Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity

Robert Pincus · David Winker · Sandrine Bony
Bjorn Stevens *Editors*

Springer

INTERNATIONAL
SPACE
SCIENCE
INSTITUTE

Space Sciences Series of ISSI

Volume 65

More information about this series at http://www.springer.com/series/6592

Robert Pincus · David Winker
Sandrine Bony · Bjorn Stevens
Editors

Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity

Editors

Robert Pincus
University of Colorado
Boulder, CO
USA

and

NOAA Earth System Research Lab
Boulder, CO
USA

David Winker
Atmospheric Sciences
NASA Langley Research Center
Hampton, VA
USA

Sandrine Bony
Institute Pierre Simon Laplace
Laboratoire de Météorologie Dynamique
Paris
France

Bjorn Stevens
Max Planck Institute for Meteorology
Hamburg
Germany

Cover illustration: Organization of shallow clouds pictured over the tropical Atlantic near Barbados during the NARVAL2 airborne field campaign (photo by Sandrine Bony).

Printed on acid-free paper

Contents

CrossMark

Preface to the Special Issue "ISSI Workshop on Shallow Clouds and Water Vapor, Circulation and Climate Sensitivity"

Robert Pincus[1] · David Winker[2] · Sandrine Bony[3] ·
Bjorn Stevens[4]

© Springer Science+Business Media B.V. 2017

Each of the 15 chapters in this work presents a paper gleaned from presentations at an International Space Science Institute Workshop on *Shallow Clouds*, *Water vapor*, *Circulation and Climate Sensitivity* organized as part of the World Climate Research Programme's Grand Science Challenge on Clouds, Circulation, and Climate Sensitivity. The workshop's somewhat awkward title reflects the practice of treating each subject separately—a state of affairs that the workshop sought to address. As the roughly 40 participating experts from around the world emphasized, the coupling of clouds and water vapor to circulation helps determine the nature of circulation systems in the past and present as well as the climate sensitivity that characterizes the response of the Earth's surface temperature to radiative forcing.

It has been known for more than a decade that an understanding of factors controlling the distribution and amount of the low-level, fair-weather, clouds over the tropical oceans is critical for determining Earth's climate sensitivity. What has become clear only recently is that these clouds do not simply respond passively to the large-scale circulations in which they form. Studies of clouds and circulations across a range of scales, enabled by increasing computational power, have shown that clouds help set these circulations through their interactions with radiation. Radiative cooling from low clouds drives low-level temperature and pressure gradients that reinforce the regions of gentle subsidence in which they prevail. This pathway is also thought to be responsible for the clustering—or self-

---

✉ Robert Pincus
   Robert.Pincus@colorado.edu

[1]   University of Colorado and NOAA/Earth System Research Lab, Physical Sciences Division, Boulder, CO 80305, USA

[2]   NASA/Langley Research Center, Hampton, VA 23681, USA

[3]   CNRS, LMD/IPSL, Université Pierre et Marie Curie, 75252 Paris, France

[4]   Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany

aggregation—of deep convection seen in simulations with spatially uniform forcing. The relevance of self-aggregation behavior to clouds on Earth was one theme emerging from the workshop.

Water vapor, like clouds, interacts powerfully with radiant energy, and there is a longstanding appreciation that water vapor influences Earth's climate sensitivity, especially through changes in water near the tropopause. But as several papers in this collection highlight, small departures in the relative humidity of the lower atmosphere can be just as important in influencing Earth's radiative balance. Moreover, perhaps no other quantity is as important for patterning the distribution of deep convection. Despite the importance of lower tropospheric humidity for a vast array of climate relevant processes it is poorly characterized in the absence of field campaign measurements, largely because current remote sensing techniques have difficulty unambiguously detecting the structure of water vapor in the tropical boundary layer.

The 15 chapters in this volume expand on these themes. The first set of papers focuses on convective self-aggregation, or the tendency of convection to organize even in the absence of external influences. Allison Wing and her co-authors provide a timely review of a rapidly expanding literature. Following this lead two groups, led by Christopher Holloway and Matthew Lebsock, respectively, explore the observational record for signatures of convective self-aggregation. A team consisting of Louise Nuijens, Kerry Emanuel, Hirohiko Masunaga and Tristan L'Ecuyer asks to what extent radiative forcing from somewhat deeper clouds, cumulus congestus, also influence the pattern of large-scale circulation. Paquita Zuidema and Giuseppe Torri consider the disaggregating aspect of deep convection, namely the cold pools that propagate away from convecting areas and initiate deeper convection well away from aggregated clusters.

Two articles address the question as to how shallow clouds influence climate sensitivity. One, a concise summary by Stephen Klein et al., synthesizes five recent studies leveraging cloud controlling factors to try and infer how low clouds may change in a warming climate. The other, by Jessica Vial, Sandrine Bony and their co-authors, explores why climate models are so sensitive to the representation of low clouds, and to what extent more fundamental modeling approaches, like large-eddy simulation, may help to constrain their behavior.

The importance of water vapor is reviewed in two articles, one led by Brian Mapes and one by Bjorn Stevens, Helene Brogniez and co-authors. The latter article, and one by Robert Pincus and his co-authors, assesses the ability of our present observing system to characterize the distribution of water in the lower troposphere.

The book's final section looks to the future. New technologies for observing lower tropospheric water vapor are becoming more widely available. An article by Christopher Kiemle demonstrates the use of one such technology in airborne field campaigns, while the contribution led by Amin Nehrir surveys the range of emerging technologies being developed for air- and space-borne measurements. The prospects of using related space-based observations of clouds to constrain the temperature-mediated dependence of clouds and circulation is reviewed by Dave Winker, Helene Chepfer and Xia Cai. Observational strategies for directly exploring how satellite observations might be used to measure shallow circulations are explored in an article by Gilles Bellon, Oliver Reitenbuch and Ann Kristin Naumann—an especially timely contribution given that the first satellite-based wind lidar will be launched just as this work appears. Finally, Sandrine Bony, Bjorn Stevens and a large international team of co-authors outline an exciting new field campaign in which they plan to measure the interaction of low clouds, water vapor and circulation with the goal of putting better bounds on Earth's climate sensitivity.

CrossMark

Convective Self-Aggregation in Numerical Simulations: A Review

Allison A. Wing[1,2] · **Kerry Emanuel**[3] · **Christopher E. Holloway**[4] · **Caroline Muller**[5]

Abstract Organized convection in the tropics occurs across a range of spatial and temporal scales and strongly influences cloud cover and humidity. One mode of organization found is "self-aggregation," in which moist convection spontaneously organizes into one or several isolated clusters despite spatially homogeneous boundary conditions and forcing. Self-aggregation is driven by interactions between clouds, moisture, radiation, surface fluxes, and circulation, and occurs in a wide variety of idealized simulations of radiative–convective equilibrium. Here we provide a review of convective self-aggregation in numerical simulations, including its character, causes, and effects. We describe the evolution of self-aggregation including its time and length scales and the physical mechanisms leading to its triggering and maintenance, and we also discuss possible links to climate and climate change.

1 Introduction

Radiative–convective equilibrium (RCE) is the statistical equilibrium state that the atmosphere and surface would reach in the absence of lateral energy transport, in which there is a balance between net radiative cooling and convective heating. It has long been

✉ Allison A. Wing
  awing@fsu.edu

1   Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

2   Present Address: Florida State University, Tallahassee, FL, USA

3   Massachusetts Institute of Technology, Cambridge, MA, USA

4   University of Reading, Reading, UK

5   Laboratoire de Météorologie Dynamique, École Normale Supérieure, Paris, France

used as an idealization of the tropical atmosphere in simulations with single-column models (e.g., Manabe and Strickler 1964; Renno et al. 1994), two- and three- dimensional cloud-resolving models (e.g., Held et al. 1993; Bretherton et al. 2005), and regional/global models with parameterized convection (e.g., Held et al. 2007).

Convective self-aggregation is the spontaneous spatial organization of convection in numerical simulations of radiative–convective equilibrium despite homogeneous boundary conditions and forcing. This instability of the RCE state arises due to interactions among convection, radiation, environmental moisture, and surface fluxes.

Aggregation has been found to occur in simulations with two-dimensional cloud-resolving models (Held et al. 1993; Grabowski and Moncrieff 2001, 2002; Stephens et al. 2008), small-domain square three-dimensional cloud-resolving models (Tompkins and Craig 1998; Bretherton et al. 2005; Khairoutdinov and Emanuel 2010; Muller and Held 2012; Jeevanjee and Romps 2013; Wing and Emanuel 2014; Abbot 2014; Muller and Bony 2015; Holloway and Woolnough 2016; Hohenegger and Stevens 2016), elongated channel three-dimensional cloud-resolving models (Tompkins 2001; Posselt et al. 2008, 2012; Stephens et al. 2008; Wing and Cronin 2016), regional/global models with parameterized clouds and convection (Su et al. 2000; Held et al. 2007; Popke et al. 2013; Becker and Stevens 2014; Reed et al. 2015; Arnold and Randall 2015; Reed and Medeiros 2016; Coppin and Bony 2015; Silvers et al. 2016; Hohenegger and Stevens 2016; Bony et al. 2016) or super-parameterized clouds and convection (Arnold and Randall 2015), and global models with explicit convection (Satoh and Matsuda 2009; Satoh et al. 2016).

