ARTICLE

Tropical precipitation clusters as islands on a rough water-vapor topography

Ziwei Li[1] | Paul A. O'Gorman[1] | Daniel H. Rothman[2]

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, 02139, United States

[2]Lorenz Center, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, 02139, United States

Correspondence
Ziwei Li, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, 02139, United States
Email: ziweili@mit.edu

Funding information
US National Science Foundation, Grant Number: AGS 1552195 and AGS 1749986; mTerra Catalyst Fund

Tropical precipitation clusters exhibit power-law frequency distributions in area and volume (integrated precipitation), implying a lack of characteristic scale in tropical convective organization. However, it remains unknown what gives rise to the power laws and how the power-law exponents for area and volume are related to one another. Here, we explore the perspective that precipitation clusters are islands above a convective threshold on a rough column-water-vapor (CWV) topography. This perspective is supported by the agreement between the precipitation clusters and CWV islands in their frequency distributions as well as fractal dimensions. Power laws exist for CWV islands at different thresholds through the CWV topography, suggesting that the existence of power-laws is not specifically related to local precipitation dynamics, but is rather a general feature of CWV islands. Furthermore, the frequency distributions and fractal dimensions of the clusters can be reproduced when the CWV field is modeled to be self-affine with a roughness exponent of 0.3. Self-affine scaling theory relates the statistics of precipitation clusters to the roughness exponent; it also relates the power-law slopes for area and volume without involving the roughness exponent. Thus, the perspective of precipitation clusters as CWV islands provides a useful framework to consider many statistical prop-

Abbreviations: CWV, column water vapor; KPZ, Kardar-Parisi-Zhang; CBL, convective boundary layer.

erties of the precipitation clusters, particularly given that CWV is well-observed over a wide range of length scales in the tropics. However, the statistics of CWV islands at the convective threshold imply a smaller roughness than is inferred from the power spectrum of the bulk CWV field, and further work is needed to understand the scaling of the CWV field.

KEYWORDS
precipitation clusters, power laws, fractals, statistical topography, self-affine scaling

1 | INTRODUCTION

Tropical convection and associated precipitation are organized in clusters of spatial scales from 10 to 1000 km (e.g., Mapes and Houze, 1993; Quinn and Neelin, 2017a). Understanding this organization is important because of the societal impacts of the spatial patterns of tropical precipitation, the influence of convective organization on the large-scale properties of the tropical atmosphere (Tobin et al., 2012), and the need to represent organization in convective parameterizations in global climate models (Mapes and Neale, 2011). Furthermore, both mean and extreme tropical precipitation are expected to experience substantial change with global warming (O'Gorman, 2012; Duffy et al., 2020), in which convective organization could play an important role (Rossow et al., 2013; Tan et al., 2015).

Many studies have investigated the cause of the spatial clumping of convection in the idealized setting of radiative convective equilibrium (Bretherton et al., 2005; Muller and Held, 2012; Craig and Mack, 2013; Emanuel et al., 2014; Wing and Emanuel, 2014; Wing and Cronin, 2015). This behavior is termed convective self-aggregation, and the physical processes that lead to self-aggregation in radiative convective equilibrium are also thought to be active in the tropical atmosphere (Holloway et al., 2017; Beucler et al., 2019). The paradigm of self-aggregation shows that convection can organize even in the absence of surface temperature gradients and background shear, but it does not by itself explain the spatiotemporal characteristics of convection and precipitation found in the tropics. In particular, numerous studies have found power-law distributions of precipitation clusters in observations (Lovejoy and Mandelbrot, 1985; Peters et al., 2010, 2012; Quinn and Neelin, 2017a; Teo et al., 2017), GCM simulations (Quinn and Neelin, 2017b), and high-resolution simulations with explicit convection (O'Gorman et al., 2021). Power-law distributions feature a probability density distribution of the form

$$\Pr(x) \propto x^{-\tau}, \tag{1}$$

where $\Pr(x)$ is the probability density of x. We refer to τ as the power-law exponent, and τ is positive in all cases throughout this paper. Eq. (1) is linear in the log-log space, and it is the only scale-invariant distribution in the sense that the distribution does not have a characteristic length scale (Turcotte, 1992). Therefore, the presence of power-law distributions suggests that precipitation is scale-free.

Scale-invariance has been associated with critical systems in statistical physics, including equilibrium critical phenomena (Pathria and Beale, 2011) and self-organized criticality (SOC) in forced-dissipative systems (Pruessner, 2012). These critical systems are characterized by divergence in correlation length and power-law distributions of quantities such as the cluster area of positive magnetization in the Ising model (Toral and Wall, 1987) and the distribution of avalanche size and duration in SOC sandpile models (Bak et al., 1987; Pruessner, 2012). Therefore, when power-law distributions were found in temporally- and spatially-connected precipitation clusters, hypotheses were made that atmospheric precipitation is an instance of SOC (Peters and Neelin, 2006; Neelin et al., 2008; Teo et al., 2017; Haerter, 2019). The power-law distribution of temporal cluster volume (precipitation integrated in time) would then correspond to the power-law distribution of avalanche size in SOC. Apart from the power laws, the analogy extends further as the atmosphere slowly builds up its water vapor via evaporation and has a sudden avalanche-like onset of precipitation once the column water vapor reaches a critical value (Peters and Neelin, 2006; Neelin et al., 2008). However, it remains unclear whether common SOC models (e.g., Pruessner, 2012, p. 82) can explain the observed power-law exponents of precipitation clusters.

In this paper, we focus on spatial precipitation clusters that are defined as groups of precipitating grid points connected in the horizontal. Cluster *area* is defined as the horizontal area of the cluster, and cluster *volume* is defined as the spatially integrated precipitation rate over the cluster following Quinn and Neelin (2017a), although they converted volume to an equivalent "power" associated with latent heating. Frequency distributions of precipitation clusters exhibit power laws with exponents in the range of 2.0 to 1.7 for area (Lovejoy and Mandelbrot, 1985; Peters et al., 2009, 2012; Quinn and Neelin, 2017a; Teo et al., 2017) and 1.7 to 1.5 for volume (Quinn and Neelin, 2017a; Teo et al., 2017). The spatial clustering of precipitation has been simulated by stochastic reaction-diffusion equations (Hottovy and Stechmann, 2015; Ahmed and Neelin, 2019). The stochastic model of Ahmed and Neelin (2019) includes representations of precipitation and lateral moisture transport, and it produces frequency-distribution exponents of 1.6 for area and 1.5 for volume, which are close to the observed exponents. To explain the exponents, Ahmed and Neelin (2019) used a stochastic branching process which gives the same exponent of 1.5 for both area and volume, although a direct connection between the branching process and precipitation processes was not provided.

From a different perspective, Pelletier (1997) proposed that the frequency distribution of tropical cumulus cloud area could be understood through the statistical properties of the convective boundary layer (CBL) height. The CBL height field was taken to be a self-affine surface, and clouds were assumed to form wherever the CBL height exceeds a certain threshold. The self-affine scaling theory of Kondev and Henley (1995) was then used to relate the area-distribution of clouds to the roughness of the CBL height field. Pelletier (1997) further hypothesized that the roughness exponent of the CBL height field has a value of 0.4 because of Kardar-Parisi-Zhang (KPZ) dynamics (Kardar et al., 1986). This roughness could also be connected to the fractal dimension of clouds, which was previously found to be 1.35 (Lovejoy, 1982), although clouds a have slightly different dimension for length scales below 1 km (Benner and Curry, 1998). The same fractal dimension of cumulus clouds has alternatively been related to three-dimensional turbulence (Siebesma and Jonker, 2000) and gradient percolation theory (Peters et al., 2009).

We take a somewhat similar approach to Pelletier (1997) in that we seek to understand clusters based on a threshold through a rough surface. However, we consider precipitation clusters rather than cumulus clouds, and we relate the clusters to the field of column-integrated water vapor (CWV) rather than CBL height. CWV has units of mm and represents the height of liquid water if all water vapor in the column is condensed onto the surface. Using CWV has the advantage that it is readily observed over a wide range of length scales. Furthermore, precipitation undergoes a rapid pickup once the column-integrated water vapor (CWV) exceeds a critical value as seen in observations (Peters and Neelin, 2006; Raymond et al., 2009; Neelin et al., 2009; Ahmed and Schumacher, 2015) and simulations (Bretherton et al., 2005; Sahany et al., 2012; Yano et al., 2012; Posselt et al., 2012). This property has been used in the stochastic models of Hottovy and Stechmann (2015) and Ahmed and Neelin (2019). The sharp pickup occurs because of the conditional instability of moist convection, as moist convection tends to occur with abundant low-level moisture through moist air parcels rising from near the surface and abundant mid-level moisture due to the effects of entrainment (Holloway and Neelin, 2009; Muller et al., 2009). We expect that the critical CWV to have weak variations in the horizontal due to the weak horizontal temperature gradients in the tropical free troposphere.

Here, we regard precipitation clusters as manifestations of CWV *islands* above a fixed *threshold* on a rough CWV topography (Fig. 1). The fixed threshold is the convective threshold of CWV above which precipitation rapidly increases. The power-law frequency distribution of precipitation cluster area is then akin to the Korčak's law, which describes a power-law distribution of island area above sea level on Earth's relief (Mandelbrot, 1982; Imre, 2015). We further assume that precipitation is linear in the excess of CWV above the threshold, such that the volume of a precipitation cluster corresponds to the volume of the CWV island above the threshold. Consistent with prior studies which show that power-law distributions for islands on a rough topography are a generic result (e.g., Olami and Zeitak, 1996), we find that the existence of power-law CWV island distributions is not dependent on the choice of the threshold and is not expected to be tied to the specific dynamics in precipitating regions.

