Radiation, Clouds, and Self-Aggregation in RCEMIP Simulations

Kieran Nicholas Pope[1], Christopher E Holloway[2], Todd Russell Jones[1], and Thorwald Hendrik Matthias Stein[1]

[1]University of Reading
[2]University of Reading, UK

November 22, 2022

Abstract

The responses of tropical anvil cloud and low-level cloud to a warming climate are among the largest sources of uncertainty in our estimates of climate sensitivity. However, most research on cloud feedbacks relies on either global climate models with parameterized convection, which do not explicitly represent small-scale convective processes, or small-domain models, which cannot directly simulate large-scale circulations. We investigate how self-aggregation, the spontaneous clumping of convection in idealized numerical models, depends on cloud-radiative interactions with different cloud types, sea surface temperatures (SSTs), and stages of aggregation in simulations that form part of RCEMIP (the Radiative-Convective Equilibrium Model Intercomparison Project). Analysis shows that the presence of anvil cloud, which tends to enhance aggregation when collocated with anomalously moist environments, is reduced in nearly all models when SSTs are increased, leading to a corresponding reduction in the aggregating influence of cloud-longwave interactions. We also find that cloud-longwave radiation interactions are stronger in the majority of General Circulation Models (GCMs), typically resulting in faster aggregation compared to Cloud-system Resolving Models (CRMs). GCMs that have stronger cloud-longwave interactions tend to aggregate faster, whereas the influence of circulations is the main factor affecting the aggregation rate in CRMs.

Radiation, Clouds, and Self-Aggregation in RCEMIP Simulations

K. N. Pope[1], C. E. Holloway[1], T. R. Jones[1], T. H. M. Stein[1]

[1]Department of Meteorology, University of Reading

Key Points:

- GCMs aggregate faster than CRMs on average due to an enhanced longwave feedback
- Feedbacks tend to decrease in magnitude as SST increases, although the rate of aggregation remains similar
- Aggregation rate in GCMs is correlated with diabatic feedbacks, while in CRMs it is more related to advection feedbacks

Corresponding author: Kieran Pope, k.n.pope@pgr.reading.ac.uk

Abstract

The responses of tropical anvil cloud and low-level cloud to a warming climate are among the largest sources of uncertainty in our estimates of climate sensitivity. However, most research on cloud feedbacks relies on either global climate models with parameterized convection, which do not explicitly represent small-scale convective processes, or small-domain models, which cannot directly simulate large-scale circulations. We investigate how self-aggregation, the spontaneous clumping of convection in idealized numerical models, depends on cloud-radiative interactions with different cloud types, sea surface temperatures (SSTs), and stages of aggregation in simulations that form part of RCEMIP (the Radiative-Convective Equilibrium Model Intercomparison Project). Analysis shows that the presence of anvil cloud, which tends to enhance aggregation when collocated with anomalously moist environments, is reduced in nearly all models when SSTs are increased, leading to a corresponding reduction in the aggregating influence of cloud-longwave interactions. We also find that cloud-longwave radiation interactions are stronger in the majority of General Circulation Models (GCMs), typically resulting in faster aggregation compared to Cloud-system Resolving Models (CRMs). GCMs that have stronger cloud-longwave interactions tend to aggregate faster, whereas the influence of circulations is the main factor affecting the aggregation rate in CRMs.

Plain Language Summary

The spatial organization of tropical rainstorms has major effects on weather and climate. This organization influences the duration and intensity of these convective storms, and alters the amount of radiation absorbed and emitted by the atmosphere. There is great uncertainty in the response of organisation to a warming climate, and this results in one of the largest sources of uncertainty in climate predictions. Climate projections rely on either General Circulation Models (GCMs) that can represent the large-scale motions, or smaller high-resolution models that represent small-scale features like cloud formations, but not the large motions. In this study, we compare convective organization in GCMs and Cloud-system Resolving Models (CRMs) across a range of sea surface temperatures (SSTs). We find that the cloud-radiation feedbacks that make the convective environment more favorable for further convection, and the non-convective environment less favorable for convection, are stronger in GCMs than CRMs on average. This is related to larger cloud amounts in GCMs, leading GCMs to have typically faster organization than CRMs. We find these feedbacks which drive aggregation decrease as SST increases, yet the aggregation rate is largely insensitive to SST because of the decrease in the effect of atmospheric motions that oppose aggregation.

1 Introduction

Convective self-aggregation is the process by which initially randomly distributed convection becomes spontaneously clustered despite homogeneous boundary conditions and forcing. It was first identified in numerical models of radiative-convective equilibrium (RCE) and has major implications for weather and climate (e.g. Wing et al., 2017). Because of this, it has been the focus of many studies in recent years (e.g. Bretherton et al., 2005; Coppin & Bony, 2015) and continues to be an active area of research. Processes that drive and maintain self-aggregation have been shown to be relevant to observed convection (Holloway et al., 2017), aiding the development of tropical cyclones (Nolan et al., 2007) and the Madden–Julian oscillation (Raymond & Fuchs, 2009; Arnold & Randall, 2015). However, there remains much debate as to the mechanisms and feedbacks responsible for controlling aggregation, which is in part due to the inter-model variability in the structures and dynamics of convection within these models (Wing et al., 2017).

Aggregation of tropical convection has significant impacts on the climate, tending to decrease the total high-cloud fraction and free-troposphere humidity (e.g. Tobin et al., 2013; Wing & Cronin, 2016), affecting the amount of shortwave radiation being absorbed by the atmosphere and surface, as well as affecting the amount of longwave radiation escaping to space. The uncertainty in the response of aggregation to a warming climate is a major source of uncertainty in our estimates for the global climate sensitivity (Sherwood et al., 2020), with models that increase in aggregation with warming tending to have a lower climate feedback parameter due to increased longwave cooling (Wing et al., 2020).

Various metrics have been proposed to characterize aggregation, many of which divide the domain into regions where convection occurs and regions of subsidence. Wing and Emanuel (2014) designed a framework to study aggregation using a variance of frozen moist static energy (FMSE) budget. FMSE, or h, is given by

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

where c_p is the specific heat capacity of dry air at constant pressure, T is temperature, g is the gravitational acceleration, z is the height above the surface, L_v is the latent heat of vaporization, q_v is the water vapor mixing ratio, L_f is the latent heat of fusion and q_i is the condensed ice mixing ratio. As aggregation increases, the spatial variance of column-integrated FMSE increases. In RCE experiments over a fixed sea surface temperature (SST), variations in humidity contribute the most to the spatial variability in FMSE as horizontal temperature gradients are weak, and the gravitational potential term is approximately uniform throughout the domain. Therefore the variance of column-integrated FMSE correlates most strongly with the variance of column relative humidity. Wing and Emanuel (2014) derive a budget equation for the rate of change of vertically-integrated FMSE variance, allowing for the quantification of the contributions of different FMSE feedbacks to the rate of change of aggregation:

$$\frac{1}{2}\frac{\partial \hat{h}'^2}{\partial t} = \hat{h}'LW' + \hat{h}'SW' + \hat{h}'SEF' - \hat{h}'\nabla_h.\hat{\mathbf{u}h} \tag{2}$$

where hats ($\widehat{\ }$) denote a density-weighted vertical integral, LW and SW are the net atmospheric column longwave and shortwave heating rates, SEF is the surface enthalpy flux, made up of the surface latent heat and sensible heat fluxes, $\nabla_h.\hat{\mathbf{u}h}$ is the horizontal divergence of the $\widehat{h}$ flux, and primes ($'$) indicate local anomalies from the instantaneous domain-mean. Each term on the right hand side is a covariance between the $\widehat{h}$ anomaly and the anomaly of a source/sink of $\widehat{h}$. If the term is positive, there is either an anomalous source of $\widehat{h}$ in a region of already high $\widehat{h}$, or an anomalous sink of $\widehat{h}$ in a region of low $\widehat{h}$, representing a positive feedback on self-aggregation. Wing and Emanuel (2014) find each of the terms are important for aggregation, with the longwave and surface flux feedback being crucial drivers of aggregation, but later decreasing and becoming negative as the convection becomes aggregated. They find the shortwave feedback to be a key maintainer of aggregation highlighting that the processes that drive aggregation are separate to the processes that maintain it.

Most research on cloud feedbacks relies on either general circulation models (GCMs) that use parameterized convection, or limited-area cloud-system resolving models (CRMs) with explicit convection that are too small to represent global-scale circulations. The climate feedback and sensitivity of aggregation are different for GCMs and CRMs in the Radiative-Convective Equilibrium Model Intercomparison Project (RCEMIP; Wing et al., 2018), with GCMs typically having a lower climate sensitivity due to convection becoming more aggregated on average at higher SSTs (Becker & Wing, 2020). This response is not seen on average in CRMs.

Despite there being debate as to the processes driving and maintaining aggregation, the majority of studies find that interactions between convection and longwave radiation are key drivers and maintainers of aggregation (Wing et al., 2017). Pope et al. (2021) quantified the contribution of radiative interactions with different cloud types to aggregation using a set of simulations from the UK Met Office Unified Model which are submitted to RCEMIP as UKMOi-vn11.0-RA1-T (referred as UKMO-RA1-T hereafter). They used a similar FMSE variance budget framework to Wing and Emanuel (2014) but normalize $\widehat{h}$ in such a way so that its SST dependence is eliminated, thus making the analysis framework insensitive to SST. They found the direct longwave interactions with high-topped cloud and clear regions to be the main drivers of aggregation. High-topped clouds typically occur in anomalously-high $\widehat{h}$ regions and drastically decrease atmospheric radiative cooling, leading to a positive longwave-FMSE feedback. Similarly, clear regions have anomalously high radiative cooling rates and tend to be found in anomalously-low $\widehat{h}$ regions, again leading to a positive longwave-FMSE feedback and driving aggregation.

Pope et al. (2021) found the main maintainers of aggregation were longwave interactions with high-topped cloud, and shortwave interactions with water vapour. Anomalously humid environments occur in positive $\widehat{h}'$ regions and are able to absorb more solar radiation leading to a positive feedback. The difference in humidity between the moist and dry regions increases with aggregation, hence the shortwave-moisture feedback has a higher impact during mature aggregation. The extents of the contributions of these feedbacks to aggregation are sensitive to SST. In their simulations, the longwave contribution to aggregation is insensitive to SST during the growth phase of aggregation, but there is a smaller longwave contribution to aggregation maintenance as SST increases due to the reduction of high-topped cloud fraction. This decrease in high-topped cloud fraction is consistent with the stability iris mechanism described by Bony et al. (2016), who describe the reduction in anvil cloud as a consequence of increased anvil stability and decreased convective outflow with increasing SST. Shortwave interactions with moisture become less important to aggregation maintenance at warmer SSTs. This is because the variability in atmospheric solar heating between humid and dry regions contributes to a smaller fraction of the total $\widehat{h}$ variability as SST increases. Despite radiative interactions with cloud and moisture being the main drivers of aggregation, the rate of aggregation was most strongly moderated by circulations that generally oppose aggregation, resulting in faster aggregation at warmer SSTs.

