Quarterly Journal of the
Royal Meteorological Society | RMetS

RESEARCH ARTICLE

Mesoscale marine tropical precipitation varies independently from the spatial arrangement of its convective cells

Matthias Brueck | Cathy Hohenegger | Bjorn Stevens

Max Planck Institute for Meteorology, Hamburg, Germany

Correspondence
C. Hohenegger, Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany.
Email: cathy.hohenegger@mpimet.mpg.de

Abstract
The relationship between mesoscale convective organization, quantified by the spatial arrangement of convection, and oceanic precipitation in the tropical belt is examined using the output of a global storm-resolving simulation. The analysis uses a 2D watershed segmentation algorithm based on local precipitation maxima to isolate individual precipitation cells and derive their properties. 10° by 10° scenes are analyzed using a phase-space representation made of the number of cells per scene and the mean area of the cells per scene to understand the controls on the spatial arrangement of convection and its precipitation. The presence of few and large cells in a scene indicates the presence of a more clustered distribution of cells, whereas many small cells in a scene tend to be randomly distributed. In general, the degree of clustering of a scene (I_{org}) is positively correlated to the mean area of the cells and negatively correlated to the number of cells. Strikingly, the degree of clustering, whether the cells are randomly distributed or closely spaced, to a first order does not matter for the precipitation amounts produced. Scenes of similar precipitation amounts appear as hyperbolae in our phase-space representation, hyperbolae that follow the contours of the precipitating area fraction. Finally, including the scene-averaged water vapour path (WVP) in our phase-space analysis reveals that scenes with larger WVP contain more cells than drier scenes, whereas the mean area of the cells only weakly varies with WVP. Dry scenes can contain both small and large cells, but they can contain only few cells of each category.

KEYWORDS
convection, object-based approaches, organization, precipitation, storm-resolving modelling

1 | INTRODUCTION

One of the striking features of deep moist convection is its ability to organize on a wide range of scales. Individual convective cells aggregate into Mesoscale Convective Systems (MCSs) which are broadly defined as cumulonimbus cloud systems with a contiguous precipitation area of at least 100 km in one direction (Houze, 2004). MCSs are building blocks for even larger-scale organizational features (Houze, 2004; Mapes *et al.*, 2006) such as hurricanes, the Madden–Julian Oscillation or the Intertropical Convergence Zone (ITCZ). Given the coarse resolution of $\mathcal{O}(100\,\text{km})$ of state-of-the-art General Circulation Models (GCMs), MCSs and any form of mesocale convective organization cannot explicitly be represented by GCMs, and convective parametrizations struggle to incorporate such effects (Moncrieff *et al.*, 2012). Despite this lack of mesoscale organization, GCMs are able to reproduce the large-scale organization of convection, for instance in an ITCZ, albeit not perfectly (Stanfield *et al.*, 2016; Hohenegger *et al.*, 2020). This raises the question as to the importance of convective organization, in particular mesoscale convective organization, for climate (Bony *et al.*, 2015). Answering this question remains challenging as it requires data with high spatial and temporal resolution on planetary-scale domains.

Answering this question first requires characterizing the degree of organization. Two approaches have been implicitly used in this respect. The first one is based on a phenomenological classification. Convection is isolated from its environment and classified into different categories generally based on the spatial extent of the convection. Typical classes are: isolated convective cells; MCSs, which may be subdivided into subclasses such as squall lines and Mesoscale Convective Complexes; and hurricanes. Isolated convective cells represent unorganized convection, whereas all the other classes are interpreted as forms of convective organization. In this view, organization is likely important for climate. The mere fact that MCSs produce more than 60% of the precipitation over the Tropics (Nesbitt *et al.*, 2006; Liu *et al.*, 2008), and even more than 90% in particular seasons (Roca *et al.*, 2014), underpins the importance of organized systems for the tropical climate.

The second approach computes indices to characterize the degree of organization. Convection is segmented into objects whose individual and collective properties are summarized into an index. Several indices have been proposed over the past years (Tobin *et al.*, 2012; Tompkins and Semie, 2017; Brune *et al.*, 2018; Kadoya and Masunaga, 2018; White *et al.*, 2018) as it remains difficult to objectively quantify organization. These indices tend to be based on objects' geometrical properties such as objects' area,

distances between objects, and/or number of objects. Distinct indices emphasize distinct aspects of organization depending on their combination of the objects' geometrical properties. These latter properties are derived with respect to a predefined geographical area of a given size, which we call a scene in the reminder of this text. In contrast to the first approach, there is no connection to a particular named convective phenomenon, a scene has to be defined, and the indices emphasize the spatial arrangement of objects in a scene rather than their extent *per se*. In this sense, a scene with isolated convective cells that are placed closer to each other, as compared to a random distribution, will be viewed as a more organized scene, which would not be the case under the first approach.

The derivation and use of organization indices have recently received renewed attention motivated by the results of radiative convective equilibrium (RCE) simulations. In such simulations (Tompkins and Craig, 1998; Bretherton *et al.*, 2005; Muller and Held, 2012; Wing and Emanuel, 2014), convection exhibits a clear phase transition from a random to an organized state, often made up of one final convective cluster. Also in such simulations, the organization of convection appears to be very important for climate as it leads to a strong drying of the non-convective region (Bretherton *et al.*, 2005; Hohenegger and Stevens, 2016). Observational evidence has confirmed that, as in RCE studies, more organized scenes are associated with a drier atmosphere, a decrease in anvil cloudiness, an increase in low clouds and an enhancement of low-level radiative cooling (Tobin *et al.*, 2012; Lebsock *et al.*, 2017; Stein *et al.*, 2017; Kadoya and Masunaga, 2018). These observational relationships have been derived by comparing scenes with distinct degrees of organization under similar large-scale conditions, such as surface temperature, vertical velocity and precipitation rate. This conditioning prevents a masking of the climatic impacts of organization by other confounding factors. Nevertheless, it does not allow determination of the actual importance of organization for climate, neither its impact on the large-scale circulation of the atmosphere nor on important climatic variables such as precipitation.