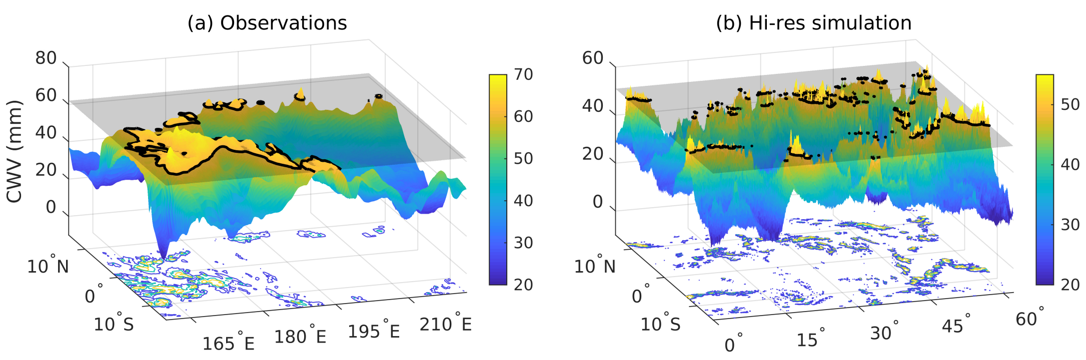

FIGURE 1 Examples of precipitation clusters as islands on a rough CWV topography using (a) observations and (b) a high-resolution simulation. The observations are taken from TRMM-3B42 for precipitation and ERA5 reanalysis for CWV. In both panels, the colored surfaces aloft show 3-hourly averaged CWV, the contours at the bottom show accumulated precipitation in the same 3-hour period, and the blue, green, and yellow contours correspond to precipitation rates of 10, 50, and 90 mm day^{-1}. The transparent planes represent convective CWV thresholds of 62 mm for (a) and 51 mm for (b), and the thick black contours highlight CWV island perimeters.

From a combination of observations and simulations, we show that tropical precipitation clusters are closely connected to CWV islands in area and volume frequency distributions and also in fractal dimensions. We further generate idealized self-affine surfaces and show that the thresholded islands on these surfaces correspond well with the statistics of CWV islands. Assuming that the CWV field is self-affine allows us to apply the self-affine scaling theories for contour loops of Kondev and Henley (1995) and Kondev et al. (2000) to predict the exponents of cluster area and volume distributions. We show how these exponents and the cluster fractal dimensions can be related to the roughness exponent of the CWV field and to each other. While the self-affine scaling theory a useful framework, its quantitative predictions are not very accurate in some cases because the CWV is not exactly self-affine, and because the scaling theory is derived for all contour loops at all thresholds, not for contour loops at a single threshold.

This paper is organized as follows. We describe the frequency distributions of precipitation clusters and their fractal dimensions in observations and simulations in section 2. We then demonstrate the similarity between the statistics of precipitation clusters and thresholded CWV islands in section 3. We further make the idealization that CWV is self-affine and apply the self-affine scaling theory to give expressions for the power-law distribution exponents and fractal dimensions of CWV islands in section 4. Lastly, we give our conclusions in section 5.

2 | DISTRIBUTIONS AND DIMENSIONS OF PRECIPITATION CLUSTERS IN DIFFERENT DATASETS

We analyze precipitation and CWV statistics in observations, a high-resolution simulation with explicit convection (hereby hi-res), and a GCM simulation. For observations, we use precipitation from TRMM-3B42 (Huffman et al., 2007) and CWV from the ERA5 reanalysis (Hersbach et al., 2020), both of which are on a 0.25° by 0.25° grid. We refer to ERA5 CWV as observations for simplicity even though it is from a reanalysis dataset. For the hi-res simulation, we use the system for atmospheric modeling (SAM) (Khairoutdinov and Randall, 2003), configured as a semi-global aquaplanet on an extended equatorial beta plane with a hemispherically- and zonally-symmetric sea surface temperature distribution. The domain spans from 78°S-78°N in latitude and 62° in longitude at the equator. The horizontal grid spacing is 12 km, and hypo-hydrostatic rescaling (Kuang et al., 2005; Garner et al., 2007; Fedorov et al., 2018) is applied to reduce the horizontal scale difference between convection and large-scale dynamics. See Yuval and O'Gorman (2020) and O'Gorman et al. (2021) for more details of hi-res. For the GCM simulation, ensemble number 1 in the CESM large ensemble dataset (Kay et al., 2015) is used as a representative coupled atmosphere-ocean GCM simulation, which has a grid spacing of 1.25° in longitude and 0.94° in latitude. The observations and GCM datasets span a period from 01/01/2002 to 12/31/2005 which is the longest overlap of the two datasets, and the hi-res simulation has a simulation length of 1200 days. The precipitation rate and the CWV field are 3-hourly averaged for observations and hi-res. For the GCM simulation, the precipitation rate is 6-hourly averaged, and the CWV field is calculated using a mass-integral of its 6-hourly instantaneous specific humidity output. All results presented are based on a region of 15°S-15°S, 160°E-222°E in the central tropical Pacific for observations and GCM, and 15°S-15°S with all available longitudes for hi-res.

We define precipitation clusters as groups of precipitating grid points that are connected via nearest-neighbor bonds, where there are four nearest-neighbors to each grid point. Precipitating grid points are grid points where the precipitation rate exceeds 0.7 mm h^{-1}. This precipitation threshold is chosen to be consistent with prior works such as Quinn and Neelin (2017a). Using a different threshold between 0.1 and 2.5 mm h^{-1} does not noticeably change the shape of the cluster distributions. Consistent with Fig. 2d in Otsuka et al. (2017), the cluster area distribution becomes lognormal-like when a much higher threshold of 20 mm h^{-1} is used, which might explain why some previous studies found that cloud clusters follow a lognormal distribution (e.g., Mapes and Houze, 1993).

Following Eq. (1), we denote the power-law exponents for cluster area and volume distributions as α and β, respectively, where α and β are positive when the log-log slope is negative. The meanings of all symbols used in the paper are summarized in Table 2. To estimate α and β, we sort cluster area and volume into 25 bins and apply linear regression in the log-log space. We use logarithmic binning because it reduces the noise in the tail of the distribution (Bauke, 2007). The widths of the bins are rounded to the nearest multiples of the smallest area or volume, following Quinn and Neelin (2017a). Each distribution's regression range is chosen based on the apparent extent of the power-law range. We report the error of each exponent in parenthesis after its estimated value. To obtain the error, we allow the starting bin to move upward by one bin, remain the same, or move downward by one bin, giving 3 choices of the starting point. The same applies to the end bin, and together these choices yield 9 exponent values. We regard the largest absolute deviation out of the 9 values from the estimated value as the measurement error of each exponent. This error dominates over the traditional standard error of regression slope, and we use it to represent the uncertainty in the measured exponents. The regression ranges of cluster volume distributions are approximately matched to those of the cluster area distributions in the sense that they cover the same fractional distrance between the smallest and largest bins of the distribution in the log space.[1]

[1] The smallest and largest bins of the distribution are determined by the minimum and maximum of the clusters, and the start and end point of the regression

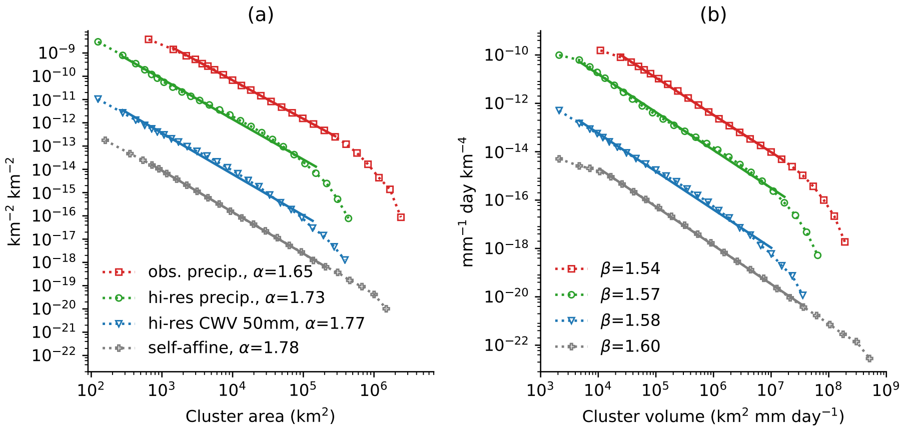

FIGURE 2 The frequency distributions of (a) cluster area and (b) cluster volume. Different colors and markers correspond to observed precipitation (red squares), hi-res precipitation (green circles), hi-res CWV islands at 51 mm (blue triangles), and islands on the self-affine surfaces with H = 0.3 at a threshold of 51 mm (gray crosses). The distributions are normalized such that the integral is the time-mean of the number of clusters per unit area of the domain, and they are consecutively shifted downwards by a factor of 50 starting from the observed precipitation for clarity. In (b), the volume of hi-res CWV islands and self-affine islands are converted to precipitation by Eq. (4) to plot island volume and precipitation volume on the same graph. The statistics are based on the region of 160°E-222°E and 15°S-15°N for observations, the same latitudinal band for hi-res, and the whole domain for self-affine. The solid lines are linear regressions in the log-log space, and their extents correspond to the regression ranges.

