

1 ~~PB: Large-Scale Clustering of Tropical Precipitation and its Implications for the Radiation~~
2 ~~Budget across Timescales~~ Can Large-Scale Clustering of Tropical
3 Precipitation Be Used to Constrain Climate
4 Sensitivity?

5 **P. Blackberg^{1,2}, M. S. Singh^{1,2}**

6 ¹School of Earth, Atmosphere, and Environment, Monash University, Victoria, Australia
7 ²Centre of Excellence for the Weather of the 21st Century, Monash University, Victoria, Australia

8 **Abstract**

9 The spatial organization of deep convection in tropical regions is posited to play an im-
 10 portant role in determining characteristics of the tropical climate such as the humidity
 11 distribution and cloudiness and may therefore be an important control on climate feed-
 12 backs. This study analyzes one aspect of convective organization, the clustering of heavy
 13 precipitation on large scales, in both interannual variability and under warming in fu-
 14 ture climate projections. *PB:* Clustering is quantified using the top 5% heaviest daily
 15 precipitation instances, and interpreted as increasing with the area covered by heavy
 16 precipitation (C , temporally clustered) and proximity of heavily rainy points to each
 17 other and key zonal and meridional reference lines (P , spatially clustered and or-
 18 thogonal to coverage). *PB:* Both observations and global climate models indicate that large-scale
 19 clustering is sensitive to the SST gradient in the Pacific, being largest during El Niño events. Under
 20 future warming, models project an increase in clustering with a large intermodel spread. The increase
 21 is associated with a narrowing of the intertropical convergence zone, while the model spread is par-
 22 tially explained by differences in projections of the SST gradient in the Pacific. Both observations
 23 and models indicate large-scale clustering influences the cloud and humidity distributions, albeit with
 24 some differences. However, the intermodel spread in changes in clustering with warming is not related
 25 to the intermodel spread in projections of tropical-mean relative humidity or low cloudiness in regions
 26 of descent, precluding attempts to provide an observational constraint on feedbacks or climate sensi-
 27 tivity. Nevertheless, the tendency for a meridional contraction of precipitation explains about 45% of
 28 the variance in projected drying, highlighting the narrowing of the ITCZ as an important aspect of
 29 large-scale convective organization in a warmer climate. Both observations and global climate
 30 models (GCMs) indicate temporal clustering of heavy rainfall is best described by the
 31 tropical-mean precipitation rate ($r \sim 0.55 - 0.85$) and spatial clustering is favored by
 32 El Niño conditions, but importantly diverge in the tropical-mean relative humidity
 33 and low cloudiness associated with spatially clustered states in deseasonalized inter-
 34 annual variability. Under future warming, all models from CMIP6 ensemble project
 35 an increase in clustering associated with a narrowing of the intertropical convergence
 36 zone and with a model-spread partially explained by differences in projections of the
 37 Pacific SST gradient ($r=-0.46$). However, unlike variability, changes in general spatial
 38 clustering with warming do not explain the climatological humidity or low-cloud re-
 39 sponse, limiting a simple observational constraint on feedbacks. Notably a few model
 40 outliers drive considerable spread in tropical-mean drying ($r=0.44$) linked to a clima-
 41 tological meridional clustering of heavy rainfall, suggesting potential for future model
 42 evaluation.

43 **Plain Language Summary**

44 The spatial distribution of rainfall in the tropics is expected to change in a warm-
 45 ing climate, with potentially important impacts on how much radiation is absorbed by
 46 water vapor and reflected by clouds. This study shows that heavy rainfall tends to move
 47 towards the equator and to the Pacific Ocean in projections with global climate mod-
 48 els, resulting in an overall increase in the "clustering" of rainfall on the large scale. Fur-
 49 ther, the results show a shift in rainfall to the equator with global warming is associated
 50 with a drying of the tropical atmosphere, which may have an influence on how much the
 51 planet warms for a given CO_2 change. However, similar observed shifts in rainfall in the
 52 current climate are not found to have the same effect on humidity and clouds as for changes
 53 with warming, suggesting caution should be exercised when using relationships derived
 54 from observations to predict future changes.

55 **1 Introduction**

56 The spatial organization of deep convection in tropical regions plays a critical role
 57 in shaping the hydrological cycle and the moisture and cloud distribution (?). Changes
 58 in organization with warming may therefore have implications for a range of climatic pro-
 59 cesses, including precipitation extremes (e.g., ???) and the radiative feedbacks that con-
 60 trol equilibrium climate sensitivity (ECS) (e.g., ???). However, because many of the rel-
 61 evant small-scale processes are not resolved in climate models, it remains unclear how
 62 convective organization will evolve in a warmer climate.

63 While there are numerous ways by which convection may organize, one important
 64 mechanism is the clumping or clustering together of convective elements (e.g., ???). Such
 65 clustering occurs on a range of scales (?), including at large scales that are resolved by
 66 climate models and at mesoscales that can typically only be resolved in high-resolution
 67 storm-resolving simulations. Recently, ? showed that the extent to which tropical pre-
 68 cipitation exhibits clustering on the large scale increases with warming in climate pro-
 69 jections from the Coupled Model Intercomparison Project phase 5 (CMIP5). This large-
 70 scale clustering is distinct from other types of organization on the mesoscale, but ide-
 71 alized simulations suggest that similar processes may act at both scales, and that both
 72 large-scale and mesoscale organization of convection may modulate the radiation bud-
 73 get (?).

74 Here we build on the work of ?, showing that increased clustering of heavy precip-
 75 itation with warming is a robust feature of the more recent CMIP6 as well as CMIP5.
 76 Further, we explore the mechanisms that lead to large-scale clustering of precipitation
 77 in the tropics and the influence of an increase in clustering on properties of the atmo-
 78 sphere that are important for the radiation budget. The analysis will compare how clus-
 79 tering varies across different timescales, from interannual variability in both models and
 80 observations to changes in the climatological clustering of convection in a warming cli-
 81 mate. This approach allows us to assess whether observational constraints of convective
 82 organization under current climate conditions can help constrain changes in organiza-
 83 tion and the associated radiative feedbacks with warming.

84 Previous studies have highlighted observed relationships between convective organi-
 85 zation and the tropical radiation budget (???). For example, ? find tropical mesoscale
 86 convective organization and Estimated Inversion Strength (EIS) in subsidence regions
 87 are the two strongest predictors of deseasonalized [*PB: monthly anomalies*](#)[*interannual vari-*](#)
 88 [*ability*](#) in net top-of-atmosphere radiation, together explaining about 60 percent of the
 89 variance. While the two predictors are significantly correlated and potentially partly mech-
 90 anistically connected (?), the authors find that both have an independent contribution
 91 in influencing the tropical radiation budget; EIS is found to have a stronger correlation
 92 with the cloudy component of the radiation budget while convective organization is found
 93 to have a stronger connection to the clear-sky component of the radiation budget. ? ar-
 94 gue that clustering of deep convective elements is associated with a tropical-mean dry-
 95 ing, resulting in increased outgoing longwave radiation due to a reduction in the green-
 96 house effect.

97 This hypothesis is supported by idealized studies of convective “self-aggregation”
 98 (e.g., ?). Both cloud-resolving and climate-model simulations run in idealized settings
 99 reminiscent of tropical conditions (i.e., low rotation rate and weak temperature gradi-
 100 ents) show increased outgoing longwave radiation when convection is more clustered within
 101 the domain (?).

102 The preceding studies suggest that increased clustering of convection with warm-
 103 ing may lead to a negative clear-sky feedback from clustering-induced drying, resulting
 104 in reduced equilibrium climate sensitivity (ECS) (?). However, recent research suggests
 105 that clustering of deep convection on the large scale may also be indirectly connected

106 to a positive shortwave feedback on warming through changes in low clouds (?). According
 107 to this argument, drying associated with increased clustering of convection is con-
 108 trolled by the large-scale overturning circulation and is most pronounced in regions of
 109 climatological descent where low-cloudiness is sensitive to changes in relative humidity.
 110 The associated cloud changes then lead to a net positive feedback.

111 As the above discussion highlights, an important control on convective organization
 112 at both large- and mesoscales comes from large-scale circulation patterns, includ-
 113 ing the Intertropical Convergence Zone (ITCZ), the South Pacific Convergence Zone (SPCZ),
 114 the Walker circulation, and convectively-coupled tropical waves (?????). Changes to the
 115 clustering of precipitation under warming may therefore be linked to, for example, a “nar-
 116 rowing” of the ITCZ due to constraints on the export of energy by the Hadley cell (?)
 117 or changes in the Walker circulation driven by changes in zonal SST gradients (?), which
 118 may be associated with “El Niño-like” shifts in the SST climatology (?).

119 *PB:* ~~In this study we explore how the spatial distribution of heavy precipitation changes with~~
 120 ~~increased large-scale clustering, demonstrating the importance of both meridional contractions and~~
 121 ~~zonal shifts in convection. We further show that the effect of large-scale clustering on low clouds and~~
 122 ~~moisture—key variables that control the radiation budget—varies with timescale and between global~~
 123 ~~climate models and observations, making it difficult to apply constraints from observed variability to a~~
 124 ~~warmer climate. The paper is structured as follows. First we describe the datasets used and our quan-~~
 125 ~~tification of large-scale clustering of precipitation (Section 2). Then we present the geographical spatial~~
 126 ~~patterns of heavy precipitation that are associated with a high level of clustering (Section 3). After~~
 127 ~~that, we connect the spatial patterns to leading mechanisms driving clustering (Section 4). Finally,~~
 128 ~~we present the moisture and low cloud distribution associated with clustering (Section 5). Section 6~~
 129 ~~gives a summary of the key findings and an outlook for future research. The existing literature~~
 130 ~~show large-scale clustering is projected to increase from the forced response of global~~
 131 ~~warming, and that a more clustered state in observed interannual variability, albeit~~
 132 ~~on a different spatial scale, is associated with clear-sky and cloud-radiative feedbacks~~
 133 ~~connected to climate sensitivity. In this study we elucidate to what extent observed~~
 134 ~~and modelled interannual variability in large-scale clustering of heavy precipitation~~
 135 ~~connects to similar radiative feedbacks and can be used to constrain climate sensitiv-~~
 136 ~~ity. To do so, we quantify and relate clustering of heavy rainfall to the distribution~~
 137 ~~and tropical-mean relative humidity and low-cloudiness in observations, a high-resolu-~~
 138 ~~tion GCM from the NextGEMS pre-final cycle, and GCMs from the CMIP6 ensemble~~
 139 ~~(method described in Section 2). The results of this study adds to the existing litera-~~
 140 ~~ture by describing the spatial preference of heavy rainfall that promote large-scale~~
 141 ~~clustering (Section 3), the associated mechanisms involving SST patterns and a nar-~~
 142 ~~rowing of the ITCZ (Section 4), and the tropical radiative feedbacks associated with~~
 143 ~~a more clustered state (Section 5) in interannual variability and for projected climati-~~
 144 ~~ological changes with warming. Section 6 gives a summary of the key findings and an~~
 145 ~~outlook for future research.~~

146 **2 Data and Methods**

147 Our analysis is focused on variations in the large-scale clustering of heavy precip-
 148 itation in the tropics and its relationship to the atmospheric state in both observations
 149 and an ensemble of global climate models (GCMs) primarily from CMIP6. We begin by
 150 describing the datasets (both model and observational) used, before we describe the quan-
 151 tification of large-scale clustering, and our analysis framework.