In a step toward addressing this latter issue, the overall goal of this paper is to investigate the relationship between mesoscale organization, precipitation and water vapour path (WVP). We include the WVP in our analysis given its strong relationship to precipitation (e.g., Bretherton *et al.*, 2004) and to organization (e.g., Bretherton *et al.*, 2005). In contrast to past studies which have used a phenomenological classification to investigate the contribution of organized convective phenomena to precipitation (e.g., Nesbitt *et al.*, 2006; Liu *et al.*, 2008), we will quantify this contribution solely based on the spatial arrangement of convective cells in a scene using an object-based approach. One

hypothesis motivating this work is that scenes with closely spaced convective cells also precipitate more. An argument for such a relationship is the idea that more closely spaced convective cells are protected from their dry hostile environment by the existence of their neighbours and thus see a moister environment. This implies a smaller reduction of updraught buoyancy through entrainment, allowing convective cells to reach deeper into the atmosphere, to be larger, to last longer and to precipitate more, as hypothesized by Lopez (1978) and Houze and Betts (1981) based on GATE (the Global Atmospheric Research Program's Atlantic Tropical Experiment) observations. RCE simulations have confirmed that strongly organized convection is protected from its environment by a moist shell and is associated with a smaller reduction of updraught buoyancy with height, as compared to randomly distributed convective cells (Becker *et al.*, 2018). The observationally based studies of Tobin *et al.* (2012) and Stein *et al.* (2017) have also suggested the presence of higher precipitation efficiency in more organized scenes, although the implications for precipitation amounts have not been quantified in their studies given the compositing of scenes on similar precipitation amounts. The higher precipitation efficiency in more organized scenes directly follows from the fact that more organized scenes are associated with lower values of WVP.

One difficulty in investigating factors controlling precipitation amounts is that it is known from past studies (e.g., Doneaud *et al.*, 1984; Nuijens *et al.*, 2009; Davies *et al.*, 2013) that precipitation scales well with the precipitating area fraction, or equivalently with the mean area of objects in a scene multiplied by their number. At the same time, these latter two characteristics likely influence our perception of organization. Larger objects may appear as more clustered than smaller objects from the mere geometrical fact that the available free space in a scene is limited. To better understand these various controls, we found it useful to analyse the degree of spatial arrangement of convection, precipitation and WVP, as well as their potential relationships, in a phase-space representation made of the mean area of objects in a scene and the number of objects. The use of such a phase-space representation is also motivated by recent developments in cumulus parametrizations with schemes now trying to predict the cumulus area fraction instead of the mass flux (e.g., Arakawa and Wu, 2013; Peters *et al.*, 2017).

Section 2 describes in more detail the simulation used for our analysis and our object-based classification algorithm with the so derived objects' properties, further justifying our analysis methodology. The relationship between the spatial arrangement of convection and precipitation is investigated in Section 3 using our phase-space representation. This phase-space representation is further applied in Section 4 to WVP. Section 5 discusses the results and conclusions are given in Section 6.

2 | SIMULATION AND OBJECT-BASED ANALYSIS

2.1 | Global storm-resolving simulation

We use a global simulation conducted with the ICOsahedral Nonhydrostatic model (ICON) at a grid spacing of 2.5 km for analysis. The advantage of such a simulation is that it resolves both the mesoscale at fine enough resolution and the planetary scale. Given the employed grid spacing, we note that the model does not use any convective parametrization, neither for shallow nor for deep convection. The simulation has been integrated following the protocol of the intercomparison project DYAMOND (DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains; Stevens *et al.*, 2019). The simulation starts on 1 August 2016 and is integrated for 40 days using prescribed daily sea surface temperature taken from the ECMWF operational analysis. A detailed description of the underlying model and a validation of the results can be found in Hohenegger *et al.* (2020). In terms of precipitation distribution, the model was found to produce 8% more tropical mean precipitation than observed. The location of the ITCZ over the eastern Pacific and over the Atlantic is nevertheless well captured; the simulation places the centre of its Atlantic and its eastern Pacific ITCZs at 8.4°N, 28°W and 9.8°N, 119.8°W, respectively, compared to 8.4°N, 24.4°W and 10.1°N, 117.2°W in observations (Hohenegger *et al.*, 2020, and their figure 5). The simulated ITCZs are wider than in observations, but the biases are smaller than 1° and thus smaller than what is obtained in low-resolution GCMs (e.g., Stanfield *et al.*, 2016). The model also captures the tropical short-wave and long-wave radiation budget at the surface within 1 and 5.7 W·m^{-2}, respectively, and within 9.9 amd 4 W·m^{-2} at the top of the atmosphere. The biases are comparable to the biases in the other storm-resolving models participating in the DYAMOND intercomparison project (table 6 in Stevens *et al.*, 2019 and table 2 in Hohenegger *et al.*, 2020).

For our analysis, we concentrate on the oceanic tropical precipitation belt, which is defined by 5°N±20° and ocean only. The latitude of 5°N corresponds to the maximum of the zonal mean precipitation averaged over the 40-day simulation period. It is thus viewed as the centre of the ITCZ. The analysis is performed on 15-min accumulated precipitation output, disregarding the first simulation day as spin-up. As ICON uses a triangular grid, the output has been regridded on a regular latitude–longitude grid of 0.05° by 0.05°. Even though the simulation extends

only for 40 days, its high spatial resolution and output frequency provides us with a sample large enough to derive meaningful statistics. The derived statistics are based on 1.37072×10^5 scenes, each scene containing 4.0×10^4 grid points (see the next subsection for the definition of a scene).

2.2 | Definition of cells and derivation of their attributes

Following Tobin *et al.* (2012), and given our interest in quantifying the mesoscale arrangement of convection using an object-based approach, we first define a scene as a 10° by 10° subdomain. The scenes are spatially not overlapping. The spatial arrangement of convection is then characterized for each of these scenes by segmenting the convective objects that populate each scene. Our segmentation procedure follows Senf *et al.* (2018), but is applied to precipitation and not to brightness temperature. The segmentation procedure is summarized below and illustrated for four distinct convective situations in Figure 1.