An example of self-aggregation in a cloud-resolving model (CRM) with no rotation is found in Fig. 1, which shows snapshots of outgoing longwave radiation, where low values indicate the presence of high, deep convective clouds. Initially, convection is quasi-randomly distributed across the domain (Fig. 1a), but after tens of days, the convection has aggregated into a single, intensely precipitating moist cluster surrounded by dry, subsiding air (Fig. 1b).

An example of self-aggregation in aquaplanet simulations of non-rotating RCE with several general circulation models (GCMs) with parameterized convection is found in Fig. 2. In GCMs, self-aggregation entails the development of a few isolated clusters of deep convection within a large area of subsidence, and the development of a large-scale circulation.

Spontaneous symmetry breaking in two-dimensional and two-column models (Raymond 2000; Nilsson and Emanuel 1999) and multiple equilibria in weak temperature gradient simulations with single column (Sobel et al. 2007; Emanuel et al. 2014; Daleu et al. 2015) and two-dimensional cloud-resolving models (Sessions et al. 2010, 2015, 2016) have also been interpreted as analogous to/consistent with convective self-aggregation. In single column or small-domain CRM weak temperature gradient simulations, in which the large-scale circulation is parameterized, the dry equilibrium corresponds to the dry, non-convecting regions of a self-aggregated state and the equilibrium with precipitation corresponds to the moist cluster in a self-aggregated state. Theory and simple models of aggregation (or its root cause, instability of the RCE state) have been presented by Bretherton et al. (2005), Craig and Mack (2013), Emanuel et al. (2014) and Beucler and Cronin (2016).

The above refers to aggregation under conditions of non-rotating RCE. Under conditions of *rotating* RCE, self-aggregation takes the form of spontaneous genesis of tropical cyclones (Bretherton et al. 2005; Nolan et al. 2007; Held and Zhao 2008; Khairoutdinov and Emanuel 2013; Shi and Bretherton 2014; Zhou et al. 2014; Boos et al. 2015; Reed and Chavas 2015; Davis 2015; Wing et al. 2016; Merlis et al. 2016; Zhou et al. 2017). While aggregation occurs across a wide variety of different models, with different dynamical

Fig. 1 Snapshot of outgoing longwave radiation (OLR) at **a** day 10 and **b** day 80 of a radiative–convective equilibrium simulation at 305 K. Reprinted from Wing and Emanuel (2014). ©2013. American Geophysical Union. All Rights Reserved

(a) Day 10 OLR (W/m^2)

(b) Day 80 OLR (W/m^2)
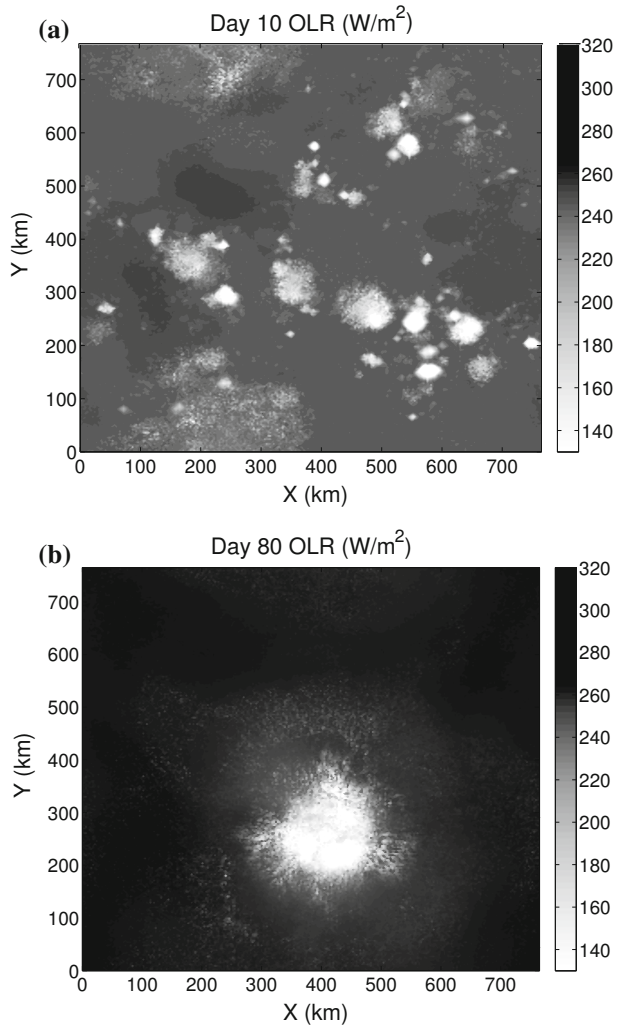
cores, different convective parameterizations (in the case of GCMs), and different radiation and microphysics schemes, the exact nature and sensitivities of aggregation vary and depend on model details. This reflects the fact that multiple processes contribute to aggregation, involving intricate interactions between clouds, moisture, radiation, and circulation. In CRMs, self-aggregation occurs more readily with large domains and coarser resolution (Muller and Held 2012), although the domain-size dependence disappears if the evaporation of rain at low levels (which causes downdrafts and cold pools) is eliminated (Jeevanjee and Romps 2013; Muller and Bony 2015). Interactive radiation and surface fluxes are generally necessary for self-aggregation to occur, a result supported by mechanism denial experiments carried out by many different authors; the details are described in Sect. 3. Some studies find that self-aggregation is favored by high sea surface temperatures (SSTs) (Held et al. 1993; Khairoutdinov and Emanuel 2010; Wing and Emanuel 2014; Emanuel et al. 2014; Coppin and Bony 2015), although it can also occur at temperatures far below current tropical SSTs (Abbot 2014; Coppin and Bony 2015; Wing and Cronin
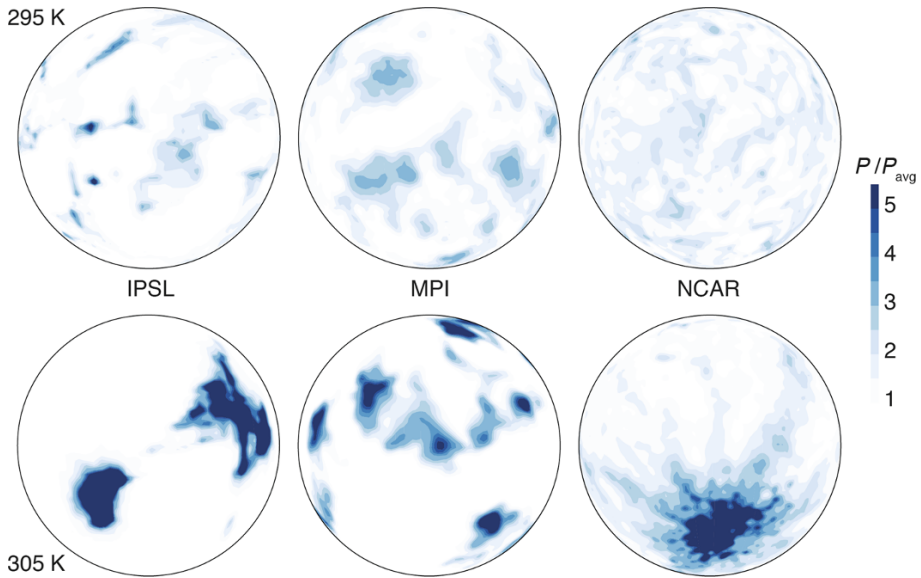
Fig. 2 Hemispheric view of monthly precipitation (normalized by its global mean value) predicted by the IPSL, MPI, and NCAR GCMs in RCE simulations forced by an SST of (*top*) 295 K and (*bottom*) 305 K. Reprinted from Bony et al. (2016)

2016; Holloway and Woolnough 2016). Vertical wind shear and/or strong mean winds make self-aggregation less likely (Held et al. 1993; Bretherton et al. 2005; Wing 2014; Abbot 2014), although much remains to be understood about the sensitivity of aggregation to vertical wind shear. Self-aggregation also exhibits hysteresis; that is, once convection occurs, it is difficult for it to disaggregate, and the aggregated state can be maintained without the feedbacks that are necessary to trigger it from homogeneous conditions (Khairoutdinov and Emanuel 2010; Muller and Held 2012; Muller and Bony 2015; Holloway and Woolnough 2016).

The structure of the remainder of this paper is as follows: First, we review the characteristics of self-aggregation, including its time and length scales and impacts on modeled climate. Next, we review the physical mechanisms leading to self-aggregation, including longwave radiation, shortwave radiation, surface fluxes, moisture feedbacks, and advective processes. We then discuss the importance of self-aggregation: Why studies of aggregation in idealized simulations might be important for understanding the tropical atmosphere and climate. Finally, we conclude with a synthesis of self-aggregation in idealized numerical models and its applicability to the real world, and what needs to be explored further. We will focus our discussion on non-rotating self-aggregation, but will note several areas in which the rotating case behaves differently.