152 **2.1 Models**

153 We use simulations from 27 GCMs from CMIP6 (?), using data from the years 1970-
 154 1999 in the historical scenario, representing the current climate, and from the years 2070-
 155 2099 under the Shared Socioeconomic Pathway 585 (SSP5-8.5), representing a warmer
 156 climate. The models are chosen based on availability of the required variables and are
 157 shown in Figure 6. We use one ensemble member from each model.

158 In addition to the CMIP6 models, we also consider a simulation using a high-resolution
 159 GCM referred to here as IFS_9_FESOM_5 (?). The Deutsches Klimarechenzentrum (DKRZ)
 160 Next Generation Earth Modelling Systems (NextGEMS) pre-final cycle provides high-
 161 resolution globally simulated atmospheric and oceanic variables for SSP3-7.0 forcing be-
 162 tween 2025-2049 using the ECMWF Integrated Forecasting System (IFS) at ~ 9 km hor-
 163 izontal grid spacing for the atmosphere and the Finite-VolumE Sea Ice-Ocean model ver-
 164 sion 2 (FESOM2) at 5 km horizontal grid spacing for the ocean (?). Although the model
 165 is at high resolution compared to the CMIP6 models, it retains a convection parameter-
 166 ization, and we therefore describe it as a GCM rather than a storm-resolving model. Be-
 167 cause the climate change signal during the simulation is small, we only use the high-resolution
 168 GCM to characterize interannual variability, using all available years.

169 **2.2 Observations**

170 Observed clustering of tropical precipitation is quantified based on daily precip-
 171 itation estimates from the National Oceanic and Atmospheric Administration Global Pre-
 172 cipitation Climatology Project (NOAA-GPCP) (?), using the method described in the
 173 next subsection. We further use NOAA-GlobalTemp (?), and Clouds and the Earth's
 174 Radiant Energy System (CERES) data (?) to provide observational estimates of surface
 175 temperature and PB : outgoing longwave radiation Top of the atmosphere radiative fluxes, re-
 176 spectively. Estimates of vertical pressure velocity and specific and relative humidity are
 177 taken from the fifth generation of the European Centre for Medium-Range Weather Fore-
 178 casts reanalysis (ERA5) (?). Apart from precipitation, all variables are taken as monthly
 179 averages.

180 Finally, we develop a simple estimate of the low-cloud fraction using the tropical
 181 weather states defined in ? based on data from the International Satellite Cloud Clima-
 182 tology Project (ISCCP) (?). ? used a clustering algorithm to categorize histograms of
 183 cloud-top pressure and optical thickness given by the ISCCP D1 dataset into a series of
 184 weather states defined in three hourly polar-orbiting satellite scans with daily global cov-
 185 erage on a $1^\circ \times 1^\circ$ grid. Each weather state is characterized by a histogram in cloud-
 186 top pressure and optical thickness that represents the centroid over all members of that
 187 weather state. Here we estimate the cloud fraction as a function of pressure for a given
 188 weather state as the total frequency of clouds of all optical thicknesses in a given range
 189 of cloud-top pressure within the corresponding centroid histogram. We then calculate
 190 the low-cloud fraction LCF_i of weather state i as the total cloud fraction below 600 hPa.
 191 The monthly low-cloud fraction is taken as

$$LCF = \sum_i f_i LCF_i, \quad (1)$$

192 where f_i is the frequency of weather state i over the month in question.

193 As described further below, all observational datasets are regridded conservatively
 194 to a common $2.8^\circ \times 2.8^\circ$ grid for analysis. We use observations covering the time pe-
 195 riod between 1998-2023 for all datasets, except for cloud fraction (based on ISCCP), which
 196 is limited to 1998-2017.

197 2.3 Quantifying Large-Scale Clustering of Heavy Precipitation

198 We quantify clustering of precipitation following ? using daily surface precipita-
 199 tion in the tropics (30°S - 30°N). To facilitate the comparison of clustering across differ-
 200 ent models and the observations, we first interpolate the daily precipitation to a $2.8^\circ \times$
 201 2.8° grid using a first-order conservative method (?) to preserve tropical-mean proper-
 202 ties from the native grid. Next we define heavily precipitating regions as gridboxes for
 203 which the precipitation rate exceeds a threshold. The threshold is calculated as the 95th
 204 spatial percentile of daily precipitation over all gridboxes in the tropics temporally av-
 205 eraged over the 30-year climatology (or 25-year in the case of observations *PB:and the*
 206 *high-resolution model*). For the GPCP observations, this threshold is 16 mm day^{-1} *PB:and*
 207 *in models the threshold range between 15-17 mm day}^{-1}*. Distinct heavy precipitation
 208 features are identified as 8-connected contiguous regions of precipitation exceeding the
 209 threshold or single grid boxes if there are no neighboring connections. *PB:While there*
 210 *are conceptual differences in what type of rainfall is represented by different percentile*
 211 *thresholds, overall conclusions of the paper is not sensitive to using the 90th, 95th, or*
 212 *97th percentile.*

213 We define our primary measure of clustering, A_m , as the mean area of heavy pre-
 214 cipitation features over the entire tropics. A_m conceptually captures clustering by dis-
 215 tinguishing scenes with many small precipitation features and scenes where precipita-
 216 tion is aggregated into fewer and/or larger precipitating features (Figure ??). The mean
 217 area of features, A_m , was chosen due to its interpretability, however, we note that it is
 218 only one aspect of the large-scale organization of precipitation, and A_m does not describe
 219 important spatial characteristics *PB: such as the total area, shape, proximity, gradients of pre-*
 220 *cipitation intensity, and location of precipitation features. such as the total area coverage, prox-*
 221 *imity, location, shape, and gradients of precipitation intensity of precipitation features.*
PB: A number of other measures of large-scale clustering are analyzed and their interrelationship is pre-
 222 *sented in Figures S1-2 in the supporting information. A number of other measures of large-scale*
 223 *clustering are analyzed and the interrelationship of a key subset of metrics is presented*
 224 *in Figures S1-2 in the supporting information.* An important aspect of our method is
 225 that, by definition, regions of heavy precipitation occupy 5% of the domain on average.
 226 Thus when comparing two climates, the *PB:meantime-mean* area *PB:fraction of covered by*
 227 heavy precipitation \bar{C} remains constant. Differences in the mean area of features, A_m ,
 228 between climates are entirely due to a reorganization of precipitation, *PB:and changes in*
 229 *A_m are inversely related to the mean number of precipitation features. where an increase in A_m*
 230 *increases the general proximity of heavily precipitating points to each other, P_g , in that*
 231 *the number of distinct precipitation features (N) is reduced (mathematically directly)*
 232 *related as reciprocals for the same total area coverage, $A_m = C/N$.* However, the
 233 above constraint does not apply to the precipitation distribution during a given month.
 234 Indeed, as we shall see, an important driver of variations in tropical precipitation clus-
 235 tering in interannual variability is the area *PB:fraction coverage* of *PB:heavy* precipitation,
 236 *PB: $A_f C$.* We therefore consider the behavior of both the mean area of heavy precipita-
 237 *tion features, A_m , and the area *PB:fraction coverage* of heavy precipitation, *PB: $A_f C$,* in our*
 238 *analysis below.*

240 2.4 Describing Relationships to Large-Scale Clustering of Heavy Pre- 241 cipitation

242 Having quantified large-scale precipitation clustering, we seek to characterize the
 243 relationships between such clustering and other large-scale climate variables. Specifically,

we consider these relationships for interannual variability in both models and observations, and for changes in climate across the CMIP6 ensemble. Throughout, we define interannual variability in a given variable by deseasonalized monthly anomalies, calculated as the monthly-mean anomaly from the climatology of the associated month after detrending the time series. The trend is estimated by a first-order linear least squares regression of the data at each location from the daily (precipitation-based metrics) or monthly (all other metrics) time series. *PB: However, relationships are consistent regardless of whether the linear trend is removed or not.*

As we will see, for interannual variability, the mean area of features, A_m , and the area *PB: fraction coverage* of precipitation, *PB: $A_f C$* , are strongly correlated with each other and to the large-scale climatic state. *PB: To estimate the individual effect of A_m , we apply the method of Pearson partial correlation. To estimate the individual effect of the general proximity of heavily precipitating points to each other (Figure 3c), P_g , analogous to the climatological change in A_m with warming, we apply the method of Pearson partial correlation* (?). The partial correlation $r(X, Y|Z)$ represents the relationship between variables X and Y after the removal of the effect of Z , and is given by

$$r(X, Y|Z) = \frac{r(X, Y) - r(X, Z)r(Y, Z)}{\sqrt{1 - r^2(X, Z)}\sqrt{1 - r^2(Y, Z)}}, \quad (2)$$

where $r(X, Y)$ is the regular correlation between X and Y . *PB: The significance of partial correlations is evaluated using the standard t-test for partial correlations. For partial correlations and simple linear regression, relationships are considered significant if the correlation coefficient is large relative to its uncertainty, as quantified by a standard two-sided t-test, such that the probability of observing a correlation by chance under the null hypothesis of no correlation is below 5%. However, further robustness testing is applied for model-spread correlations, in which potential outliers are removed before calculating the correlations.*

