RCE run with the same SST. The large runs have been run for only 20 days, mainly because they already showed signs of significant aggregation and at least an approach toward equilibrium by the end of that time. Figure 13 shows that there is an aggregation process that occurs over the first 15 days, although this process may still be slowly continuing at the end of the runs, particularly for the 295 K run (see also Figures S10–S11 in the supporting information). Autocorrelation length scales of CWV for the colder SST runs (Figure 13c) reach values that only slightly smaller than those in the 300 K control (Figure 9c). While the actual values of peak IQR and UQR, and the magnitude of the reduction of domain-mean CWV, are smaller for lower SSTs, this is a consequence of the much smaller column-integrated saturation specific humidity for moist adiabats lifted from moist boundary layers above these SSTs. Figure 13f shows that the saturation-normalized UQR, which is similar to a UQR based on CRH, actually reaches levels that are a little higher (over an even shorter time period) than the 300 K control (Figure 9f).

For the same reason that the IQR and UQR cannot reach values that are as large as those in the control even for an aggregated state, the magnitude of $\{\hat{h}'^2\}$ (Figure 13e) is smaller for an aggregated state for lower SSTs compared to the control at 300 K. For this reason, we show the domain-mean $\hat{h}'^2$ budget terms normalized by $\{\hat{h}'^2\}$ in Figures 14a and 14b for the 295 and 290 K simulations, respectively. Compared with the analogous plot for the 300 K control run (Figure 4a), the normalized diabatic term is slightly larger in both runs, with the normalized shortwave term in the 295 K run roughly 1.5 times the equivalent term in the control run and even larger in the 290 K run. The normalized longwave terms are similar in magnitude to the control run for the 295 K run and slightly smaller in the 290 K run, and they are dominated by cloud effects as in the control run. The shortwave effect also has a significant cloud component in these colder-SST runs, especially in the 290 K run. The normalized surface flux terms are similar in the first 5 days and then smaller (more negative) at later times. The normalized convergence terms are fairly similar to the control overall.

Figure 15 shows Ψ and vertical profiles of h and radiation for 48 km × 48 km blocks ranked by CRH for the 290 K run at 7.5 days, when the convergence term is near zero. There is a low-level circulation around 1 km which likely has a net transport of h toward moist columns (a positive contribution). However, an upper-level circulation centered around 5–7 km likely has the opposite effect. Radiatively driven Ψ_{rad} contours (not shown) reveal that the circulation centered near 1 km is approximately half radiatively driven, although it is right near the undefined lower 0.6 km (where WTG does not apply). The lower part of the upper-level circulation also has a substantial radiatively driven component, but it has a positive gross moist stability, exporting MSE from the moist columns. As discussed above, the net circulation effect for this run is small.

Emanuel et al. [2014] suggest that the nature of moisture-radiation feedback implies that the shortwave feedback (as measured here by the covariance term in the $\hat{h}'^2$ budget) should indeed be large and positive at cold SSTs (as we found), since subsidence drying further reduces shortwave absorption in dry regions (although we also have a significant positive shortwave feedback in moist regions that is colocated with thick cloudy layers). However, they argue that the longwave feedback becomes increasingly negative at colder SSTs (unlike in our simulations), since subsidence drying strongly reduces cooling by longwave emission in dry regions. Since their argument concerns only clear-sky effects, our simulations agree with this for clear-sky longwave for most days, but again cloud effects dominate longwave effects in our runs after the first day.

As discussed above, longwave radiation is still clearly a positive feedback for self-aggregation even at colder SSTs in our simulations, unlike in *Emanuel et al*. [2014] but similar to *Wing and Cronin* [2016]. It is likely that the thick anvil ice clouds (reaching very low down in the moist regions) are important for this positive feedback, allowing for reduced longwave emission in the moist areas (Figures 15d and 15f). Low-level cloud in parts of the dry regions may also help increase radiative cooling at the top of the boundary layer there. Overall, the assumption in *Emanuel et al*. [2014] that clear-sky effects dominate radiation feedbacks in self-aggregation likely break down in our simulations in general, and particularly in these colder SST runs.