The precipitation field is used to define precipitation cells based on two precipitation thresholds. We use a lower precipitation threshold of $0.1 \, \text{mm} \cdot \text{hr}^{-1}$ to segment precipitating objects from their non-precipitating environment. We then use an upper threshold of $10 \, \text{mm} \cdot \text{hr}^{-1}$ to find precipitation cores inside those objects. Due to possible regridding artifacts, only precipitation cores that are larger than three pixels are considered. Also, precipitation objects without precipitation cores are disregarded. The precipitation cores are used to subdivide the precipitation objects into what we call cells in the reminder of the text. The local maxima inside the precipitation cores serve as starting markers for the watershed segmentation procedure. The watershed procedure essentially fills the precipitating neighbourhood surrounding a local maximum until it gets into contact with another neighbourhood. Each so identified convective cell gets an identifier. Local maxima cannot be located directly next to each other. Also, in order to find the seed points for the watershed segmentation, the precipitation field is filtered with a multidimensional Gaussian filter with a standard deviation for the Gaussian kernel of 0.05. This value is very low and merely serves to compensate regridding artifacts which increases the robustness of the seed points. The domain average precipitation rate and the total precipitating area are not affected by this filtering.

Having segmented the precipitation field into its individual cells, the total area fraction f_A of the cells in a scene, the number N of cells in a scene, the mean area $\overline{A}$ of the cells in a scene and the spatial arrangement of the cells in a scene can be computed. These are the cells' attributes that will be used for our analysis. To quantify the spatial

arrangement of the cells in a scene, we measure the deviation of the spatial distribution of the cells from the distribution resulting from randomly distributing these same cells, following Weger *et al.* (1992). This approach was used by Tompkins and Semie (2017) to build an index called I_{org} that we apply in our study as well. Formally, I_{org} is defined as the integral below the curve of the nearest-neighbour cumulative density function (NNCDF) of the cells plotted as a function of the NNCDF for a random distribution of these cells. A value of 0.5 corresponds to a random distribution, values larger than 0.5 indicate clustering, whereas values smaller than 0.5 indicate regularly distributed cells. To derive the NNCDFs, we follow Pscheidt *et al.* (2019). To compute the nearest-neighbour distance between the cells, the cells are approximated as disks and the edge-to-edge distance d_{e} between two cells i and j is computed as: $d_{\text{e}} = d - r_i - r_j$ with d the distance between the two cells' centres and r_i the equivalent radius of a cell i ($r_i = \sqrt{A_i/\pi}$). The theoretical random distribution of the cells is obtained by randomly distributing disks with the same areas and same number as the cells present in a scene. If the disks overlap, their position is shifted so that they become adjacent as overlapping cells are not possible in our simulation. The procedure is repeated one hundred times and the theoretical random distribution results by taking the mean over these hundred realizations. If few cells are present in a scene, it is difficult to distinguish random from clustered or regular states. For our analysis, we thus only consider scenes with at least three cells. Note also that using an object-based approach requires defining a scene and that the finite size of a scene imposes some geometrical constraints on the possible arrangement of cells. As such, our results are only valid for the chosen scene's size and do not generalize.

Quantifying the impact of convective clustering on precipitation is not trivial. To better illustrate our analysis strategy and our end choice of the explanatory variables, Figure 1 shows the precipitation field for four scenes (Figure 1a–d) as well as the resulting identified cells with typical scene's properties (Figure 1e–h). Subjectively looking at the four scenes, Figure 1a seems to be associated with random or maybe regularly distributed isolated cells. The three remaining scenes appear clustered in the form of a MCS (Figure 1b), confined isolated cells (Figure 1c) and a hurricane (Figure 1d). These differences are directly captured by I_{org}, with values close to 0.5 in the first scene (Figure 1e) and values larger than 0.5 in the remaining three scenes (Figure 1f–h). But a closer look at Figure 1 also reveals some surprises and explains why only considering I_{org} may make the results difficult to interpret. I_{org} classifies the MCS scene (Figure 1f) as less clustered than the scene with confined isolated cells (Figure 1g). The difficulty arises from the fact that we subjectively associate

clustering with scenes with large convective area fraction (f_A), closely spaced cells (I_{org}) and few objects (N). How these attributes contribute to our perception of clustering is likely subjective. Attempting to summarize these attributes in one index has proven difficult as the weight of each attribute toward organization is unknown. Distinct indices emphasize distinct aspects of clustering and can lead to different ordering. The Convective Organization Potential (COP) index (White *et al.*, 2018), which includes both area, distance and number of objects, for instance does classify the MCS scene (Figure 1f) as more clustered than the scene with confined isolated cells (Figure 1g). But we also know that precipitation itself correlates well with the precipitating area fraction f_A (e.g., Doneaud *et al.*, 1984; Davies *et al.*, 2013) or equivalently with $\overline{A}$ multiplied by N. From these considerations it seems judicious to use an index to quantify the degree of convective clustering that does not mix different cells' attributes and use area and number of cells as supplementary explanatory variables given their link to precipitation and organization.

Our analysis methodology, even though being able to segment a precipitation field into its individual cells, is not able to track individual cells. Tracking individual cells is not straightforward as subjective and non-trivial decisions have to be made on merging and splitting events (Fiolleau and Rocca, 2013; Moseley *et al.*, 2019). Moreover, the use of fixed scenes makes the tracking even less straightforward as a scene might contain a mix of cells in different stages of their life cycle, and with the possibility for cells to move from one scene to another one. Given these difficulties, our analysis methodology takes a pure snapshot view on precipitation. It does not provide any information on the life cyle of the segmented convective cells. As a consequence, scenes with convective cells in their growing phase are mixed with scenes with convective cells in their mature or decaying phase, even though the underlying physics may be quite different. This could blur out existing relationships between precipitation, area, number and arrangement of cells.