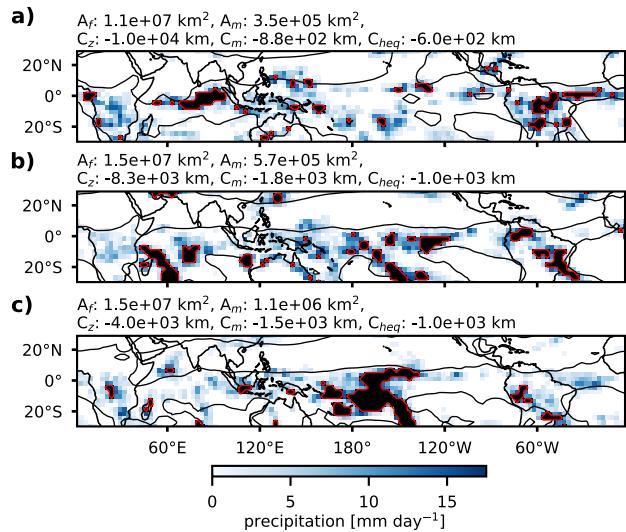


Figure 1. *PB:* February daily snapshots of GPCP precipitation (blue colors) and regions of heavy precipitation (black shading) with monthly contour of ERA5 500 hPa relative humidity representing the median over the tropics (black line) and monthly ISCCP low cloud fraction (red colors). The panel titles show the total area of heavy precipitation as a fraction of the tropical domain area (A_f), the mean area (A_m) and number (N) of heavy precipitation features, and the Oceanic Niño Index (ONI) taken from NOAA Global Temp. From (a-c), clustering increases according to A_m . The clustering from (a-b) is primarily due to an increase in A_f , whereas the clustering from (b-c) is primarily due to the closer proximity of heavy precipitation to the central Pacific. *February daily snapshots of GPCP precipitation (blue colors) and heavy precipitation (black shading) features (red contour), with monthly specific humidity representing the median over the tropics (black contour). The panel titles show the total area coverage of heavy precipitation, C , and the mean area of precipitation features, A_m , with the proximity of heavily precipitation points to the central pacific, equator, and hydrological equator below (described in greater detail in Figure 4).*

268 3 Spatial Patterns of Heavy Rainfall Clustering

269 The purpose of this section is to elucidate the spatial patterns of precipitation that
 270 produce a high degree of clustering. We first consider how clustering changes in inter-
 271 annual variability before we investigate the spatial patterns associated with ^{PB: strong} in-
 272 creases in clustering with warming across the CMIP6 ensemble.

273 Figure ?? shows the regression of monthly anomalies in the frequency of occurrence
 274 of heavy precipitation, ^{PB: C FOO}, onto the mean area of heavy precipitation features,
 275 A_m , for the observations. When tropical precipitation is observed to be highly clustered
 276 on the large scale, heavy precipitation tends to occur more frequently in the central equa-
 277 torial Pacific. Figure ?? is calculated for all months, but similar spatial patterns can be
 278 seen in individual months, strongest in DJF (Figure S ^{PB: 53} in the supporting informa-
 279 tion). Other notable spatial characteristics of the regression include a decrease in heavy
 280 precipitation over the maritime continent and a small but statistically significant (crosses)
 281 decrease in heavy precipitation over the Amazon and Atlantic.

282 ^{PB:} ~~To quantify the shift of heavy precipitation to the central equatorial Pacific, we define the~~
 283 ~~distance metrics C_z , which represents the mean distance of heavily precipitating points within the~~
 284 ~~tropics to the meridian given by the longitude $180^\circ E$ ($-\infty, 0$], and C_m , which represents a similar~~
 285 ~~metric defined based on distance to the equator. To quantify the shift of heavy precipitation~~
 286 ~~to the central equatorial Pacific, we define the proximity metrics P_z , which represents~~
 287 ~~the mean distance of heavily precipitating points within the tropics to the meridian~~
 288 ~~given by the longitude $180^\circ E$, and P_{eq} and P_{heq} , which represents a similar metric~~
 289 ~~defined based on distance to the geographic equator and hydrological equator (Figure~~
 290 ~~4). Here the hydrological equator is defined as the latitude of highest specific humidity~~
 291 ~~at 700 hPa as a function of longitude and time in months, highlighting the location of~~
 292 ~~the “centre of the ITCZ”.~~

293 These are somewhat arbitrary zonal and meridional reference lines with which to
 294 describe the zonal and meridional shifts, and the clustering redistribution of heavy pre-
 295 cipitation may be better characterized relative to the climatological distribution. For ex-
 296 ample, in most models, with warming heavy precipitation moves south relative to the
 297 northern hemisphere climatological convergence zone or north relative the climatolog-
 298 ical SPCZ or both. However, for simplicity, we use ^{PB: C_z and C_m} P_z , P_m , and P_{heq} de-
 299 fined based on $180^\circ E$ and the equator.

300 As expected from the regression map, there is a strong relationship between the
 301 mean area of features, A_m , and ^{PB: $C_z P_z$} ^{PB: (Figure S2 in the supporting information)}. How-
 302 ever, as shown in Figure ??a, there is also a strong observed relationship in interannual
 303 variability between A_m and the total area ^{PB: fractioncoverage} of heavy precipitation, ^{PB: A_f}
 304 $[r^2(A_f, A_m) \sim 0.7]$ $[r^2(C, A_m) \sim 0.7]$, with greater ^{PB: $A_f C$} favoring greater A_m . This sug-
 305 gests a large part of the observed regression pattern is due to the effect of changes in the
 306 total precipitating area, rather than a pure spatial redistribution of a fixed number of
 307 heavily precipitating points. This complicates the interpretation of increased clustering
 308 in internal variability, since when comparing climates, the mean area ^{PB: fractioncoverage}
 309 of heavy precipitation ^{PB: $A_f C$} remains fixed at ^{PB: 0.05%}. To address this, we estimate
 310 the variations in the distribution of heavy precipitation that contribute to variations in
 311 A_m independent of changes in ^{PB: $A_f C$} using Pearson partial correlation (see Methods).

312 ^{PB:} ~~Both C_z and C_m are negatively correlated with~~ Both P_z and P_{eq}/P_{heq} are pos-
 313 ~~itively correlated with~~ the mean area of heavy precipitation features, A_m , when the
 314 effect of changes in the area ^{PB: fractioncoverage} of heavy precipitation, ^{PB: $A_f C$} , is removed
 315 (star in Figure ??b-c), suggesting that there is a shift of ^{PB: convectionheavy rainfall} to-
 316 toward the equator and the central Pacific when the tropics are highly clustered. For a given
 317 ^{PB: $A_f C$} , this contraction of heavily precipitating regions explains about 5-10 percent of
 318 the remaining variance after the effect of changes in ^{PB: $A_f C$} is removed.

319 Both the CMIP ensemble and the high-resolution model generally show similar spatial
 320 patterns associated with increasing large-scale clustering (Figure S^{PB: 6a-b}4a-b in the
 321 supporting information). All models show significant relationships between A_m and ^{PB:} $A_f C$,
 322 although ^{PB:}generally on average somewhat weaker than in observations (Figure ??b),

323 ~~PB: and most models show significant negative partial correlations between the distance met-~~
 324 ~~ries C_z (24/27 CMIP6 models) and C_m (22/27 CMIP6 models) after removing the effect of changes~~
 325 ~~in A_f and most models show significant positive partial correlations between the~~
 326 ~~proximity metrics P_z , P_{eq} , and P_{heq} with 24, 22, and 21 out of 27 CMIP6 models~~
 327 ~~exhibiting statistically significant relationships after removing the effect of changes~~
 328 ~~in C (Figure 4).~~ That is, as for the observations, most models agree that higher clus-
 329 ~~tering, as measured by a larger mean area ^{PB:}for a given area coverage of heavy pre-~~
 330 ~~cipitation features, is associated with a shift in ^{PB:}convectionheavy precipitation to the~~
 331 ~~central equatorial Pacific. Interestingly, the high-resolution GCM did not show a signif-~~
 332 ~~icant partial correlation with ^{PB:} $C_m P_m$, suggesting clustering in that model is not sen-~~
 333 ~~sitive to the meridional contraction of heavy precipitation to the equator.~~

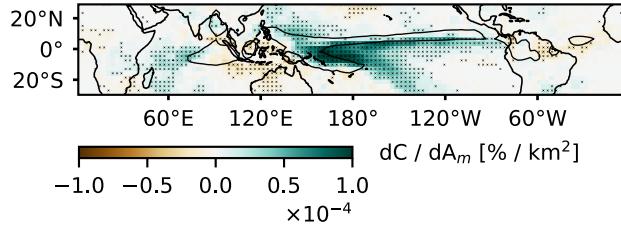


Figure 2. GPCP frequency of occurrence of heavy precipitation, $\text{PB:}\underline{\text{CFOO}}$, regressed onto the mean area of heavy precipitation features, A_m , for interannual variability. The contour shows the 90th percentile of the climatological $\text{PB:}\underline{\text{CFOO}}$ and crosses indicate whether correlations are statistically significant.

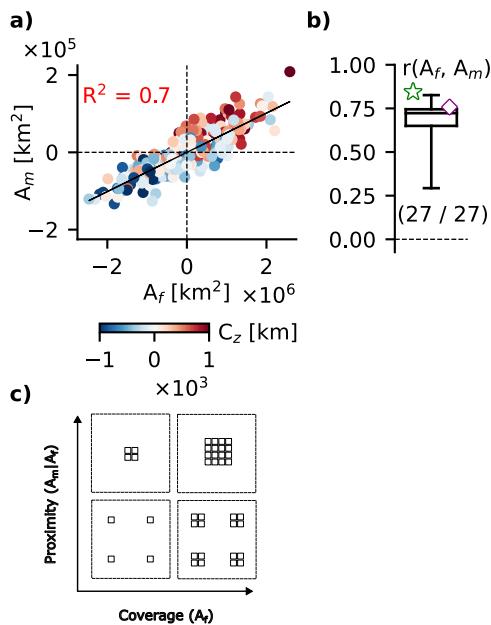


Figure 3. Scatter plot of monthly anomalies in area $\text{PB:}\underline{\text{fractioncoverage}}$ of heavy precipitation, $\text{PB:}\underline{\text{AfC}}$, and the mean area of heavy precipitation features, A_m , colored by the zonal proximity of heavy precipitation to the longitude 180° E in the central pacific, $\text{PB:}\underline{\text{CzP}_z}$, for GPCP observations (a). Boxplot of the correlations between A_f and A_m $\text{PB:}\underline{\text{and partial correlations of Cz and the proximity of heavy precipitation to the equator, Cm, with Am outside the influence of Af}}$ for the CMIP6 ensemble (b). The fraction of CMIP models with statistically significant correlations is indicated below $\text{PB:}\underline{\text{eachthe}}$ box, and the GPCP (star) and high-resolution GCM (diamond) correlations are shown in lighter colors if not statistically significant. $\text{PB:}\underline{\text{A schematic of the change in mean area of heavy precipitation features, A_m, for a given coverage of heavy precipitation, C}}$ (c).