2 Characteristics of Self-Aggregation

2.1 General Evolution of Aggregation

Simulations of convection in non-rotating RCE using three-dimensional cloud-resolving models initially produce distributions of convection that are nearly random in space and

time. Aggregation generally begins with the emergence of one or several dry regions in which convection is suppressed. The dry regions have strong radiative cooling, weaker surface enthalpy fluxes, and subsidence, which yields further drying and suppression of convection. Over time, these persistent dry regions amplify and expand such that convection, clouds, and precipitation are increasingly confined to one (or several, depending on the domain setup) moist area. As in observations (e.g., Bretherton et al. 2004), there is a strong correlation between moister columns and more active deep convection. Throughout the evolution of aggregation, the dry regions get drier and the moist regions get moister, such that the distribution of precipitable water widens considerably. Usually the dry regions amplify first, but there are at least a few simulations in which several dry and moist regions amplify roughly at the same time during the early stages of aggregation (Holloway and Woolnough 2016). This general picture of the evolution of aggregation is notably different in simulations of rotating RCE, in which a large moist region is formed which then spins up into a tropical cyclone (Wing et al. 2016) or multiple moist vortices merge (Davis 2015), rather than dry regions amplifying and expanding (the online supplemental videos[1] from Wing et al. (2016) show this distinction nicely). Non-rotating RCE simulations in which rain evaporation is suppressed also are characterized by growth and merger of moist regions (e.g., Wing 2014).

2.2 Identifying Metrics

To study self-aggregation, we need to identify it and quantify its strength. As suggested by the large number of changes that occur during the evolution of self-aggregation in idealized simulations, there are a number of different metrics that have been used to characterize the aggregation state.

One category of metrics includes measurements of the drying of the non-convective environment and the associated increase in the variance of precipitable water (PW) and in column-integrated moist static energy (MSE). The non-convective environment grows and dries, while the convective region stays moist and often becomes moister, but the domain-mean PW decreases with aggregation. Many studies use the decrease in domain-mean PW and, in particular, the increase in the spread in the PW distribution as measured by, for instance, the interquartile range (IQR) of PW to quantify self-aggregation (e.g., Bretherton et al. 2005; Muller and Held 2012; Arnold and Randall 2015; Holloway and Woolnough 2016). Since self-aggregation in non-rotating RCE should have very weak horizontal temperature gradients, the horizontal variability in PW is also the dominant source of horizontal variability in column-integrated MSE, $\hat{h}$. Wing and Emanuel (2014) use the horizontal variance of $\hat{h}$ as a measure of aggregation state (Fig. 3a), and they further derive an equation for the budget of this quantity that allows for the quantification of the contribution of different processes to the growth of the total variance. Craig and Mack (2013) use the distribution of free tropospheric PW instead of total PW, and they are also one of several studies which normalize PW distributions (or measures of spread such as IQR) by saturation PW to allow for fairer comparisons between simulations with different SSTs. An example of one of these normalized metrics, the spatial variance of column relative humidity, is shown in Fig. 3b. Note that while the horizontal variance of $\hat{h}$ increases with SST because of the Clausius–Clapeyron exponential dependence of water vapor on

---

[1] Video S1 and Video S2 from Wing et al. (2016) can be found at http://dx.doi.org/10.1175/JAS-D-15-0380.1.

temperature (Fig. 3a), the horizontal variance of column relative humidity is roughly the same across simulations with different SSTs (Fig. 3b).

The increase in domain-mean outgoing longwave radiation has also been used as a measure of aggregation (e.g., Khairoutdinov and Emanuel 2010; Wing and Emanuel 2014). Global RCE studies have also used the increase in "subsiding fraction," the fractional area of the domain covered by large-scale mid-tropospheric subsidence, as a measure of aggregation (Coppin and Bony 2015). A limitation of the metrics mentioned above is that they do not quantify the horizontal scale of the aggregated convective (or non-convective) regions. Methods of quantifying the horizontal length scales for convective aggregation and what may define them are discussed more in Sect. 2.4 below.

For most smaller square domains, aggregation in non-rotating RCE appears to be binary, either on or off (e.g., Muller and Held 2012; Wing and Emanuel 2014) although this is not true for at least one study with smaller square domains (Holloway and Wool-nough 2016). For long-channel experiments or larger domains, aggregation appears to exist more on a gradual continuum as boundary conditions are varied or mechanisms are suppressed (e.g., Wing and Cronin 2016; Coppin and Bony 2015). This suggests that small domain size, at least for some models, may prevent key aggregation feedbacks from occurring at all for certain setups (or enhance the influence of negative feedbacks opposing aggregation (Jeevanjee and Romps 2013)). Even for simulations that do show strong aggregation, the time scale for aggregation to fully develop can vary greatly for different simulations, as discussed in the next section.

2.3 Time Scale

The time scale for idealized self-aggregation from homogeneous initial conditions to reach a relatively stable state can vary from 15 to 100 days or more. For instance, Bretherton et al. (2005) found this time scale to be approximately 40 days for a $(576 \text{ km})^2$ domain, 301 K SST, and a 3-km grid. Holloway and Woolnough (2016) had 16 days for a fairly similar setup but a different model. Muller and Held (2012) found it took 20–25 days or more, with less time for coarser grid spacing. Wing and Emanuel (2014) found a time scale of 60 days for a $(768 \text{ km})^2$ domain, 305 K SST, and 3-km grid. Holloway and Woolnough (2016) found a time scale of only 8 days when rain evaporation and hence downdrafts and cold pools were suppressed, suggesting that downdrafts and/or cold pools slow or suppress aggregation as proposed by Jeevanjee and Romps (2013) and Muller and Bony (2015). Muller and Held (2012) and Holloway and Woolnough (2016) both found a disaggregation time scale (which is the time needed to return to a less aggregated equilibrium) as small as 10 days when simulations were initialized with an aggregated state and then interactive radiation was suppressed.

Over a range of SSTs in the square-domain setup of Wing and Emanuel (2014), the e-folding time for growth of spatial variance of column-integrated MSE was ∼11–13 days (Wing 2014). Wing and Cronin (2016) found the times to reach a stable aggregated state ranged from 15 days for 280 K to 50 days for 310 K, though rapid growth and most of the organization occurred in the first 10–20 days for all temperatures and the initial e-folding growth time scale from logistic fits was 2–6 days (Fig. 3); they used long-channel domains of 12,228 km × 192 km with a 3-km grid. This kind of exponential growth will lead to much larger horizontal scales in a given amount of time when starting from larger initial clustering, as is typically found in nature. The horizontal scale of aggregation is addressed in the next section.

(a)

(b)
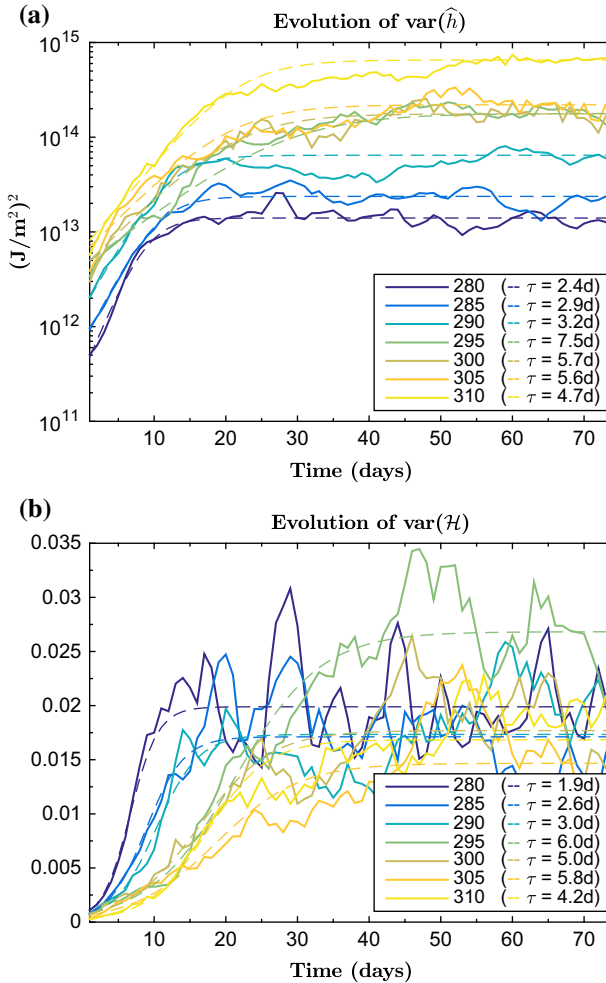
Fig. 3 Evolution of **a** spatial variance of column frozen moist static energy, var($\widehat{h}$) (J^2 m^{-4}), and **b** spatial variance of column relative humidity, var($\mathcal{H}$), in *solid lines*, and logistic fits to each, in *thin dashed lines*, for the channel RCE simulations at different values of SST (K). The vertical scale is **a** logarithmic and **b** linear, and the *legend* indicates the initial e-folding growth time-scale from each logistic fit. Reprinted from Wing and Cronin ([2016](#)). ©2015 Royal Meteorological Society

The time scale for self-aggregation results from the growth rates associated with different feedbacks that favor or oppose aggregation. Bretherton et al. ([2005](#)) developed a simple semi-empirical model to predict the initial e-folding rate of self-aggregation. They used physically motivated curve fits of the advective, surface flux, and radiative forcing from the initial stages of aggregation in their CRM simulations as parameters in an ordinary differential equation for column relative humidity. Using this semi-empirical model, Bretherton et al. ([2005](#)) found an e-folding time of the instability of 9 days. The contribution of different processes to the growth rate of column moist static energy variance is also quantified by the budget introduced by Wing and Emanuel ([2014](#)).