Consistent with previous studies, we find power-law frequency distributions with exponential upper cutoffs for precipitation cluster area and volume in observations and hi-res (Fig. 2). The cluster area exponent α is 1.65 (0.04) for observations with a similar value of 1.73 (0.05) for hi-res. The parentheses after each exponent indicates the regression error as described above. The values and errors of all exponents for different datasets in this paper are summarized in table 1. The cluster volume exponent β is lower at 1.54 (0.04) for observations with a similar value of 1.57 (0.04) for hi-res. These values for α and β are similar to values in previous studies that also analyzed TRMM-3B42 (Quinn and Neelin, 2017a; Teo et al., 2017). We regard the hi-res simulation as having an idealized yet faithful representation of tropical precipitation (O'Gorman et al., 2021), and we will focus on hi-res from here on. The cluster distributions for GCM are different, and they are discussed in Appendix B.

We use the area-perimeter scaling to estimate the fractal perimeter dimension of precipitation clusters (Fig. 3a). This self-similar scaling was first adopted to study fractal cloud dimensions by Lovejoy (1982). For a set of two-dimensional self-similar fractal objects, their perimeter length is related to area and radius by

$$l \propto A^{D_l/2} \propto R^{D_l}, \tag{2}$$

where l is perimeter length, A is area, D_l is the perimeter dimension, and R is the radius which can be thought of as the edge of the smallest square that can cover the object. The perimeters are traced out using *find_contours()* in the *scikit-image* library which implements a two-dimensional version of the marching cubes algorithm (Lorensen and

has the same linear location in the log space relative to the largest and smallest bins. In practice, we use the same set of consecutive bins out of all 25 bins as the regression range for cluster area and cluster power distributions.

TABLE 1 Precipitation cluster area (α) and cluster volume (β) exponents, distribution scaling relation (Eq. 10), perimeter dimension (D_l), volume dimension (D_V), and dimension scaling relation (Eq. 13) for different datasets. The results for hi-res CWV and the self-affine surface are calculated for islands cut by a threshold at 51 mm, about 2.0σ above the mean.

	Obs. precip.	Hi-res precip.	Hi-res CWV	Self-affine, H = 0.3	Theory, H = 0.3
α	1.65 (0.04)	1.73 (0.05)	1.77 (0.05)	1.78 (0.01)	1.85 (Eq. 6)
β	1.54 (0.04)	1.57 (0.04)	1.58 (0.04)	1.60 (0.03)	1.74 (Eq. 9)
$\alpha + 2/\beta$	2.95 (0.07)	3.00 (0.08)	3.04 (0.09)	3.03 (0.03)	3 (Eq. 10)
D_l	1.37 (0.02)	1.41 (0.02)	1.35 (0.02)	1.39 (0.01)	1.35 (Eq. 11)
D_V	2.32 (0.02)	2.33 (0.04)	2.32 (0.03)	2.41 (0.01)	2.3 (Eq. 12)
$2D_l + D_V$	5.07 (0.06)	5.14 (0.08)	5.02 (0.07)	5.18 (0.02)	5 (Eq. 13)

Cline, 1987). D_l is determined by binning $\sqrt{A}$ in the log space, taking the average of l in each bin, and regressing the l averages against $\sqrt{A}$ in the log-log space. The regression ranges used are indicated by the extents of the solid lines in Fig. 3. The uncertainty in the regression slopes are estimated in the same way as for the frequency distributions by varying the start and end point upwards or downwards by one bin and finding the maximum deviation. The D_l of precipitation clusters is 1.37 (0.02) for observations and has a similar value of 1.41 (0.02) for hi-res. These values are also broadly consistent with previous findings that the fractal dimension of cloud perimeter is 1.35 for radii from 1 to 1000 km (Lovejoy, 1982).

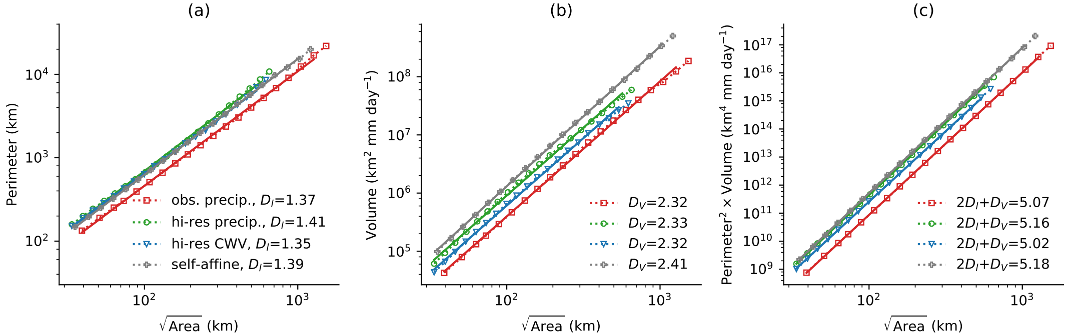

FIGURE 3 (a) Perimeter, (b) volume and (c) perimeter squared multiplied by volume as functions of the square root of area for observed precipitation clusters (red squares), hi-res precipitation clusters (green circles), hi-res CWV islands at 51 mm (blue trangles), and islands on a self-affine surface with H = 0.3 at 51 mm (gray crosses). Solid lines show linear regressions in the log-log space with the estimated slopes in the legends. In (b) and (c), the volume of hi-res CWV islands and self-affine islands are converted to precipitation by Eq. (4).

We also investigate the scaling of cluster volume with area (Fig. 3b). We introduce a volume fractal dimension, D_V, such that

$$V \propto A^{D_V/2} \propto R^{D_V}. \tag{3}$$

The precipitation clusters in observations and hi-res have similar D_V values of 2.32 (0.02) and 2.33 (0.04), respectively, with the dimensions and errors estimated using the same approach as for D_l, α, and β.

3 | PRECIPITATION AND THRESHOLDED CWV CLUSTERS

To better understand the statistical properties of precipitation clusters, we envision them as islands above a convective threshold on a rough CWV topography. Denoting CWV as Q, we define a CWV *convective threshold*, Q_c, which quantifies convective inhibition. We assume that the precipitation rate is zero when CWV is below Q_c, and the precipitation rate scales linearly with the excess of CWV when CWV is above Q_c:

$$P(\mathbf{r}) = \begin{cases} C(Q - Q_c) & \text{when } Q > Q_c, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

$P(\mathbf{r})$ is the 3-hourly precipitation rate at location $\mathbf{r}$, and C is a proportionality factor. The value of C does not affect the analytical results of the power-law exponents or fractal dimensions in later sections. Eq. (4) can be thought of as a first-order parameterization that captures the onset of precipitation once CWV exceeds a threshold. Fig. 4 shows the mean precipitation rate conditioned on CWV, in which CWV values are binned with constant intervals in the linear space, and the precipitation rate is averaged in each bin. We find that Eq. (4) works well for the hi-res simulation as shown in Fig. 4, while noticing the fact that the exact functional form relating precipitation to CWV differs to some extent across different observational and modeling studies (Neelin et al., 2009; Sahany et al., 2012; Yano et al., 2012; Posselt et al., 2012; Ahmed and Schumacher, 2015)

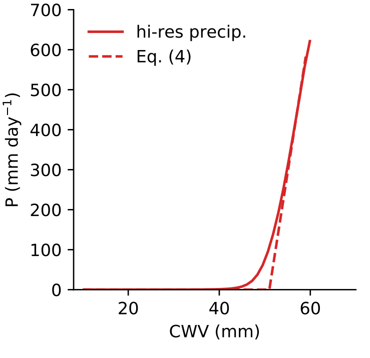

FIGURE 4 Tropical precipitation binned by CWV in the hi-res simulation (solid) and the estimate of Eq. (4) with $C = 72.8$ day^{-1} and $Q_c = 51$ mm (dashed). Bins which have less than a millionth of the total number of data instances are not shown.

The threshold Q_c cuts through the CWV field and gives a collection of distinct islands above the threshold (Fig. 1). With Eq. (4), each CWV island has an associated hypothetical precipitation cluster. We define the volume of the CWV island as the volume of the hypothetical precipitation cluster, which is the spatial integral of $C(Q - Q_c)$ within the island. The island's projected area is the area of the hypothetical precipitation cluster.

We choose a CWV threshold of 51 mm for hi-res throughout the paper because this is the integer threshold that gives CWV islands with both α and β closest to the precipitation clusters as discussed below. This threshold is also close to the value of 48 mm that gives the best match for the mean area fraction of precipitation in the hi-res simulation. We determine the proportionality factor, C, by regressing the linear part of Eq. (4) against the bin-averaged precipitation rates in Fig. 4. For hi-res, the CWV threshold of 51 mm gives C = 72.8 day^{-1}, meaning that 1 mm in CWV above the threshold corresponds to 72.8 mm day^{-1} in precipitation. A higher threshold of 62 mm is chosen for the case of observations, but it is only used for illustration in Fig. 1(a). We don't match the distribution of CWV islands to the distribution of precipitation clusters in observations because the moisture field from the ERA5 reanalysis is smooth at small length scales, making its CWV island distributions not power-law-like at high CWV thresholds. The threshold for hi-res is lower than for observations because the average sea surface temperature in hi-res is lower than that in observations in the selected central tropical Pacific domain.

To support the notion that precipitation clusters are manifestations of thresholded CWV islands, we first directly compare the pattern of thresholded CWV islands to that of precipitation clusters in observations and hi-res in Fig. 1. CWV islands in both observations and hi-res have very similar shapes to precipitation clusters. There is a dominant CWV island accompanied by multiple smaller islands in observations (Fig. 1 a), whereas multiple medium-area islands prevail in hi-res (Fig. 1 b); the same pattern also goes for precipitation clusters. This difference in CWV island (and precipitation cluster) configuration is due to the tropical Pacific warm pool being located on the western side in the domain of observations while the sea surface temperature is zonally uniform for hi-res.