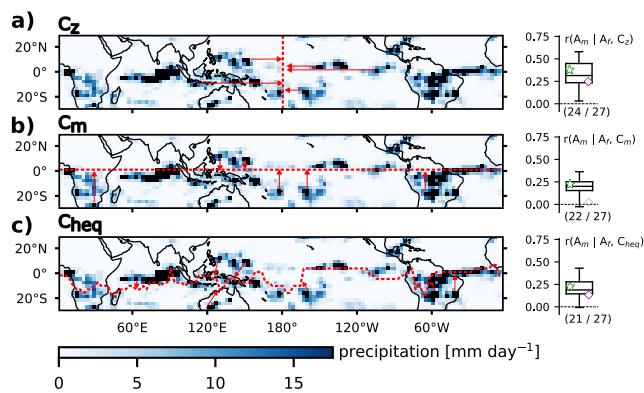


Figure 4. Schematic of metric and boxplot of partial correlations with A_m outside the influence of C , for P_z (a), P_{eq} (b), P_{heq} (c) for the CMIP6 ensemble. The fraction of CMIP models with statistically significant correlations is indicated below each box, and the GPCP (star) and high-resolution GCM (diamond) correlations are shown in lighter colors if not statistically significant.

334 We now consider how large-scale clustering of heavy precipitation changes under
 335 climate change. ? noted that large-scale clustering increases as the climate warms in pro-
 336 jections of climate change by the CMIP5 ensemble. Figure ??a shows that this is also
 337 true for CMIP6; all models project an increase in A_m with warming to varying degrees.
 338 However, there is a wide spread in climatological A_m and wide spread in climatological
 339 increase in A_m across the ensemble, suggesting the degree of large-scale clustering and
 340 the change in large-scale clustering with warming is poorly constrained in the models.

341 Given the climatologically fixed $PB: A_f C$ in this framework, changes in the mean area
 342 of precipitation features, A_m , with warming are driven entirely by a spatial reorganiza-
 343 tion of convection to larger features. The present analysis evaluates whether the spatial
 344 patterns associated with a high degree of clustering of precipitation in interannual vari-
 345 ability may also be relevant to the redistribution of heavy precipitation causing increased
 346 clustering under warming. Indeed, all but one model show $PB: a reduction in C_m with warm-$
 347 $ing an increase in P_{eq}$ with warming, indicating a contraction of heavy precipitation to-
 348 wards the equator (Figure ??b) $PB: Further, and$ all models contract heavy precipitation
 349 towards the hydrological equator, or "ITCZ center". $PB: which is defined here as the lati-$
 350 $ture of highest specific humidity at 700 \text{ hPa as a function of longitude and time in months}$ (Figure 5c).
 351 One possible interpretation of this is that a narrowing of the ITCZ provides a mecha-
 352 nism for the overall increase in clustering with warming. However, we note that the change
 353 in $PB: C_m P_{eq}$ and P_{heq} with warming is uncorrelated with the change in A_m across the
 354 CMIP6 ensemble; models exhibiting stronger ITCZ narrowing with warming do not show
 355 stronger increases in large-scale clustering of precipitation. On the other hand, there is
 356 a significant correlation between increases in mean area of precipitation features, A_m ,
 357 and changes in $PB: C_z P_z$ across the CMIP6 ensemble. That is, models that show a greater
 358 clustering with warming also show a $PB: more zonal shift zonally smaller expansion / greater$
 359 $contraction$ of heavy precipitation to $PB: / from$ the central Pacific (note that the zonal and
 360 meridional contractions are somewhat anticorrelated, as also identified by ?).

361 Going beyond the simple distance metrics, Figure ?? regresses the projected in-
 362 crease in frequency of occurrence of heavy precipitation onto the projected increase in
 363 A_m across the CMIP ensemble. This reveals a spatial pattern of precipitation changes
 364 with several similarities to the pattern for interannual variability shown in Figure ??.
 365 However, an important distinction is that, for changes with warming, the redistribution
 366 of heavy precipitation is a conserved property, due to the climatologically fixed $PB: A_f C$.
 367 Similarities between the regression patterns include a regression coefficient largest in the
 368 central Pacific close to the equator, and a redistribution of precipitation away from the
 369 maritime continent, Amazon, and Atlantic. While under interannual variability heavy
 370 precipitation shifts southward in the Pacific for highly clustered states, the changes with
 371 warming show a northward shift of heavy precipitation in the Pacific.

372 We note that the spatial patterns associated with a climatologically high degree
 373 of clustering of heavy precipitation across the CMIP6 ensemble $PB: (\text{Figure S5 in the}$
 374 $\text{supporting information})$ are rather different from those of internal variability and changes
 375 with warming $PB: (\text{Figure S6c in the supporting information})$. Models with high climatologi-
 376 cal values of A_m tend to have more convection in the warm pool, over tropical continents,
 377 and at the edges of the northern and southern convergence zones over the Pacific Ocean.
 378 These relationships cannot be summarized by a single value of either $PB: C_z$ or $C_m P_z$, P_{eq} ,
 379 or P_{heq} . Rather, we find that high climatological values of A_m are associated with high
 380 month-to-month variability in $PB: A_f C$. This indicates that high $PB: \text{spatial climatological}$
 381 clustering $PB: \text{also}$ corresponds to high clustering in time, with some days producing a large
 382 $PB: \text{amount of area covered by}$ heavy precipitation across the tropics and other days pro-
 383 ducing much $PB: \text{less smaller area coverage}$.

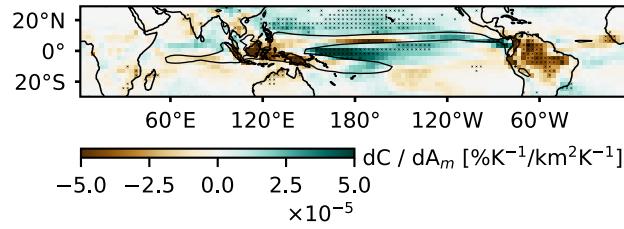


Figure 5. Scatter plot of change in climatological mean area of heavy precipitation features, A_m , with change in climatological mean distance of heavy precipitation to the central pacific, $PB:C_z P_z$ (a) and change in mean heavy precipitation proximity to the equator, $PB:C_m P_m$ (b). Boxplot of change in climatological mean distance to the hydrological equator, C_{heq} , (c) where the hydrological equator is defined as the latitude of highest specific humidity at 700 hPa as a function of longitude and time in months. All quantities are normalized by the tropical surface temperature warming $PB:(\text{land and ocean})$ from the historical to the SSP585 scenario simulation period in CMIP6 models.

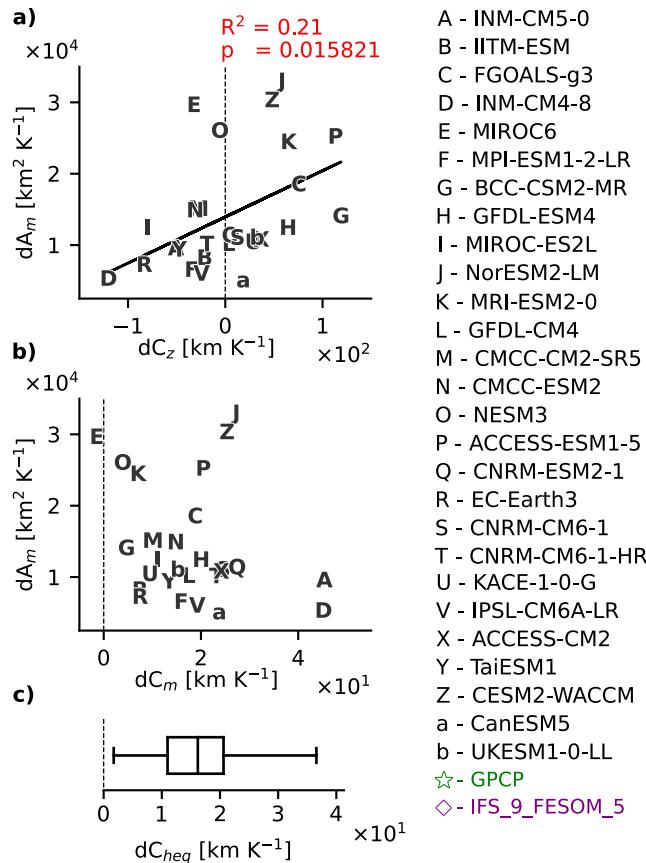


Figure 6. Increase in frequency of occurrence of heavy precipitation, PB:GFOO , regressed onto increase in A_m per kelvin tropical warming from the historical to the SSP585 scenario simulation period across the CMIP6 ensemble. Contour shows the ensemble-mean 90th percentile of PB:GFOO in the historical period and crosses indicate whether correlations are statistically significant.

In summary, the spatial patterns of heavy precipitation associated with highly clustered states vary across timescales, but there are important common threads. In both internal variability and for changes with warming, higher clustering as measured by A_m is associated with more heavy precipitation in the central equatorial Pacific. In particular, models with stronger increases in large-scale clustering of precipitation under warming also exhibit ^{PB:}[a relatively greater zonal shift](#)^{PB:}[in convection to the central Pacific.](#) This potentially suggests the Walker circulation, and the East-West SST gradient in the Pacific, as an important control on the magnitude of changes in large-scale clustering of precipitation with warming. ^{PB:}[We next investigate the relationships between SST changes and changes in \$A_m\$](#) [We next investigate the mechanisms associated with large-scale spatial clustering, including SST changes, cloud-radiative feedbacks, and a narrowing of the ITCZ](#) in both internal variability and under climate change.

396 4 ~~PB:SST Drivers of Heavy Rainfall Clustering~~Mechanisms associated with
397 Heavy Rainfall Spatial Clustering

398 This section investigates the extent to which changes in the El Niño-Southern Os-
399 cillation can explain changes in clustering ~~across timescales~~, whether top of the atmo-

400 sphere radiative fluxes can be argued to feedback on the highly clustered states, and
401 if proximity measures defined in section 2 are related to the area of ascent.

402 We use ~~PB:~~the Southern Oscillation Index (SOI)

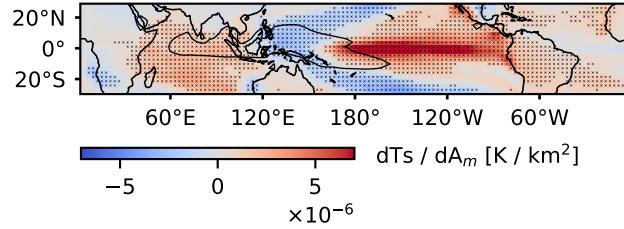


Figure 7. NOAA-GlobalTemp surface temperature, T_s , regressed onto A_m for interannual variability. The contour shows the climatological 90th percentile of T_s , and crosses indicate whether correlations are statistically significant.