2.4 Length Scale

One of the limitations of square-domain simulations of non-rotating aggregation conducted thus far is that, when such simulations reach a fully aggregated state, they only contain one moist, precipitating cluster. This suggests that the size of the domain constrains the size of the cluster, and so it is difficult to define the length scale of aggregation or develop a theory for it. In square-domain simulations, the absolute size of the aggregated area increases with domain size, such that the aggregated area is 20–25% of the total domain area (Muller and Held 2012). There is no correlation between the area of the aggregated region and resolution (Muller and Held 2012) or SST (Wing 2014). The precise size of the cluster is somewhat sensitive to the metric used to define it; for example, Muller and Held (2012) used various threshold values for precipitable water. Defining the cluster size as the area where the precipitable water is greater than one standard deviation above the domain-mean value, Wing (2014) found that, across a range of SSTs at a given domain size, the cluster covered 15–17% of the horizontal area of the domain. In rotating RCE, where the nonzero Coriolis parameter introduces the Rossby radius of deformation and the ratio of tropical cyclone potential intensity over the Coriolis parameter as important horizontal scales, large enough domains allow for multiple cyclonic features with measurable average size and separation (e.g., Held and Zhao 2008; Khairoutdinov and Emanuel 2013).

Simulations using an elongated channel domain geometry in non-rotating RCE (e.g., Posselt et al. 2012; Wing and Cronin 2016) have the advantage of containing multiple aggregated areas, allowing a spatial scale to be more easily defined. In such simulations, the average spacing between convective regions in the longer horizontal dimension is a simple measure of scale (e.g., Stephens et al. 2008). The autocorrelation length scale of PW, which is the largest horizontal scale at which the average horizontal autocorrelation coefficient is $\geq e^{-1}$, is a more objective measure of this scale and has been shown to grow with self-aggregation in idealized models (Craig and Mack 2013; Wing and Cronin 2016; Holloway and Woolnough 2016). The power spectrum can also be calculated to characterize dominant scales of horizontal variability of PW or $\widehat{h}$ (Bretherton and Khairoutdinov 2015; Wing and Cronin 2016).

Over a 30 K range of SSTs, Wing and Cronin (2016) found that, using average wavenumber and correlation length metrics, the spatial scale of the aggregation varied from ~ 1000 to ~ 4000 km, with simulations at higher SSTs having smaller spatial scales. They presented a theory for the separation distance between convectively active regions based on boundary layer remoistening. A length scale resulting from this theory, proportional to the boundary layer height divided by the surface enthalpy exchange coefficient, was highly correlated with the spatial scale of aggregation across the main set of Wing and Cronin (2016)'s simulations; however, attempts to confirm the scaling were inconclusive. Further, this theory related to the maximum size of a dry region, and it is unknown whether the size of a moist region scales with that of a dry region or is controlled by other mechanisms. In addition, there could be a (perhaps temperature-dependent) minimum length scale of aggregation, below which the instability does not emerge. Bretherton and Khairoutdinov (2015) examined the scale dependence of self-aggregation feedbacks in near-global aquaplanet simulations of realistic tropical variability; similar analysis in idealized simulations may lead to insights on what controls the intrinsic length scale of self-aggregation. As of now, though, the question of what sets the spatial scale of self-aggregation in non-rotating RCE remains largely unsolved.

2.5 Impacts

As alluded to above, self-aggregation is not solely a spatial reorganization of the convection, but has dramatic impacts on the domain-mean climate. Figure 4 shows horizontally averaged profiles of humidity, MSE, and saturation MSE averaged over day 1 (disorganized convection) and day 50 (aggregated convection) from a non-rotating RCE simulation from Bretherton et al. (2005). The entire depth of the free troposphere is substantially drier, with relative humidities near 20%, when convection is aggregated. This is because the domain mean is dominated by the dry, non-convecting areas. Mean drying is found in all numerical simulations of self-aggregation, to our knowledge, although some simulations exhibit a less extreme form of aggregation, and consequently, a more muted (though still substantial) amount of drying (e.g., Wing and Cronin (2016), their Figure 2). Associated with this decrease in humidity, the domain-mean outgoing longwave radiation is ~ 10–30 Wm^{-2} larger when convection is aggregated (Wing and Cronin 2016), which is comparable to the increase of ~ 20–30 Wm^{-2} found in observations of aggregated regimes (Tobin et al. 2013).

Aggregation is also associated with domain-mean warming in the free troposphere, as indicated by the increase in saturation MSE in Fig. 4, which corresponds to several degrees of temperature increase (also see Wing and Cronin (2016), their Figure 1). The temperature increase is consistent with the fact that, when aggregated, the convecting environment is moister, which reduces the influence of entrainment and drives the troposphere closer to a
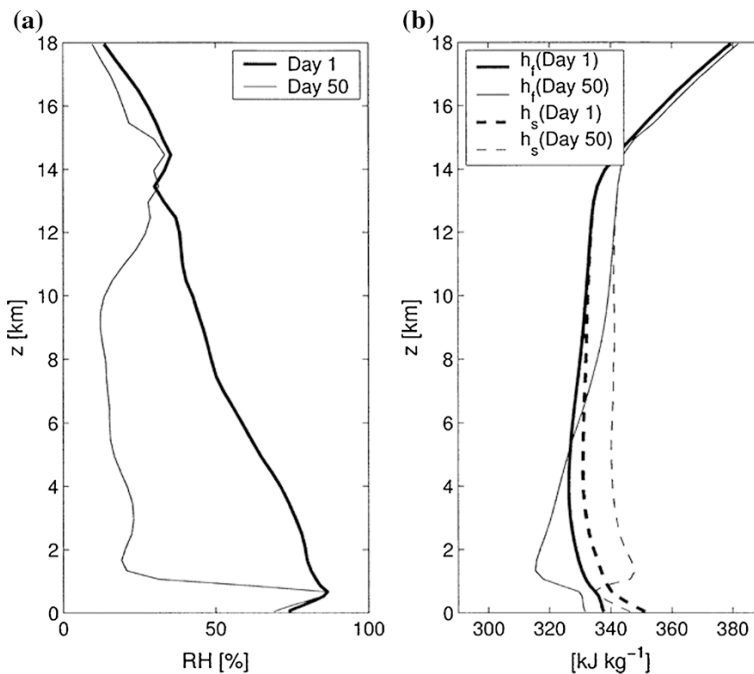

Fig. 4 Horizontally averaged profiles of **a** relative humidity and **b** moist static energy (h_f) and saturation moist static energy (h_s) averaged over days 1 and 50 from an RCE simulation. Reprinted from Bretherton et al. (2005). ©2005 American Meteorological Society

moist adiabat. It is also consistent with an increase in boundary layer moisture and MSE in the convective region, shifting buoyant parcels toward a warmer moist adiabat.

Changes in radiative fluxes by aggregation are strongly influenced by changes in cloudiness. A decrease in high clouds with aggregation is found in CRM simulations (Fig. 5), GCM simulations of RCE with parameterized convection (Bony et al. 2016), and in observations (Tobin et al. 2012, 2013). Bony et al. (2016) argued that the decrease in anvil cloud fraction with aggregation was a result of an increase in upper tropospheric stability due to the increase in temperature. By mass conservation, the amount of mass divergence in the convecting regions corresponds to the maximum clear-sky radiatively driven divergence in the upper troposphere. As the upper troposphere warms with aggregation, it becomes more stable, reducing the amount of divergence necessary to balance the same clear-sky radiative cooling. This mechanism is similar to arguments based on subsidence regions which can explain the weakening of the overturning circulations in the tropics as a response to greenhouse warming (Knutson and Manabe 1995; Held and Soden 2006). The reduction in anvil cloud fraction is then linked to the reduction in convective outflow. If the frequency or degree of aggregation changes with warming, the reduction in high clouds (and the increased drying) could affect cloud feedbacks on surface warming and climate sensitivity (Mauritsen and Stevens 2015).