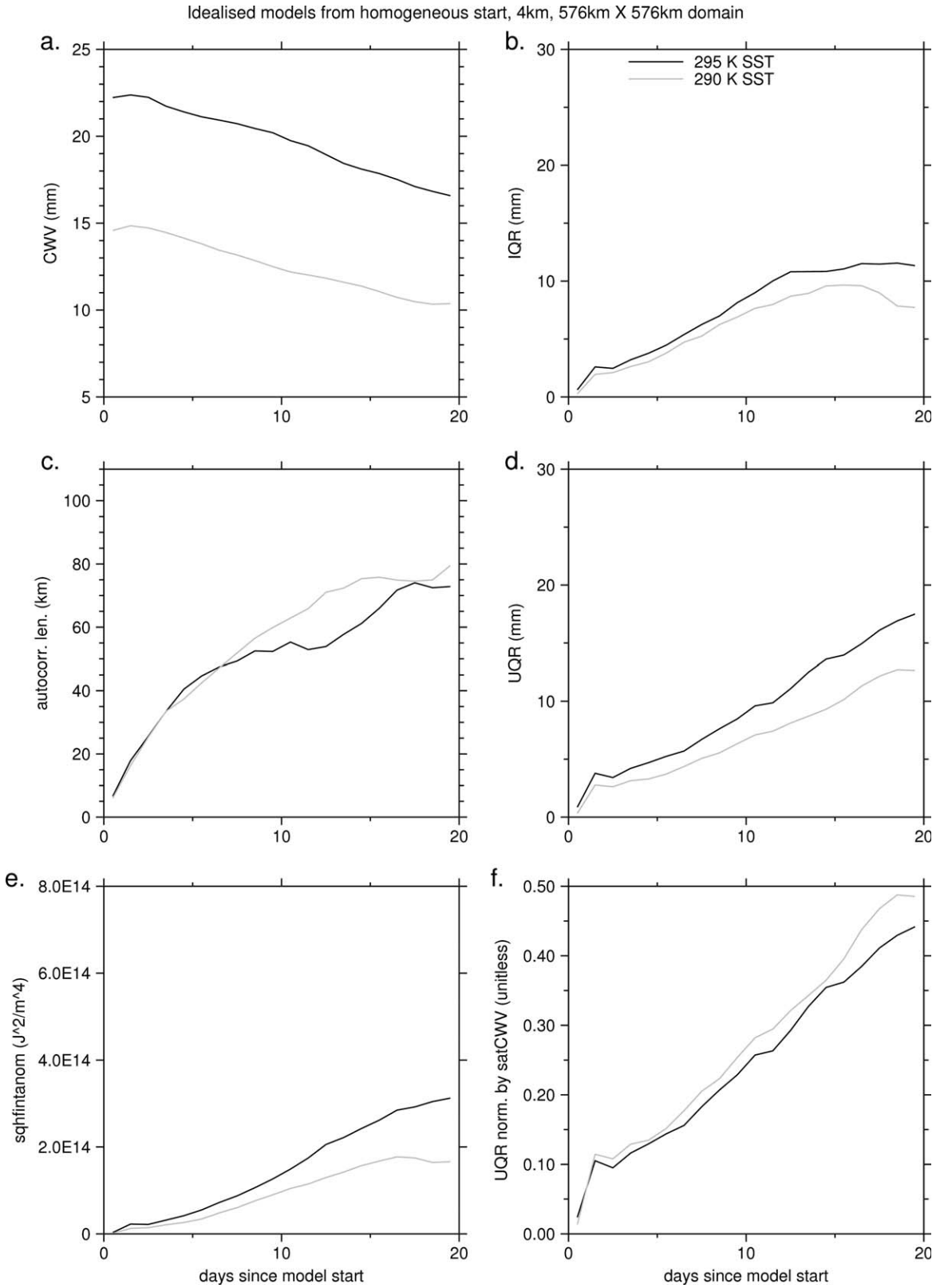

Figure 13. Daily mean values of (a) domain-mean CWV (mm), (b) interquartile range of CWV over domain at each hour (mm), (c) mean of autocorrelation length scale for all rows and columns at each hour (km), (d) "upper-quantile" range (UQR, 95th percentile minus median) of CWV over domain at each hour (mm), (e) $\{\hat{h}'^2\}$ ($J^2\ m^{-4}$), and (f) UQR divided by daily mean domain-mean saturation CWV (unitless), for two idealized simulations which start from homogeneous conditions and constant SST of 295 K (black lines) and 290 K (grey lines).