3 | THE PHASE SPACE OF ORGANIZATION AND PRECIPITATION

We start our analysis by examining the variations of I_{org} as a function of $\overline{A}$ and N, before moving to precipitation amounts and finally linking the two analyses. Computing I_{org} for all 1.37072×10^5 scenes and displaying it as a function of $\overline{A}$ and N gives Figure 2. Looking at the colour bar, it is evident that convection is almost always clustered over the tropical oceans. The values range between 0.56 and
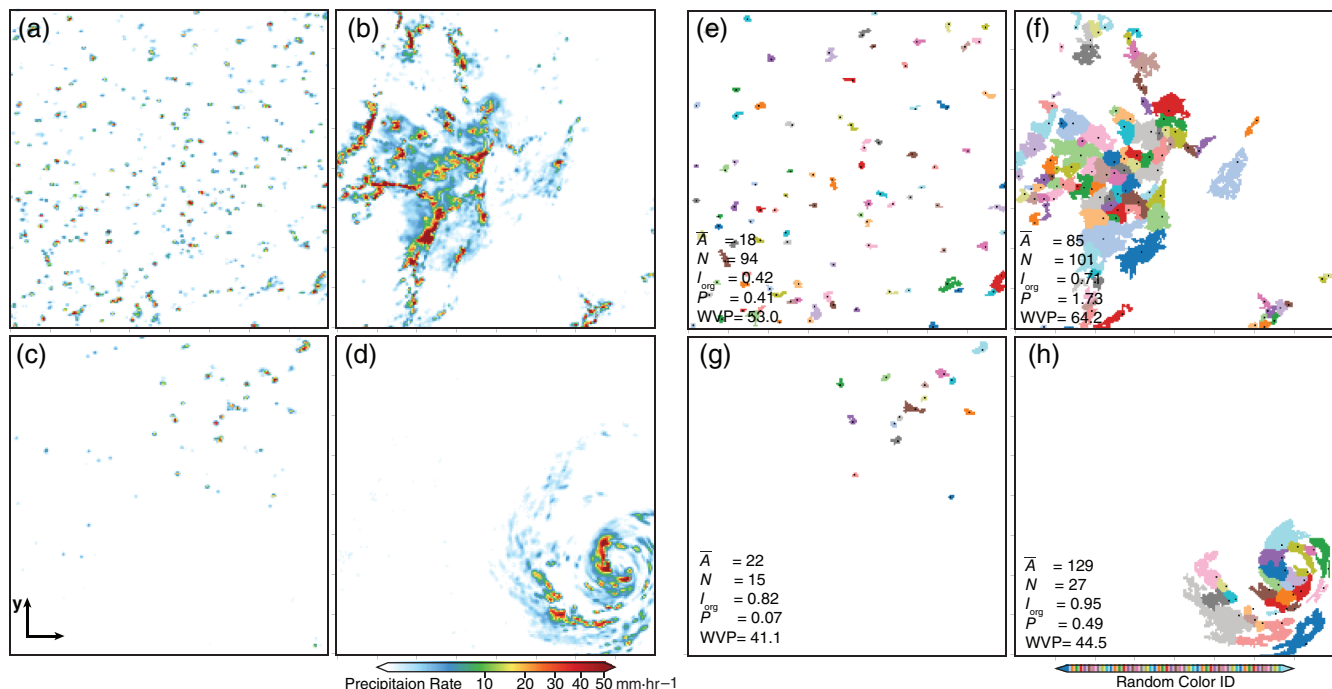

FIGURE 1 Precipitation cell watershed segmentation examples for four prototype synoptic cases. The (a)–(d) 2 × 2 scenes show the precipitation rate and the (e)–(h) 2 × 2 scenes the resulting cells (the ID of each cell is randomly coloured), together with the scene's properties $\overline{A}$ (pixel), N, I_{org}, precipitation (P, mm·hr^{-1}) and water vapour path (WVP, mm). The black dot in each labelled cell indicates the local maximum precipitation rate. One pixel corresponds to the grid box size of our regular latitude–longitude grid (0.05° by 0.05°)

0.9. The largest values of I_{org} are found for scenes with large $\overline{A}$ and small N, whereas small values of I_{org} are found in the opposite corner, for scenes with small $\overline{A}$ and large N. Few large cells in a scene (e.g., Figure 1h) tend to be more closely spaced than many small cells in a scene (e.g., Figure 1e).

Figure 2 further reveals that for a given N, I_{org} increases with increasing $\overline{A}$. In other words, larger cells in a scene appear as more clustered than smaller cells. This indicates that larger cells tend to not exist in isolation but as part of a larger convective object; compare Figure 1e to Figure 1f or Figure 1g to Figure 1h, where each of these figure pairs have a similar N. The positive relationship between I_{org} and $\overline{A}$ is especially pronounced for $\overline{A}$ values smaller than 50 pixels, whereas it flattens for larger values. The sharp increase in I_{org} with $\overline{A}$ by $\overline{A}$ values smaller than 50 pixels might express the transition from scenes with isolated cells to scenes being predominantly populated by precipitating objects made up of several cells whose sizes then just keep on expanding. The finite size of our scenes may be another explanation for the flattening of the I_{org} curve. When $\overline{A}$ is large, there is less free area and hence less possible arrangements of cells.

If $\overline{A}$ is instead held fixed, Figure 2 reveals that I_{org} increases with decreasing N. Few cells in a scene are unlikely to be randomly distributed. Thermodynamical considerations can help explain this behaviour. In a homogeneous scene, if there is one cell, there will be many other randomly distributed cells as well, as seen when conducting idealized simulations starting from homogeneous thermodynamic conditions. The only way to limit the number of cells in a scene is through the presence of inhomogeneities. Those inhomogeneities have to be large,
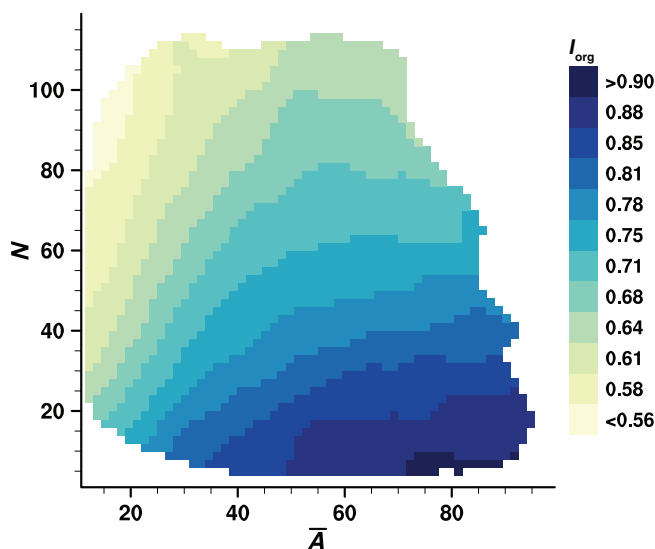
at least compared to any convective threshold, and will prevent the development of convection in some area of a scene. The scene will thus be associated with larger I_{org}. This is clearly visible by comparing Figure 1e to Figure 1g. Putting together the variations of I_{org} with $\overline{A}$ and N, we conclude that the degree of clustering mostly increases with decreasing number of cells for scenes containing cells larger than 50 pixels, whereas it depends both on the number of cells and on the mean area of cells for the remaining scenes. In this latter case, scenes with fewer cells or larger cells will appear as more clustered.