In CRM simulations, this decrease in high clouds is largely offset by an increase in low clouds, such that the reflected shortwave radiation changes little (Fig. 5, Wing 2014; Wing and Cronin 2016). This result is uncertain, however, because the horizontal resolution of 3 km used in those studies is too coarse to model low clouds accurately. The response of the top-of-atmosphere net radiation budget in CRM simulations differs from observations of aggregated convection, which find that the reflected shortwave radiation is reduced due to a reduced total cloud fraction, which largely cancels the increase in outgoing longwave radiation (Tobin et al. 2012, 2013). Both numerical simulations and observations agree that the domain-mean tropospheric radiative cooling increases with aggregation, due to the drier troposphere.

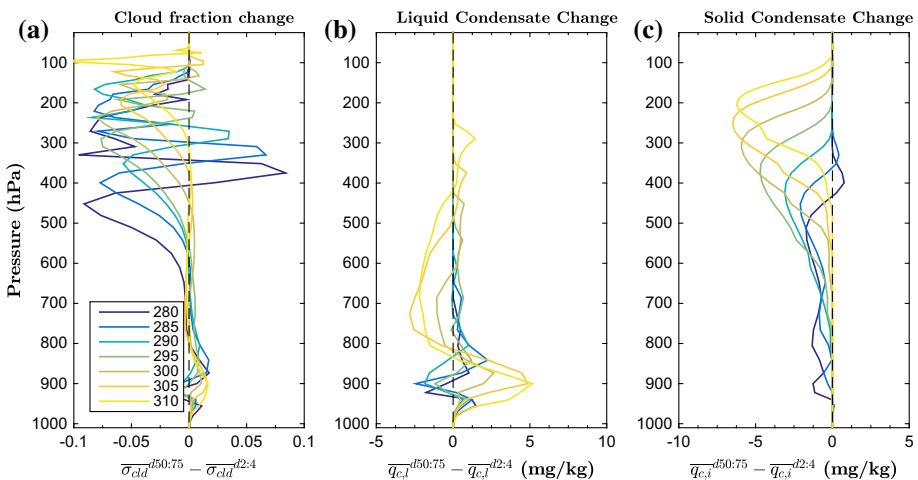**Fig. 5** Domain-average profiles of change in **a** cloud fraction, **b** liquid condensate $q_{c,l}$ and **c** solid condensate $q_{c,i}$ between days 2 and 4 and 50 and 75 of RCE channel simulations. *Colors* indicate the sea-surface temperature (K) of the channel simulation, and an *overbar* indicates a mean over the time range indicated. Reprinted from Wing and Cronin (2016). ©2015 Royal Meteorological Society

3 Mechanisms of Self-Aggregation

Numerical studies of self-aggregation have identified multiple processes involving convection-moisture-radiation feedbacks that are capable of creating an area around clouds that is more favorable for future convection than areas further away. These processes have been identified using both mechanism denial experiments and diagnostic frameworks. One diagnostic framework that has been employed is a budget for the spatial variance of column-integrated moist static energy (Wing and Emanuel 2014), which is given by:

$$\frac{1}{2}\frac{\mathrm{d}\widehat{h}'^2}{\mathrm{d}t} = \widehat{h}' F_K' + \widehat{h}' N_S' + \widehat{h}' N_L' - \widehat{h}' \nabla_h \cdot \widehat{\mathbf{u}h}. \tag{1}$$

where h is the moist static energy (or frozen moist static energy), F_K is the surface enthalpy flux, N_S is the column shortwave flux convergence, N_L is the column longwave flux convergence, and $-\nabla_h \cdot \widehat{\mathbf{u}h}$ is the "advective term," the horizontal convergence of the density-weighted vertical integral of the flux of frozen moist static energy. A primed quantity, $(\cdot)'$, denotes the spatial anomaly from the horizontal mean, $\langle(\cdot)\rangle$, and $\widehat{\cdot}$ indicates a density-weighted vertical integral.

The advantage of this (or similar) frameworks is that, since self-aggregation is associated with an increase in the spatial variance of MSE, the budget enables the quantification of the each feedback associated with a process that is a source or sink of MSE, hence contributing or opposing self-aggregation. Each feedback can be quantified across the entire evolution of the simulation. The magnitude of these feedbacks can be compared to each other within a given simulation and across simulations using different boundary conditions, parameters, and models. However, while this budget diagnoses the direct effect of radiative processes [the second and third terms on the right-hand side of Eq. (1)], it does not explicitly diagnose the indirect effect of a radiatively driven circulation [this, among other dynamical contributions, is a part of the fourth term on the right-hand side of Eq. (1)]. Since it is a vertically integrated budget, the impact of shallow circulations or the sensitivity to the profile of radiative heating anomalies is also not explicitly included (they are indirectly included, insofar as they determine the circulation response which is a part of the advective term). On the other hand, mechanism denial experiments, in which a feedback process is disallowed, test the sensitivity of self-aggregation to all aspects of that feedback. For example, horizontally homogenizing the radiative heating rates removes both the direct and indirect feedbacks associated with radiation. If self-aggregation still occurs without a particular feedback enabled, this indicates that that feedback is not necessary for aggregation. However, since many different processes can contribute to aggregation, caution must be taken to interpret the results of such sensitivity tests, as the importance of a particular feedback could vary depending on what parameters are used and what other feedback processes are active. In addition, if there is a critical SST for aggregation to occur, aggregation will be sensitive to virtually everything when one is near the critical point.

One fundamental aspect of moist convection in the tropics that underpins the mechanisms of aggregation discussed here is that, in a weak temperature gradient environment, deep convection is more active in moister tropospheric columns, as shown by precipitation observations, for example (Bretherton et al. 2004). This should be kept in mind when interpreting the feedbacks on aggregation identified in this section.

In this section, we review the various processes leading to the self-aggregation of convection in RCE simulations. This includes longwave radiation (Bretherton et al. 2005;

Muller and Held 2012; Posselt et al. 2012; Wing and Emanuel 2014; Emanuel et al. 2014; Muller and Bony 2015; Coppin and Bony 2015; Arnold and Randall 2015; Wing and Cronin 2016; Holloway and Woolnough 2016), shortwave radiation (Wing and Emanuel 2014; Wing and Cronin 2016), surface fluxes (Bretherton et al. 2005; Wing and Emanuel 2014; Coppin and Bony 2015; Wing and Cronin 2016), moisture feedbacks (Tompkins 2001; Craig and Mack 2013; Muller and Bony 2015; Holloway and Woolnough 2016), and advective processes (Bretherton et al. 2005; Muller and Held 2012; Muller and Bony 2015). At the end of this section, we address the relative importance of some of these processes for the maintenance of an aggregated state as opposed to the initial stages of self-aggregation. We focus on non-rotating simulations, although we note a few instances where the mechanisms differ if instead rotating RCE is simulated.

3.1 Surface Fluxes

Surface flux feedbacks favor self-aggregation, though they are not always necessary for aggregation to occur. Tompkins and Craig (1998) and Bretherton et al. (2005) both found that sensitivity runs without radiative feedbacks or without surface flux feedbacks did not aggregate. Muller and Held (2012) extended these sensitivity runs to cover a large range of domains and resolutions. They also found that surface flux feedbacks favor self-aggregation, but convection could still aggregate without them as long as radiative feedbacks are active and the domain is large enough. The converse is not true. Holloway and Woolnough (2016) confirmed that sensitivity runs with homogeneous surface fluxes can self-aggregate or not depending on the strength of the surface fluxes imposed. Non-rotating RCE simulations without radiative feedbacks do not aggregate (unless rain evaporation is artificially removed; see below section on moisture feedbacks). Therefore, at current temperatures, it seems that surface flux feedbacks are not sufficient on their own, without longwave radiative feedbacks, for non-rotating aggregation to occur. Rotating RCE simulations, on the other hand, have stronger surface flux feedbacks and can aggregate without radiative feedbacks (Wing et al. 2016).

Physically, there are two opposing contributions to the surface flux feedback (Wing and Emanuel 2014). The air-sea enthalpy disequilibrium is smaller in the moist region than in the dry region, which would tend to suppress surface fluxes in the moist region, a negative feedback on aggregation. On the other hand, the surface winds are stronger in the moist, convecting region, which would tend to enhance surface fluxes there, a positive feedback on aggregation. The latter dominates in initial triggering of aggregation, yielding an overall positive surface flux feedback.

3.1.1 Sensitivity to SST

In the GCM simulations of Coppin and Bony (2015), the surface flux-wind feedback was the leading mechanism of aggregation at high temperatures. In the high-temperature regime, strong surface winds in the convective region yield strong surface fluxes, moistening the high-MSE convective region, thereby enhancing the MSE gradient and favoring self-aggregation. In contrast, Wing and Cronin (2016) found that the total surface flux feedback, as diagnosed from the MSE variance budget, was approximately constant in magnitude across a wide range of temperatures (280–310 K). Both of these studies used fixed SSTs and thus do not have surface energy balance. Therefore, caution must be taken in interpreting these results; with fixed SST, there is no guarantee that the relationship

between surface evaporation and SST is correct and so the behavior of the surface flux feedback should not be taken as general.