In this study, we test the robustness of the conclusions from Pope et al. (2021) by applying their analysis framework to the CRM and GCM simulations in RCEMIP. We quantify the contributions of cloud-radiation interactions to self-aggregation at different stages of organisation and study their SST dependence. We investigate whether the differences in cloud-radiation interactions between models and model types can explain the differences in the behaviour of self-aggregation.

2 Methods

The CRMs and GCMs of RCEMIP are configured using a strict protocol which is described in Wing et al. (2018). CRMs perform ∼100-day, non-rotating, long channel simulations on a domain of ∼6,000 km × 400 km with a 3 km horizontal grid spacing, doubly periodic boundary conditions, and explicit convection. GCMs perform ∼1,000-day, non-rotating, global-scale aquaplanet simulations with parameterized convection. They have a mean grid spacing of $\mathcal{O}$ (1°) varying between ∼100 km and ∼170 km, with the average of all GCMs being ∼120 km. Every model in RCEMIP has constant solar forcing and performs simulations with three fixed SSTs of 295 K, 300 K and 305 K to compare how convection in RCE may be affected by a warming climate.

We study aggregation using the variance of normalized frozen moist static energy budget framework that is described by Pope et al. (2021) (referenced as P21 hereafter).

The framework is based on that used in Wing and Emanuel (2014), however vertically-integrated FMSE is normalized between hypothetical upper and lower limits based on SST in an attempt to eliminate the strong temperature dependence of FMSE. This approach uses the variance of normalized FMSE ($\text{var}(\widehat{h}_n)$) as the aggregation metric because it is approximately insensitive to SST. The budget equation for the rate of change of $\text{var}(\widehat{h}_n)$ is:

$$\frac{1}{2}\frac{\partial \widehat{h}_n'^2}{\partial t} = \hat{h}_n' LW_n' + \hat{h}_n' SW_n' + \hat{h}_n' SEF_n' - \hat{h}_n' \nabla_h.\mathbf{u}\hat{h}_n \tag{3}$$

Here, $\widehat{h}_n'$ and each of the three normalized flux anomalies on the RHS (LW_n', SW_n', and SEF_n') is equal to the original flux anomaly in equation (2) divided by the difference between the upper and lower limits of $\widehat{h}$ ($\widehat{h}_{\max}$ and $\widehat{h}_{\min}$). $\widehat{h}_{\max}$ is defined as the vertically-integrated FMSE of a fully saturated moist pseudoadiabatic profile from the surface to the tropopause, plus the integrated FMSE of the initial profile above the tropopause. For $\widehat{h}_{\min}$, the vertically-integrated FMSE of a dry adiabatic profile with zero moisture is used within the troposphere, and integrated FMSE above the tropopause from the initial profile is added. The SST is used as the temperature at sea-level pressure to initiate both adiabatic profiles. The tropopause is defined as the lowest level in the initial profile at which the lapse rate decreases to 2°C/km or less, which has some variability in height between model simulations.

$\text{Var}(\widehat{h}_n)$ is not only dependent on spatial aggregation, but it is also sensitive to grid spacing, particularly while convection is well-scattered. This is because small-scale features, e.g. convective updrafts and downdrafts that tend to have strong positive and negative $\widehat{h}_n'$ respectively, are not resolved at coarser resolutions. This leads to a smaller $\text{var}(\widehat{h}_n)$ for coarser horizontal resolutions. As the size of the convective clusters increase and $\widehat{h}_n$ anomalies are strong over large areas, $\text{var}(\widehat{h}_n)$ becomes less sensitive to grid spacing (analysis not shown). To make the comparison between CRMs and the 40× coarser GCMs as fair as possible, we horizontally smooth the raw output fields of the CRMs so that every grid box is the mean of the 40 × 40 grid boxes surrounding it (accounting for the periodic boundary conditions). When using this smoothing technique in the analysis, we refer to the CRMs as Smoothed CRMs.

In a similar way to P21, we define Growth and Mature phases of aggregation by two ranges of $\text{var}(\widehat{h}_n)$ for which convection, in the majority of models, is randomly scattered or well clustered, respectively. The Growth phase is identified as any time after day 2 (to neglect spin-up effects) when $\text{var}(\widehat{h}_n)$ for GCMs and Smoothed CRMs is between 0.8×10^{-4} and 2.4×10^{-4}. The Mature phase is identified as any time when $\text{var}(\widehat{h}_n)$ for GCMs and Smoothed CRMs is between 0.8×10^{-3} and 2.4×10^{-3}. Given our previous notion that $\text{var}(\widehat{h}_n)$ is sensitive to grid spacing, we use the times of the Growth and Mature phases identified from the Smoothed CRMs to also analyse the (non-Smoothed) CRMs.

Since $\widehat{h}_n'$ is a factor of every term in Equation 3, one might expect the magnitude of the terms to increase with aggregation. By dividing each term by the instantaneous horizontal standard deviation of $\widehat{h}_n$, we can eliminate the dependence of the terms on the magnitude of $\widehat{h}_n'$. After dividing by this standard deviation, the sensitivity of the terms to aggregation will depend on the sensitivity of the other variable in the term and its correlation with $\widehat{h}_n'$.

A drawback of the $\text{var}(\widehat{h}_n)$ budget framework is that it is a vertically-integrated framework that is not able to quantify the effects of processes occurring at specific vertical levels. Studies have shown that there are many low-level processes that are important for aggregation. For example, Muller and Held (2012) highlight the importance of shallow, radiatively-driven circulations caused by cooling atop shallow clouds in dry regions, yielding an upgradient transport of FMSE, inducing a positive aggregation feedback. Jeevanjee and Romps (2013) describe how cold pools are responsible for the domain size dependence of self-aggregation. Boundary layer processes are key for the production of available potential energy that is associated with the development of self-aggregation (Yang, 2018a), and are theorized to determine the length scale of aggregation (Yang, 2018b). The use of our vertically-integrated framework means the effects of these processes are not directly studied. Circulations that are induced by diabatic forcing are included in the vertically-integrated advection term in the var($\widehat{h}_n$) budget framework. So the radiation and surface flux terms only account for the *direct* diabatic feedbacks.

2.1 Cloud Classification Scheme

We use a cloud classification scheme to define a cloud type at each grid point in the simulations. The contribution of radiative interactions with these cloud types to aggregation are calculated by multiplying each cloud type's fraction by the mean covariance between its radiative and FMSE anomalies. This analysis technique is based on that used by P21, however the cloud type definitions in this study are different. In RCEMIP, 3D data are only available for the final 25 days of CRMs and GCMs, so we are not able to define cloud based on the vertical profile of condensed water for the full simulation as in P21. Instead, we define clouds using top of atmosphere (TOA) fluxes, using the same method as Becker and Wing (2020) (referenced as BW20 hereafter). This method produces four different cloud types: Clear, Shallow, Deep, and Other. The outgoing shortwave radiation (OSR) and outgoing longwave radiation (OLR) thresholds used to define the four cloud types are shown in Table 1.

Table 1. OSR and OLR thresholds used to define the cloud types.

Cloud type	OSR (W m^{-2})	OLR (W m^{-2})
Clear	< 100	N/A
Shallow	≥ 100	> 250
Other	≥ 100	173 - 250
Deep	≥ 100	< 173

A comparison of the cloud type classification schemes between that used in P21 and this study is shown in Figure 1(a-d). These figures show the P21 cloud distributions for each of the BW20 cloud types across all of the CRMs. Approximately 80% of this study's Clear category is made up of the Clear type defined in P21, meaning the condensed water content is less than 10^{-6} kg m^{-3} everywhere in the column. The remainder of the BW20 Clear category is mostly made up of optically-thin High and Low cloud. The Shallow cloud type is mostly made up of Low cloud, and the Deep cloud is almost entirely made up of the high-topped cloud (High, High & Mid, High & Low, and Deep). The Other cloud type is made up of approximately two thirds high-topped cloud that is perhaps too optically thin or having too small a vertical extent to lead to an OLR less than 173 W m^{-2} and be classed as Deep.

Cloud types are redefined using the Smoothed radiative fluxes in order to make a fairer comparison to GCMs. The distribution of non-Smoothed clouds for each Smoothed cloud type is shown in Figure 1(e-h). The Smoothed Clear and Deep categories are mainly made up of the non-Smoothed Clear and Deep categories respectively. The Smoothed Shallow cloud is only about one quarter made up of non-Smoothed Shallow cloud, and mostly made up of Clear. The Smoothed Other cloud type is mostly a combination of Clear, Other and Deep regions.

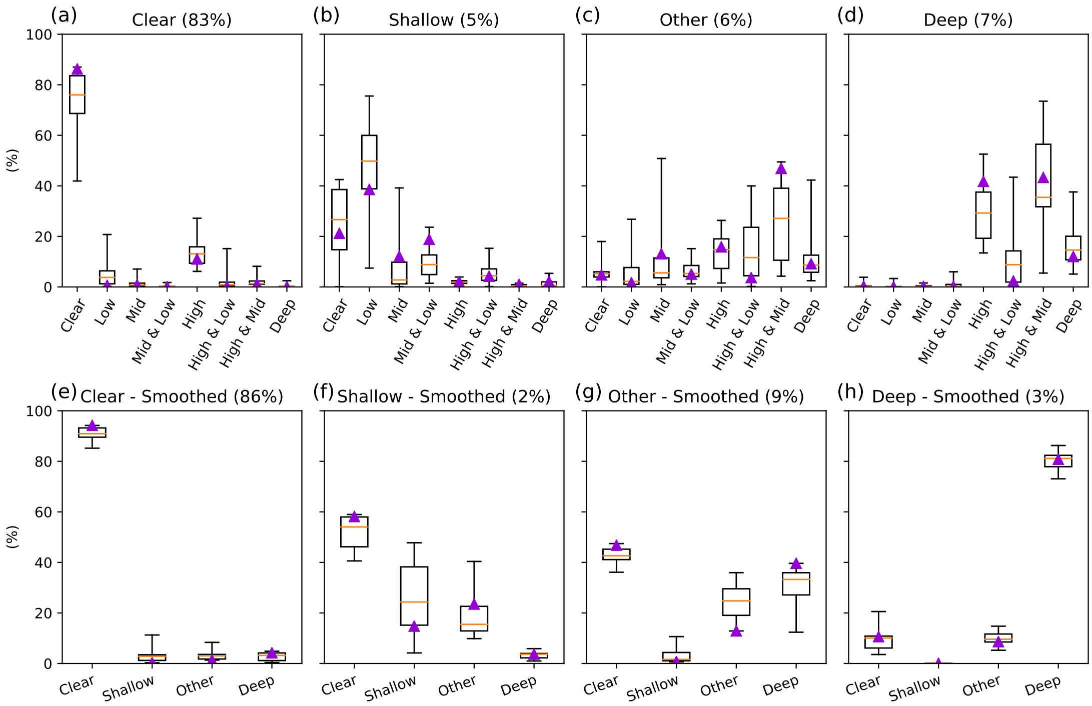

Figure 1. (a-d) Distributions of the cloud categories used in P21 for each of the four cloud types used in this study. Data is averaged over the final 25 days of the CRMs for all SSTs. (e-h) Distributions of this study's cloud types for each of the Smoothed cloud types. Data is averaged over the full duration of the CRMs (neglecting the 2-day spin-up period) for all SSTs. Orange lines represent the median, boxes represent the interquartile range, and whiskers represent the full range of the models. The UKMO-RA1-T model is shown in purple triangles. Average domain fraction is shown in the subplot titles.