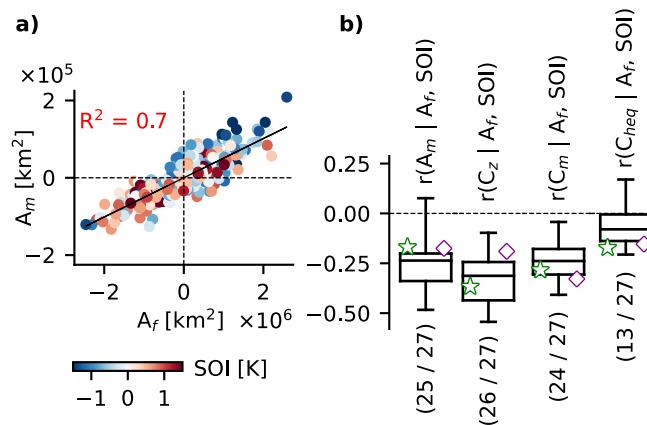


Figure 8. Same as Figure 3, but with the Oceanic Niño Index (ONI) in scatter colors and as explanatory variable in boxplot.

The strong connection between El Niño conditions and A_m in interannual variability motivates the investigation of changes in the climatological Pacific SST gradient to a more El Niño-like state as a mechanism for explaining model-spread in projected changes to clustering. Consistent with expectations, the magnitude of the weakening of the East-West Pacific SST gradient explains a similar amount of variance in projected changes in clustering as the zonal shift in heavy precipitation (Figure ??). Models that have more El Niño-like warming patterns tend to exhibit larger increases in large-scale clustering of precipitation. Regressing the SST changes against projected changes in the mean area of precipitation features, A_m , also shows an El-Niño-like pattern, with a relative warming in the east and relative cooling in the west (Figure ??). In addition, the regression pattern has a noticeable north-south gradient, consistent with the positive regression coefficients for heavy precipitation frequency north of the equator in Figure ??.

We have shown that El Niño-like states tend to result in a higher degree of clustering in both interannual variability and across the CMIP6 ensemble under climate change. Note, however, that there are changes in the zonal SST gradient, T_z , and the mean distance of heavy precipitation to the central Pacific of both signs across the 27 CMIP6 models we analyse, suggesting that zonal shifts in convection are not the primary reason for the ensemble-mean increase in large-scale clustering with warming that we document. We hypothesize that the ensemble-mean increase in A_m is instead associated with a meridional shift in convection, potentially related to a narrowing of the ITCZ (?). All but one model exhibit negative changes in C_m with warming, and this is associated with an increase in the large-scale clustering of precipitation in natural variability.

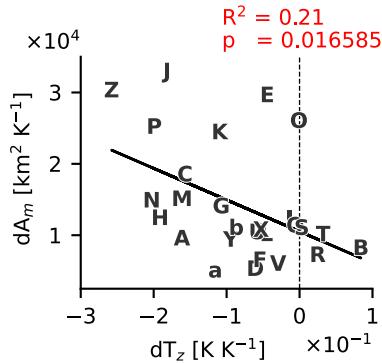


Figure 9. Same as Figure 4a, but with the change in the climatological Pacific SST gradient, T_z , per Kelvin tropical warming from the historical to the SSP585 scenario simulation period across the CMIP6 ensemble as explanatory variable. Models are as given in the legend in Figure 4.

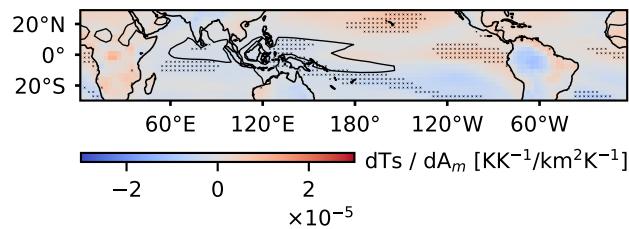


Figure 10. Change in surface temperature, T_s , regressed onto change in mean area of heavy precipitation features, A_m , per Kelvin tropical warming from the historical to the SSP585 scenario simulation period across the CMIP6 ensemble. Contour shows ensemble-mean 90th percentile climatological T_s and crosses indicate if correlations are statistically significant.

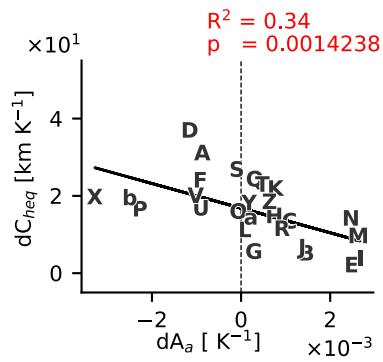


Figure 11. Same as Figure 4a, but with the change in the climatological Pacific SST gradient, T_z , per Kelvin tropical warming from the historical to the SSP585 scenario simulation period across the CMIP6 ensemble as explanatory variable. Models are as given in the legend in Figure 4.

427 **5 The Effect of Heavy Rainfall Clustering on Clouds and Humidity**
 428 **Heavy Rainfall Clustering and Clouds and Humidity**

429 We now consider how large-scale clustering of precipitation influences the cloud
 430 and humidity distribution. Our motivation is to understand how such clustering may
 431 influence radiative feedbacks. Previous authors have found that the degree of clustering
 432 on different spatial scales has an effect on the radiation budget and clouds (e.g., ???).
 433 The literature suggests changes in clustering under warming may lead to different
 434 cloud feedbacks and may therefore affect equilibrium climate sensitivity (ECS) (?).
 435 This section investigates this hypothesis for large-scale clustering across the CMIP6
 436 ensemble. As for previous sections, the analysis assesses whether relationships in
 437 interannual variability can be used to infer the response to climate change, raising
 438 the possibility of an observational constraint on particular radiative feedbacks or ECS
 439 itself.

440 Rather than focusing on changes in radiative fluxes or calculating feedback
 441 strength directly, we focus on changes in mid-tropospheric relative humidity, which
 442 has been argued to cause a negative longwave feedback associated with changes in
 443 convective organization (??), and changes in low-cloud fraction in regions of subsi-
 444 dence, which have been argued to cause a positive shortwave feedback associated with
 445 changes in convective organization (?). Changes in low clouds in regions of subsidence
 446 are also known to be important for understanding model spread in ECS (?). Corre-
 447 lations between measures of large-scale clustering of heavy precipitation and various
 448 other metrics commonly used to assess changes to the radiation budget on interannual
 449 and climatological timescales are presented in Figure S1-4 in the supporting informa-
 450 tion.

451 For our analysis, the mid-tropospheric relative humidity, RH, is taken as the 500
 452 hPa value, but the conclusions are not sensitive to using proximate pressure levels
 453 down to 700 hPa. Observed low-cloud fraction, LCF, is calculated using the ISCCP
 454 weather states (?) as described in Section ??, and taken as the cloud fraction below
 455 600 hPa. CMIP6 low-cloud fraction is calculated analogously, with cloud fraction
 456 pre-processed by interpolating hybrid-sigma coordinates to 19 pressure levels if not
 457 already available on pressure levels. We also consider the mean low-cloud fraction in
 458 regions of descent, denoted by a subscript d and calculated as the mean of gridpoints
 459 for which the monthly-mean vertical pressure velocity at 500 hPa is positive. Later we
 460 will consider variables in regions of ascent, defined analogously for negative 500 hPa
 461 vertical velocity and identified by a subscript a .

462 Figure ??a and Figure ??a show observational estimates of the regression pat-
 463 terns of RH and LCF against the mean area of precipitation features, A_m , for inter-
 464 annual variability. The regressions show a clear El-Niño-like pattern, with increases in
 465 RH and decreases in LCF in the central and eastern Pacific, and opposite changes over
 466 the warm pool. This suggests the changes in RH and low clouds with increased tropical
 467 clustering are caused at least in part by variations associated with El Niño-Southern
 468 Oscillation.

469 From a tropics-wide perspective, when the observed degree of clustering is high
 470 according to A_m , the tropical mean is drier (Figure ??b) while LCF increases, both
 471 when averaged over descending grid points (LCF_d in Figure ??b) and in regions of
 472 time-mean descent (contour on Figure ??a). The environmental signature associated
 473 with large-scale clustering is therefore consistent with a negative longwave feedback
 474 identified for large-scale clustering in idealized simulations (?) and a longwave- and
 475 low-cloud cooling signature found associated with interannual variations in mesoscale
 476 organization (?).

477 However, partial correlations of A_m with RH excluding the influence of the total
 478 area fraction of heavy precipitation, A_f , are insignificant in the observations (Figure
 479 ??b). This suggests that the influence of A_m on relative humidity is almost entirely
 480 due to increasing A_f . Observed correlations of relative humidity and the distance
 481 metrics C_m and C_z , representing proximity of heavy rainfall to the equator and the
 482 central Pacific, respectively, are also generally insignificant. Observed LCF_d on the
 483 other hand increases for all three forms of spatial clustering, outside the influence of
 484 A_f (Figure ??b).

485 The CMIP6 ensemble generally agrees on the strong association between A_f and
 486 the aforementioned tropical environmental signatures. However, unlike the observa-
 487 tional estimates, about half of the models also show a significant relationship between
 488 relative humidity and spatial shifts of heavy precipitation (Figure ??b). In particular,
 489 in a subset of models, zonal shifts in heavy precipitation to the central Pacific are
 490 (independent of variations in A_f) associated with a moistening in the tropics, while
 491 meridional shifts to the equator result in domain-mean drying. These relationships
 492 are also present in the high-resolution GCM. The models generally do not capture
 493 the observed LCF_d signature for clustering outside the influence of A_f , except for a
 494 subset of CMIP models producing increases in LCF_d for meridional shifts in heavy
 495 precipitation (Figure ??b).

496 Other notable independent effects of the spatial preference of heavy precipitation
 497 on environmental conditions include a reduction in high cloud fraction above 400 hPa
 498 in regions of ascent, HCF_a , with meridional shifts in precipitation ($r(C_m, HCF_a | A_f) \sim$
 499 0.25 in observations, the high-resolution GCM, and the model-mean of CMIP model
 500 correlations).

501 Our analysis of interannual variability has revealed strong relationships between
 502 RH and low-cloud fraction and large-scale clustering of precipitation in observations.
 503 However, in observations, the RH relationships are primarily driven by changes in the
 504 area fraction of convection A_f . A_f also influences low cloud fraction, LCF_d , but spatial
 505 shifts in heavy precipitation retains a connection to LCF_d outside the influence of A_f
 506 in observations. We now consider relationships between RH and low-cloud fraction
 507 changes and changes in the clustering of precipitation in climate projections.