We now repeat a similar analysis, but considering the variations of precipitation P as a function of $\overline{A}$ and N (Figure 3). The precipitation is averaged over each scene. Note that this means that weak isolated precipitation objects, which are disregarded by the watershed segmentation procedure when deriving $\overline{A}$ and N, are included when computing P. The contribution of these objects to the overall precipitation is nevertheless small, amounting to 12%. Neglecting these objects in the computation of P yields a similar phase-space representation as displayed in Figure 3. If anything, the contours of precipitation even become more parallel to the contours of f_A.

Figure 3 indicates that largest P values are found for largest $\overline{A}$ and largest N. In general, contours of precipitation are hyperbolae that roughly follow contours of f_A. This is consistent with past studies (e.g., Doneaud *et al.*, 1984; Nuijens *et al.*, 2009; Davies *et al.*, 2013) which have shown that precipitation amounts correlate very well with the precipitating area and gives us confidence in our segmentation and analysis technique. As $f_A = N \cdot \overline{A}/A_{scene}$, with A_{scene} the area of a scene, it follows that, for a given N,
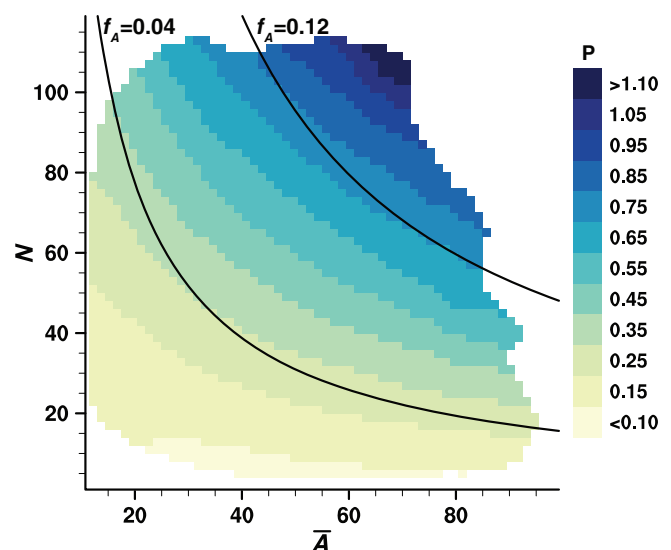

FIGURE 2 I_{org} index plotted in the phase space of mean cell area $\overline{A}$ (x-axis) and number of cells N (y-axis). Units for $\overline{A}$ are pixels

FIGURE 3 Scene-averaged precipitation P (mm·hr^{-1}) plotted in the $\overline{A}$–N phase space. The two black lines show contours of the total precipitating area fraction f_A in a scene

P should increase with $\overline{A}$, as indeed displayed by Figure 3. Likewise and as confirmed by Figure 3, for a given $\overline{A}$, P increases with N. The variations of P as a function of $\overline{A}$ and N as displayed by Figure 3 look also very similar to the precipitation variations obtained by Louf *et al.* (2019) based on 13-year of radar observations over Darwin.

More interesting is the comparison of Figures 2 and 3. Visual inspection and our previous discussion of Figures 2 and 3 already indicate that I_{org} and precipitation are not well correlated. The largest P amounts are found in the upper right corner of the $\overline{A}$–N phase space, by large $\overline{A}$ and large N, whereas the largest I_{org} values are found in the bottom right corner of the $\overline{A}$–N phase space, by large $\overline{A}$ but small N. The largest scene-averaged precipitation amounts do not occur in scenes with the most closely spaced cells, but actually occur in scenes with a weak degree of clustering of around 0.65.

To ease the comparison between P and I_{org}, we overlay the two fields in Figure 4. Figure 4 indicates that contours of P and contours of I_{org} cross each other. Scenes with vastly differing values of I_{org} can thus be associated with same P. Starting from the upper left corner of Figure 4 as an example and moving along a precipitation contour indicates that, keeping P constant in a scene, N decreases and $\overline{A}$ increases, so as to keep f_A constant. Given the phase space of I_{org} and its dependencies on N and $\overline{A}$, these changes in N and $\overline{A}$ induce corresponding changes in I_{org}. In that case, I_{org} increases. Tobin *et al.* (2012), who composited their observational scenes on similar precipitation amounts, could also observe a wide range of organizational states (quantified by SCAI) for a given precipitation amount. The fact that scenes with vastly differing spatial arrangements of their cells are associated with similar scene-averaged precipitation amounts is confirmed by the visual examples of Figure 1. Both the scene with randomly distributed isolated cells in Figure 1e and the one with closely packed cells in a hurricane (Figure 1h) display similar P.