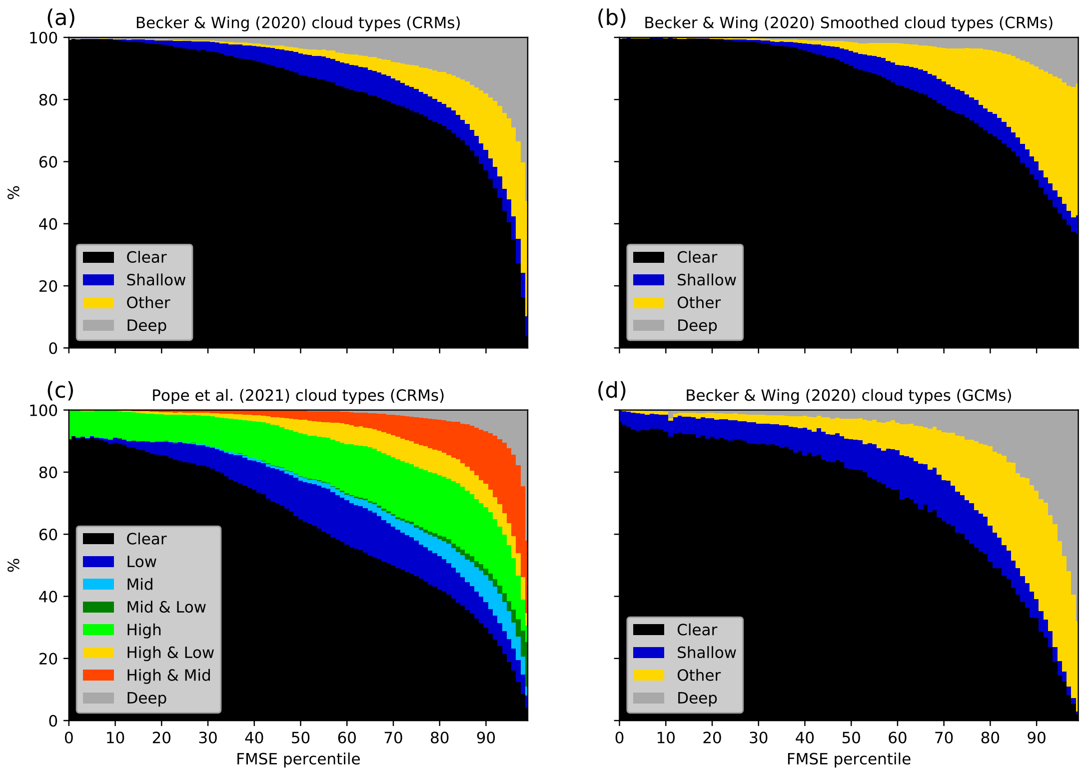

Figure 2. Cloud type fraction vs FMSE percentile for the (a) BW20 cloud types for all CRMs, (b) Smoothed BW20 cloud types for all CRMs, (c) P21 cloud types for all CRMs, and (d) BW20 cloud types for the GCMs during the final 24 hours of the simulations.

Figure 2 shows the fraction of different cloud types as a function of FMSE percentile during the final 24 hours of the simulations. Differences in the BW20 and P21 cloud classification schemes within the CRMs can be seen by comparing Figures 2a and 2c. Cloud fraction increases with FMSE percentile regardless of the cloud classification scheme used. There is a lower cloud fraction in the BW20 cloud types compared to the P21 cloud types at all FMSE percentiles except for the extremely moist environments in which the cloud fraction is close to 100%. There is greater high-topped cloud in the P21 classification scheme compared to the BW20 Deep cloud which may be due to the presence of optically-thin High cloud that has OSR < 100 W m^{-2}. There is also a greater fraction of P21 Low cloud compared to BW20 Shallow cloud at all FMSE percentiles, again due to the presence of optically thin Low cloud with OSR < 100 W m^{-2}.

The effect of Smoothing is shown by comparing Figures 2a with 2b. Smoothing reduces the total cloud fraction in the lower 40% and upper 10% of FMSE values. The fraction of Deep cloud is reduced and the fraction of Other cloud is increased at all FMSE percentiles. The difference between Smoothed CRMs and GCMs can be seen by comparing Figures 2b and 2d. There is a greater cloud fraction in GCMs at all FMSE percentiles, which is largely due to the increase in Deep cloud fraction. There is also a greater Shallow cloud fraction particularly at lower FMSE values, and a lower Other cloud fraction at higher FMSE values. The effects of Smoothing, and comparisons between CRMs and GCMs are discussed further in section 4. The cloud type fractions of the non-Smoothed CRMs are most similar to the fractions of the GCMs, suggesting GCMs may be tuned to have a more accurate cloud fraction in a discrete grid box sense rather than on subgrid scales. Yet GCMs still have a greater average cloud fraction particularly at higher $\widehat{h}'_n$ regions.

Radiative interactions with high-topped cloud and Clear regions are shown to have the largest role in aggregation in P21. With the majority of BW20 Clear and Deep clouds being collocated with P21 Clear and high-topped cloud respectively, results from P21 can be fairly compared to results from this study.

3 Variance of Normalized FMSE

The RCEMIP CRMs simulate a wide range of convective characteristics (Wing et al., 2020). All models display aggregation to some degree except for the UKMO-CASIM model at 305 K, whose convection remains scattered throughout the entire simulation. Figure 3 shows 24-hour running averages of var($\widehat{h}_n$) for each Smoothed CRM and SST. Also shown are the var($\widehat{h}_n$) limits for the Growth and Mature phase of aggregation (introduced in section 2), which will be discussed later. There is much variability in the rate of aggregation amongst the CRMs as well as the maximum degree of aggregation, with no consistent SST dependence. The inconsistent SST dependence of aggregation is seen regardless of aggregation metric used (Wing et al., 2020). Not all models reach both the Growth and Mature stages of aggregation at all three SSTs. These models are marked with an asterisk in Figure 3 and do not contribute to model-mean calculations to prevent skewing the results.

Figure 4 shows 24-hour running averages of var($\widehat{h}_n$) for each GCM and SST. Also shown are the var($\widehat{h}_n$) limits for the Growth and Mature phase of aggregation. All of the GCMs aggregate, again displaying a wide range of characteristics (Wing et al., 2020). Unlike the CRMs, aggregation increases with SST in the majority of GCMs. GCMs that reach a more aggregated state at warmer SSTs do not usually aggregate faster as SST increases, but they tend to continue aggregating for a longer duration. As with the CRMs, we do not include all GCMs in the model-mean calculations as not all models have data in both the Growth and Mature phases of aggregation for each of the SSTs. These models are marked with an asterisk. Note CAM5 and CAM6 have FMSE data only for the final 25 days of the 1095-day simulation. ICON-GCM at 300 K already has a variance

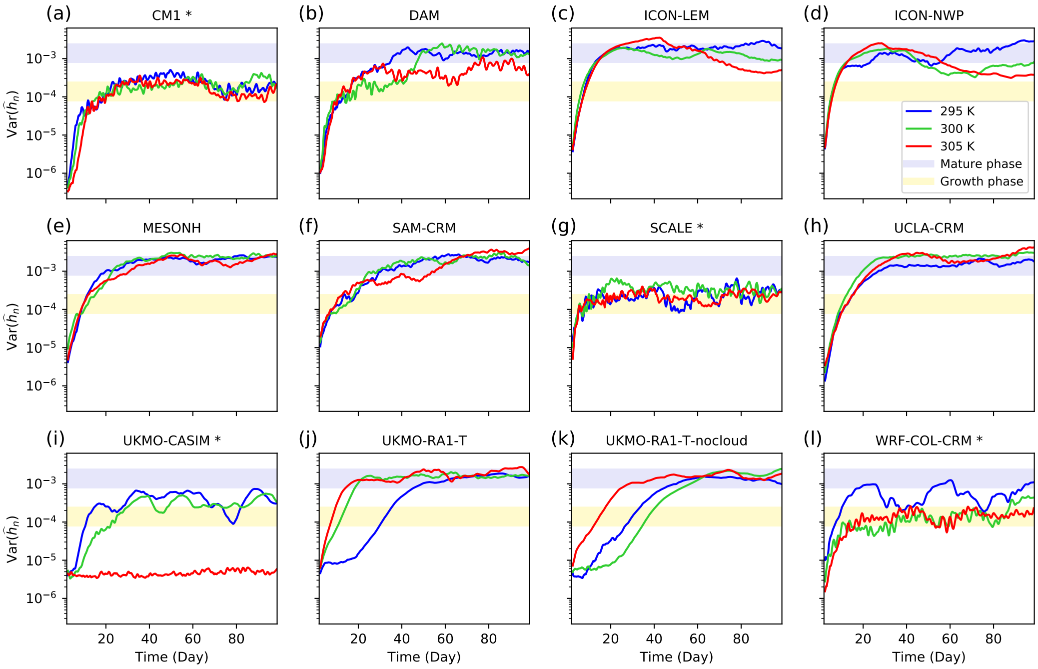

Figure 3. Time series of var($\widehat{h}_n$) for each Smoothed CRM and SST neglecting the first two days accounting for model spin-up (24-hour running averages). The Growth and Mature phases are indicated by the yellow and blue shaded regions respectively. Models marked with an asterisk (*) are excluded in future model-mean calculations as not all of their simulations reach the Growth and Mature phase for all SSTs.

greater than the upper limit for the Growth phase after two days (which we consider the spin-up period) so is not included in model-mean calculations. SP-CAM and SPX-CAM are also excluded from all further analysis because of abnormally-large longwave cooling rates across the entire domain. Domain-mean longwave cooling within the 300 K simulations of both the GCMs and CRMs range between 150 and 230 W m^{-2}, whereas the cooling rates for SP-CAM and SPX-CAM are around 325 W m^{-2}. This has knock-on effects, affecting the longwave heating anomalies of clouds, their longwave-FMSE anomaly covariance and their contribution to aggregation (analysis not shown). ECHAM6 and GEOS *are* included in the model-mean calculations because the 295 K simulations reach the Mature stage after the 100 days shown in Figure 4.