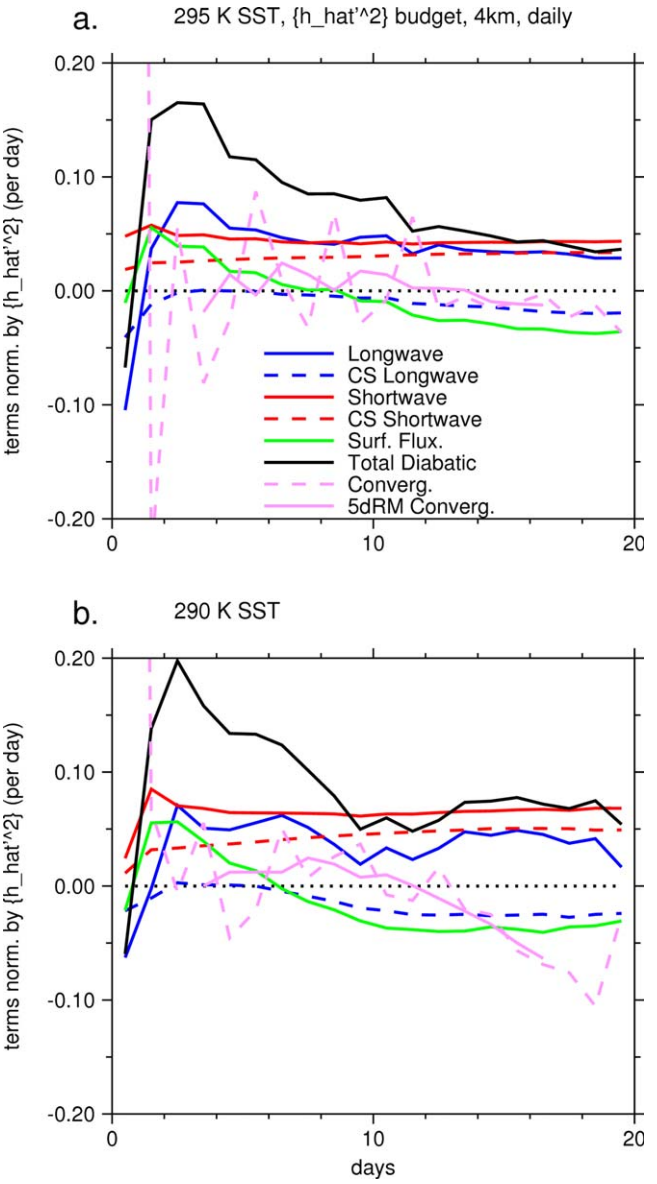

Figure 14. Daily mean values of domain-mean terms in the $\hat{h}'^2$ budget, each normalized by $\{\hat{h}'^2\}$, in unit day^{-1}. The lines show the sum of all diabatic terms (black) as well as contributions from longwave radiation (blue solid), clear-sky-only longwave (blue dashed), shortwave radiation (red solid), clear-sky-only shortwave (red dashed), surface heat flux (green), and horizontal convergence of h flux (pink dashed), with a 5 day running mean of the latter shown as pink solid, for the runs started from homogeneous conditions and (a) 295 K SST and (b) 290 K SST.

8. Discussion and Conclusions

Using the Met Office Unified Model at 4 km grid spacing in idealized RCE mode, we reproduce several findings of previous studies [e.g., *Bretherton et al*., 2005; *Muller and Held*, 2012; *Wing and Emanuel*, 2014] regarding the evolution of self-aggregation in idealized models, including the reduction of domain-mean CWV with increasing aggregation, the increase of domain-mean OLR, the lack of change in reflected shortwave radiation, and the increase of IQR. This is valuable for the community because the model used here is different from the model used in the studies above, the SAM. Self-aggregation has also been found in several other models including the DAM [*Jeevanjee and Romps*, 2013] and RAMS [*Stephens et al*., 2008], but the present study attempts to make close comparisons to several recent studies, specifically regarding mechanisms involved in self-aggregation.

Using WE14's column-integrated frozen moist static energy variance ($\hat{h}'^2$) budget analysis, we find that diabatic covariance terms (radiation and turbulent surface fluxes) are most important for self-aggregation at early times in our control run, and that surface fluxes are especially important in the first few days, as in WE14. While interactive surface fluxes in WE14 (and here to a lesser extent) appear to be important for self-aggregation at early times, and may be necessary to get self-aggregation from a homogeneous state at high SSTs [*Wing*, 2014; D. Coppin, personal communication, 2015], we know of no study in which surface flux feedbacks are necessary to *maintain* aggregation that is already in place. This suggests that surface flux feedbacks may not be important for the occurrence of large-scale convective organization in nature, where there are many inhomogeneities in both forcing terms and dynamical effects on large scales which could allow small-scale clustering. On the other hand, *Tobin et al*. [2012] find that satellite analyses of precipitating regions reveal surface fluxes that are more than 20 W m^{-2} larger within deep convective regions than within the environment (regardless of aggregation state), which would be a positive feedback for aggregation even in aggregated states, unlike the near-zero surface flux term in the $\hat{h}'^2$ budget for mature aggregation in our control run.