Hi-res CWV islands also have power-law distributions in area and volume, and the power-law exponents are close to those of the precipitation clusters (Fig. 2). To generate CWV island distributions, we randomly sample 500 snapshots of 3-hourly averaged CWV field of hi-res. The hi-res simulation is used instead of observations or GCM because hi-res has the highest resolution and doesn't show evidence of smoothing in the CWV field at small length scales. We set the CWV island volumes that are smaller than the minimum precipitation cluster volume, 2419.2 km^2 mm day^{-1}, to 2419.2 km^2 mm day^{-1}.[2] Otherwise, the plotting of CWV island distributions is exactly the same as for precipitation clusters.

For the CWV threshold of 51 mm in hi-res, the frequency distributions of area and volume of the CWV islands are a good match to those of precipitation clusters for the power law ranges (Fig. 2). The measured power law exponents are α = 1.77 (0.05) and β = 1.58 (0.04) for the CWV islands as compared to α = 1.73 (0.05) and β = 1.57 (0.04) for the precipitation clusters. The areas of the largest precipitation clusters and the largest CWV islands at the convective threshold of 51 mm are similar at about 4×10^5 km^2. The mean and standard deviation of the CWV field are 35.3 mm and 8.0 mm, respectively, so that 51 mm is roughly 2.0σ above the mean value.

The fractal dimensions of the hi-res CWV islands at 51 mm are also in good agreement with those of the hi-res precipitation clusters (Fig. 3a, b). The D_l for CWV clusters at a threshold of 51 mm is 1.35 (0.02), slightly lower than the D_l of 1.41 (0.02) for hi-res precipitation clusters. Similarly, the D_V for CWV clusters at the 51 mm threshold is 2.32 (0.03), close to the D_V of 2.33 (0.04) for precipitation clusters.

Interestingly, power-law distributions of CWV island area and volume exist for a wide range of CWV thresholds (Fig. 5). This is the case even for thresholds like 35 mm far below the convective threshold (51 mm), such that the precipitation rate over most of the island coverage is close to zero. The area and volume distributions for 35 mm have a much larger maximum area and volume than the distributions for 51 mm because 35 mm cuts through a larger portion of the CWV topography as compared to 51 mm. Thus, the CWV island distributions at 35 mm has a local maximum at very large area and volume due to the presence of continents in the domain, and the maximum values in area and volume of precipitation and CWV islands at the 51 mm convective threshold are smaller. Similar to the case of precipitation clusters (Fig. 2), α is larger than β for the distributions of CWV islands at different thresholds

[2]The minimum volume, 2419.2 km^2 mm day^{-1}, is equal to having a precipitation rate of 0.7 mm h^{-1} at a single grid point of size 144 km^2. The 0.7 mm h^{-1} rate is the minimum precipitation rate used to define precipitation in section 2.

(Fig. 5). α and β are not constant for different CWV thresholds. Rather, both exponents follow a similar trend where they decrease and then increase as the CWV threshold is raised from near the mean level of 35 mm to the convective threshold of 51 mm (Figs. 5 and S1). The reasons for this variation are discussed in section 4.

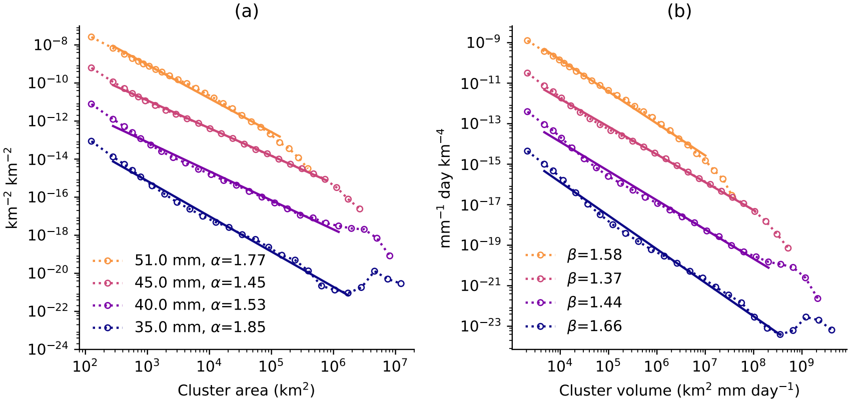

FIGURE 5 Dotted lines with circles show the frequency distributions of (a) area and (b) volume of hi-res CWV islands at different thresholds of 51 mm, 45 mm, 40 mm, and 35 mm (from orange to navy, top to bottom). Solid lines show linear regressions in the log-log space, with the corresponding exponents in the legends. The regression ranges are indicated by the horizontal extent of the solid lines. The distributions are normalized as those in Fig. 2 and, starting from 51 mm, are consecutively shifted downwards by two decades for clarity.

That the frequency distributions for the area and volume of CWV islands at the convective CWV threshold are very similar to those of precipitation clusters and that their fractal dimensions are also in good agreement suggest that tropical precipitation clusters are manifestations of thresholded CWV islands and are in turn related to the CWV field. This allows us to use the geometric properties of the CWV field to understand the existence of power laws and the relationships between α, β, D_l, and D_V. The fact that power-law frequency distributions exist for CWV islands at different thresholds also implies that the existence of power laws does not depend on local precipitation dynamics such as gust fronts or cold pools, but is more related to the scale-free nature of CWV dynamics which occurs in both precipitating and non-precipitating regions in the tropics. On the other hand, precipitation dynamics may affect the roughness of the CWV field and thus influence the power-law exponents of the frequency distributions and fractal dimensions. In the next section, we use a self-affine scaling theory to obtain analytical expressions that help explain the power-law exponents and fractal dimensions.

4 | APPLYING SELF-AFFINE SCALING THEORY TO THE CWV TOPOGRAPHY

We seek theories that can predict the CWV island frequency distributions and fractal dimensions from the statistical properties of the CWV field, which in turn give predictions for the corresponding properties of precipitation clusters. The traditional percolation problem on a two-dimensional lattice does not explain the slope of the island area distribution because it only has a power-law area distribution at a single threshold, i.e., the percolation threshold, and the area distribution at this threshold has a power-law exponent of $187/91 \approx 2$ (see table 2 in Stauffer and Aharony, 1994), which is steeper than the area-distribution exponents (α) in this paper.

We observe that the perimeter and volume of CWV islands in hi-res exhibit scaling relationships with area (Fig. 3a, b), and that the power spectrum of CWV approximately follow a power-law over a wide range of wavenumbers (Fig. 9). These properties suggest that CWV may be modeled as a *self-affine* surface (e.g., Mandelbrot, 1985; Barabási and Stanley, 1995). An isotropic self-affine surface, $h(\mathbf{r})$, satisfies

$$h(\mathbf{r}) \sim b^{-H} h(b\mathbf{r}), \tag{5}$$

where $h(\mathbf{r})$ is surface height at location $\mathbf{r}$, b is a rescaling factor, H is the roughness exponent, and $\sim$ means statistical equivalence. Eq. (5) states that the statistical properties of a subset of the surface (left side of the equation assuming $b > 1$) is the same as that of the surface itself (right side), subject to a rescaling of b^{-H} in height. Typically, H takes values between 0 and 1. For a fixed vertical width (standard deviation) at the largest horizontal scale of the system, the surface has less small-scale variations for larger H (Krim and Indekeu, 1993).

4.1 | Idealized self-affine surfaces

We first generate idealized self-affine surfaces to assess whether the islands on these surfaces correspond well to the CWV islands. The self-affine surfaces are generated in a square domain with 512 points in each direction. A grid spacing of 13.5km is chosen such that the side of the square domain has the same extent as the hi-res simulation in the zonal direction. The mean and standard deviation are chosen to match those of the hi-res CWV field. The self-affine surfaces are statistically isotropic with a power-law power spectrum $S(k) \propto k^{-\mu}$ where k is the wavenumber and $\mu = 2H + 1$ (Eq. 7.48 in Turcotte, 1992). We generate 500 surfaces, and for each surface, the phases of its Fourier components are randomly sampled in $[0, 2\pi)$ with a uniform distribution. The resulting surfaces also belong to Gaussian random surfaces because the height field has a Gaussian distribution.

We test a range of H values and find that H = 0.3 gives the best overall agreement with the hi-res CWV field in terms of island frequency distributions and island fractal dimensions at the convective threshold of 51 mm, or 2.0σ above the mean. Interestingly, H = 0.3 is close to the surface growth model of KPZ (Kardar et al., 1986) which measures $H \simeq 0.39$ in numerical simulations[3] and was used to relate cumulus cloud distribution to convective boundary layer height (Pelletier, 1997). Fig. 6(b) shows an example of the generated self-affine surface. The area and volume frequency distributions of self-affine islands at the 51 mm threshold follow power laws (Fig. 2). The exponents are α = 1.78 (0.01) and β = 1.60 (0.03), respectively, which are close to the exponents of the 51 mm CWV islands: α = 1.77 (0.05) and β = 1.58 (0.04). The perimeter dimension also agrees well with D_l = 1.39 (0.01) for self-affine islands and D_l = 1.35 (0.02) for the 51 mm CWV islands, whereas the agreement in volume dimension is not quite as good with D_V = 2.41 (0.01) for self-affine islands and D_V = 2.32 (0.03) for the CWV islands (Fig. 3a, b).