Figure 5 shows the spatiotemporal mean of the budget terms during the Growth phase and Mature phase of aggregation for Smoothed CRMs and GCMs and for each SST. From this figure, we can see which FMSE covariances are enhancing or opposing aggregation at these different stages. The var($\widehat{h}_n$) tendency is calculated using a second-order finite difference approximation from 6-hourly calculated var($\widehat{h}_n$). The diabatic terms are explicitly calculated, and the advection term is calculated as a residual of the other terms. By comparing GCMs to the Smoothed CRMs, we remove biases that may be a result of the small-scale features that cannot be resolved in the larger grid spacing in GCMs.

Figure 5 shows that for all model types, and at all SSTs, FMSE feedbacks with longwave radiation and surface fluxes are typically the main drivers of aggregation in the Growth phase, however the magnitude of each feedback is highly variable from model to model. The shortwave term is consistently small and positive and has little inter-model variabil-

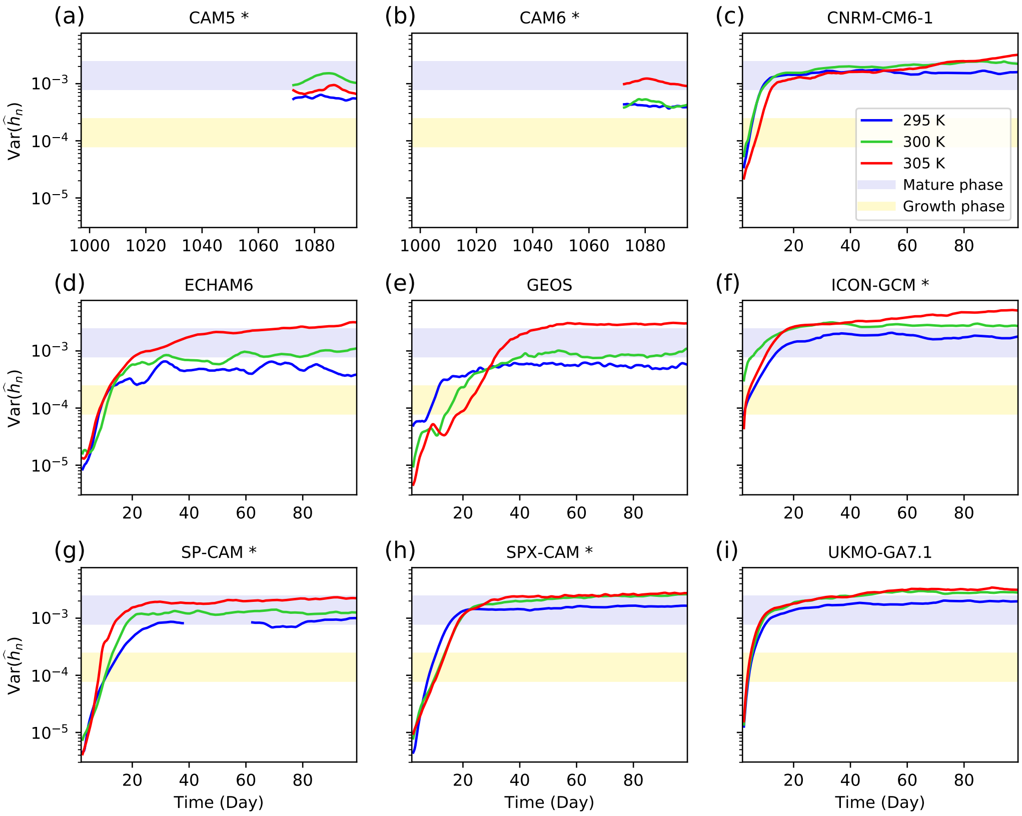

Figure 4. Time series of var($\widehat{h}_n$) for each GCM and SST for the first 100 days, neglecting the first two days accounting for model spin-up (24-hour running averages). Note CAM5 and CAM6 output FMSE for the final 25 days only and so we only show that time period for those models. The Growth and Mature phases are indicated by the yellow and blue shaded regions respectively. Models marked with an asterisk (*) are excluded in future model-mean calculations.

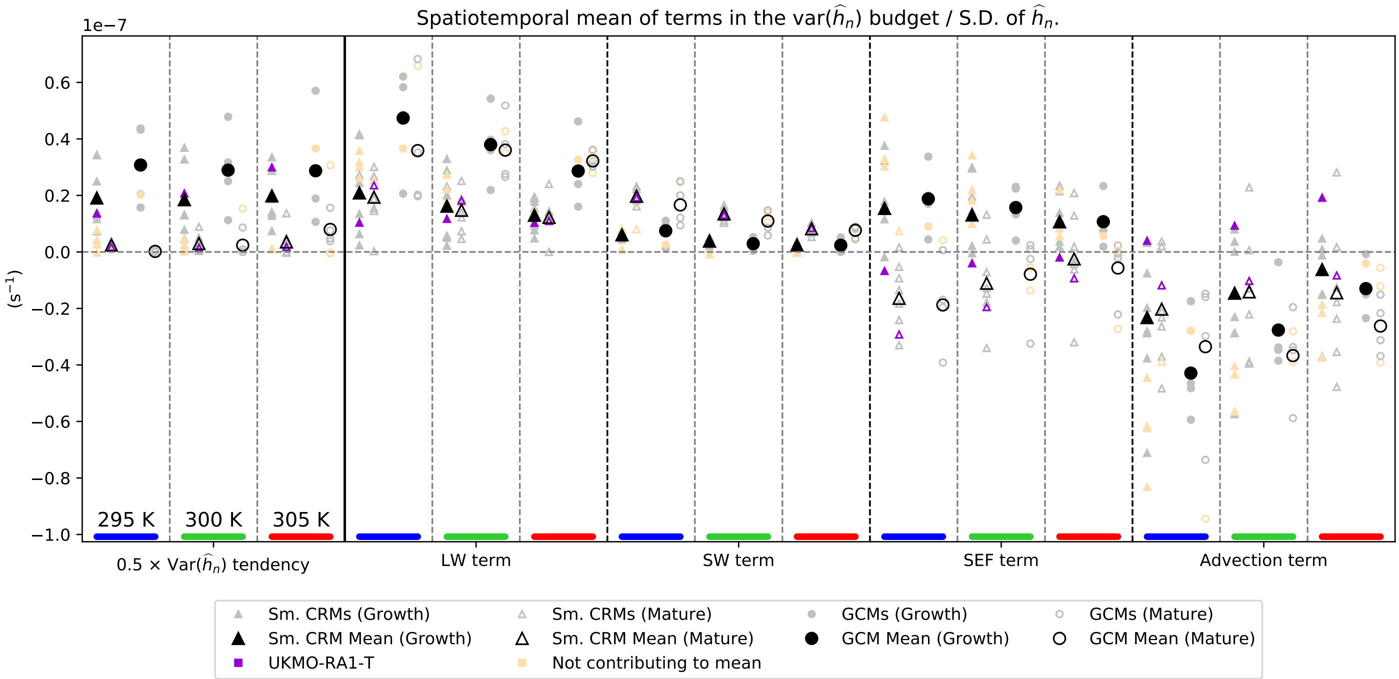

Figure 5. Spatiotemporal mean of terms in the var($\widehat{h}_n$) budget equation divided by the instantaneous standard deviation of $\widehat{h}_n$ for Smoothed CRMs (triangles) and GCMs (circles) at each SST during the Growth phase (filled markers) and Mature phase (open markers) of aggregation. For each term, SST increases to the right. The mean for the Smoothed CRMs and GCMs for each SST are shown in black markers. Models that do not reach both the Growth and Mature phase at all three SSTs are shown with orange markers and do not contribute to the mean. SP-CAM and SPX-CAM are excluded from the figure. UKMO-RA1-T is shown in purple.

ity. The advection term typically opposes aggregation and is the greatest source of variability for the rate of aggregation across the models.

During the Mature phase of aggregation, both the longwave and shortwave feedbacks maintain aggregation, and are balanced by the typically-negative surface flux and advection feedbacks. On average, the magnitude of the longwave feedback has little dependence on the degree of aggregation, whereas the shortwave feedback increases with aggregation as moist and dry regions amplify, leading to larger differences in shortwave absorption between positive and negative $\widehat{h}'_n$ regions. The surface flux feedback is usually positive during the Growth phase as higher surface wind speeds in moist convective regions leads to a positive feedback. During the mature phase, the wind speed-surface flux feedback becomes overcompensated by the negative air-sea disequilibrium feedback, whereby surface evaporation rates are enhanced in drier environments (Wing & Emanuel, 2014). The surface flux feedback during the Mature phase at higher SSTs may be less negative due to the wind-evaporation feedback being relatively stronger (Coppin & Bony, 2015).

As noted by Wing et al. (2020), GCMs tend to reach a higher degree of aggregation at higher SSTs. With little SST dependence of the rate of aggregation in our defined Growth phase, aggregation rates increase with SST for $\mathrm{var}(\widehat{h}_n)$ greater than the upper limit of the Growth phase. This can be seen in many of the models in Figure 4 and to some extent in Figure 5 by looking at the $\mathrm{var}(\widehat{h}_n)$ tendency of GCMs during the Mature phase which increases slightly with SST. However, the greatest SST dependence of the rate of change of $\mathrm{var}(\widehat{h}_n)$ is during the times in between the Growth and Mature phase (not shown). For GCMs during the Growth phase, the sum of the diabatic terms decrease in magnitude with SST, yet the advection term becomes more positive with SST, resulting in little SST dependence in the rate of aggregation in the Growth phase. After the Growth phase however, the sum of the diabatic feedbacks becomes less SST dependent, while the advection term remains more positive with SST. This results in a greater rate of aggregation after our defined Growth phase. In CRMs, the sum of the diabatic terms also becomes less sensitive to SST after the Growth phase, though they still have a more negative SST dependence than the average of the GCMs. The main difference between GCMs and CRMs is the SST sensitivity of the longwave term after the Growth phase, which remains more constant on average with SST in GCMs. This will be explored further in the following section.

The longwave feedback is on average a factor 2 greater in GCMs compared to CRMs for all stages of aggregation. The larger longwave feedback in GCMs is the main difference in terms of the diabatic feedbacks between CRMs and GCMs. This results in GCMs having an overall larger diabatic feedback, corresponding to a more negative advection feedback and/or a higher rate of aggregation in the Growth phase. There is, however, a large spread in the models' advection term and aggregation rate. The difference between the mean advection term between GCMs and Smoothed CRMs is not statistically significant at the 95% confidence level for a given SST, even when including the models that are neglected from the model-mean comparisons. The increase in mean aggregation rate from the Smoothed CRMs to the GCMs is only significant at each SST when we include the models neglected from the model-mean comparisons. The difference in the longwave feedbacks in CRMs and GCMs is significant and will be discussed further in the next section.