The convergence term, representing advective or mixing processes transporting h between regions, is not as large or as dominant during intermediate stages of our control run as it is in WE14. Nevertheless, we do see evidence that low-level circulations between dry and moist regions are important for this convergence

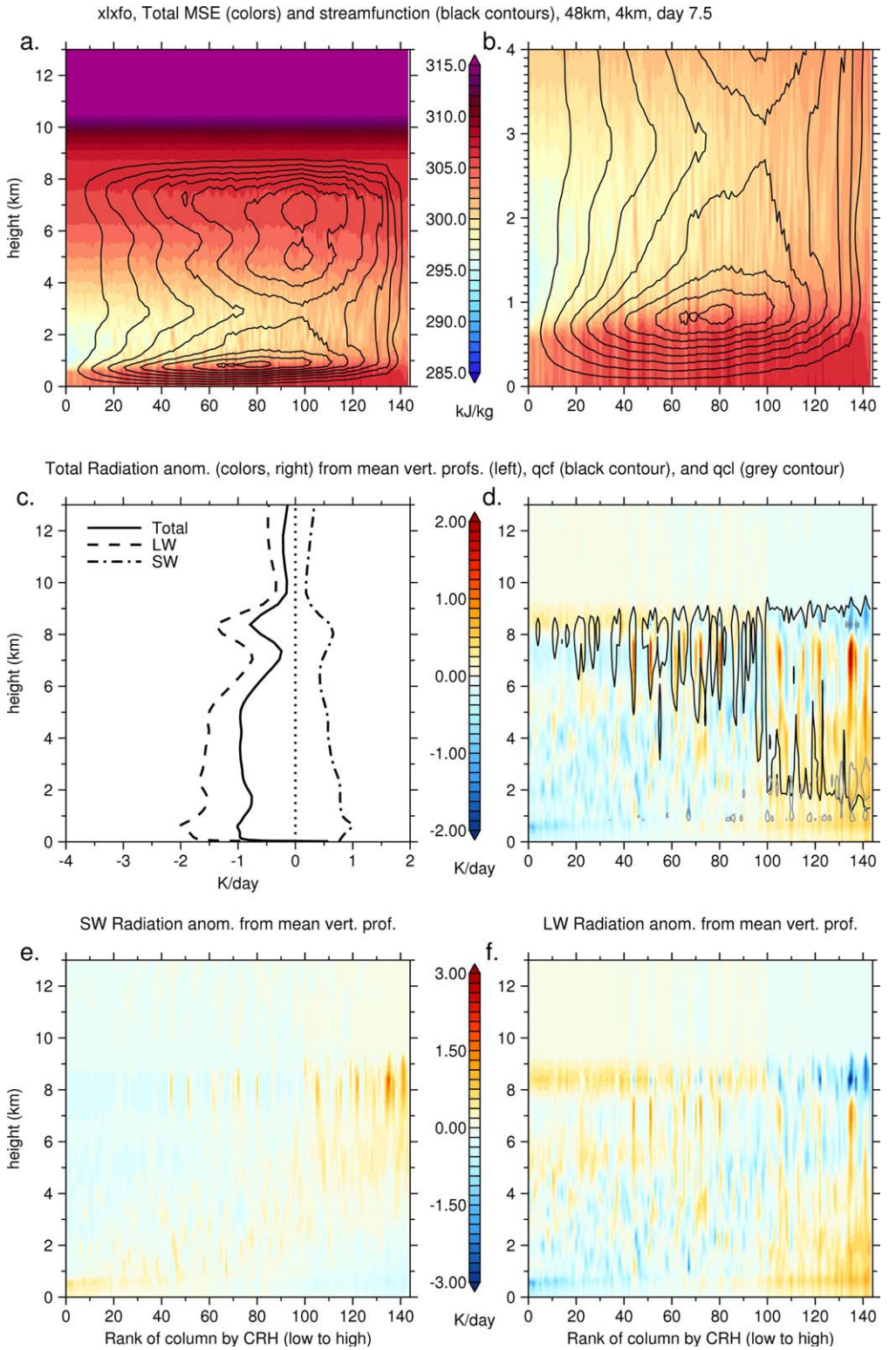

Figure 15. The 290 K SST simulation, day 7.5 mean values of 48 km × 48 km blocks, sorted in the *x* dimension by CRH, for (a) $\hat{h}$ (kJ kg^{-1}, colors) and Ψ (black lines, negative lines solid, contour interval is 0.06 kg m^{-2} s^{-1} starting from ±0.06), (b) lowest 4 km heights zoomed from Figure 15a, (c) profiles of domain-day-mean total, longwave, and shortwave radiative heating (K d^{-1}), (d) horizontal anomaly (from domain-day-mean profile) of total radiation (K d^{-1}) (colors) with 5% cloud fraction contours of ≥1 mg kg^{-1} cloud ice/snow (black line) and ≥1 mg kg^{-1} cloud liquid (grey line), (e) horizontal anomaly of shortwave radiation (K d^{-1}), and (f) horizontal anomaly of longwave radiation (K d^{-1}). The *x* axis shows the rank of the boxes sorted by column relative humidity (CRH), driest at the left; the sorting is done for each day.

term. Unlike previous studies, however, we find that these circulations are not primarily driven by low-level gradients in radiative cooling between these regions. These circulations are presumably driven mainly by anomalies in low-level diabatic heating from convection and related microphysical processes, at least in the control run at 300 K. Relatedly, there are somewhat analogous low-level circulations, of magnitude similar to that of the deep circulations, in the constant radiation simulations. These tend to have low-level inflow at somewhat higher altitudes and are less embedded in the very moistest regions compared with the control (and of course there is less of a CWV and *h* difference between the moist and dry columns since aggregation is much weaker in these constant radiation runs).

We also find a mid-level circulation in the control run (with a significant contribution from anomalous radiation) that acts as a negative feedback for aggregation. The deep tropospheric circulation also evolves during the aggregation process so as to contribute a less positive (or more negative) component to the convergence term as the upper-troposphere warms and the upper-level *h* increases more than the lower-level *h* in the moist regions. When the total convergence term becomes significantly negative the aggregation stops. This is somewhat a consequence of a closed budget, so it is a bit hard to argue cause and effect, but it seems important for developing aggregation in our model that the convergence term is at least near zero or larger.