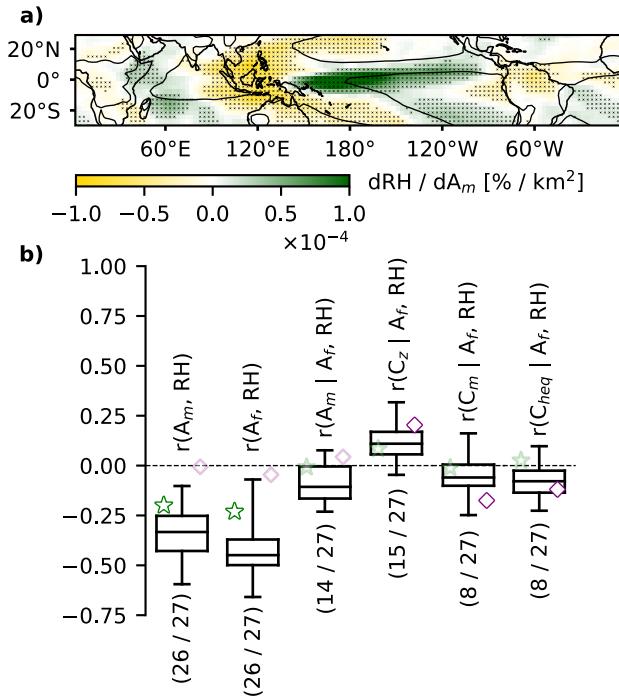


Figure 12. Relative humidity at 500 hPa, RH, regressed onto mean area of heavy precipitation features, A_m , in interannual variability (a). Boxplots of correlations and partial correlations outside the influence of the total area of heavy precipitation, A_f , of RH and A_m , mean distance of heavy precipitation to 180°E, C_z , and mean distance of heavy precipitation to the equator, C_m (b). Star and diamond show results for observations and high-resolution GCM, respectively, shown in lighter colors if not statistically significant. The numbers below the boxplots gives the fraction of models with statistically significant correlations.

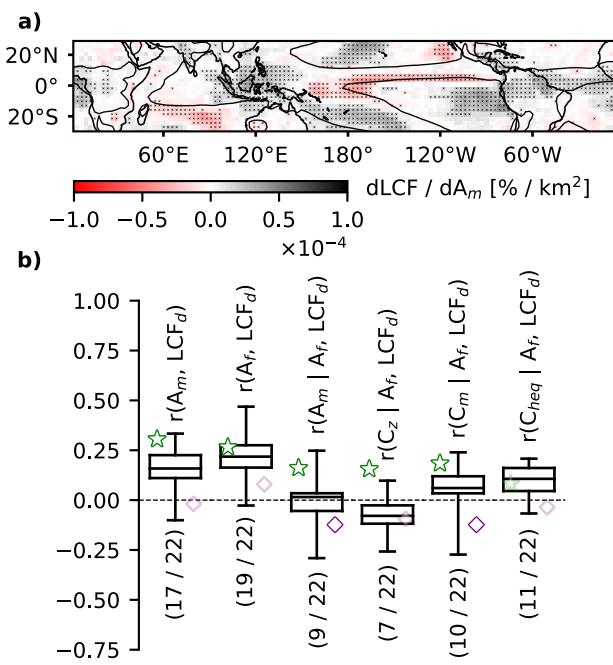


Figure 13. Same as Figure 10, but with the low cloud fraction, LCF, and low cloud fraction in regions of descent, LCF_d , as response variable.

508 Consistent with the results from the previous section, models with larger increases
 509 in large-scale clustering under warming tend to have changes in relative humidity and
 510 clouds consistent with an El Niño-like shift in the tropical circulation. The regression
 511 patterns of RH and LCF onto changes in A_m under warming across the CMIP6 en-
 512 semble (Figure S10a-b) are similar to those for interannual variability presented above.
 513 However, in contrast to the case for interannual variability, changes in large-scale clus-
 514 tering with warming have little connection to changes in tropical-mean mean RH or
 515 low-cloud fraction in regions of descent, LCF_d , across the CMIP6 ensemble (Figure
 516 ??a, b). Given this, it is perhaps not surprising that there is no correlation between
 517 the increase in A_m within a model under warming and the model's ECS (Figure ??c).
 518 Here we take ECS from the supplementary material of ? and ?.

519 The results therefore indicate that changes in large-scale clustering under warm-
 520 ing do not strongly affect radiative feedbacks, despite indications from observations
 521 that more clustered states are drier with more low-clouds in regions of large-scale de-
 522 scent. One reason for this result appears to be the different ways in which large-scale
 523 clustering can manifest at different timescales. In interannual variability, increases in
 524 clustering are associated with increases in the fractional area of heavy precipitation,
 525 defined here by A_f . Under climate change, increases in A_f in one month must be
 526 balanced by decreases in another month such that the overall average must remain
 527 constant. When the effects of changes in area fraction are removed, the observed re-
 528 lationship to RH becomes weak. However, this explanation is not the whole story,
 529 as many of the models do exhibit changes in RH associated with increased clustering
 530 independent of changes in the area fraction of heavy precipitation. Even among this
 531 subset of models, however, future increases in A_m are not a good predictor of future
 532 changes in clouds or relative humidity. This suggests that caution should be used in
 533 extrapolating relationships—either observed or simulated—between large-scale clus-
 534 tering and other properties of the climate in internal variability to those for climate
 535 change.

536 Finally, we note that there does exist a relationship between a tropical-mean
 537 drying and the proximity of heavy rainfall to the equator. Under warming, variations
 538 in the meridional contraction of heavy rainfall, as measured by the mean distance
 539 of heavily precipitating gridpoints to the equator C_m , explain about 45 percent of
 540 the variance in tropical-mean drying (Figure ??c and Figure ??). This relationship
 541 is consistent with the sign of the relationship between RH and C_m in interannual
 542 variability in a subset of CMIP6 models (Figure ??b). This result potentially highlights
 543 the importance of ITCZ narrowing as a specific manifestation of large-scale clustering
 544 that appears to be important for setting the tropical-mean relative humidity. However,
 545 we note that most of the spread in projected drying is due to the result of four models
 546 with dramatically different drying trends, and thus further work is required to confirm
 547 if this relationship is robust and physical.

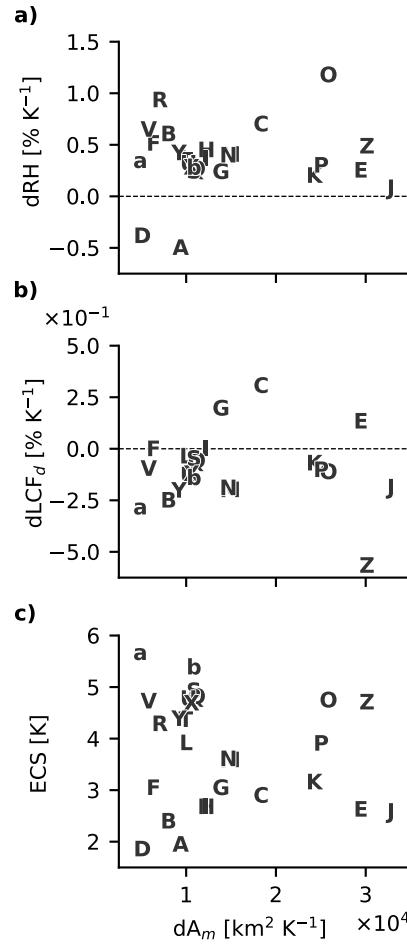


Figure 14. Scatter plot of change in A_m between historical period and SSP585 period per Kelvin tropical warming and associated tropical-mean change in 500 hPa relative humidity, RH, (a), low cloud fraction in regions of descent, LCF_d (b), and equilibrium climate sensitivity (ECS) in the CMIP6 ensemble (c). Models are as given in the legend in Figure 4.

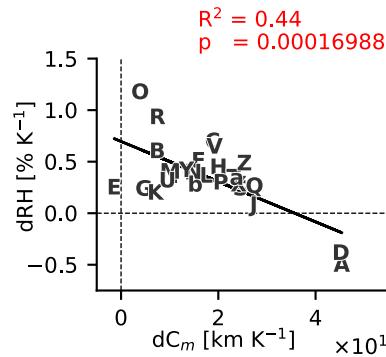


Figure 15. Change in relative humidity at 500 hPa, RH, regressed onto changes in mean distance of heavy precipitation to the equator, C_m between the historical and SSP585 periods per Kelvin warming.

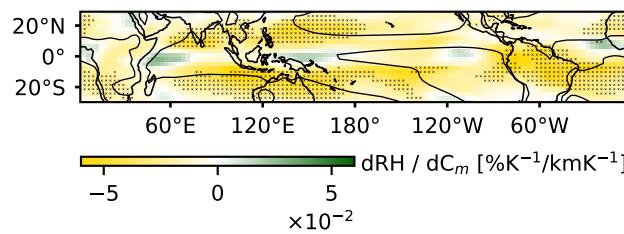


Figure 16. Scatter plot of change in mean distance of heavy precipitation to the equator, C_m , between historical period and SSP585 period per Kelvin tropical warming and associated tropical-mean change in 500hPa relative humidity, RH. The values are plotted as anomalies from the ensemble mean and models are as given in the legend in Figure 4.

548 **6 Summary and Discussion**

549 In this paper we have (1) presented the dominant spatial patterns of heavy pre-
 550 cipitation that produce a high degree of clustering on the large scale (Section 3); (2)
 551 tied the associated spatial patterns to mechanisms driving clustering through large-
 552 scale SST patterns (Section 4); and (3) evaluated the associated changes in properties
 553 of the atmosphere that are important for the radiation budget (Section 5) in both
 554 interannual variability and for projected changes with warming. We have defined the
 555 degree of clustering of precipitation based on the spatial distribution of the top 5 per-
 556 cent heaviest daily rainfall instances, with high clustering corresponding to scenes in
 557 which the mean area of individual precipitation features is large. A challenge in any
 558 definition of convective organization is in how one measures organization consistently
 559 as the total amount of precipitation changes (??). In the present study, the use of
 560 a percentile precipitation threshold accounts for changing mean precipitation rates in
 561 different climates. However, in internal variability, the area of precipitation features
 562 is affected by both spatial shifts in the precipitation distribution and variations in the
 563 total area fraction of heavy precipitation, A_f . This is addressed here by using Pearson
 564 partial correlations to evaluate the independent contributions of different measures of
 565 the spatial distribution of precipitation while controlling for the effect of A_f .

566 When tropical precipitation is observed to be highly clustered on the large scale
 567 in interannual variability, heavy precipitation gravitates meridionally to the equator
 568 and zonally towards the central Pacific. In climate projections, large-scale clustering
 569 of precipitation is found to increase in all models, and this coincides with a shift of
 570 precipitation toward the equator across the ensemble. We therefore hypothesize that
 571 a narrowing of the ITCZ may be an important contributor to increases in large-scale
 572 clustering of precipitation under warming. This implicates mechanisms related to the
 573 transport of energy by the Hadley circulation that have been argued to control changes
 574 in ITCZ width (?).