There is little difference in the budget terms between the non-Smoothed and Smoothed CRMs (not shown). After dividing the terms by the standard deviation of $\widehat{h}_n$, the rate of aggregation, longwave term, and shortwave term remain similar on average. The most significant change is the surface flux term during the Growth phase, which decreases by about 40% after smoothing. With the surface flux term decreasing in the Growth phase, and the other diabatic terms and $\mathrm{var}(\widehat{h}_n)$ tendency term remaining similar, the advection term becomes more positive after smoothing as it is calculated as a residual of the other terms.

If FMSE feedbacks in CRMs and GCMs are represented similarly despite the different grid spacings, the budget terms in GCMs should be similar to the budget terms in the Smoothed CRMs. For both CRMs and GCMs, each of the diabatic terms are typically positive during the Growth phase but on average decrease in magnitude as SST increases (Figure 5). P21 studied the UKMO-RA1-T model simulations which are represented by the purple, triangular data points in Figures 5, 7 & 8. They analysed this SST dependence of the UKMO-RA1-T CRM and found the longwave feedback decreases with SST due to the reduction of high-cloud fraction at higher SSTs. However in their study, this SST dependence was only found in the Mature phase. We explore how high-cloud fraction affects the longwave feedback in the RCEMIP CRMs and GCMs in the following section. P21 found the decrease in the shortwave feedback to be inversely proportional to the difference between $\widehat{h}_{\max}$ and $\widehat{h}_{\min}$. Physically, this means that the shortwave heating anomalies contribute similar amounts to increasing the non-normalized FMSE variance at different SSTs. However, since FMSE anomalies are higher at warmer SSTs, the shortwave heating anomalies contribute to a smaller fraction of FMSE variance. For both CRMs and GCMs in RCEMIP, the advection term becomes less negative with SST on average and is inversely proportional to the sum of the diabatic terms. The result is that the rate of aggregation during the Growth phase for both CRMs and GCMs does not depend strongly on SST.

Some of the results from the mean of the models are in contrast to the results found in P21. According to the model means, the surface flux feedback is almost as important as the longwave feedback in driving aggregation, which is in stark contrast to the UKMO-RA1-T model that shows the surface flux feedback to be slightly negative even during the Growth phase. This suggests the air-sea disequilibrium feedback in the UKMO-RA1-T model dominates over the wind speed-surface flux feedback to a larger degree than in the majority of models. The sum of the diabatic terms decreases with SST for the model means, yet it is more constant with SST in the UKMO-RA1-T simulations and is also more negative. Despite the more negative diabatic feedback in UKMO-RA1-T, the rate of aggregation is faster than the model means at 300 K and 305 K. This is because the UKMO-RA1-T model has the most positive advection feedback of all models. This feedback increases with SST despite the diabatic terms remaining similar, resulting in faster aggregation at higher SSTs in UKMO-RA1-T, but there is little change in aggregation rate with SST for the model mean.

Previous literature has shown the diabatic terms to be essential drivers of aggregation, so we would expect that a greater diabatic-FMSE feedback would lead to an increased rate of aggregation. Despite the diabatic terms driving aggregation in the Growth phase of the RCEMIP simulations (Figure 5), we cannot conclude that the magnitude of the sum of the diabatic terms is correlated to the rate of aggregation. Figure 6a shows the correlation between the longwave term and the $\mathrm{var}(\widehat{h}_n)$ tendency term in Equation 3 during the Growth phase for Smoothed CRMs and GCMs. We find there is a significant correlation between the longwave term and rate of aggregation in the GCMs, but there is no significant correlation between the longwave term and rate of aggregation in the CRMs (regardless of Smoothing). Figure 6b shows the correlation between the sum of the diabatic terms and the $\mathrm{var}(\widehat{h}_n)$ tendency term. Again there is a significant positive correlation between the diabatic feedbacks and rate of aggregation in the GCMs, but not for the CRMs. A greater diabatic feedback is associated with a more negative advection feedback (Figure 6c). In the CRMs, the sum of the diabatic terms is, on average, proportional to the magnitude of the advection feedback, hence there is no significant relationship between the diabatic feedbacks and aggregation rate. There is a less negative relationship between the sum of the diabatic terms and the advection term in the GCMs, allowing GCMs with a higher diabatic feedback to aggregate faster. The rate

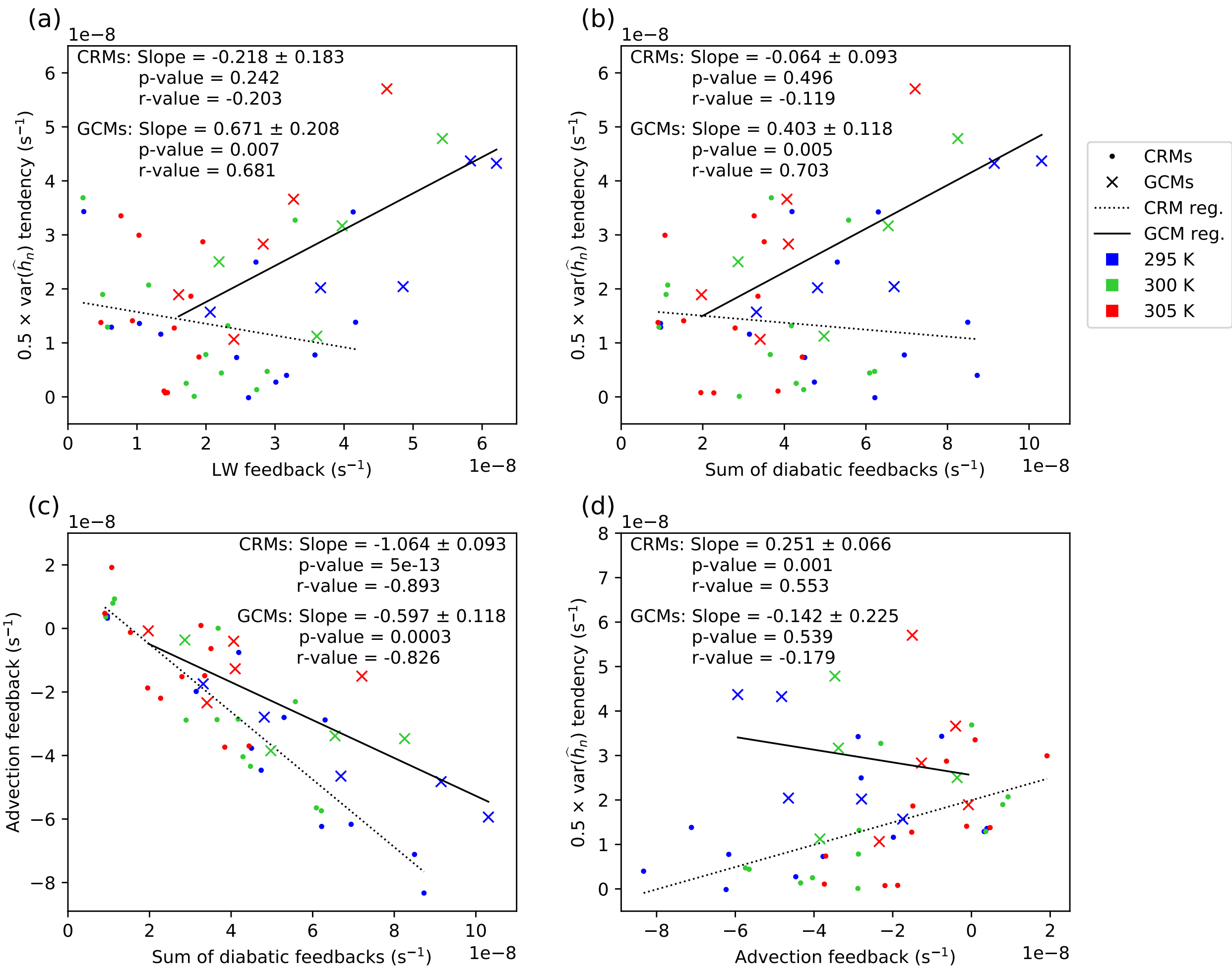

Figure 6. (a) Average of the var($\widehat{h}_n$) tendency term vs the longwave term in Equation 3, (b) average of the var($\widehat{h}_n$) tendency term vs the sum of the three diabatic terms (longwave, shortwave & surface flux), (c) average of the advection term vs the sum of the diabatic terms, and (d) average var($\widehat{h}_n$) tendency term vs the advection term, for each Smoothed CRM (points) and GCM (crosses) averaged over the Growth phase. Also shown is the regression line for CRMs (dotted) and GCMs (solid line), as well as their slope, p-value and r-value.

of aggregation in CRMs is most strongly correlated with the advection feedback (Figure 6d), with no significant correlation between the advection feedback and aggregation rate in the GCMs.

The longwave feedback is a key driver and maintainer of aggregation in the majority of models at each SST. It is typically a larger feedback in GCMs, resulting in largely faster aggregation rates compared to CRMs. The longwave feedback is a key factor in determining the model spread in the rate of aggregation, as well as the sensitivity of the degree of aggregation to SST in GCMs.

4 Contributions of Cloud-Radiation Interactions to Aggregation

In this section, we compare longwave-cloud interactions within the CRMs and GCMs. We first study these interactions in the CRMs to test the robustness of the conclusions in P21. We then compare CRMs to GCMs by first seeing how cloud-longwave interactions are affected by coarsened grid spacing using the Smoothed CRMs. Then we compare the Smoothed CRMs to GCMs to study why the longwave feedback tends to be stronger in GCMs.

4.1 Cloud-Radiation Interactions within CRMs

The contributions of longwave interactions for the different cloud types in the CRMs and Smoothed CRMs during the Growth and Mature phase of aggregation for each SST are shown in Figure 7a. Each model that contributes to the mean is shown in grey, the model mean shown in black, UKMO-RA1-T is shown in purple, and models that do not contribute to the mean are shown in light orange. We first focus on the (non-Smoothed) CRMs.

For the CRMs during the Growth phase of aggregation, longwave interactions with the Clear and Deep regions contribute most to the longwave feedback. The Clear regions have a large contribution mainly because of their large domain-fraction (Figure 7b) and positive $LW'_n \times \widehat{h}'_n$ covariance (Figure 7c), despite the covariance being on average the lowest in magnitude out of all cloud types. Deep clouds are the next most abundant cloud type on average and typically have the largest $LW'_n \times \widehat{h}'_n$ covariance of all cloud types. They have the largest LW'_n due to their cold cloud tops (Figure 7e) and have the second highest $\widehat{h}'_n$ of the cloud types (Figure 7d). A large portion of the Deep category comes from thin anvil cloud which often extend a great distance from the high-FMSE updraft that they originated from. This transport of high cloud to lower-FMSE regions lowers the average $\widehat{h}'_n$ of the Deep category. The Shallow and Other cloud types have an insignificant contribution to the longwave feedback in comparison because their $LW'_n \times \widehat{h}'_n$ covariance is small in magnitude (mostly due to a small-magnitude LW'_n) and they have a small fraction (although the fraction is highly variable between models).