3.2 Longwave Radiation

Mechanism denial experiments have shown that the longwave radiative feedback is essential for non-rotating aggregation to occur (Bretherton et al. 2005; Muller and Held 2012; Wing 2014; Holloway and Woolnough 2016). Consistent with this, interactive radiation drastically increases the parameter range that supports multiple equilibria (the analog to self-aggregation) in CRM simulations employing the weak temperature gradient approximation, although there are some limited conditions under which multiple equilibria can exist with fixed radiation (Sessions et al. 2016). There is also broad agreement that the formation of one or several dry areas that are driven by enhanced longwave radiative cooling, termed "radiatively driven cold pool" by Coppin and Bony (2015), is important in driving non-rotating aggregation.

Longwave radiative feedbacks also contribute substantially to spontaneous tropical cyclone genesis in simulations of rotating RCE, but are not strictly necessary for it to occur (Wing et al. 2016).

Enhanced longwave radiative cooling in the dry regions triggers aggregation in two ways: the direct diabatic effect, where the enhanced cooling relative to the moist regions decreases the MSE in the dry regions; and the indirect effect mediated by a circulation, where the enhanced longwave cooling in the dry regions drives a shallow circulation between the dry and moist regions and this circulation transports MSE upgradient. Both effects act to suppress convection in the dry regions and enhance convection in the moist regions. Note that there is no guarantee that there will be enhanced longwave cooling in the drier regions; this depends on temperature (Emanuel et al. 2014), clouds, and the vertical structure of the moisture perturbation (Beucler and Cronin 2016). In some circumstances, the opposite may occur (perhaps at cold temperatures), which would yield a negative feedback on aggregation (Emanuel et al. 2014; Wing and Cronin 2016). There is also no guarantee that the circulation driven by radiative heating anomalies transports moist static energy upgradient (this depends on the vertical structure of the heating anomalies (Muller and Bony 2015)).

In simulations that aggregate, the direct diabatic effect of the longwave feedback in the dry regions at the beginning of the aggregation process, as diagnosed with the MSE variance budget, is large and positive and results from both clear sky and cloud effects (Wing and Emanuel 2014; Holloway and Woolnough 2016; Wing and Cronin 2016). However, while the direct longwave feedback is important at amplifying the MSE anomalies in the early stages of aggregation, as aggregation proceeds, it switches to become a negative feedback in the dry regions, at least at temperatures near current tropical SSTs (Wing and Emanuel 2014). The partitioning of the enhanced longwave cooling in the dry regions between clear sky and cloud effects is sensitive to the choice of radiation scheme (Wing and Cronin 2016), and it would not be surprising if this was also sensitive to the cloud microphysics or, in the case of GCMs, the cloud parameterization. Cloud amounts and hence cloud radiative effects can also be sensitive to resolution and domain size (Muller and Held 2012; Muller and Bony 2015).

The shallow radiatively driven circulation is largely induced by strong longwave cooling from low-level clouds in the dry region (Muller and Held 2012; Muller and Bony 2015; Coppin and Bony 2015; Holloway and Woolnough 2016). This low-level cooling in the dry region yields low-level subsidence and outflow from dry to moist regions near the

surface where the MSE is large. This results in a divergence of moist static energy out of the dry region, causing upgradient MSE transport that reinforces MSE gradients and further drives aggregation. The role of advective processes in aggregation is discussed further below.

3.2.1 Sensitivity to SST

In the GCM simulations of Coppin and Bony (2015), the enhanced low-level cooling in the dry region and its associated circulation were found to be most efficient in driving aggregation at temperatures near current tropical SSTs and colder. On the other hand, instability driven by the direct clear-sky longwave feedback is favored by higher temperatures, because the rapid increase in water vapor concentration with temperature causes the lower troposphere to be very opaque in the longwave at high temperature (Emanuel et al. 2014). This dependence was interpreted as the cause of the temperature dependence of non-rotating aggregation in the square simulations of Wing and Emanuel (2014), but subsequent results have cast doubt on that conclusion. Although the direct longwave feedback is initially large and negative in the cold (SST < 295 K) simulations of Wing and Cronin (2016), this is not sufficient to prevent aggregation and, moreover, the negative longwave feedback is nearly entirely a result of clouds, not clear-sky processes (the clear-sky longwave feedback is near zero). Wing and Cronin (2016) hypothesized that this is due to the fact that a low-temperature atmosphere is so optically thin that the presence of clouds (in the moist regions) would actually increase the longwave cooling of the atmosphere by increasing the number of longwave emitters. We note, though, that the initial negative longwave cloud feedback in cold simulations does not persist; after a few days, the longwave cloud feedback is positive (Wing and Cronin 2016; Holloway and Woolnough 2016).

Nearly all simulations of self-aggregation have used fixed sea-surface temperature, but a few studies that have employed a slab ocean have found that coupling between the SST and the net surface energy may disrupt self-aggregation or delay its onset, or even prevent it if the slab is thin enough (Bretherton et al. 2005; Reed et al. 2015; Hohenegger and Stevens 2016). Hohenegger and Stevens (2016) found that, in a coupled simulation, SST gradients develop which tend to oppose the development of the radiatively driven low-level circulation, therefore delaying self-aggregation. However, air-sea coupling could also allow other instabilities to be realized (Beucler and Emanuel 2016); more work is needed to fully understand the behavior of self-aggregation with an interactive surface.

3.3 Shortwave Radiation

Shortwave feedbacks can contribute to self-aggregation, but do not appear to be essential for it to occur. The direct, diabatic effect is positive as measured by its contribution to the MSE variance budget (only accounting for diabatic warming/cooling in moist/dry regions), although its magnitude is smaller than the longwave and surface flux terms at current temperatures (Wing and Emanuel 2014; Holloway and Woolnough 2016; Wing and Cronin 2016). This direct positive shortwave feedback is due to a reduction in shortwave heating from clear-sky atmospheric absorption in dry regions. The shortwave cloud feedback can be either positive or negative; it is generally negative in the moist regions after convection has aggregated, where deep clouds reflect shortwave radiation before it can penetrate into the column and be absorbed.

In sensitivity experiments including both direct (diabatic cooling/heating) and indirect (dynamic response to this diabatic forcing) effects, shortwave radiative feedbacks slightly oppose aggregation (Muller and Held 2012). This was clarified in Holloway and Woolnough (2016) as resulting from the positive anomalous heating at high levels by high clouds in the moist regions, favoring upward motion and yielding higher MSE divergence at high levels from the moist region. This transports MSE down-gradient and damps the convective aggregation. Overall, the impact of shortwave radiation is controlled by this MSE transport, at least in the simulations of Muller and Held (2012) and Holloway and Woolnough (2016).

3.3.1 Sensitivity to SST

The direct shortwave feedback was found to be much stronger at low (SST < 295 K) temperatures in the channel CRM simulations of Wing and Cronin (2016), and, along with the surface flux feedback, is the dominant initial driver of aggregation in those low temperature simulations. They found that the positive shortwave feedback at low temperatures was a result of clouds. They proposed several hypotheses for a positive shortwave cloud feedback, including direct shortwave absorption by cloud water and ice, reflected shortwave by low clouds in the moist regions (back to the atmosphere, where it could be absorbed by water vapor), and increased atmospheric absorption because of a higher fraction of diffuse radiation in cloud regions. It is not known which, if any, of these effects dominates, or why they should be stronger at lower temperature. Holloway and Woolnough (2016) found that the clear-sky shortwave feedback was smaller than but of comparable magnitude to the total shortwave feedback in a simulation at 290 K. This further indicates that the relative importance of clouds and clear-sky process to radiative feedbacks depends on the model and radiation package used.

3.4 Advective Processes

As alluded to in Sect. 3.2, advective processes may also contribute to self-aggregation. Bretherton et al. (2005) first showed that upgradient transport of MSE by the circulation consistent with a negative gross moist stability occurred during self-aggregation, as diagnosed from the MSE budget and visualized with a moisture-sorted streamfunction. Muller and Held (2012) and Muller and Bony (2015) further emphasized the importance of upgradient advection and specified that strong radiative cooling at the top of low clouds in the dry region was responsible for driving a shallow circulation that transported MSE upgradient (an indirect effect of radiation on aggregation). This is shown in Fig. 6, which displays the moisture-sorted stream function introduced by Bretherton et al. (2005), radiative cooling rates, MSE, and clouds. The bottom two panels show a simulation without low clouds and the resulting difference in the circulation. The total contribution of all advection processes, integrated over the entire column, as quantified with the MSE variance budget is a positive feedback during the intermediate stages of aggregation, but a negative feedback during other times (Wing and Emanuel 2014; Holloway and Woolnough 2016). Using the same metric, Wing and Cronin (2016) found that, in an elongated channel geometry, advective processes always damped the MSE variance tendency, exporting MSE from the moist regions.