It is the direct diabatic radiative covariance terms (both longwave and shortwave) that we find to be most important for self-aggregation in the control simulation and in the two colder SST runs. The positive longwave term is mainly due to cloud effects for all runs, and the shortwave term has significant cloud contributions in the two colder SST runs. This agrees with *Wing and Cronin* [2016]. In a domain-mean sense, both radiation terms appear to maintain a mature aggregated state against a negative convergence term (whereas in WE14, it was mainly the shortwave term that was positive near the end of the control run). Furthermore, we find that the moist, cloudy columns contribute the largest proportion of positive radiative feedback (mainly because of the longwave term) even at fairly early stages of aggregation, in contrast to the preponderance of clear-sky effects stemming from the driest columns found at early stages in WE14. This may be because we do not observe self-aggregation beginning as a single dry patch that expands but rather as both moist and dry patches growing and merging simultaneously. In the mature stages, our results agree with WE14 that the strongest local positive feedbacks come from longwave effects in the moistest columns.

We note that radiation may have a direct diabatic positive effect in the column-integrated sense but, depending on its vertical location and the mean *h* stratification, can have different signs and magnitudes of convergence term impacts. For instance, *Chikira* [2014] discusses the various impacts of radiation anomalies at different vertical levels on different phases of the MJO. *Muller and Bony* [2015] found that self-aggregation was sensitive to the vertical structure of anomalous radiative cooling applied only in dried regions rather than just the vertical mean cooling: this could be due to circulations that have a positive enough GMS to result in a large enough negative convergence term which offsets any positive radiation anomaly term in their simulations.

Initial aggregation state is not important in determining the final state for our runs with constant radiation (but with normal rain evaporation). Initial state is important for the constant surface flux (but interactive radiation) experiments, with an ability for these runs to hold on to strong aggregation when it is in the initial conditions but for a lack of self-aggregation from homogeneous conditions (although even initial homogeneous conditions lead to strong self-aggregation when we used larger constant surface flux values). The lack of self-aggregation for fixed surface fluxes and homogeneous initial conditions is partly related to weaker positive radiation feedbacks, including weaker longwave feedbacks in cloudy regions and negative clear-sky contributions in dry regions. Another potential reason is that there is very vigorous cold pool activity in this run, causing a negative convergence term even at early times. This is because cold pools cannot be strongly damped without interactive surface fluxes; the higher the constant surface flux values, the more damped the cold pools will be, which would explain the sensitivity to the constant values.