The negative SST dependence of the longwave feedback, as seen in Figure 5, can be explained by the negative SST dependence of the longwave interactions with the Deep and Clear regions as follows, in agreement with P21. During both the Growth and the Mature phases, the $LW'_n \times \widehat{h}'_n$ covariance of the Deep regions remains similar with SST (Figure 7c) while the Deep cloud fraction steadily decreases (Figure 7b), so the SST dependence of the Deep cloud's longwave contribution to aggregation is primarily due to the decrease in Deep cloud fraction.

The contribution of the Clear regions decreases with SST due to the decrease in the Clear $LW'_n \times \widehat{h}'_n$ covariance. There are multiple factors that influence this SST dependence: the change in longwave heating rates of the different cloud types, the change in their fraction, the increase in the range of $\widehat{h}_{\max}$ and $\widehat{h}_{\min}$, and the change in correlation between longwave and FMSE anomalies in the Clear regions. The correlation be-

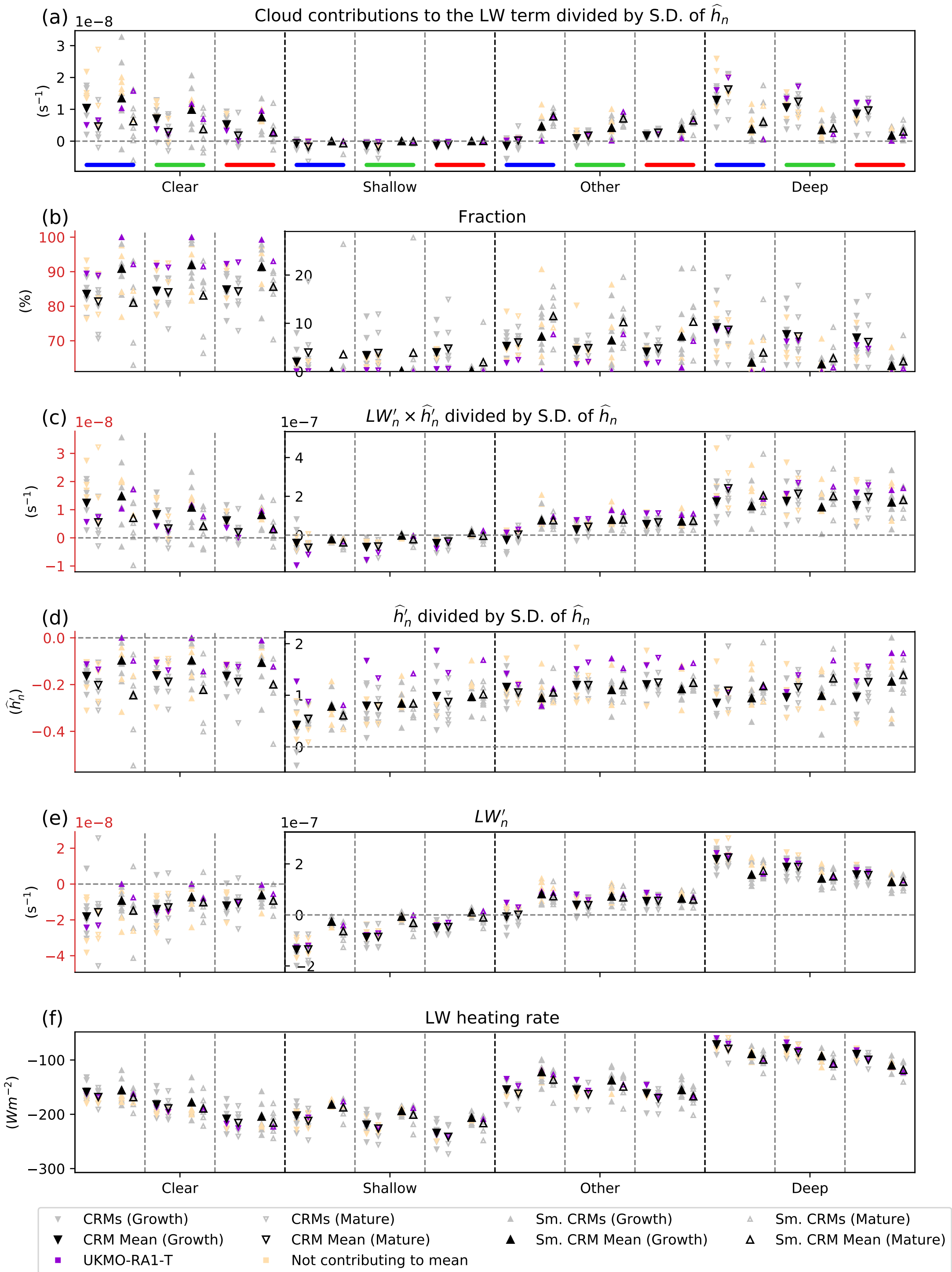

Figure 7. Non-Smoothed CRMs (downward triangles) vs Smoothed CRMs (upward triangles): (a) Contributions of longwave interactions for each cloud type to the longwave term in equation 3 divided by the standard deviation of $\widehat{h}_n$, (b) Fraction of each cloud type, (c) $LW'_n \times \widehat{h}'_n$ covariance divided by the standard deviation of $\widehat{h}_n$, (d) $\widehat{h}'_n$ divided by the standard deviation of $\widehat{h}_n$, (e) LW'_n, and (f) absolute longwave heating. Data points and layout follow the same protocol as in Figure 5. Note different y-axis ranges for Clear in b, c, d & e.

tween LW'_n and $\widehat{h}'_n$ remains similar with SST (15% decrease in the correlation coefficient from 0.173 at 295 K to 0.147 at 305 K), as does the mean $\widehat{h}'_n$ (Figure 7d). The change in the Clear $LW'_n \times \widehat{h}'_n$ covariance is therefore mainly due to the change in LW'_n.

To isolate the effects of the changing longwave heating rates with SST on the Clear longwave feedback, we use the average cloud type fractions at 295 K with the average cloud type longwave heating rates at 305 K. From these, we calculate a hypothetical new domain-mean longwave cooling rate and cloud type LW', and find that the average Clear LW' becomes 74% more negative compared to the values at 295 K. However, after normalising LW' to account for the changing SST, we find this hypothetical new Clear LW'_n is largely insensitive to SST. We next isolate the effect of the changing cloud fraction with SST by using the average cloud type longwave heating rates at 295 K with the average cloud type fractions at 305 K to calculate the cloud types' LW'. We find the domain-mean longwave cooling rate increases by approximately 3 W m^{-2} compared to the value at 295 K, and is mainly a result of the decreasing Deep cloud fraction allowing for enhanced radiative cooling. The increased domain-mean cooling rate is closer to the mean cooling rate of the Clear regions, making their LW' 37% less anomalously negative. This is close to the actual 30% decrease in the mean LW'_n of the Clear regions. This shows that the SST sensitivity of the Clear LW'_n is primarily due to changes in cloud fraction with SST.

Next, we look at the effects of smoothing on cloud-longwave interactions in the CRMs to see how a coarser grid spacing affects cloud-longwave interactions. After smoothing the TOA radiative fluxes and reclassifying the cloud types using the smoothed radiation, there is a large difference in the fraction of the different cloud types (Figure 7b). Firstly, there is an almost complete elimination of Shallow cloud in the Smoothed CRMs during the Growth phase, with a large reduction in Deep cloud in the Growth and Mature phases. This is because the Shallow and Deep clouds are often small in area, particularly during the Growth phase, meaning that after averaging the TOA radiative fluxes across the surrounding 120 km × 120 km area, these clouds are often reclassified as either Clear or Other clouds. This results in an increase in Other cloud, although there is an approximate halving of the total cloud fraction during the Growth phase. During the Mature phase, all cloud types increase in fraction in the Smoothed CRMs as a likely result from increased cloud clustering. The total cloud fraction in the Mature phase is similar to the non-Smoothed CRMs.

Smoothing also has an effect on the average $LW'_n \times \widehat{h}'_n$ covariance of the cloud types (Figure 7c). The covariance remains similar for Deep cloud, but increases slightly for the Other cloud, perhaps a result of a significant proportion of the non-Smoothed Deep cloud regions becoming reclassified as Other after Smoothing, as can be inferred by comparing Figures 2a & b. The combined effects of the change in cloud fraction and $LW'_n \times \widehat{h}'_n$ covariance after Smoothing is a reduction in the contribution from Deep cloud with subsequent increases in the contributions from the Other and Clear cloud types during all stages of aggregation.

4.2 Comparison of Cloud-Radiation Interactions within CRMs and GCMs

In Figure 8 we compare the longwave-cloud interactions between the Smoothed CRMs and GCMs. Figure 8a shows that during the Growth phase, longwave interactions with the Clear regions and Deep regions are the main drivers of aggregation for GCMs, with interactions with Other clouds also having a significant contribution. Contributions of each of these cloud types to the total longwave feedback are higher in GCMs compared to the Smoothed CRMs. This is largely due to the increased fraction of the Other and Deep cloud types (Figure 8b), but also the increased $LW'_n \times \widehat{h}'_n$ covariance of the Deep and Clear cloud types (Figure 8c).