To further quantify the dependency between P and I_{org}, Figure 4 also displays the deciles of the two distributions. To compute the precipitation decile, as an example, we rank the precipitation distribution and determine its deciles. As several scenes can have the same precipitation amounts, we then take the values of $\overline{A}$ and N averaged over those scenes as x and y coordinate to plot the deciles in Figure 4. The deciles in precipitation and I_{org} are perpendicular to each other (Figure 4). Following the decile line of precipitation in Figure 4 indicates that an increase in precipitation, from one decile to the next one, if anything is associated with a small decrease in I_{org}. An increase in the amount of precipitation in a scene namely comes with an increase in f_A. This implies an increase in the number of cells, which will tend to reduce I_{org}, and an increase in the mean area of the cells, which will tend to
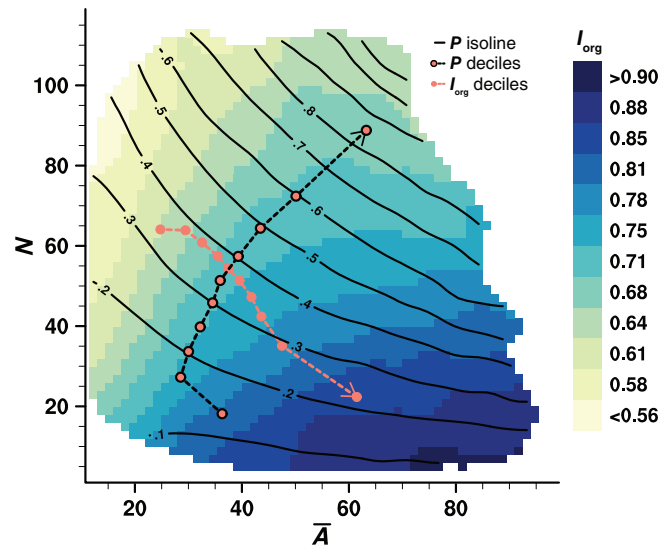

FIGURE 4 Combination of Figure 2 (shading) and Figure 3 (contours) with additional lines (dashed) indicating the deciles of the precipitation and of the I_{org} distribution (see text for more detail)

increase I_{org}. The two effects tend thus to compensate and leave I_{org} mostly unaffected. Likewise, following the decile line of I_{org} reveals first a slight increase of precipitation from the first to the second decile, then constant precipitation up to the fourth decile, and finally a decrease in precipitation. Hence, Figure 4 generally does not support our initial hypothesis that more closely spaced convective cells in a scene are associated with more precipitation. The dependency of scene-averaged precipitation on the spatial arrangement of its cells is weak. If anything, scenes with more closely spaced cells tend to rain less as they tend to be associated with a lower precipitating area fraction.

Even though precipitation and the spatial arrangement of its cells are not well correlated, the latter could still be important for the tropical climate if the scenes with the most highly packed cells (bottom right corner of Figure 4) would end up contributing the most to total precipitation. To investigate this, Figure 5 displays the frequency of occurrence of the scenes and their contribution to the precipitation averaged over all scenes. The frequency of occurrence rapidly decreases with increasing $\overline{A}$ and increasing N. The most frequent scenes have yet to be found in another corner of the $\overline{A}$–N phase space, namely in the bottom left corner by small N and small $\overline{A}$. Those scenes are characterized by weak clustering, with I_{org} values between 0.6 and 0.7, and very weak scene-averaged precipitation, around 0.15 mm·hr^{-1}. Given these very weak precipitation amounts, the scenes contributing the most to the total precipitation are vertically displaced in the $\overline{A}$–N phase space. Scenes with a moderate number of cells, between 40 and 60, and a small mean area, between 20 and 30 pixels, end up dominating the precipitation distribution. These scenes
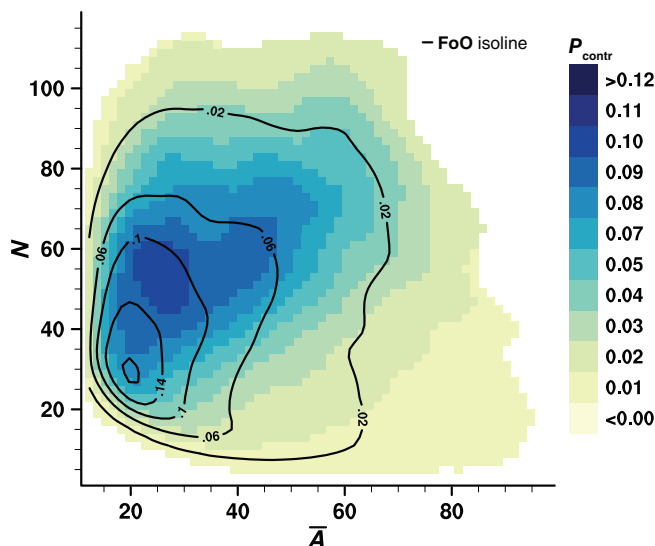
FIGURE 5 $\overline{A}$–N phase space showing the contribution of a particular $\overline{A}$–N pair to the precipitation averaged over all 1.37072×10^5 scenes (colour shading). Black contours indicate the frequency of occurrence (FoO) of a particular $\overline{A}$–N pair

FIGURE 6 $\overline{A}$–N phase space showing the scene-averaged water vapour path (WVP, solid contours, mm) and the deviation of the cell-averaged WVP from its scene mean value (colour shading, mm). Only scenes with precipitating cells are considered. The decile line of the scene-averaged water vapour path (dashed) is also included

precipitate between 0.2 and 0.3 mm·hr^{-1} and are weakly clustered with I_{org} being around 0.65. As can also be seen by comparing Figures 4 and 5, the contribution to the total precipitation increases in a direction roughly parallel to the curve of the precipitation deciles and perpendicular to the one of I_{org}. This means that an increase in the contribution to the total precipitation requires inclusion of scenes that precipitate more and not of scenes with more closely spaced cells. Lastly, Figure 5 indicates that both the scenes with extreme precipitation or with a high degree of clustering rarely occur. The contribution of the highly clustered scenes to the total precipitation is negligible given their rarity and their weak scene-averaged precipitation. In contrast, the scenes with extreme precipitation benefit from their high scene-averaged precipitation amount and contribute only slightly to the total precipitation.