However, we also see deviations of the CWV field from self-affine scaling. In particular, the power spectrum of CWV in hi-res has μ = 2.51 (0.39) as shown in Fig. 9, which would imply a larger value of $H \approx 0.75$ compared to the roughness of H = 0.3 of self-affine surfaces that gives the best match for the CWV islands at 51 mm. The difference in power spectrum manifests in the differences in spatial patterns between hi-res CWV and the self-affine surface with H = 0.3. Because the total variance is the same, hi-res CWV has less small-scale variability than the self-affine surface due to the steeper slope in the hi-res CWV power spectrum (compare Figs. 6a and b). Similarly, we find that α and β vary differently as the threshold changes for hi-res CWV as compared to the self-affine surface (Fig. S1a). Thus, we speculate that precipitation dynamics may be decreasing H for high values of the CWV threshold as compared

[3] There has not been an exact calculation of H for KPZ in 2 dimensions. Numerical simulations in 2 dimensions seem to converge to $H \simeq 0.39$ (Pagnani and Parisi, 2015).

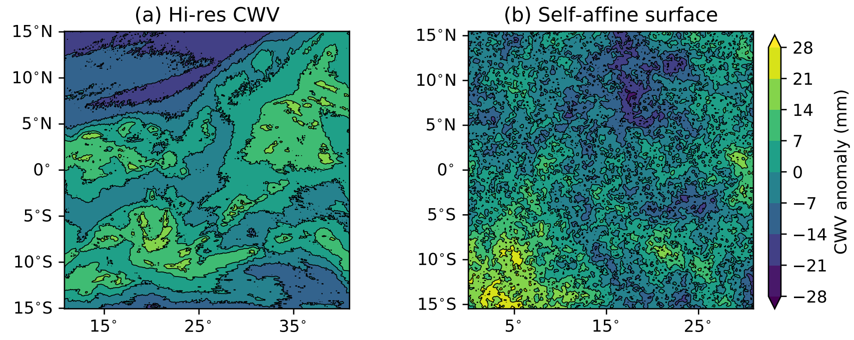

FIGURE 6 Snapshots (shading) and the corresponding level sets (contours) of the anomalies of (a) hi-res CWV and (b) an idealized self-affine surface with $H = 0.3$. Only a subset of the domain is shown in each case, and the spatial mean in each panel is removed for a better comparison.

to the appropriate H for the bulk CWV field measured from the power spectrum. A more general scaling form than self-affine scaling may be needed to capture all the statistical properties of the turbulent CWV field.

Although the CWV field is not exactly self-affine, islands on a self-affine surface at $H = 0.3$ do provide a good match to the CWV islands in hi-res at 51 mm for all of the statistical properties we investigate in this study. Thus, in the next section, we connect analytical results based on self-affine scaling theory to the measured frequency distributions and fractal dimensions.

4.2 | Theoretical predictions of frequency distributions

Suppose that a series of evenly-spaced thresholds cuts through a self-affine topography and generates an ensemble of contour loops and the encircled islands at different levels (Fig. 6b). The frequency distribution of the loop length in the contour ensemble is a power law whose slope is related to the roughness exponent, H (Kondev and Henley, 1995). Pelletier (1997) then showed that the frequency distribution of area within the contour loops also follows a power law as $\mathrm{Pr}(A) \propto A^{-\alpha}$, where Pr denotes frequency distribution, A denotes loop area, and

$$\alpha = 2 - \frac{H}{2}. \tag{6}$$

Eq. (6) shows a reverse dependence of α on H, consistent with Fig. 14 in Wood and Field (2011) which is based on a one-dimensional bounded cascade model for clouds. It's important to note that these power-law distributions of contour length and area apply to contours at all levels rather than at one particular threshold, and they also include the contours and areas of lakes within islands. For contours and islands at single levels near the mean level, Eq. (6) still holds (Rajabpour and Vaez Allaei, 2009), but when the level is raised far above the mean, Eq. (6) overestimates α (Olami and Zeitak, 1996)[4].

From Eq. (6), we derive a formula for the frequency distribution of island volume. The volume of an island scales as $V \propto Ah$, where A is the area and h is the peak height of the island above the threshold. We assume that the area of

[4] Olami and Zeitak (1996) neglected contours and areas associated with lakes within islands, whereas Rajabpour and Vaez Allaei (2009) considered all contours including contours within an island. We find that considering lakes inside islands reduces the bias in Eq. (6) at thresholds close to the mean (0σ-1σ), but does not diminish the overall decreasing trend in α at high thresholds.

lakes within the island is small compared to its total area, so that $A \propto R^2$ where R is the island's radius. Define vertical width, $W(R)$, as the root-mean-square fluctuation of the surface height where the mean is taken over R. For a self-affine surface, it follows that $W^2(R) \propto R^{2H}$. We further assume that the peak height of each island is proportional to the vertical width of the surface within the island's area coverage: $h \propto W(R) \propto R^H$, such that the volume scales as

$$V \propto AR^H \propto R^{2+H}. \tag{7}$$

Let $\Pr(V)$ be the frequency distribution of island volume, and assume that it has a power law form $\Pr(V) \propto V^{-\beta}$. Substituting $\Pr(A) \propto A^{-\alpha}$ and Eq. (7) into $\Pr(A)\mathrm{d}A = \Pr(V)\mathrm{d}V$ yields

$$\beta = \frac{2\alpha + H}{2 + H}. \tag{8}$$

Substituting for α using Eq. (6) gives

$$\beta = \frac{4}{2 + H}. \tag{9}$$

Therefore, the distributions of island area and volume both follow power laws for a self-affine topography, and the exponents of the power laws are controlled by the roughness exponent of the topography, H. Similar to α, larger values of H lead to smaller values of β, suggesting that both α and β should follow similar trends when H is varied. Since Eq. (6) overestimates α for a single threshold far above the mean, we expect Eq. (9) would also overestimate β in that case since we have used Eq. (6) in our derivation above.

The numerically generated self-affine surfaces in section 4.1 suggest that self-affine surfaces with $H = 0.3$ are an appropriate match to the CWV field for islands at the 51 mm convective threshold. For this H value, Eqs. (6) and (9) predict that $\alpha = 1.85$ and $\beta \approx 1.74$, as compared to $\alpha = 1.78$ (0.01) and $\beta = 1.60$ (0.03) measured from the generated self-affine surfaces at 51 mm (2.0σ above the mean). Thus, the theory correctly predicts that α is larger than β, but it over-predicts both values when applied to a single threshold high above the mean, consistent with previous work on α at different single thresholds (Olami and Zeitak, 1996).[5]

The theoretical predictions for α and β are related to each other via a scaling relation upon eliminating H from Eqs. (6) and (9):

$$\alpha + \frac{2}{\beta} = 3. \tag{10}$$

This relation allows the prediction of β given α and vice versa without knowing the value of H. Furthermore, for all α values between 1 and 2, β is always smaller than α by Eq. (10), which explains why β is generally found to be smaller than α for precipitation clusters in prior works. Despite the inaccuracies in the individual estimates of α and β, Eq. (10) holds well for observations and hi-res precipitation clusters (table 1) and also for hi-res CWV islands and self-affine islands under a wide range of thresholds (Fig. S1b) [6].

[5] The numerically generated self-affine surfaces give $\alpha = 1.84$ for all contours at the mean threshold including lakes within islands (Fig. S3), and this value is in better agreement with the theoretical prediction of $\alpha = 1.85$. We do not report β here because the volume is not well-defined for lakes within islands.

[6] Although we don't focus on GCM in the main text, it is interesting to note that Eq. (10) holds with $\alpha + 2/\beta = 2.96$ (0.14) for the very different values of α and β that occur for GCM as compared to hi-res and observations (α=1.10, β=1.07 as shown in Fig. 8).

4.3 | Theoretical predictions of fractal dimensions

For self-affine surfaces, the scaling theory also predicts the fractal dimension of contour loops,

$$D_l = \frac{3 - H}{2}, \tag{11}$$

which was derived by Kondev and Henley (1995) (partly based on a conjecture) and numerically confirmed by Rajabpour and Vaez Allaei (2009) and Nezhadhaghighi and Rajabpour (2011). Note that this dimension is the fractal dimension of a single contour loop, not the fractal dimension of all contours at the same level ($D = 2 - H$ as in Mandelbrot, 1975). For the volume fractal dimension, comparing its definition in Eq. (3) and the volume scaling in Eq. (7) gives that

$$D_V = 2 + H. \tag{12}$$

For $H = 0.3$, these theoretical predictions give $D_l = 1.35$ and D_V=2.3. These values are in good agreement with the results for the self-affine surface with $H = 0.3$ which have $D_l = 1.39$ (0.01) and $D_V = 2.41$ (0.01) and CWV islands at the threshold of 51 mm which have $D_l = 1.35$ (0.02) and $D_V = 2.32$ (0.03), shown in Fig. 3. Unlike for α and β, D_l and D_V for the islands on self-affine surfaces do not vary strongly as the threshold is varied, but there is some evidence for systematic variations in hi-res CWV island dimensions (Figs. 7 and S2a).