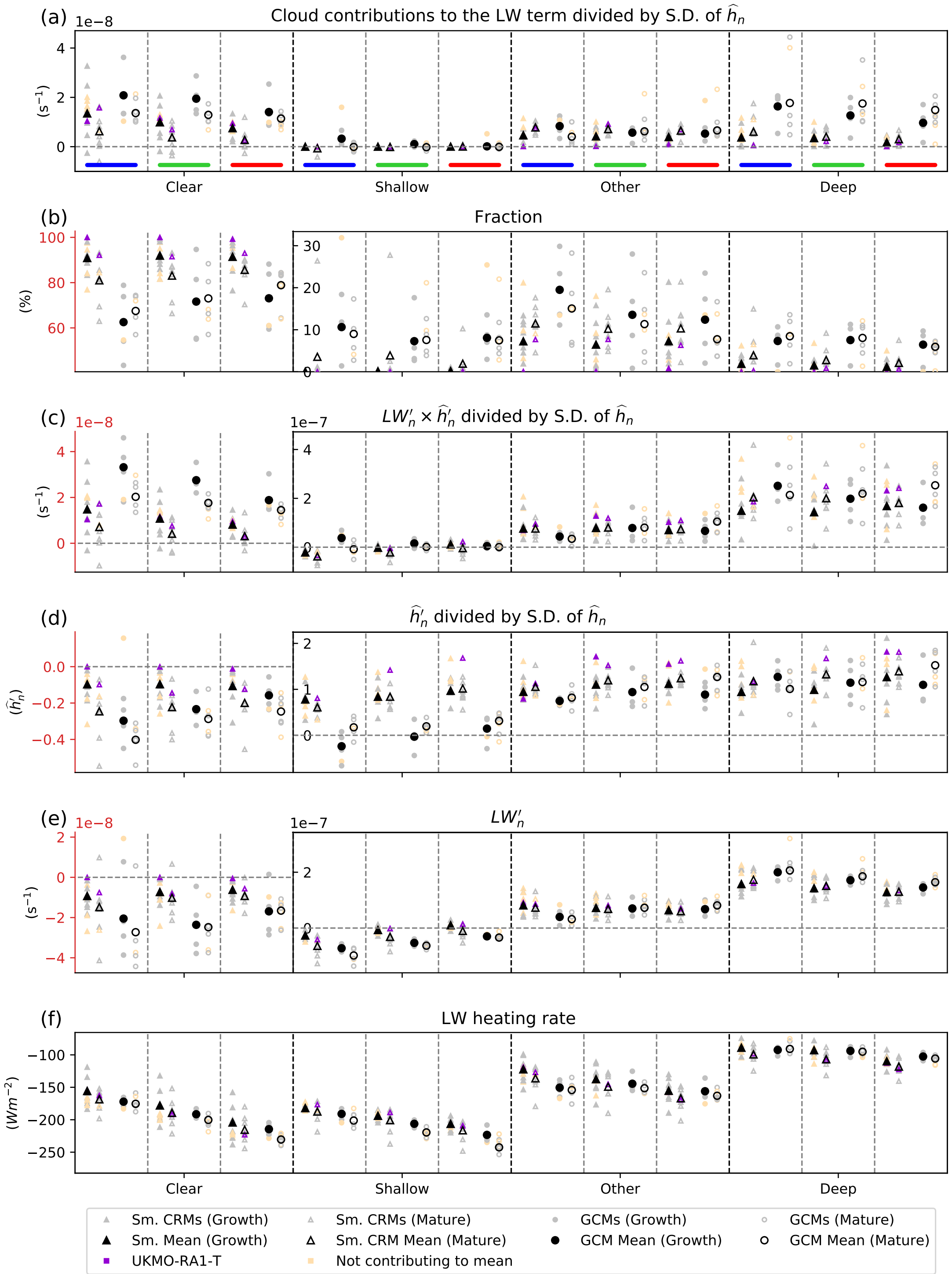

Figure 8. Smoothed CRMs (upward triangles) vs GCMs (circles): (a) Contributions of longwave interactions for each cloud type to the longwave term in equation 3 divided by the standard deviation of $\widehat{h}_n$, (b) Fraction of each cloud type, (c) $LW'_n \times \widehat{h}'_n$ covariance divided by the standard deviation of $\widehat{h}_n$, (d) $\widehat{h}'_n$ divided by the standard deviation of $\widehat{h}_n$, (e) LW'_n, and (f) absolute longwave heating. Data points and layout follow the same protocol as in Figure 5. Note different y-axis ranges for Clear in b, c, d & e.

The absolute longwave heating rate of Deep cloud is similar in the Smoothed CRMs and GCMs, but in the Clear regions, the longwave heating rate is more negative on average for GCMs (Figure 8f). Clear regions occupy the majority of the domain, meaning the domain-mean longwave emission is closely linked to that of the Clear regions. This makes the Deep clouds in GCMs have a more positive LW'_n (Figure 8e), helping increase their $LW'_n \times \widehat{h}'_n$ covariance.

The $LW'_n \times \widehat{h}'_n$ covariance of the Clear regions is more than double that of the Smoothed CRMs. This is in part because Clear regions in GCMs typically occur in more negative $\widehat{h}'_n$ compared to Smoothed CRMs (Figure 8d), which is a likely consequence of the greater cloud fraction in GCMs, confining the Clear regions to drier environments. The LW'_n is also more negative in GCMs partially due to the mean absolute longwave heating rates being more negative on average, but mainly because of the difference in cloud fraction between the model types. To isolate the effect of the difference in cloud fraction between CRMs and GCMs on the Clear longwave feedback, we use the mean longwave heating rates of the cloud types in the Smoothed CRMs with the cloud fractions of the GCMs. We then calculate a hypothetical new domain-mean longwave cooling and cloud type LW', and find that the LW'_n of the Clear regions becomes approximately 2.5 times more negative. This is thanks to the Deep clouds lowering the domain-mean longwave cooling rate in GCMs, hence making the Clear regions more anomalously negative. These effects suggest that the greater Deep cloud fraction in GCMs is a key factor in the enhanced total longwave-FMSE feedback, and therefore rate of aggregation in GCMs compared to CRMs. The non-Smoothed CRMs have a similar Deep cloud fraction and Deep $LW'_n \times \widehat{h}'_n$ covariance to the GCMs, yet the contributions from Other and Clear cloud types remain larger in GCMs thanks to the increase in the Other cloud fraction in GCMs. The increase in Other cloud fraction, with their positive LW', helps further lower the (negative) LW' of the Clear regions in GCMs compared to non-Smoothed CRMs, helping increase these cloud types' contributions to the longwave feedback.

As the convection reaches the Mature phase, longwave interactions in the Clear, Other and Deep cloud types maintain aggregation in the Smoothed CRMs. For GCMs, longwave interactions with the Clear and Deep cloud types are the key maintainers of aggregation. Despite the GCMs having a larger Shallow fraction, these clouds have a similarly insignificant contribution to the longwave feedback as in the Smoothed CRMs. Their $LW'_n \times \widehat{h}'_n$ covariance is consistently close to 0 because both their LW'_n and $\widehat{h}'_n$ is small.

The SST sensitivity of the longwave feedback in GCMs is less straightforward than CRMs with multiple factors playing a role. During the Growth phase, the longwave feedback decreases with SST, and this is due to the decrease in the contributions of the Clear and Deep cloud types. This in turn, is mainly due to their decreasing $LW'_n \times \widehat{h}'_n$ covariance since the fractions of these cloud types remain relatively insensitive to SST. The decrease in the Clear covariance with SST is mainly due to the Clear regions occurring in less anomalously-negative $\widehat{h}'_n$ regions. The main factor responsible for the decreasing contribution from Deep cloud is the increase in the range of $\widehat{h}_{\max}$ and $\widehat{h}_{\min}$ that is used to normalize the longwave heating anomalies. During the Mature phase of aggregation, the longwave feedback has little SST sensitivity for GCMs.

In GCMs, the change in the SST dependence of the longwave term from negative during the Growth phase to more neutral after the Growth phase is one of the main factors causing GCMs to be more aggregated at higher SST, since the advection feedback remains less negative with SST throughout the majority of the simulations. For GCMs during the Growth phase, we find a negative SST dependence of the contribution of each cloud type to the longwave feedback. During the Mature phase, these SST dependencies are more positive. The contributions from the Deep and Other clouds have a more positive SST dependence after the Growth phase because their $LW' \times \widehat{h}'_n$ covariance increases with SST (Figure 8c). This is because these clouds form in more anomalously positive $\widehat{h}'_n$ regions as SST increases (Figure 8d).

5 Conclusions

In this study, we compare the effects of cloud-radiation interactions on convective self-aggregation within the CRMs and GCMs submitted to RCEMIP (Wing et al., 2018). We use the normalized vertically-integrated FMSE variance ($\text{var}(\widehat{h}_n)$) budget framework to study aggregation (Pope et al., 2021, referred to as P21.), and define "Growth" and "Mature" phases of aggregation to compare how FMSE feedbacks contribute to aggregation at similar stages of aggregation across the range of models. We define four different cloud types based on the top of atmosphere radiative fluxes following the method from Becker and Wing (2020) and calculate the contribution of radiative interactions with these cloud types to aggregation. These cloud types are: Clear, Shallow, Deep and Other. GCMs have on average a 40 times larger grid spacing than CRMs. When comparing these two model types we account for biases in our analysis technique due to the resolution difference by horizontally smoothing the CRMs so that each grid point is an average of the 40 $\times$ 40 grid points surrounding it, referred to as Smoothed CRMs.

The goals of the study are to:

- Validate the robustness of the results in P21 who studied the effects of cloud-radiation interactions on self-aggregation within the Met Office Unified Model version 11.0 CRM (submitted to RCEMIP and referred to as "UKMO-RA1-T").
- Investigate to what extent differences in cloud-radiation interactions affect self-aggregation within CRMs and GCMs, and how these are sensitive to SST.

5.1 Robustness of P21 results

We consider the robustness of the following five conclusions from P21:

1. Key **drivers** of aggregation are longwave interactions with high-topped clouds and Clear regions. *(Robust)*

   Most CRMs and GCMs are in agreement with this conclusion when considering that Deep cloud are mostly equivalent to high-topped clouds in P21. Deep clouds have strong longwave heating anomalies and occur in anomalously moist regions. Clear regions typically have negative longwave heating anomalies and tend to occur in anomalously dry regions. Both of these radiative interactions result in a strongly positive longwave feedback.

2. The main **maintainers** of aggregation are longwave interactions with high-topped clouds and shortwave interactions with water vapor. *(Robust)*

   Most CRMs and GCMs are in agreement that these radiative interactions are key maintainers of aggregation. The shortwave feedback increases with aggregation as moist and dry regions amplify, leading to a greater contrast in shortwave absorption by water vapor between the moist and dry regions, resulting in an enhanced shortwave-FMSE feedback.

3. The main **resistors** of aggregation are negative surface flux and advection feedbacks. *(Not Robust for surface flux in the Growth phase)*

   In the majority of models, the surface flux feedback is actually a key *driver* of aggregation, with the UKMO-RA1-T model having the most negative surface flux contribution during the Growth phase. In most models, this is likely due to a strong wind speed-induced surface flux feedback outweighing the air-sea disequilibrium feedback during the Growth phase of aggregation (unlike in UKMO-RA1-T where the opposite is true). As aggregation matures, the models are in agreement that the surface flux feedback becomes increasingly negative and often opposes aggregation. The advection feedback is typically negative and highly variable between models.

4. The **SST-dependence** of the longwave feedback is absent during the Growth phase, but is negative in the Mature phase. *(Not Robust for Growth phase)*

   For the RCEMIP models, the domain-mean longwave feedback decreases with SST at *all stages* of aggregation, which is primarily due to the decrease in Deep and/or Other cloud fraction at warmer SSTs. P21 also find the high-topped cloud fraction decreases with SST, however this is compensated by an increase in their mean longwave-FMSE covariance in the Growth phase. We do not find the longwave-FMSE covariance of the Deep and Other clouds increasing with SST in the majority of RCEMIP models, hence their domain mean longwave feedback tends to decrease with SST.

   The RCEMIP CRMs and GCMs differ in the processes leading to the decrease in the longwave feedback with SST. For the CRMs, the average longwave-FMSE covariance of these clouds remains similar with SST, so the decrease in their cloud fraction reduces their total aggregating influence. A secondary effect of the decreased Deep cloud fraction is an increase in the magnitude of domain mean longwave cooling. This makes the typically-negative longwave heating anomalies of the Clear regions less anomalous, also decreasing the Clear regions' aggregating influence at warmer SSTs. In GCMs, the longwave feedback decreases with SST because the normalized longwave heating anomalies of Deep clouds decreases, reducing their aggregating influence. In addition, the Clear regions occur in less anomalously dry regions due to the reduced total cloud fraction, also reducing their average aggregating influence as SST increases.