575 On the other hand, the intermodel spread in changes in clustering with warming
 576 across the CMIP6 ensemble is related to zonal rather than meridional shifts in the
 577 precipitation. This motivated an investigation of the role played by Pacific SST gradi-
 578 ents in changes in large-scale clustering of precipitation. In interannual variability, El
 579 Niño-Southern Oscillation linked variability appears to be a major driver of variability
 580 in large-scale clustering of precipitation, with precipitation during El Niño events more
 581 clustered than during La Niña events.

582 Under warming, changes in zonal SST gradients appeared to explain the sensi-
 583 tivity of projected clustering to zonal shifts in heavy rainfall; those models with more
 584 El Niño-like warming patterns tended to exhibit stronger increases in precipitation
 585 clustering. This is important given the large disagreement between observed and sim-
 586 ulated SST trends in the topical Pacific (e.g., ?). Observations show a strengthening
 587 of the SST gradient, suggesting a weaker increase in large-scale clustering compared
 588 to simulations, which tend to show a weakening of tropical SST gradients.

589 Finally, we assessed if the changes in clustering with warming may have an influ-
 590 ence on climate sensitivity. In observed interannual variability, a greater area fraction
 591 of heavy precipitation, A_f , is associated with a drier domain-mean and an increase in
 592 low-cloudiness in subsidence regions, LCF_d . The connection between clustering for a
 593 given A_f persists for LCF_d , but changes in the mean area of precipitation features and
 594 meridional and zonal shifts in heavy precipitation generally have weak relationships to
 595 the tropical-mean relative humidity, RH, independent of their relationship to A_f .

596 GCMs from the CMIP ensemble generally capture the observed tropical environ-
 597 ment signatures associated with changes in A_f , but often have different RH and LCF_d
 598 connections to shifts in heavy precipitation independent of A_f . In contrast to obser-

vations, RH in several models is sensitive to both meridional and zonal shifts in heavy precipitation. In CMIP6 models, zonal shifts of precipitation to the central Pacific tend to moisten, whereas meridional shifts to the equator tend to dry. Realistically represented or not, these sensitivities appear to affect how these models project relative humidity into the future; the subset of models sensitive to drying from meridional contraction of heavy precipitation create considerable spread in the model ensemble relative humidity changes under warming.

The study includes several limitations that are worth highlighting. Perhaps most importantly, the models we examined do not resolve the processes leading to organization of convection on mesoscales, which in turn may affect how they simulate heavy precipitation associated with large-scale convective features (?). This includes the high-resolution GCM, which still employs a parameterized convection scheme (?). Another limitation is that monthly anomalies from the climatology of the associated month obscure variations in diurnal and daily clustering tendencies and seasonal differences in the strength of relationships. Similarly, the climatological values do not control for variations in the contribution from different timescales, including diurnal up to seasonal and decadal variations in clustering. Finally, we note that our model ensemble is one of opportunity, and the models used were dictated by the available data. Correlation across the ensemble is not guaranteed to be produced by a physical relationship, and the extent to which such relationships arise by chance rise the more variables are examined. Nevertheless, the relationships between SST gradients and shifts in the precipitation distribution we highlight here are based on well-established physical relationships that provides some confidence in their robustness.

Future research is encouraged to adopt the control for total convective area, or other similar controls for changes in the mean precipitation rate, as used in the present framework. One avenue for further investigation is to identify models with realistic clustering compared to observations. The CMIP6 models considered here show a wide range in climatological clustering and internal variability in clustering, and perhaps a subset of models with more realistic clustering characteristics should be given more weight in projections of climate. In a similar way, investigating the connection between large-scale clustering and mesoscale clustering in high-resolution observations and storm resolving models may further constrain the model spread in projections by identifying unrealistic behavior. Further developing these research endeavors would allow for increased confidence and reduce the model uncertainty in aspects of projections that could be influenced by changes in convective organization, ultimately allowing for improvement in mitigation and adaptation strategies for a warming climate.

635 Open Research Section

636 CMIP models used in this study are listed in Table S1 and model output is avail-
637 able through the Earth System Grid Federation (ESGF) at <https://esgf-node.llnl.gov/search/cmip6/>.
638 Observational datasets and access are listed here; GPCP precipitation dataset: <https://doi.org/10.5065/ZGJD-9B02>, ERA5 dataset: <https://doi.org/10.24381/cds.adbb2d47>, NOAA surface temper-
639 ature dataset: <https://psl.noaa.gov/data/gridded/>, CERES outgoing longwave radia-
640 tion dataset: <https://ceres.larc.nasa.gov/data/>, ISCCP: ISCCP cloud states dataset:
641 <https://isccp.giss.nasa.gov/analysis/climanal5.html>. The IFS_9_FESOM_5 model data
642 is available via the World Data Center for Climate (WDCC) at DKRZ: <https://doi.org/10.26050/WDCC/nextGEMS>.
643 Code examples for reproducing key metric calculations and figures are available at ?.
644 Key metrics are available at ?.

646 **Acknowledgments** We acknowledge the World Climate Research Programme's Working
647 Group on Coupled Modelling, which is responsible for CMIP, and we thank the
648 climate modeling groups (listed in Table S1 in the supporting information) for producing
649 and making available their model output. Further, we acknowledge the Deutsches
650 Klimarechenzentrum (DKRZ) for providing resources for using the Levante supercom-
651 puter, which facilitated the analysis of the nextGEMS Cycle 2 simulation: IFS_9_FESOM_5
652 used in this study. We also extend a special thank you to Dr. Cathy Hohenegger at
653 Max Planck Institute for Meteorology for hosting the first author during a research
654 visit to Hamburg, for providing guidance on model selection, and for facilitating ac-
655 cess to institutional resources that greatly supported the analysis of the high-resolution
656 GCM used in this study. Finally, we acknowledge support from the Australian Re-
657 search Council (ARC) (grant no. DP230102077), the ARC Centre of Excellence for
658 the Weather of the 21st Century (grant no. CE170100023), and computational re-
659 sources and services from the National Computational Infrastructure, all funded by
660 the Australian Government.

661 **References**

- 662 Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N.-C., & Scott,
 663 J. D. (2002). The atmospheric bridge: The influence of enso teleconnec-
 664 tions on air-sea interaction over the global oceans. *Journal of Climate*,
 665 15(16), 2205 - 2231. Retrieved from https://journals.ametsoc.org/view/journals/clim/15/16/1520-0442.2002.015.2205.tabtio.2.0.co_2.xml doi:
 666 10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2
 667 Arnold, N. P., & Randall, D. A. (2015). Global-scale convective aggrega-
 668 tion: Implications for the madden-julian oscillation. *Journal of Advances*
 669 *in Modeling Earth Systems*, 7(4), 1499-1518. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015MS000498> doi:
 670 <https://doi.org/10.1002/2015MS000498>
 671 Bao, J., Sherwood, S. C., Colin, M., & Dixit, V. (2017, 10). The robust relationship
 672 between extreme precipitation and convective organization in idealized numerical
 673 modeling simulations. *Journal of Advances in Modeling Earth Systems*, 9, 2291-
 674 2303. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017MS001125> doi: 10.1002/2017MS001125
 675 Blackberg, P. (2025a). *Large-scale clustering of tropical precipitation: Code repos-
 676 itory*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.17070290>
 677 ([Software]) doi: 10.5281/zenodo.17070290
 678 Blackberg, P. (2025b). *Metrics for large-scale clustering of tropical precipita-
 679 tion*. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.16946243>
 680 ([Dataset]) doi: 10.5281/zenodo.16946243
 681 Bläckberg, C. P., & Singh, M. S. (2022, 5). Increased large-scale convective aggre-
 682 gation in cmip5 projections: Implications for tropical precipitation extremes. *Geo-
 683 physical Research Letters*, 49. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021GL097295> doi: 10.1029/2021GL097295
 684 Bony, S., Semie, A., Kramer, R. J., Soden, B., Tompkins, A. M., & Emanuel, K. A.
 685 (2020). Observed modulation of the tropical radiation budget by deep con-
 686 vective organization and lower-tropospheric stability. *AGU Advances*, 1(3),
 687 e2019AV000155. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019AV000155> (e2019AV000155 10.1029/2019AV000155) doi:
 688 <https://doi.org/10.1029/2019AV000155>
 689 Bretherton, C. S., Blossey, P. N., & Khairoutdinov, M. (2005, December). An
 690 energy-balance analysis of deep convective self-aggregation above uniform sst.
 691 *Journal of the Atmospheric Sciences*, 62(12), 4273-4292. Retrieved from
 692 <http://dx.doi.org/10.1175/JAS3614.1> doi: 10.1175/jas3614.1
 693 Byrne, M. P., & Schneider, T. (2016). Energetic constraints on the width of the
 694 intertropical convergence zone. *Journal of Climate*, 29(13), 4709 - 4721. Re-
 695 tried from <https://journals.ametsoc.org/view/journals/clim/29/13/jcli-d-15-0767.1.xml> doi: 10.1175/JCLI-D-15-0767.1
 696 Doelling, D. R., et al. (2013). *Ceres syn1deg-day terra-aqua modis edition 4a*. NASA
 697 Langley Atmospheric Science Data Center DAAC. doi: 10.5067/CERES/SYN1deg
 698 -Day_Terra-Aqua_Edition4A
 699 Emanuel, K., Wing, A. A., & Vincent, E. M. (2014, 2). Radiative-convective insta-
 700 bility. *Journal of Advances in Modeling Earth Systems*, 6(1), 75-90. Retrieved
 701 from <http://dx.doi.org/10.1002/2013MS000270> doi: 10.1002/2013ms000270
 702 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., &
 703 Taylor, K. E. (2016). Overview of the coupled model intercomparison project
 704 phase 6 (cmip6) experimental design and organization. *Geoscientific Model De-
 705 velopment*, 9(5), 1937-1958. Retrieved from <https://gmd.copernicus.org/articles/9/1937/2016/> doi: 10.5194/gmd-9-1937-2016
 706 Hartmann, D. L., Hendon, H. H., & Houze, R. A. (1984). Some implications
 707 of the mesoscale circulations in tropical cloud clusters for large-scale dynam-