Colder SSTs (down to 290 K) still lead to self-aggregation at similar rates as in the 300 K control, unlike runs which failed to aggregate below about 300 K in some previous studies [*Khairoutdinov and Emanuel*, 2010; *Wing and Emanuel*, 2014] but in agreement with others [*Abbot*, 2014; *Wing and Cronin*, 2016]. *Wing and Cronin* [2016] suggest that this difference is due to their "bowling alley" geometry, but the present study has

square geometry like the simulations in the studies which did not find self-aggregation at colder SSTs. The shortwave term (normalized by $\{\hat{h}'^2\}$) is stronger than in the control run, while the longwave term is similar or slightly smaller. These colder-SST results contradict the *Emanuel et al*. [2014] hypotheses and simple model results, particularly because we have a strong longwave feedback. This discrepancy is most likely due to the fact that we have strong cloud effects for the radiative terms in our model, with positive contributions mainly from the moist, cloudy region, whereas *Emanuel et al*. [2014] focus on clear-sky feedbacks. We note that *Emanuel et al*. [2014] mainly address the initial linear growth phase of self-aggregation, whereas it is possible that the simulations in the present study largely bypass this linear phase.

We find that a simulation similar to the control but with suppression of rainfall evaporation (at all model levels in our simulation) results in faster self-aggregation, agreeing with *Jeevanjee and Romps* [2013] that cold pools interfere with (and presumably delay) self-aggregation. They hypothesized that this occurs because cold pools mix low-level air masses and reducing h gradients at low levels; in our simulation with no rain evaporation (but with interactive radiation), it appears that the initial longwave radiative effect is especially important, with the convergence term not large but slightly larger than the control for the first 6 days, but this may be model (or domain or resolution) dependent.

Emanuel et al. [2014] suggest that moisture-convection feedback is not crucial for self-aggregation. On the other hand, *Craig and Mack* [2013] argue that this is the fundamental feedback. Although we find that aggregation is very weak in our constant radiation simulations, those simulations still have significant autocorrelation length scales of CWV. This may relate to the "moisture-memory" feedback discussed most recently in *Muller and Bony* [2015], which seems to cause self-aggregation in their simulations without cold pools even without interactive radiation. We find that we do get strong self-aggregation, even in the absence of interactive radiation, when we suppress the evaporation of rainfall, agreeing with *Muller and Bony* [2015]. We also find evidence of this process in other runs without interactive radiation, but as mentioned above it is much weaker. Indeed, one result of our simulations is that aggregation occurs more along a spectrum, with various degrees of organization, than in previous studies, and this may be partly due to a "moisture-memory" feedback. Such a feedback would stem from moister environments favoring convection by resulting in less reduction of updraft buoyancy due to entrainment, although these processes are likely to be somewhat poorly represented by models with grid spacing on the order of one to a few kilometers and may be sensitive to numerics and resolution. Perhaps this feedback is also easily disrupted by mixing processes, which would mean it may not be as important in the real world as it is in idealized models.

These simulations were carried out in large part because we want to know whether processes important for self-aggregation in idealized models are also important for organized tropical convection in the real world. C. E. Holloway (manuscript in preparation, 2016) looks at limited-area simulations of real case studies of equatorial convection using the same model (with slight differences in a few physics settings) and includes sensitivity tests using mechanism denial setups that are analogous to those used in the present study. That study finds that the overall evolution of convective organization is similar in the different simulations, suggesting the importance of the large-scale forcing at the boundaries, but there are systematic increases in domain-mean CWV and decreases in IQR and UQR for models with constant radiative cooling, which suggests some similarities in how radiation affects the mean state in realistic cases and idealized cases for this model. Future work should continue to test mechanisms found in idealized self-aggregation in more realistic simulations and in observations of the real world.

Appendix A: Bands Versus Circular Patches in Biperiodic Idealized RCE

Here we present a geometric argument for the existence of different shapes (bands versus patches) for aggregated convection in idealized simulations with doubly periodic domains. We can define aggregation in idealized models as a separation between moist patches and the dry environment (or vice versa, for which the arguments below would also hold). The moist patches could be defined by choosing an appropriate contour of CWV (or CRH or $\hat{h}$), preferably one which is embedded in strong CWV gradients. There is no obvious a priori constraint on the size of such a patch, and evidence from this study and others suggest a range of possible areas of moist or dry patches both at different times in a single run and for different equilibrium states of different runs. Since these are vertically integrated quantities, we refer to shapes in the horizontal dimensions only and ignore the vertical dimension. Since lateral mixing would reduce these

gradients, it seems plausible that these moist patches would develop in a way to form shapes that minimize their perimeter-to-area ratio.