4 | THE PHASE SPACE OF WATER VAPOUR PATH

Past studies (Bretherton *et al.*, 2004; Holloway and Neelin, 2009) have revealed a high degree of correlation between WVP (or column relative humidity) and precipitation, with precipitation picking up once the WVP reaches a given threshold. Also, the self-aggregation of convection has been shown to yield a moistening of the convective region and a drying of the surrounding environment in idealized studies of radiative convective equilibrium (e.g., Bretherton *et al.*, 2005; Muller and Held, 2012), a signature that has been confirmed in observations (Tobin
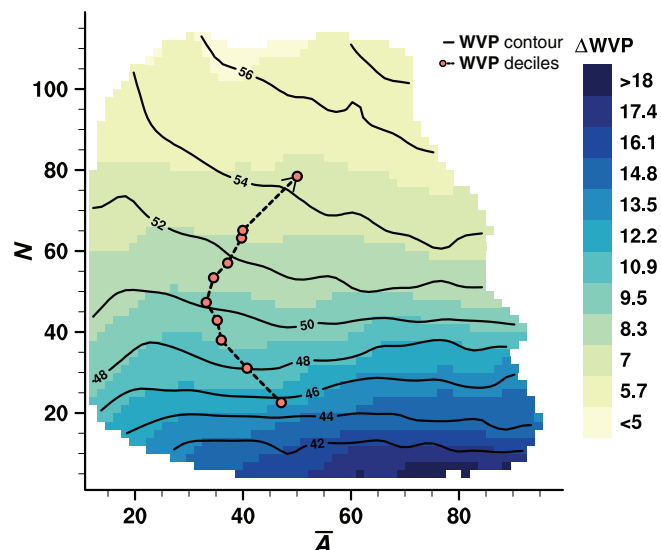
et al., 2012). We investigate these relationships using our simulation output and our $\overline{A}$–N phase-space representation (Figure 6). In particular, we investigate variations of the WVP averaged over a scene, which may be seen as more closely linked to the pick-up of precipitation, and variations of the deviation of the WVP between cells and scene, which may be more sensitive to clustering effects.

The most striking feature of Figure 6 is the fact that the number of cells, in contrast to the mean area of the cells, is strongly correlated to the WVP or, *vice versa*, that the WVP especially varies with the number of cells. This can be recognized from the small dependency of the WVP deciles upon $\overline{A}$ and from the flatness of the WVP contours in Figure 6, in particular for N values smaller than 50. Dry scenes can contain both small and large cells, but they can only contain few cells of each category. Larger cells likely benefit here from their higher degree of clustering (e.g., Figure 2), which will protect them from their hostile dry environment (e.g., Becker *et al.*, 2018). The moister a scene, the more cells it contains.

Given the small precipitating area fraction f_A, in particular in the bottom part of our phase space, contours of WVP firstly captures the behaviour of the WVP of the non-precipitating region. The WVP of the precipitating region can be implicitly deduced from the sum between the colour shading and the contour lines of Figure 6. This WVP increases both with N for a given $\overline{A}$ and with $\overline{A}$ for a given N. Its variation of about 7 mm across the $\overline{A}$–N phase space is smaller than the variation in environmental WVP

(about 12 mm). In fact, contours of the WVP averaged over the precipitating area follow contours of f_A (not shown).

Concerning the deviation of the WVP of the cells from the scene-averaged WVP, Figure 6 indicates that, to a first approximation, the deviations are anticorrelated to the moisture state of the atmosphere. The drier the scene, the larger the deviation. This follows from the fact that a cell has to be saturated to precipitate. Nevertheless closer inspection of Figure 6 reveals discrepancies from this general rule, discrepancies which might express the effects of clustering on the WVP, as further examined below by combining the results of Figures 4 and 6.

Combining the results of these two figures, a few more insights can be gained on the dependencies between precipitation, clustering, WVP, area of cells and number of cells. Not surprisingly given the results of studies on precipitation pick-up (e.g., Bretherton *et al.*, 2004; Holloway and Neelin, 2009), precipitation and WVP are correlated and both maximize in the upper right corner of our phase space. In contrast, both the degree of clustering and the deviations in WVP maximize in the bottom right corner of our phase space. This is consistent with the idea that a clustering of convection leads to a moistening of the convective region and to a drying of the non-convective region (e.g., Bretherton *et al.*, 2005; Tobin *et al.*, 2012).

Taking two scenes with the same mean precipitation amount but different WVP, the combination of Figures 4 and 6 also indicates that the drier scene has to contain larger and hence more clustered cells. The increase in $\overline{A}$ compensates for the smaller number of cells present in drier scenes, so as to keep f_A and P constant. In contrast, the strong increase in scene-averaged precipitation with WVP is primarily associated with a strong increase in the number of cells per scene. An increase in the number of cells in a scene, if anything, moves the distribution of cells closer to a random distribution. Finally, for two scenes with the same WVP, the larger scene-averaged precipitation amount is to be found in the scene with larger, more clustered cells, and potentially fewer cells.

5 | DISCUSSION

Our previous analysis concentrated on the tropical ocean. To assess the generality of our findings, we repeat our $\overline{A}$–N phase-space analysis for the tropical land areas (Figure 7), which confirms Figure 4 and the results previously obtained over the oceanic regions. Scenes with strongest precipitation are located in the upper right corner of the $\overline{A}$–N phase space, scenes with highest I_{org} values
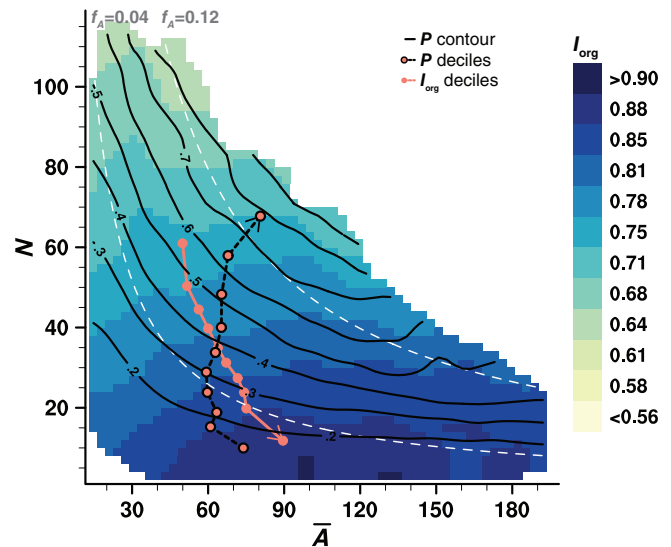

FIGURE 7 As Figure 4, but for land scenes, also including the two contours of $f_A = 0.04$ and 0.12 (white dashed). The total number of scenes is 4.9099×10^4. Note that the *x*-axis range is double that used in Figure 4

are found in the bottom right corner, and precipitation is poorly correlated to the degree of clustering. Pscheidt *et al.* (2019) also suggested a potential poor correlation between organization and precipitation by looking at simulations, radar and satellite observations over Germany.