However, even when the total column-integrated advective feedback across the domain, as expressed by the MSE variance budget, is negative, there could still be local upgradient transport (Coppin and Bony 2015), and the shallow component of the circulation could still

be acting to further aggregation. In addition, this metric includes all advective processes, so it is possible that the indirect effect of radiative anomalies on the circulation could have a positive influence on aggregation, but is counteracted by other processes. Indeed, Holloway and Woolnough (2016) found that a low-level circulation did appear to transport MSE from drier to moister regions, but that this circulation was mostly balanced by other advective effects of the opposite sign and was forced primarily by horizontal anomalies of convective heating (leading to low-level upward motion in the moist region), rather than radiation. Note that Holloway and Woolnough (2016) used the weak temperature gradient approximation to diagnose circulation components caused by different diabatic processes,
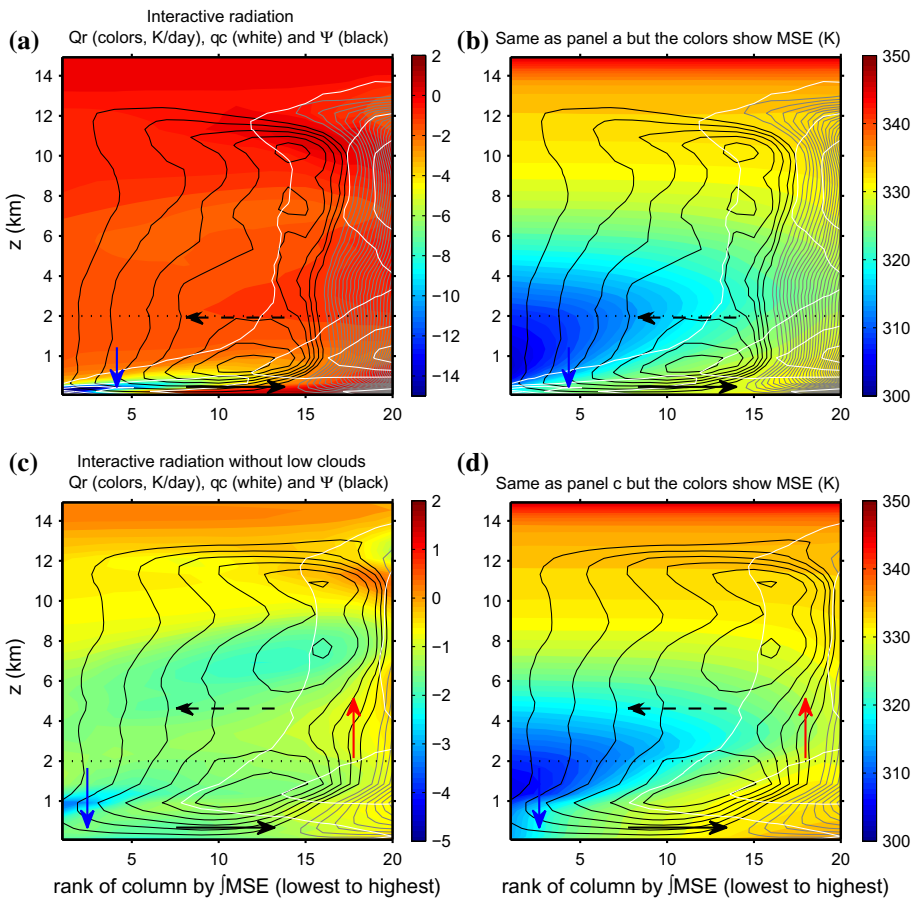
Fig. 6 Radiative cooling rates (*colors* in **a** and **c**), moist static energy (*colors* in **b** and **d**), cloud water content (liquid + ice, white contours every 5×10^{-2} g/kg, starting at 5×10^{-3} g/kg), and stream function (*black contours* for counterclockwise, *gray contours* otherwise, every 8×10^{-3} kg m^{-2} s^{-1}), averaged over the last 20 days of aggregated RCE simulations, plotted as a function of height z and vertically integrated MSE. Note the stretched vertical coordinate z below 2 km. **a**, **b** Simulations with fully interactive radiation; **c**, **d** similar simulation but without the low-cloud radiative effects. The *arrows* schematically represent the subsidence generated by the radiative cooling (*blue*) and rising motion by the warming (*red*), as well as the low-level (*solid black*) and midlevel (*dashed black*) flows induced. Reprinted from Muller and Bony (2015). ©2015. American Geophysical Union. All Rights Reserved

and so they neglected the potential effects of radiative cooling within the boundary layer on the circulation.

Overall, it is clear that advective processes contribute to non-rotating self-aggregation, but there remains some disagreement in the literature as to whether they trigger aggregation on their own or amplify it once direct diabatic feedbacks have started the process.

3.5 Moisture Feedbacks

Moisture feedbacks, which result from the interaction between convection and humidity, are known to organize convection (Tompkins 2001; Grabowski and Moncrieff 2004; Mapes and Neale 2011). Perhaps surprisingly, it was recently found that those feedbacks could lead to the full convective aggregated state, even in the absence of radiative feedbacks (Muller and Bony 2015; Holloway and Woolnough 2016). In that case, the aggregation process is different from the radiatively driven dry cold pool expansion discussed earlier. Instead, it develops similarly to the coarsening process described in the theoretical model of Craig and Mack (2013), with moist areas growing and merging or dying out with little horizontal drift of moist and dry regions. This occurs when the evaporation of rain is artificially suppressed, hence when evaporation-driven downdrafts and cold pools below clouds are weak. This implies that the rain falls without evaporation, which is not realistic in standard conditions, but may occur when the boundary layer is nearly saturated and the precipitation efficiency approaches 100%.

The positive moisture feedback is one in which more moisture favors convection, which in turn yields more moisture. The details of the physical process are still unclear, though several processes have been proposed. First, convection is favored where the boundary layer is anomalously moist (hence boundary layer parcels are more buoyant). Without downdrafts advecting dry air into the boundary layer, it remains moist and the upward motion remains above the boundary layer moisture anomaly. Second, the absence of cold pools in this case may also be important (both downdrafts and cold pools are absent when the evaporation of rain is suppressed), consistent 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. Third, the moisture feedback could be due to entrainment, since a parcel ascending in a moister environment will be less cooled by entrainment, leading to a larger parcel buoyancy and stronger convection (Tompkins 2001; Holloway and Neelin 2009; Mapes and Neale 2011). In a version of the simple model of Emanuel et al. (2014), the sensitivity of convection to free tropospheric water vapor can boost the otherwise radiatively driven instability. Emanuel et al. (2014) also found that aggregation is favored by increased precipitation efficiency; in fact, when the lower troposphere is opaque in the infrared (i.e., at high temperature) and the precipitation efficiency is unity, their two-layer model is always unstable.

3.6 Triggering Versus Maintenance

Several studies have now confirmed that some feedbacks which are not sufficient to trigger self-aggregation from homogeneous conditions may still be able to maintain aggregation once it is established (Khairoutdinov and Emanuel 2010; Muller and Held 2012; Holloway and Woolnough 2016). This is consistent with the feedback analysis from the MSE variance budget, which shows a strong time evolution of the leading feedback throughout the aggregation process (Wing and Emanuel 2014; Wing and Cronin 2016). The strongest

positive feedbacks are typically found in the dry region at early times, while at later times strong positive feedbacks are found in the moist region.

Although strong longwave cooling in the dry regions, at least partially due to low clouds, was found to be crucial for the onset of aggregation, for maintenance, low-cloud longwave cooling is not necessary (Muller and Held 2012; Muller and Bony 2015). Instead, high-level clouds in the moist regions and clear-sky longwave cooling in the dry regions can maintain aggregation. The direct diabatic effect of high-level clouds is a strong longwave positive feedback in the very moistest regions where all the deep clouds are concentrated, primarily because the column longwave cooling is strongly reduced by the longwave opacity and low temperature of high clouds. This is the strongest positive feedback that maintains the high-MSE region during the mature phase of self-aggregation (Wing and Emanuel 2014).

Surface flux feedbacks are not sufficient to maintain aggregation (Holloway and Woolnough 2016), at least not at current climate temperatures. While the surface flux feedback is positive during the early stages of aggregation, later in the evolution of aggregation, as the boundary layer in the dry regions gets drier, the total surface flux feedback becomes negative (Wing and Emanuel 2014). This is not the case in simulations of rotating RCE, in which the surface flux feedback remains positive throughout and in fact dominates over the radiative feedbacks once a broad vortex has formed (Wing et al. 2016).

4 Importance of Self-Aggregation

Self-aggregation of moist convection represents an important phase transition in moist convective systems, at least those that have been modeled in cloud system permitting models and in aquaplanet GCMs. In some simulations (e.g., Wing and Emanuel 2014), the phase transition is discrete, occurring above some threshold temperature, while in others (e.g., Bony et al. 2016) it is gradual. Either way, the transition is accompanied by a substantial drying of the free troposphere (Bretherton et al. 2005), an effect which, if the surface temperature was allowed to vary, would cool the system by reducing the greenhouse effect of water vapor. Khairoutdinov and Emanuel (2010) proposed that this drying, combined with the temperature dependence of self-aggregation, could strongly regulate tropical climate. In a simple model, they proposed that this feedback would result in a self-organized critical state in which the system is attracted to the critical temperature for aggregation. The general idea that aggregation can act as a kind of thermostat was extended to the rotating case by Khairoutdinov and Emanuel (2013) and has been described as a kind of "iris" effect by Mauritsen and Stevens (2015). The temperature dependence of self-aggregation remains uncertain, however, as it can occur at SSTs far below current tropical values (e.g., Wing and Cronin 2016) and it is not obvious how or whether the degree of aggregation depends on temperature in those or other simulations.