If patches of high CWV (or $\hat{h}$) tend to form geometric shapes that minimize their perimeter-to-area ratio, then the preferred shape would be a circle for patches of any area in an infinite plane. However, in an idealized model with lateral boundaries that are periodic in both horizontal dimensions, this is not the case: a patch that forms a "band" shape, which looks like a rectangle that spans the entire domain in one dimension and some fraction of the domain in the other dimension, will have an effective perimeter (that is, the length of the boundary between the moist patch and the dry environment) of $2L$ for any band area, where L is the length of the side of the band that spans the entire domain length in that dimension. If the domain is square as in this paper, then L is equal to both the length and the width of the domain. If the domain is rectangular but not square, then the relevant L for our purposes will be equal to the shorter dimension of the domain, since any band with a given area that stretches across the longer dimension will always have a larger perimeter-to-area ratio than some suitably constructed band of the same area that stretches across this shorter dimension of length L.

Since the perimeter of our hypothetical band is always $2L$, we can then find the radius and area of a circle with a perimeter equal to this. For a patch with this "critical area" A_c, a circle or band will have the same perimeter and thus the same perimeter-to-area ratio. For a patch with an area smaller than A_c, the circle's perimeter will be smaller while the band's perimeter will stay the same, so the circle will be the preferred shape for a patch that minimizes its perimeter-to-area ratio. Accordingly, a patch with an area larger than A_c will have a preferred shape that is a band. Setting $2L = 2\pi r$, we find that $r = L/\pi$ is the radius for a circle with the same perimeter as a band, and therefore $A_c = \pi(L/\pi)^2 = L^2/\pi$. For a square domain, L^2 is the area of the whole domain, so A_c is about one third of the total area of a square domain.

Acknowledgments

We thank Allison Wing, George Craig, Sandrine Bony, Alison Stirling, Bob Plant, Julia Mack, Caroline Muller, and Jun-Ichi Yano for useful discussions. We also thank Kerry Emanuel and an anonymous reviewer for helpful suggestions that have improved the manuscript. We acknowledge support on preliminary work related to this topic by Mike Blackburn and Guiying Yang. Grenville Lister, Carol Halliwell, Peter Clark, and Adrian Lock provided guidance on running the model in idealized mode. This work made use of the facilities of HECToR (and its successor, ARCHER), the UK's national high-performance computing service, which is provided by UoE HPCx Ltd at the University of Edinburgh, Cray Inc and NAG Ltd, and funded by the Office of Science and Technology through EPSRC's High End Computing Programme. Preliminary work for this paper was carried out with funding from UK NERC grant NE/E00525X/1, and C.E.H. was funded from UK NERC grant NE/I021012/1. The MetUM simulation data will be made available through the NERC Centre for Environmental Data Archival (CEDA) at catalogue.ceda.ac.uk: in the meantime, the data can be made available by the authors on request.

References

Abbot, D. S. (2014), Resolved snowball earth clouds, *J. Clim.*, *27*(12), 4391–4402, doi:10.1175/JCLI-D-13-00738.1.

Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov (2005), An energy-balance analysis of deep convective self-aggregation above uniform SST, *J. Atmos. Sci.*, *62*, 4273–4292.

Chikira, M. (2014), Eastward-propagating intraseasonal oscillation represented by Chikira–Sugiyama cumulus parameterization. Part II: Understanding moisture variation under weak temperature gradient balance, *J. Atmos. Sci.*, *71*(2), 615–639, doi:10.1175/JAS-D-13-038.1.

Craig, G. C., and J. M. Mack (2013), A coarsening model for self-organization of tropical convection, *J. Geophys. Res. Atmos.*, *118*, 8761–8769, doi:10.1002/jgrd.50674.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood (2005), A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Q. J. R. Meteorol. Soc.*, *131*(608), 1759–1782.

Emanuel, K., A. A. Wing, and E. M. Vincent (2014), Radiative-convective instability, *J. Adv. Model. Earth Syst.*, *6*, 75–90, doi:10.1002/2013MS000270.

Held, I. M., R. S. Hemler, and V. Ramaswamy (1993), Radiative-convective equilibrium with explicit two-dimensional moist convection, *J. Atmos. Sci.*, *50*(23), 3909–3927.