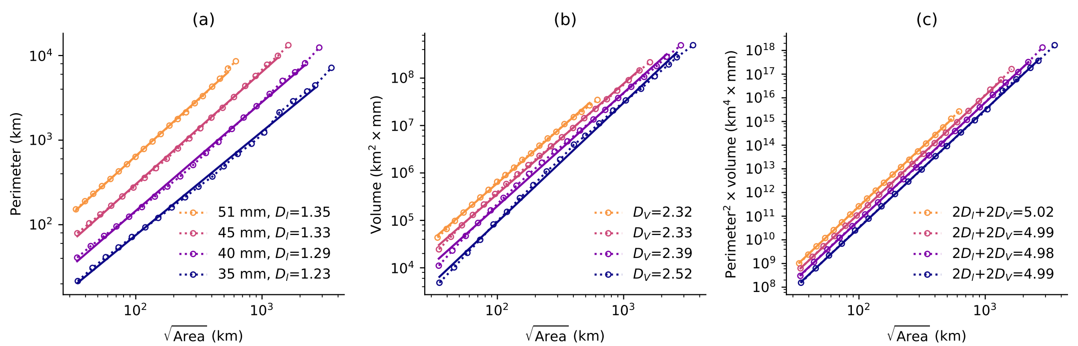

FIGURE 7 Same as in Fig. 3 but for hi-res CWV islands at thresholds of 51 mm, 45 mm, 40 mm, and 35 mm (from orange to navy, top to bottom). Starting from 51 mm, the scalings are consecutively shifted downwards by a factor of 2 for clarity.

Similar to the spirit of Eq.(10), we can eliminate H by combining Eqs.(11) and (12) and obtain

$$2D_l + D_V = 5. \tag{13}$$

Eq. (13) holds approximately for the precipitation clusters in observations and hi-res (Table 1 and Fig. 3c). Note that Table 1 shows $2D_l + D_V$ based on individual D_l and D_V from different datasets, whereas Fig. 3(c) shows the scaling exponent measured from regressing $R^{2D_l + D_V} \sim l^2 V$ in the log-log space. Eq. (13) also holds approximately for the self-affine surface and CWV islands at 51 mm (Table 1) and also for a wide range of thresholds (Fig. 7c and Fig. S2b).

Overall, the predictions based on the self-affine scaling theory provide considerable insight into how the roughness of the CWV field controls the statistical properties of the CWV islands, even though there are some inaccuracies related to the intrinsic limitations in the theory (which overestimates α and β for thresholds high above the mean) and related to the deviation of the CWV field from self-affine scaling.

5 | CONCLUSIONS AND DISCUSSION

We have shown from observations and a high-resolution simulation with explicit convection that tropical precipitation clusters can be seen as islands on a rough CWV topography cut by a convective threshold, analogous to the actual islands above sea level on Earth's relief. The physical basis for this link between precipitation clusters and CWV islands is the onset of precipitation at a critical CWV level, which has been widely found in observations and simulations of the tropical atmosphere. Using the hi-res simulation as an idealized representation of the tropical atmosphere, we find that the CWV islands at a convective threshold match precipitation clusters in the power-law frequency distributions of area and volume and also in their fractal dimensions. The frequency distributions of CWV island also follow power laws at a wide range of other CWV thresholds, suggesting that the existence of power-law distributions is not related to specific precipitation dynamics such as gust fronts within the precipitation clusters, but is instead a general property of thresholded islands on the CWV field.

We further assume that the CWV field is self-affine which allows us to apply the self-affine scaling theory. By numerically generating self-affine surfaces, we find that the CWV islands at the convective threshold are well-matched by islands on a self-affine surface with a roughness exponent of $H = 0.3$ at the same threshold. Within the self-affine framework, the roughness exponent of the topography governs the statistical properties of the islands. Previous work gave analytical expressions for the area distribution exponent (α) and the perimeter fractal dimension (D_l). Here, we further derive expressions for the volume distribution exponent (β) and the volume fractal dimension (D_V). While the expressions for the fractal dimensions are accurate, the expressions for α and β are overestimates. The overestimation is likely due to the scaling theory being applicable to all contours at all levels, not contours at the convective threshold which is high above the mean level.

The roughness of idealized self-affine surfaces that gives the best correspondence to CWV islands ($H = 0.3$) is lower than the roughness directly measured from the CWV power spectrum ($H \approx 0.75$). We speculate that the roughness may effectively be lower in regions of precipitation, but it is also possible that the turbulent CWV field would be better described by a more general scaling (e.g., multifractals). Hence, deviations from the simple self-affine scaling in the CWV field should be investigated in future work. Nonetheless, we derive a scaling relation from the scaling theory that directly relates α to β, and a similar relation that connects D_l and D_V. These scaling relations are approximately satisfied by the precipitation clusters and CWV islands across different thresholds. Given the discrepancies between the H-value best corresponding to CWV islands and the H-value measured from power spectra, these scaling relations are particularly useful as they don't involve H.

The framework presented here connects precipitation clusters to the properties of the CWV field, but the question of what determines the roughness of the CWV field has not been addressed. Horizontal diffusion and noise play important roles in existing stochastic models of the CWV field (Craig and Mack, 2013; Hottovy and Stechmann, 2015; Ahmed and Neelin, 2019). In addition, horizontal advection by rotational winds (e.g., as in two-dimensional turbulence) and gravity wave dynamics (Stiassnie et al., 1991) may also contribute to the scaling behavior of CWV. One complication with associating precipitation clusters with CWV islands is that precipitation itself reduces the local volume of CWV islands, but this issue can be avoided by considering the column moist static energy (CMSE) which is not affected by condensation and precipitation. Under the weak-temperature-gradient approximation (e.g., Neelin and Held, 1987), the spatial patterns of water vapor and moist static energy are similar, and we expect CMSE islands to behave similarly to the CWV islands. The deviation of the CWV field from self-affinity is also worthy of further research. The distributions of CWV islands best-matching the distributions of precipitation clusters are explained by self-affine surfaces with $H = 0.3$, which is close to $H \simeq 0.39$ as given by the KPZ universality class (Pagnani and Parisi, 2015). Therefore, more work is needed to confirm whether tropical CWV displays KPZ-type behavior, and to identify the physical mechanism in precipitation dynamics that may give rise to the observed scaling relations. Such mechanism may be responsible for the smaller roughness exponent associated with the statistics of CWV islands at a high threshold, which is different from the larger H value of the bulk CWV field as measured from its power spectrum.

An additional future avenue for research is to examine the response of precipitation cluster statistics to climate change (cf. Quinn and Neelin, 2017b), particularly in high-resolution simulations that have extensive power-law ranges. Eq. (10) suggests that any changes in the power-law exponent for the area distribution under warming would be directly related to changes in the exponent for the volume distribution, and thus affect the spatially integrated impacts of strong precipitation events.

Acknowledgements

We acknowledge support from NSF AGS 1552195 and 1749986 and from the mTerra Catalyst Fund. We thank David Neelin, Tom Beucler, Tim Cronin, William Boos, and Yi Ming for helpful discussions. We thank William Boos for providing the hi-res output, and we thank Marat Khairoutdinov for making SAM available to the community. We acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

Conflict of interest

The authors declare no conflict of interest.

A | MEANINGS OF SYMBOLS

TABLE 2 Meanings of symbols in the main text.

Symbol	Meaning
α	Cluster area exponent
β	Cluster volume exponent
σ	Standard deviation of CWV
μ	Power spectrum exponent
A	Cluster area
C	Proportionality factor from CWV to precipitation
D_l	Perimeter fractal dimension
D_V	Volume fractal dimension
H	Roughness (Hurst) exponent
l	Cluster perimeter length
$P(\mathbf{r})$	Precipitation at location **r**
$\Pr(X)$	Frequency distribution of X
R	Cluster radius
V	Cluster volume

B | RESULTS FOR GCM AND FOR CWV POWER SPECTRA

In the GCM simulation, the precipitation cluster area and volume also follow power laws, but the exponents are shallower than those with the observations and hi-res simulation. As shown in Fig. 8, the GCM simulation has α = 1.10 (0.07) for cluster area and β = 1.07 (0.05) for cluster volume compared to α = 1.65 (0.04) and β = 1.54 (0.04) in observations. This discrepancy remains if 6-hourly averaged precipitation is used for observations and hi-res to be consistent with the 6-hourly precipitation used for GCM.

One-dimensional spectra in the zonal direction of CWV for observations, hi-res, and GCM are shown in Fig. 9. The spectra are binned in the log wavenumber space with the bin widths rounded to multiples of the smallest wavenumber, $k_0 = 2\pi/L_x$, where L_x is the domain width. The same tropical domains as in the main text are used to calculate the spectra, and the spectra are calculated at each latitude and then averaged in latitude and time. We apply the Hann window in the zonal direction of the CWV fields of observations and GCM to reduce spectral leakage, and although not necessary, we also apply it in the case of hi-res for consistency. Similarly as for α and β, we measure the spectrum slope (μ) by applying linear regression on the binned power spectrum. The regression ranges of the power spectra are matched to those of cluster area distributions as follows. R is approximately related to A by $R^2 \approx 5A$ when averaged across all clusters for all datasets, and thus $\sqrt{A}$ corresponds to wavenumber $k = 2\pi/R \approx 2\pi/\sqrt{5A}$. We use this conversion between A and k to match the regression ranges, with the exception of the ERA5 CWV spectrum due to smoothing at small scales.

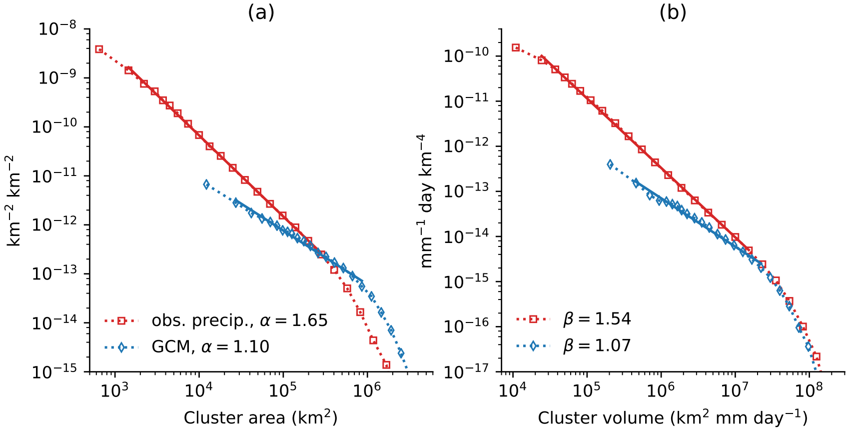

FIGURE 8 The frequency distributions of (a) cluster area and (b) cluster volume for the GCM (blue diamonds) and observations (red squares). The selected region, regression method and normalization are the same as in Fig. 2.