When aggregation takes the form of tropical cyclones, an additional set of feedbacks comes into play, involving turbulent mixing of the upper ocean (e.g., Bender et al. 1993). The mixing cools the surface waters and warms deeper waters, conserving the ocean column enthalpy. But the surface cold wakes recover over a period of weeks, and this represents a net warming of the column. So, ironically, although tropical cyclones operate by extracting heat from the ocean, their net effect, after a few weeks, is to transport enthalpy from the atmosphere to the ocean. This may have effects on ocean circulation (Emanuel 2001), although the magnitude of this effect is disputed (Jansen and Ferrari

2009; Jansen et al. 2010). Mixing of nutrients and dissolved CO_2 to the surface may also affect marine biology and the carbon cycle (Lin et al. 2003).

Whether and to what degree all these feedbacks operate in the natural world remains an open question, although attempts to bridge the gap between self-aggregation in idealized simulations and real organized convection are beginning to be made. For example, Bretherton and Khairoutdinov (2015) investigated feedbacks related to self-aggregation in near-global aquaplanet cloud-resolving simulations of realistic tropical variability, finding that radiative feedbacks amplify humidity variance at all scales, consistent with idealized CRM simulations. Vertical shear of the large-scale horizontal wind is known to be destructive to the formation of tropical cyclones and, given what we know about the physics of non-rotating self-aggregation, it seems likely that shear would inhibit this as well. The modeling work described here leaves little doubt that the character and perhaps even the existence of self-aggregation depends on how clouds, radiation, convection, and the boundary layer are modeled. This casts into some doubt whether current climate models can simulate aggregation or, if they do, whether it is simulated accurately. Given that aggregation physics may be important for such phenomena as tropical cyclones (Bretherton et al. 2005; Nolan et al. 2007; Davis 2015; Wing et al. 2016) and the Madden-Julian Oscillation (Bretherton et al. 2005; Arnold and Randall 2015), the sensitivity to physics may help explain why such phenomena have been notoriously difficult to simulate with global models. If aggregation does indeed have an important negative feedback on climate change, it is not clear how well this is handled by current GCMs.

Precipitation efficiency is much higher in aggregated convection, because rain falls through humid air and loses less mass to evaporation. Evaporation tends to concentrate heavier isotopes of hydrogen and oxygen in the raindrops. Thus, we expect to find smaller ratios of the heavier to lighter isotopes in rain from aggregated convection. This is a well-known signal of tropical cyclone rainfall (Lawrence and Gedzelman 1996) and may provide a proxy for aggregation that could be used as an aggregation metric by measuring the isotopic composition of rain. Since this composition is recorded in, for example, tree rings (Miller et al. 2006) and cave deposits (Frappier et al. 2007), there is some hope that one could detect past variations in aggregation in past climates. This might help test the hypothesis that aggregation of moist convection serves as a brake on tropical climate change.

5 Conclusions

5.1 What Aspects of Self-Aggregation do Modeling Studies Agree on?

In the 20+ years since self-aggregation was first described by Held et al. (1993), a growing body of literature has investigated its characteristics, mechanisms, and impacts. In particular, a great deal of progress has been made in the last ~ 5 years, as there has been a resurgence of interest in radiative–convective equilibrium as an idealization of the tropical atmosphere which, despite its simplicity, exhibits a rich spectrum of behavior that is yet to be completely understood. Several aspects of non-rotating self-aggregation have emerged as robust across these modeling studies; these common features are noted here.

1. Moist static energy variance is dominated by the variance in humidity above the boundary layer, a consequence of the maintenance of weak temperature gradients in the tropical atmosphere.

2. Convection preferentially occurs in humid, high moist static energy regions.
3. As convection aggregates, there is an increase in humidity variance and, in most cases, not only do the dry regions become drier, but the moist regions also become moister.
4. Self-aggregation is not merely a spatial reorganization of convection, it entails large changes to the domain-mean climate. Most dramatically, there is a decrease in the domain-mean humidity.
5. Feedbacks between longwave radiation and water vapor and/or clouds are essential for triggering and maintaining aggregation.
6. Surface flux feedbacks favor the development of aggregation. In the rotating case, surface flux feedbacks dominate.
7. The amplification and expansion of dry regions, in which convection is suppressed, is important in the triggering of self-aggregation.
8. The self-aggregated state exhibits strong hysteresis.

5.2 What Remains Uncertain?

Although the fundamentals of self-aggregation have been established, there are many details that remain uncertain. In particular, there is disagreement in the literature regarding the following issues:

1. The relative importance of cloud versus clear-sky radiative processes. This is likely dependent on the treatment of radiative transfer and cloud microphysics, and the fact that shallow convection is not well represented at typical CRM resolutions.
2. The relative contributions of the direct (diabatic) and indirect (circulation mediated) effects of radiative forcing on the growth of moist static energy anomalies and evolution of self-aggregation.
3. The role of advective processes. Is advective transport of MSE by the circulation essential for triggering self-aggregation, or does it only contribute after diabatic processes have started the process?
4. The temperature dependence of self-aggregation. Some studies find it to be favored by high temperatures, while others find that it occurs across a wide range of temperatures including those much colder than current tropical SSTs.

5.3 What Could be Explored More?

In addition to reconciling the disagreements between studies noted above, there are many aspects of self-aggregation that need to be explored further to achieve a complete understanding of its physics and importance for climate. Several of them are noted here:

1. How does self-aggregation operate when subjected to mean winds and/or vertical wind shear? Does wind shear affect the initiation and maintenance of aggregation differently? Does unidirectional shear simply change the form of aggregation to be more squall line-like, or does it prevent it?
2. How are the mechanisms of self-aggregation altered when the sea-surface temperature is interactive (i.e., calculated from surface energy balance over a slab ocean) versus fixed?
3. Does self-aggregation occur over land surfaces? If it does, how are its behavior and dependencies altered?
4. What controls the spatial scale of self-aggregation?

5. How and why does the degree of aggregation depend on temperature?
6. How does self-aggregation impact climate and climate sensitivity?
7. What is the sensitivity of self-aggregation to boundary layer processes? Most CRM simulations of self-aggregation are too coarse to fully resolve boundary layer turbulence yet lack a boundary layer parameterization. Does this have a fundamental effect on the aggregation of convection?
8. What is the sensitivity of self-aggregation to the dynamical model? A model intercomparison study in which the simulation design and model configuration is controlled would enable a better understanding of the robustness of self-aggregation. "RCEMIP," a recently proposed model intercomparison of radiative–convective equilibrium involving both cloud-resolving models and GCMs with convective parameterizations, may be able to answer this question.
9. How does the self-aggregation found in idealized simulations of radiative–convective equilibrium relate to organized convection in the real world? For what observed convective phenomenon is the self-aggregation of convection in RCE the best simple starting point for understanding? Which aspects of self-aggregation are found in nature, and which are unrealistic? An overview of observational work on self-aggregation and ways forward in this area is presented in Holloway et al. (2017).

5.4 Synthesis

Self-aggregation of moist convection represents a new frontier in meteorology and climate, not simply because a new phenomenon has been added to the panoply of atmospheric processes, but because it also represents a novel intellectual endeavor, breaking the classical stove pipes of, e.g., dynamics vs. radiation physics vs. cloud microphysics. Rapid progress is being made largely by a new generation of atmospheric scientists who are well versed in traditional dynamics, convective and cloud physics, thermodynamics, and radiative transfer.

The novelty of self-aggregation is reflected by the many remaining unanswered questions about its character, causes and effects. It is clear that interactions between longwave radiation and water vapor and/or clouds are critical: non-rotating aggregation does not occur when they are omitted. Beyond this, the field is in play, with the relative roles of surface fluxes, rain evaporation, cloud versus water vapor interactions with radiation, wind shear, convective sensitivity to free atmosphere water vapor, and the effects of an interactive surface yet to be firmly characterized and understood. The sensitivity of simulated aggregation not only to model physics but to the size and shape of the numerical domain and resolution remains a source of concern about whether we have even robustly characterized and simulated the phenomenon. While aggregation has been observed in models (e.g., global models) in which moist convection is parameterized, it is not yet clear whether such models simulate aggregation with any real fidelity. The ability to simulate self-aggregation using models with parameterized convection and clouds will no doubt become an important test of the quality of such schemes.

Understanding self-aggregation may hold the key to solving a number of obstinate problems in meteorology and climate. There is, for example, growing optimism that understanding the interplay among radiation, surface fluxes, clouds, and water vapor may lead to robust accounts of the Madden Julian oscillation and tropical cyclogenesis, two long-standing problems in atmospheric science. Indeed, the difficulty of modeling these phenomena may be owing in part to the challenges of simulating them using