Holloway, C. E., S. J. Woolnough, and G. M. S. Lister (2013), The effects of explicit versus parameterized convection on the MJO in a large-domain high-resolution tropical case study. Part I: Characterization of large-scale organization and propagation, *J. Atmos. Sci.*, *70*(5), 1342–1369, doi:10.1175/JAS-D-12-0227.1.

Jeevanjee, N., and D. M. Romps (2013), Convective self-aggregation, cold pools, and domain size, *Geophys. Res. Lett.*, *40*, 994–998, doi:10.1002/grl.50204.

Khairoutdinov, M., and K. Emanuel (2010), Aggregation of convection and the regulation of climate. preprints, in *29th Conference on Hurricanes and Tropical Meteorology*, pp. P2.69, Am. Meteorol. Soc., Tucson, Ariz.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell (2008), Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom, *Mon. Weather Rev.*, *136*, 3408–3424.

Muller, C., and S. Bony (2015), What favors convective aggregation and why?, *Geophys. Res. Lett.*, *42*, 5626–5634, doi:10.1002/2015GL064260.

Muller, C. J., and I. M. Held (2012), Detailed investigation of the self-aggregation of convection in cloud-resolving simulations, *J. Atmos. Sci.*, *69*(8), 2551–2565, doi:10.1175/JAS-D-11-0257.1.

Raymond, D. J., and X. Zeng (2005), Modelling tropical atmospheric convection in the context of the weak temperature gradient approximation, *Q. J. R. Meteorol. Soc.*, *131*(608), 1301–1320, doi:10.1256/qj.03.97.

Reed, K. A., B. Medeiros, J. T. Bacmeister, and P. H. Lauritzen (2015), Global radiative–convective equilibrium in the Community Atmosphere Model, version 5, *J. Atmos. Sci.*, *72*(5), 2183–2197, doi:10.1175/JAS-D-14-0268.1.

Roberts, N. M. (2003), The impact of a change to the use of the convection scheme to high-resolution simulations of convective events, *Tech. Rep. 407*, 30 pp., U. K. Met Off., Exeter, U. K.

Sobel, A. H., and C. S. Bretherton (2000), Modeling tropical precipitation in a single column, *J. Clim.*, *13*(24), 4378–4392, doi:10.1175/1520-0442(2000)013<4378:MTPIAS>2.0.CO;2.

Stephens, G. L., S. van den Heever, and L. Pakula (2008), Radiative–convective feedbacks in idealized states of radiative–convective equilibrium, *J. Atmos. Sci.*, *65*(12), 3899–3916, doi:10.1175/2008JAS2524.1.

Su, H., C. S. Bretherton, and S. S. Chen (2000), Self-aggregation and large-scale control of tropical deep convection: A modeling study, *J. Atmos. Sci.*, *57*(11), 1797–1816, doi:10.1175/1520-0469(2000)057<1797:SAALSC>2.0.CO;2.

Tobin, I., S. Bony, and R. Roca (2012), Observational evidence for relationships between the degree of aggregation of deep convection, water vapor, surface fluxes, and radiation, *J. Clim.*, *25*(20), 6885–6904, doi:10.1175/JCLI-D-11-00258.1.

Tobin, I., S. Bony, C. E. Holloway, J.-Y. Grandpeix, G. Sèze, D. Coppin, S. J. Woolnough, and R. Roca (2013), Does convective aggregation need to be represented in cumulus parameterizations?, *J. Adv. Model. Earth Syst.*, *5*, 692–703, doi:10.1002/jame.20047.

Tompkins, A. M. (2001), Organization of tropical convection in low vertical wind shears: The role of water vapor, *J. Atmos. Sci.*, *58*, 529–545.

Tompkins, A. M., and G. C. Craig (1998), Radiative-convective equilibrium in a three-dimensional cloud-ensemble model, *Q. J. R. Meteorol. Soc.*, *124*, 2073–2097.

Wilson, D. R., and S. P. Ballard (1999), A microphysically based precipitation scheme for the UK Meteorological Office Unified Model, *Q. J. R. Meteorol. Soc.*, *125*, 1607–1636.

Wing, A. A. (2014), Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations, PhD thesis, Mass. Inst. of Technol., Cambridge.

Wing, A. A., and T. W. Cronin (2016), Self-aggregation of convection in long channel geometry, *Q. J. R. Meteorol. Soc.*, doi:10.1002/qj.2628, in press.

Wing, A. A., and K. A. Emanuel (2014), Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations, *J. Adv. Model. Earth Syst.*, *6*, 59–74, doi:10.1002/2013MS000269.