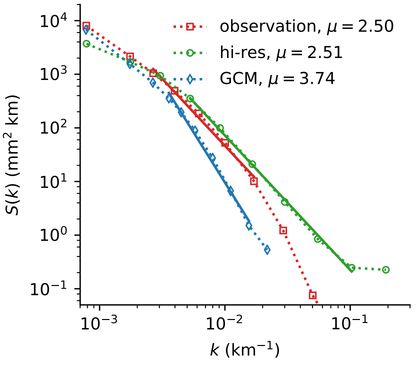

FIGURE 9 One-dimensional power spectra of CWV as a function of wavenumber. The spectra are based on observations (red squares), hi-res simulation (green circles), and GCM simulation (blue diamonds) in the respective equatorial regions. The solid lines are linear regressions in the log-log space, and their extents correspond to the regression ranges.

For the hi-res simulation which has the best resolved CWV field of the three datasets, we find $\mu = 2.51$ (0.39). The relation $\mu = 2H + 1$ then implies $H \approx 0.75$. Measurements of H using the second-order structure function and detrended fluctuation analysis (Bakke and Hansen, 2007) for hi-res CWV yield somewhat smaller values for H of 0.62 and 0.69, respectively. According to Eqs. (6) and (9), the self-affine scaling theory predicts that the steeper CWV power spectrum in GCM, $\mu = 3.74$ (0.48), leads to a greater value of H and thus smaller α and β compared to observations and hi-res simulation. Indeed, the α and β exponents for the GCM are much smaller than those for the observations (Fig. 8).

references

Ahmed, F. and Neelin, J. D. (2019) Explaining scales and statistics of tropical precipitation clusters with a stochastic model. *Journal of the Atmospheric Sciences*, **76**, 3063–3087.

Ahmed, F. and Schumacher, C. (2015) Convective and stratiform components of the precipitation-moisture relationship. *Geophysical Research Letters*, **42**, 10,453–10,462.

Bak, P., Tang, C. and Wiesenfeld, K. (1987) Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, **59**, 381–384.

Bakke, J. Ø. H. and Hansen, A. (2007) Accuracy of roughness exponent measurement methods. *Physical Review E*, **76**, 031136.

Barabási, A.-L. and Stanley, H. E. (1995) *Fractal Concepts in Surface Growth*. Cambridge: Cambridge University Press.

Bauke, H. (2007) Parameter estimation for power-law distributions by maximum likelihood methods. *European Physical Journal B*, **58**, 167–173.

Benner, T. C. and Curry, J. A. (1998) Characteristics of small tropical cumulus clouds and their impact on the environment. *Journal of Geophysical Research Atmospheres*, **103**, 28753–28767.

Beucler, T., Abbott, T. H., Cronin, T. W. and Pritchard, M. S. (2019) Comparing convective self-aggregation in idealized models to observed moist static energy variability near the equator. *Geophysical Research Letters*, **46**, 10589–10598.

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

Craig, G. C. and Mack, J. M. (2013) A coarsening model for self-organization of tropical convection. *Journal of Geophysical Research Atmospheres*, **118**, 8761–8769.

Duffy, M. L., O'Gorman, P. A. and Back, L. E. (2020) Importance of Laplacian of low-level warming for the response of precipitation to climate change over tropical oceans. *Journal of Climate*, **33**, 4403–4417.

Emanuel, K., Wing, A. A. and Vincent, E. M. (2014) Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, **6**, 75–90.

Fedorov, A. V., Muir, L., Boos, W. R. and Studholme, J. (2018) Tropical cyclogenesis in warm climates simulated by a cloud-system resolving model. *Climate Dynamics*, **52**, 107–127.

Garner, S. T., Frierson, D. M., Held, I. M., Pauluis, O. and Vallis, G. K. (2007) Resolving convection in a global hypohydrostatic model. *Journal of the Atmospheric Sciences*, **64**, 2061–2075.

Haerter, J. O. (2019) Convective self-aggregation as a cold pool-driven critical phenomenon. *Geophysical Research Letters*, **46**, 4017–4028.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J. N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049.

Holloway, C. E. and Neelin, D. J. (2009) Moisture vertical structure, column water vapor, and tropical deep convection. *Journal of the Atmospheric Sciences*, **66**, 1665–1683.

Holloway, C. E., Wing, A. A., Bony, S., Muller, C., Masunaga, H., L'Ecuyer, T. S., Turner, D. D. and Zuidema, P. (2017) Observing convective aggregation. *Surveys in Geophysics*, **38**, 1199–1236.

Hottovy, S. and Stechmann, S. N. (2015) A spatiotemporal stochastic model for tropical precipitation and water vapor dynamics. *Journal of the Atmospheric Sciences*, **72**, 4721–4738.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F. and Wolff, D. B. (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, **8**, 38–55.

Imre, A. R. (2015) Description of the area distribution of landmasses by Korcak exponent—the importance of the Arabic and Indian subcontinents in proper classification. *Arabian Journal of Geosciences*, **8**, 3615–3619.

Kardar, M., Parisi, G. and Zhang, Y.-C. (1986) Dynamic scaling of growing interfaces. *Physical Review Letters*, **56**, 889–892.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L. and Vertenstein, M. (2015) The community earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, **96**, 1333–1349.

Khairoutdinov, M. F. and Randall, D. A. (2003) Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, **60**, 607–625.

Kondev, J. and Henley, C. L. (1995) Geometrical exponents of contour loops on random Gaussian surfaces. *Physical Review Letters*, **74**, 4580–4583.

Kondev, J., Henley, C. L. and Salinas, D. G. (2000) Nonlinear measures for characterizing rough surface morphologies. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*.

Krim, J. and Indekeu, J. O. (1993) Roughness exponents: A paradox resolved. *Physical Review E*, **48**, 1576–1578.

Kuang, Z., Blossey, P. N. and Bretherton, C. S. (2005) A new approach for 3D cloud-resolving simulations of large-scale atmospheric circulation. *Geophysical Research Letters*, **32**, L02809.

Lorensen, W. E. and Cline, H. E. (1987) Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*, 163–169.

Lovejoy, S. (1982) Area-perimeter relation for rain and cloud areas. *Science*, **216**, 185–187.

Lovejoy, S. and Mandelbrot, B. B. (1985) Fractal properties of rain, and a fractal model. *Tellus A*, **37A**, 209–232.

Mandelbrot, B. B. (1975) Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines, and the number area rule for islands. *Proceedings of the National Academy of Sciences of the United States of America*, **72**, 3825–3828.

— (1982) *The Fractal Geometry of Nature*. New York: W. H. Freeman and Co.

— (1985) Self-affine fractals and fractal dimension. *Physica Scripta*, **32**, 257–260.

Mapes, B. and Neale, R. (2011) Parameterizing convective organization to escape the entrainment dilemma. *Journal of Advances in Modeling Earth Systems*, **3**, M06004.

Mapes, B. E. and Houze, R. A. (1993) Cloud clusters and superclusters over the oceanic warm pool. *Monthly Weather Review*, **121**, 1398–1415.

Muller, C. J., Back, L. E., O'Gorman, P. A. and Emanuel, K. A. (2009) A model for the relationship between tropical precipitation and column water vapor. *Geophysical Research Letters*, **36**, L16804.

Muller, C. J. and Held, I. M. (2012) Detailed investigation of the self-aggregation of convection in cloud-resolving simulations. *Journal of the Atmospheric Sciences*, **69**, 2551–2565.

Neelin, J. D. and Held, I. M. (1987) Modeling tropical convergence based on the moist static energy budget. *Monthly Weather Review*, **115**, 3–12.

Neelin, J. D., Peters, O. and Hales, K. (2009) The transition to strong convection. *Journal of the Atmospheric Sciences*, **66**, 2367–2384.

Neelin, J. D., Peters, O., Lin, J. W., Hales, K. and Holloway, C. E. (2008) Rethinking convective quasi-equilibrium: Observational constraints for stochastic convective schemes in climate models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **366**, 2581–2604.

Nezhadhaghighi, M. G. and Rajabpour, M. A. (2011) Contour lines of the discrete scale-invariant rough surfaces. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, **83**, 021122.

O'Gorman, P. A. (2012) Sensitivity of tropical precipitation extremes to climate change. *Nature Geoscience*, **5**, 697–700.

O'Gorman, P. A., Li, Z., Boos, W. R. and Yuval, J. (2021) Response of extreme precipitation to uniform surface warming in quasi-global aquaplanet simulations at high resolution. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **379**, 20190543.

Olami, Z. and Zeitak, R. (1996) Scaling of island distributions, percolation, and criticality in contour cuts through wrinkled surfaces. *Physical Review Letters*, **76**, 247–250.

Otsuka, S., Trilaksono, N. J. and Yoden, S. (2017) Comparing simulated size distributions of precipitation systems at different model resolution. *Scientific Online Letters on the Atmosphere*, **13**, 130–134.