5. The **SST-dependence** of the aggregation rate is positive because the advection feedback becomes increasingly positive with SST. *(Not Robust)*

   P21 find the sum of the diabatic feedbacks are insensitive to SST during the Growth phase, however for the RCEMIP CRMs and GCMs, each diabatic feedback tends to decrease with SST during the Growth phase. Despite the sum of these diabatic feedbacks decreasing with SST, the rate of aggregation remains similar on average. The sum of the diabatic feedbacks tends to be proportional to the magnitude of the (negative) advection feedback, resulting in no significant change in aggregation rate with SST.

5.2 Differences between GCMs and CRMs

Using var($\widehat{h}_n$) as our aggregation metric, we find there is much variability in the rate of aggregation and the maximum degree of aggregation within the CRMs, with no consistent SST dependence on the rate of aggregation and the maximum degree of aggregation. GCMs, on the other hand, aggregate faster than CRMs on average, and tend to be more aggregated at higher SSTs.

Both the contributions of shortwave-FMSE and surface flux-FMSE feedbacks to aggregation are similar in magnitude in Smoothed CRMs and GCMs. However, the longwave-FMSE feedback is, on average, approximately twice as strong in GCMs compared with CRMs. This results in typically faster rates of aggregation in GCMs. This is primarily due to GCMs having a larger cloud fraction than Smoothed CRMs, but more crucially a larger Deep cloud fraction. However, if GCMs are instead compared to the non-Smoothed CRMs, GCMs have a similar Deep fraction but a larger Other fraction, which still results in a greater total longwave-FMSE feedback. The longwave-FMSE feedback is strongest for Deep clouds because they typically occur in anomalously-high FMSE regions, and have anomalously strong positive longwave heating rates. Like with the SST sensitivity of cloud fraction in CRMs, a secondary effect of the increased Deep cloud fraction in GCMs is an increase in the longwave-FMSE feedback in the Clear regions. This is because an increased cloud fraction reduces the magnitude of domain-mean longwave cooling. With Clear regions occupying the majority of the domain, their typically-negative longwave heating anomalies become more negative, increasing their longwave-FMSE feedback. The increase in the contributions from Deep and Clear regions to the longwave-FMSE feedback accounts for the doubling of the total feedback.

As previously mentioned, the sum of the diabatic feedbacks with FMSE tend to decrease with SST during the Growth phase, yet the aggregation rate remains insensitive to SST thanks to the increasingly positive advection feedback. After the Growth phase however, the sum of the diabatic feedbacks in GCMs becomes less SST dependent, yet the advection feedback remains more positive at higher SSTs, resulting in GCMs being more aggregated at higher SSTs. Their diabatic terms become less SST dependent after the Growth phase in part because the Deep and Other cloud types tend to occur in more anomalously moist environments at higher SSTs, increasing their longwave-FMSE feedback. This finding, and the point made above about differences in cloud amount between GCMs and CRMs, suggests that GCMs should be compared more systematically to CRMs to investigate their total cloud amount, and their tendency to place high-topped clouds in more anomalously moist environments as SSTs increase.

Despite the difference in the diabatic feedbacks between GCMs and CRMs accounting for the difference in the aggregation rate between these model types, there is no evidence that the model spread in the magnitude of the diabatic feedbacks can explain the model spread in the rate of aggregation in CRMs. For CRMs, the model spread in the rate of aggregation is mostly determined by the magnitude of the advection term due to it having the highest inter-model variability compared to the other diabatic terms. The advection term may be largely influenced by circulations induced by strong radiative cooling from low cloud in dry regions that result in an upgradient transport of FMSE, helping aid aggregation (Muller & Held, 2012; Muller & Bony, 2015). This effect is not investigated in this study. Unlike in CRMs, the diabatic feedbacks are significantly correlated with aggregation rate in GCMs, and this may suggest that GCMs are not capturing key circulations that would otherwise mediate aggregation.

We have shown that the production of cloud in CRMs and GCMs, in terms of quantity and distribution, is very different. This in turn, results in largely different longwave-FMSE feedbacks that alter the rate and degree of aggregation. Not only are the longwave-FMSE interactions enhanced in GCMs, but there is a less negative correlation between the diabatic and advection feedbacks in GCMs than CRMs. This suggests that GCMs are not resolving circulations the may otherwise export FMSE away from moist regions, mediating aggregation, as seen in CRMs. These factors highlight our limitations to accurately represent the cloud response to warming in climate studies. CRMs are often used to study the cloud response to warming, but are too small to capture the large-scale circulations that affect the total cloud feedback. GCMs are used in climate modelling studies because they are complete representations of the climate system, and they can perform hundreds of years of global-scale simulations. However, there are discrepancies between cloud-radiation interactions and circulations between GCMs and CRMs.

We might expect that CRMs are better at representing smaller-scale convective processes and circulations, but systematic comparisons of these attributes with observed cases of organised convection, would help us understand the discrepancies between GCMs and CRMs, and might lead to improvements in these simulations.

Acknowledgments

This work was supported by the Natural Environment Research Council SCENARIO DTP (NE/L002566/1). The simulations of the UKMOi-vn11.0-RA1-T model have been produced by Todd Jones, supported by the Natural Environment Research Council (NERC) under the joint NERC/Met Office ParaCon program's Circle-A project (NE/N013735/1), as well as the ParaCon Phase 2 project: Understanding and Representing Atmospheric Convection across Scales (NE/T003871/1). The simulations have been conducted using Monsoon2, a High Performance Computing facility funded by the Met Office and NERC, the NEXCS High Performance Computing facility funded by NERC and delivered by the Met Office, and JASMIN, the UK collaborative data analysis facility. We thank the German Climate Computing Center (DKRZ) for hosting the standardized RCEMIP data, which is publicly available at http://hdl.handle.net/21.14101/d4beee8e-6996-453e-bbd1-ff53b6874c0e. All data used for plotting each figure, as well as the original python scripts are available on Zenodo at: https://doi.org/10.5281/zenodo.6869033

References

Arnold, N. P., & Randall, D. A. (2015). Global-scale convective aggregation: Implications for the Madden-Julian Oscillation. *Journal of Advances in Modeling Earth Systems*, *7*, 1499–1518. doi: 10.1002/2015MS000498

Becker, T., & Wing, A., Allison. (2020). Understanding the Extreme Spread in Climate Sensitivity within the Radiative-Convective Equilibrium Model Intercomparison Project. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002165. doi: 10.1029/2020MS002165

Bony, S., Stevens, B., Coppin, D., Becker, T., Reed, K. A., Voigt, A., & Medeiros, B. (2016). Thermodynamic Control of Anvil Cloud Amount. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(32), 8927–8932. doi: 10.1073/pnas.1601472113

Bretherton, C. S., Blossey, P. N., & Khairoutdinov, M. (2005). An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST. *Journal of the Atmospheric Sciences*, *62*(12), 4273–4292. doi: 10.1175/JAS3614.1

Coppin, D., & Bony, S. (2015). Physical mechanisms controlling the initiation of convective self-aggregation in a General Circulation Model. *Journal of Advances in Modeling Earth Systems*, *7*, 2060–2078. doi: doi:10.1002/2015MS000571

Holloway, C. E., Wing, A. A., Bony, S., Muller, C., Masunaga, H., L'Ecuyer, T. S., ... Zuidema, P. (2017). Observing Convective Aggregation. *Surveys in Geophysics*, *38*(6), 1199–1236. doi: 10.1007/s10712-017-9419-1

Jeevanjee, N., & Romps, D. M. (2013). Convective self-aggregation, cold pools, and domain size. *Geophysical Research Letters*, *40*(5), 994-998. doi: 10.1002/grl.50204

Muller, C., & Bony, S. (2015). What favors convective aggregation and why? *Geophysical Research Letters*, *42*(13), 5626–5634. doi: 10.1002/2015GL064260

Muller, C., & Held, I. M. (2012). Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences*, *69*(8), 2551–2565. doi: 10.1175/JAS-D-11-0257.1

Nolan, D., Rappin, E., & Emanuel, K. (2007). Tropical cyclogenesis sensitivity to environmental parameters in radiative-convective equilibrium. *Quarterly Journal of the Royal Meteorological Society*, *133*(629 B), 2085–2107. doi: 10.1002/qj.170

Pope, K. N., Holloway, C. E., Jones, T. R., & Stein, T. H. M. (2021). Cloud-radiation interactions and their contributions to convective self-aggregation. *Journal of Advances in Modeling Earth Systems*, *13*(9), e2021MS002535. doi: https://doi.org/10.1029/2021MS002535

Raymond, D. J., & Fuchs, Z. (2009). Moisture modes and the madden–julian oscillation. *Journal of Climate*, *22*(11), 3031 - 3046. doi: 10.1175/2008JCLI2739.1

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., ... Zelinka, M. D. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*, e2019RG000678. doi: https://doi.org/10.1029/2019RG000678

Tobin, I., Bony, S., Holloway, C. E., Grandpeix, J.-Y., Sèze, G., Coppin, D., ... Roca, R. (2013). Does convective aggregation need to be represented in cumulus parameterizations? *Journal of Advances in Modeling Earth Systems*, *5*(4), 692–703. doi: 10.1002/jame.20047

Wing, A. A., & Cronin, T. W. (2016). Self-aggregation of convection in long channel geometry. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 1–15. doi: 10.1002/qj.2628

Wing, A. A., Emanuel, K., Holloway, C. E., & Muller, C. (2017). Convective Self-Aggregation in Numerical Simulations: A Review. *Surveys in Geophysics*, *38*(6), 1173–1197. doi: 10.1007/s10712-017-9408-4

Wing, A. A., & Emanuel, K. A. (2014). Physical Mechanisms Controlling Self-Aggregation of Convection in Idealized Numerical Modeling Simulations. *Journal of Advances in Modeling Earth Systems*, *6*(1), 59–74. doi: 10.1002/2013MS000269

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., & Ohno, T. (2018). Radiative-Convective Equilibrium Model Intercomparison Project. *Geoscientific Model Development*, *11*(2), 793–813. doi: 10.5194/gmd-11-793-2018

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M.-s., Arnold, N., ... Silvers, L. (2020). Clouds and Convective Self-Aggregation in a Multi-Model Ensemble of Radiative-Convective Equilibrium Simulations. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2020MS0021380. doi: 10.1029/2020MS0021380,

Yang, D. (2018a). Boundary layer diabatic processes, the virtual effect, and convective self-aggregation. *Journal of Advances in Modeling Earth Systems*, *10*(9), 2163-2176.

Yang, D. (2018b). Boundary layer height and buoyancy determine the horizontal scale of convective self-aggregation. *Journal of the Atmospheric Sciences*, *75*(2), 469 - 478. doi: 10.1175/JAS-D-17-0150.1