- 715 ics and climate. *Journal of Atmospheric Sciences*, 41, 113 - 121. Retrieved
 716 from https://journals.ametsoc.org/view/journals/atsc/41/1/1520-0469_1984_041_0113_siotsmc_2.0.co_2.xml doi: 10.1175/1520-0469(1984)041(0113:
 717 SIOTMC)2.0.CO;2
- 718 Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., & Zelinka, M.
 719 (2022). Climate simulations: Recognize the ‘hot model’ problem. *Nature*, 605,
 720 26–29. doi: 10.1038/d41586-022-01192-2
- 721 Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., ...
 722 Thépaut, J.-N. (2023). *Era5 hourly data on pressure levels from 1940 to present*.
 723 Copernicus Climate Change Service (C3S) Climate Data Store. Retrieved from
 724 <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels> (Accessed: 2025-05-07) doi: 10.24381/cds.bd0915c6
- 725 Holloway, C. E., Wing, A. A., Bony, S., Muller, C., Masunaga, H., L'Ecuyer, T. S.,
 726 ... Zuidema, P. (2017, November). Observing convective aggregation. *Surveys
 727 in Geophysics*, 38(6), 1199–1236. Retrieved from <https://link.springer.com/article/10.1007/s10712-017-9419-1> doi: 10.1007/s10712-017-9419-1
- 728 Huang, B., Yin, X., Menne, M. J., Vose, R. S., & Zhang, H.-M. (2024). *Noaa global
 729 surface temperature dataset (noaaglobaltemp), version 6.0*. NOAA National Cen-
 730 ters for Environmental Information. Retrieved from <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc%3AC01704>
 731 doi: 10.25921/rzxg-p717
- 732 Huffman, G. J., Adler, R. F., Behrangi, A., Bolvin, D. T., Nelkin, E. J., Gu, G.,
 733 & Ehsani, M. R. (2023). *The new version 3.2 global precipitation climatol-
 734 ogy project (gpcp) monthly and daily precipitation products* (Vol. 36) (No. 21).
 735 American Meteorological Society. Retrieved from <https://doi.org/10.1175/JCLI-D-23-0123.1> doi: 10.1175/JCLI-D-23-0123.1
- 736 Jones, P. W. (1999). First- and second-order conservative remapping schemes
 737 for grids in spherical coordinates. *Monthly Weather Review*, 127(9), 2204
 738 - 2210. Retrieved from https://journals.ametsoc.org/view/journals/mwre/127/9/1520-0493_1999_127_2204_fasocr_2.0.co_2.xml doi: 10.1175/
 739 1520-0493(1999)127(2204:FASOCR)2.0.CO;2
- 740 Koldunov, N., Kölling, T., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R.,
 741 Sidorenko, D., ... Ziemen, F. A. (2023). *nextgems: output of the model de-
 742 velopment cycle 3 simulations for icon and ifs*. World Data Center for Cli-
 743 mate (WDCC) at DKRZ. Retrieved from https://doi.org/10.26050/WDCC/nextGEMS_cyc3 doi: 10.26050/WDCC/nextGEMS_cyc3
- 744 Maddox, R. A. (1980, November). Meoscale convective complexes. *Bulletin of
 745 the American Meteorological Society*, 61(11), 1374–1387. Retrieved from [http://dx.doi.org/10.1175/1520-0477\(1980\)061<1374:MCC>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1980)061<1374:MCC>2.0.CO;2) doi: 10.1175/
 746 1520-0477(1980)061<1374:mcc>2.0.co;2
- 747 Mapes, B. E. (1993, July). Gregarious tropical convection. *Journal of the At-
 748 mospheric Sciences*, 50(13), 2026–2037. Retrieved from [http://dx.doi.org/10.1175/1520-0469\(1993\)050<2026:gtc>2.0.co;2](http://dx.doi.org/10.1175/1520-0469(1993)050<2026:GTC>2.0.CO;2) doi: 10.1175/1520-0469(1993)
 749 050<2026:gtc>2.0.co;2
- 750 Mapes, B. E., & Houze, R. A. (1993, May). Cloud clusters and superclusters
 751 over the oceanic warm pool. *Monthly Weather Review*, 121(5), 1398–1416.
 752 Retrieved from [http://dx.doi.org/10.1175/1520-0493\(1993\)121<1398:ccasot>2.0.co;2](http://dx.doi.org/10.1175/1520-0493(1993)121<1398:CCASOT>2.0.CO;2) doi: 10.1175/1520-0493(1993)121<1398:ccasot>2.0.co;2
- 753 Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. London:
 754 Academic Press.
- 755 Pendergrass, A. G., Reed, K. A., & Medeiros, B. (2016). The link between extreme
 756 precipitation and convective organization in a warming climate: Global radiative-
 757 convective equilibrium simulations. *Geophysical Research Letters*, 43(21), 11,445-
 758 11,452. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/>

- 769 10.1002/2016GL071285 doi: <https://doi.org/10.1002/2016GL071285>
- 770 Popp, M., & Bony, S. (2019). Stronger zonal convective clustering associated with a
771 wider tropical rain belt. *Nature Communications*, *10*, 4261. doi: 10.1038/s41467
772 -019-12167-9
- 773 Quan, H., Zhang, B., Wang, C., & Fueglistaler, S. (2025). The sea surface tem-
774 perature pattern effect on outgoing longwave radiation: The role of large-scale
775 convective aggregation. *Geophysical Research Letters*, *52*(11), e2024GL112756.
776 Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2024GL112756> (e2024GL112756) doi: <https://doi.org/10.1029/2024GL112756>
- 777 Retsch, M. H., Jakob, C., & Singh, M. S. (2020, 4). Assessing convective organiza-
778 tion in tropical radar observations. *Journal of Geophysical Research: Atmospheres*,
779 *125*. doi: 10.1029/2019JD031801
- 780 Ropelewski, C., & Jones, P. (1987). An extension of the tahiti–darwin southern os-
781 cillation index. *Monthly Weather Review*, *115*, 2161–2165.
- 782 Schiro, K., Su, H., Ahmed, F., Dai, N., Singer, C., Gentine, P., ... Neelin, J.
783 (2022, 11). Model spread in tropical low cloud feedback tied to overturning
784 circulation response to warming. *Nature Communications*, *13*, 7119. Retrieved
785 from <https://www.nature.com/articles/s41467-022-34787-4> doi:
786 10.1038/s41467-022-34787-4
- 787 Semie, A. G., & Bony, S. (2020). Relationship between precipitation extremes
788 and convective organization inferred from satellite observations. *Geophysical
789 Research Letters*, *47*(9), e2019GL086927. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086927> (e2019GL086927)
790 10.1029/2019GL086927 doi: <https://doi.org/10.1029/2019GL086927>
- 791 Sobel, A. H., Held, I. M., & Bretherton, C. S. (2002). The enso signal in tropical
792 tropospheric temperature. *Journal of Climate*, *15*(18), 2702 - 2706. Retrieved
793 from https://journals.ametsoc.org/view/journals/clim/15/18/1520-0442_2002_015_2702_tesitt_2.0.co_2.xml doi: 10.1175/1520-0442(2002)015<2702:
794 TESITT>2.0.CO;2
- 795 Tobin, I., Bony, S., Holloway, C. E., Grandpeix, J.-Y., Sèze, G., Coppin, D., ...
796 Roca, R. (2013). Does convective aggregation need to be represented in cumulus
797 parameterizations? *Journal of Advances in Modeling Earth Systems*, *5*(4), 692-
798 703. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jame.20047> doi: <https://doi.org/10.1002/jame.20047>
- 800 Tselioudis, G., Tromeur, E., Rossow, W. B., & Zerefos, C. S. (2010, 7). Decadal
801 changes in tropical convection suggest effects on stratospheric water vapor. *Geo-
802 physical Research Letters*, *37*. doi: 10.1029/2010GL044092
- 803 Watanabe, M., Kang, S. M., Collins, M., Hwang, Y.-T., McGregor, S., & Stuecker,
804 M. F. (2024). Possible shift in controls of the tropical pacific surface warming
805 pattern. *Nature*, *630*(8016), 315–324. Retrieved from <https://www.nature.com/articles/s41586-024-07452-7> doi: 10.1038/s41586-024-07452-7
- 806 Wheeler, M., & Kiladis, G. N. (1999). Convectively coupled equatorial waves: Anal-
807 ysis of clouds and temperature in the wavenumber-frequency domain. *Journal
808 of the Atmospheric Sciences*, *56*, 374-399. Retrieved from https://journals.ametsoc.org/view/journals/atsc/56/3/1520-0469_1999_056_0374_ccewao_2.0.co_2.xml doi: 10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2
- 809 Williams, A. I. L., Jeevanjee, N., & Bloch-Johnson, J. (2023). Circus tents, con-
810 vective thresholds, and the non-linear climate response to tropical sssts. *Geophys-
811 ical Research Letters*, *50*(6), e2022GL101499. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL101499> (e2022GL101499)
812 2022GL101499) doi: <https://doi.org/10.1029/2022GL101499>
- 813 Wills, R. C. J., Dong, Y., Proistosecu, C., Armour, K. C., & Battisti, D. S. (2022,
814 September). Systematic climate model biases in the large-scale patterns of recent

- 823 sea-surface temperature and sea-level pressure change. *Geophysical Research Letters*, 49(17). Retrieved from <http://dx.doi.org/10.1029/2022GL100011> doi:
824 10.1029/2022gl100011
- 825 Wing, A. A., Emanuel, K., Holloway, C. E., & Muller, C. (2018). Convective
826 self-aggregation in numerical simulations: A review. In R. Pincus, D. Winker,
827 S. Bony, & B. Stevens (Eds.), *Shallow clouds, water vapor, circulation, and*
828 *climate sensitivity* (pp. 1–25). Cham: Springer International Publishing.
829 Retrieved from https://doi.org/10.1007/978-3-319-77273-8_1 doi:
830 10.1007/978-3-319-77273-8_1
- 831 Wing, A. A., & Emanuel, K. A. (2014). Physical mechanisms controlling self-
832 aggregation of convection in idealized numerical modeling simulations. *Journal*
833 *of Advances in Modeling Earth Systems*, 6(1), 59–74. Retrieved from
834 <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013MS000269>
835 doi: <https://doi.org/10.1002/2013MS000269>
- 836 Wodzicki, K. R., & Rapp, A. D. (2016). Long-term characterization of
837 the pacific itcz using trmm, gpcp, and era-interim. *Journal of Geophysical*
838 *Research: Atmospheres*, 121(7), 3153–3170. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD024458> doi:
839 <https://doi.org/10.1002/2015JD024458>
- 840 Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., & Rossow, W. B. (2018).
841 The international satellite cloud climatology project h-series climate data record
842 product. *Earth System Science Data*, 10(1), 583–593. Retrieved from <https://essd.copernicus.org/articles/10/583/2018/> doi: 10.5194/essd-10-583-2018
- 843 Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppli,
844 P., ... Taylor, K. E. (2020). Causes of higher climate sensitivity in cmip6 models.
845 *Geophysical Research Letters*, 47(1), e2019GL085782. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GL085782>
- 846
- 847
- 848
- 849