Pagnani, A. and Parisi, G. (2015) Numerical estimate of the Kardar-Parisi-Zhang universality class in (2+1) dimensions. *Physical Review E*, **92**, 010101.

Pathria, R. K. and Beale, P. D. (2011) *Statistical Mechanics*. Cambridge, MA: Academic Press, third edn.

Pelletier, J. D. (1997) Kardar-Parisi-Zhang scaling of the height of the convective boundary layer and fractal structure of cumulus cloud fields. *Physical Review Letters*, **78**, 2672–2675.

Peters, O., Christensen, K. and Neelin, J. D. (2012) Rainfall and dragon-kings. *European Physical Journal: Special Topics*, **205**, 147–158.

Peters, O., Deluca, A., Corral, A., Neelin, J. D. and Holloway, C. E. (2010) Universality of rain event size distributions. *Journal of Statistical Mechanics: Theory and Experiment*, **2010**, P11030.

Peters, O. and Neelin, J. D. (2006) Critical phenomena in atmospheric precipitation. *Nature Physics*, **2**, 393–396.

Peters, O., Neelin, J. D. and Nesbitt, S. W. (2009) Mesoscale convective systems and critical clusters. *Journal of the Atmospheric Sciences*, **66**, 2913–2924.

Posselt, D. J., van den Heever, S., Stephens, G. and Igel, M. R. (2012) Changes in the interaction between tropical convection, radiation, and the large-scale circulation in a warming environment. *Journal of Climate*, **25**, 557–571.

Pruessner, G. (2012) *Self-Organised Criticality: Theory, Models and Characterisation*. Cambridge University Press.

Quinn, K. M. and Neelin, J. D. (2017a) Distributions of tropical precipitation cluster power and their changes under global warming. Part I: Observational baseline and comparison to a high-resolution atmospheric model. *Journal of Climate*, **30**, 8033–8044.

— (2017b) Distributions of tropical precipitation cluster power and their changes under global warming. Part II: Long-term time dependence in coupled model intercomparison project phase 5 models. *Journal of Climate*, **30**, 8045–8059.

Rajabpour, M. A. and Vaez Allaei, S. M. (2009) Scaling relations for contour lines of rough surfaces. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, **83**, 021122.

Raymond, D. J., Sessions, S. L., Sobel, A. H. and Fuchs, Ž. (2009) The mechanics of gross moist stability. *Journal of Advances in Modeling Earth Systems*, **1**, 20.

Rossow, W. B., Mekonnen, A., Pearl, C. and Goncalves, W. (2013) Tropical precipitation extremes. *Journal of Climate*, **26**, 1457–1466.

Sahany, S., Neelin, J. D., Hales, K. and Neale, R. B. (2012) Temperature–moisture dependence of the deep convective transition as a constraint on entrainment in climate models. *Journal of the Atmospheric Sciences*, **69**, 1340–1358.

Siebesma, A. P. and Jonker, H. J. (2000) Anomalous scaling of cumulus cloud boundaries. *Physical Review Letters*, **85**, 214–217.

Stauffer, D. and Aharony, A. (1994) *Introduction to Percolation Theory*. London, UK: Taylor & Francis, second edn.

Stechmann, S. N. and Neelin, J. D. (2011) A stochastic model for the transition to strong convection. *Journal of the Atmospheric Sciences*, **68**, 2955–2970.

— (2014) First-passage-time prototypes for precipitation statistics. *Journal of the Atmospheric Sciences*, **71**, 3269–3291.

Stiassnie, M., Agnon, Y. and Shemer, L. (1991) Fractal dimensions of random water surfaces. *Physica D*, **47**, 341–352.

Tan, J., Jakob, C., Rossow, W. B. and Tselioudis, G. (2015) Increases in tropical rainfall driven by changes in frequency of organized deep convection. *Nature*, **519**, 451–454.

Teo, C. K., Huynh, H. N., Koh, T. Y., Cheung, K. K., Legras, B., Chew, L. Y. and Norford, L. (2017) The universal scaling characteristics of tropical oceanic rain clusters. *Journal of Geophysical Research*, **122**, 5582–5599.

Tobin, I., Bony, S. and Roca, R. (2012) Observational evidence for relationships between the degree of aggregation of deep convection, water Vapor, surface fluxes, and radiation. *Journal of Climate*, **25**, 6885–6904.

Toral, R. and Wall, C. (1987) Finite-size scaling study of the equilibrium cluster distribution of the two-dimensional Ising model. *Journal of Physics A: Mathematical and General*, **20**, 4949–4965.

Turcotte, D. L. (1992) *Fractals and Chaos in Geology and Geophysics*. Cambridge, UK: Cambridge University Press, second edn.

Wing, A. A. and Cronin, T. W. (2015) Self-aggregation of convection in long channel geometry. *Quarterly Journal of the Royal Meteorological Society*, **142**, 1–15.

Wing, A. A. and Emanuel, K. A. (2014) Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations. *Journal of Advances in Modeling Earth Systems*, **5**, 1–14.

Wood, R. and Field, P. R. (2011) The distribution of cloud horizontal sizes. *Journal of Climate*, **24**, 4800–4816.

Yano, J.-I., Liu, C. and Moncrieff, M. W. (2012) Self-organized criticality and homeostasis in atmospheric convective organization. *Journal of the Atmospheric Sciences*, **69**, 3449–3462.

Yuval, J. and O'Gorman, P. A. (2020) Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nature Communications*, **11**, 3295.

Supplementary figures for article "Tropical precipitation clusters as islands on a rough water-vapor topography"

Ziwei Li[1], **Paul A. O'Gorman**[1], **Daniel H. Rothman**[2]

[1] Department of Earth, Atmospheric and Planetary Science
Massachusetts Institute of Technology
Cambridge, MA, United States

[2]Lorenz Center, Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA, United States

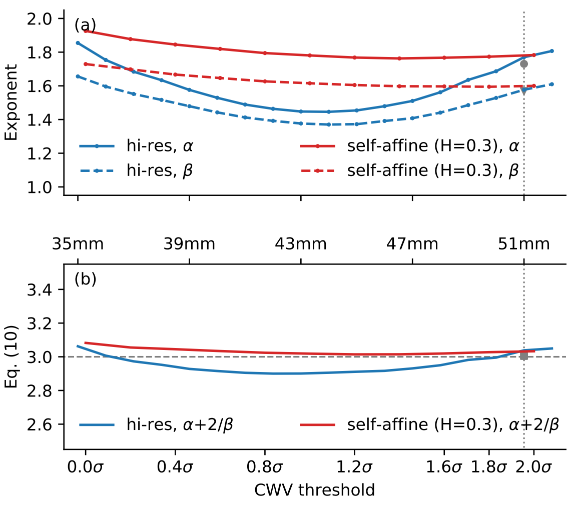

Figure S1: (a) Exponents of power-law frequency distributions and (b) scaling relation $\alpha + 2/\beta$ (Eq. 10) as functions of CWV threshold for hi-res CWV islands (blue) and islands on self-affine surfaces with H = 0.3 (red). In (a), the solid lines show the island area exponent (α), and the dashed lines show the island volume exponent (β). The single circle, triangle, and square markers in grey show α, β, and $\alpha + 2/\beta$, respectively, for hi-res precipitation clusters. The self-affine surfaces have the same mean and variance as the hi-res CWV field. The vertical dotted line shows the convective threshold of 51 mm.

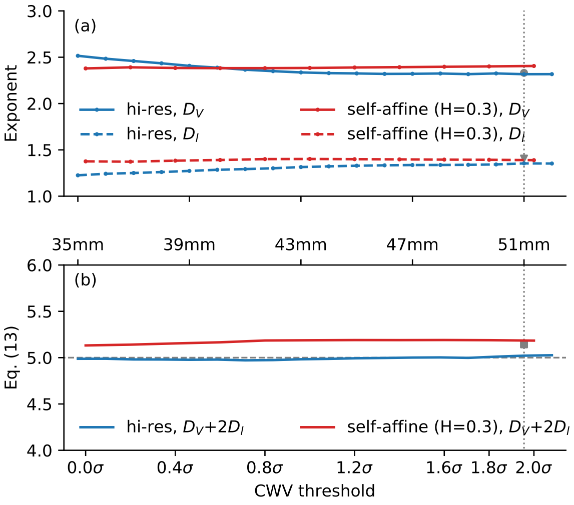

Figure S2: (a) Fractal dimensions and (b) their scaling relation $2D_l + D_V$ (Eq. 13) as functions of CWV threshold for hi-res CWV islands (blue) and islands on self-affine surfaces with H = 0.3 (red). In (a), the solid lines show the cluster volume dimension (D_V), and the dashed lines show the cluster perimeter dimension (D_l). The single circle, triangle, and square markers in grey show D_V, D_l, and $2D_l+D_V$, respectively, for hi-res precipitation clusters. The self-affine surfaces have the same mean and variance as the hi-res CWV field. The vertical dotted line shows the convective threshold of 51 mm.

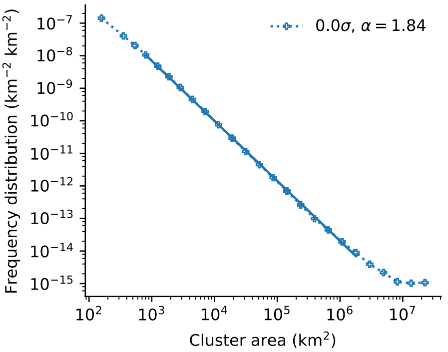

Figure S3: Frequency distribution of contour areas, including lakes within islands, at the mean threshold of the self-affine surfaces with H = 0.3. The solid line represents linear regression in log-log space.