Comparing in more detail the phase spaces over land and over ocean, a few interesting differences become apparent. In agreement with the known more frequent occurrence of MCSs over land than over ocean (e.g., Nesbitt *et al.*, 2006), the convection appears much more clustered over land (Figure 7) than over ocean (Figure 4) in our analysis as well. The systems should be larger over land than over ocean (Liu and Zipser, 2013), which remains true in our analysis, even if we split precipitation objects in their individual cells. This leads to a rather triangular $\overline{A}$–N phase space over land, in contrast to the squared population of the phase space over ocean. More importantly, over land, precipitation is not only poorly correlated to I_{org}, but it is almost negatively correlated to it, as can be seen by comparing the two decile lines in Figure 7. This results from the fact that I_{org} only very weakly depends upon $\overline{A}$ and mainly increases with decreasing N over land. This itself results from the generally higher values of I_{org} and of $\overline{A}$ over land as compared to ocean. Even over ocean, for high values of I_{org} and $\overline{A}$, I_{org} becomes independent of $\overline{A}$ (Figure 4).

Our previous analysis also required a segmentation of the precipitation field into objects and their individual cells, based on subjectively chosen precipitation thresholds (Section 2.2). The chosen thresholds were 0.1 mm·hr^{-1} to define the precipitation objects and

$10\,\mathrm{mm\cdot hr^{-1}}$ to split objects into their individual cells. To assess the robustness of our results, other thresholds of 0.01 and $1\,\mathrm{mm\cdot hr^{-1}}$ for the lower precipitation threshold (with an upper threshold of $10\,\mathrm{mm\cdot hr^{-1}}$), as well as of 5 and $15\,\mathrm{mm\cdot hr^{-1}}$ for the upper precipitation threshold (with a lower threshold of $0.1\,\mathrm{mm\cdot hr^{-1}}$) have been tested. These changes affect the sampled number of cells and the mean area of the cells in a scene, but do not affect the main conclusions of the study. Maximum precipitation is always to be found on the upper right corner of the $\overline{A}$–N phase space, maximum I_{org} on the lower right corner, and the decile lines of the two distributions tend to be perpendicular to each other. This robustness of the results follows from the fact that the precipitation amounts will follow to a first degree the precipitating area fraction, and hence will always maximize in the upper right corner of the $\overline{A}$–N phase space, independently of the chosen precipitation thresholds. Likewise, and as shown in Section 3, I_{org} maximizes for scenes with few and large cells, which will also remain true independently of the chosen thresholds.

6 | CONCLUSIONS

To help assess the importance of mesoscale organization for climate, we investigated in this study the relationship between mesoscale organization, precipitation and WVP. Based on a 40-day global storm-resolving simulation conducted with a grid spacing of 2.5 km, we quantified potential relationships using an object-based approach and analysed our results making use of a phase-space representation. To that aim, we defined 10° by 10° scenes and segmented each scene into its precipitation objects. The precipitation objects were further split into their individual cells using a watershed segmentation algorithm. Cells' attributes in a scene, namely mean area $\overline{A}$, number N, distances between cells, were recorded. The distances between cells were used to compute the clustering index I_{org} (Tompkins and Semie, 2017) which quantifies the degree of spatial arrangement of cells in a scene by measuring the deviation of their distribution from a random distribution. The cells' attributes $\overline{A}$ and N were used to display the results in an $\overline{A}$–N phase space so as to understand the variations of and relationships between I_{org}, precipitation and WVP.

Not surprisingly, the computed I_{org} values revealed that precipitation is almost always clustered over the tropical oceans. Looking at individual scenes, many small cells in a scene tend to be randomly distributed whereas few large cells in a scene appear as more clustered than expected from a random distribution. Also

in agreement with past studies which have shown that precipitation amounts are mostly determined by the precipitating area, scene-averaged precipitation appears as hyperbolae of constant precipitating area fraction in our $\overline{A}$–N phase-space representation.

Our initial hypothesis was that more closely spaced convective cells in a scene would be associated with a larger scene-averaged precipitation amount. Our results did not support that claim. First, the stronger precipitation amounts are found in the upper-right corner of the $\overline{A}$–N phase space, by high $\overline{A}$ and high N, whereas the scenes with more closely packed cells are found in the bottom-right corner of the $\overline{A}$–N phase space, by high $\overline{A}$ and low N. Second, scenes with extreme precipitation are characterized by a weak degree of clustering of around 0.65. Third, the scenes that contribute the most to the total precipitation of the tropical belt are scenes characterized by a similarly weak degree of clustering of 0.65. This last result may be seen at odds with past studies which have shown that mesoscale convective systems (MCSs), which might be interpreted as a typical example of organized convection, contribute predominantly to tropical precipitation. Nevertheless it should be noted that, by using an object-based rather than a phenomenological classification as used in those past studies, the link to the underlying convective phenomenon is lost. It is unclear where MCSs predominantly lie in our $\overline{A}$–N phase space. Moreover, I_{org} only measures organization based on the distances between cells, whereas phenomenological classifications emphasize the spatial extent of precipitation objects.

Determining the deciles of the I_{org} and precipitation distributions further confirmed the poor correlation between the two distributions. An increase in scene-averaged precipitation namely comes with an increase in the precipitating area fraction and hence with an increase in N and $\overline{A}$. When $\overline{A}$ increases, the degree of clustering increases, but when N increases, the degree of clustering decreases. The two effects mostly compensate so that moving along the curve of the precipitation deciles in our $\overline{A}$–N phase space does not result in significant changes in the spatial arrangement of cells in the scenes. In other words, scenes with randomly distributed cells or highly packed cells can precipitate the same amount.

Past studies have shown that precipitation exponentially increases with WVP. Our phase-space decomposition revealed that this is primarily a result from an increase in the number of cells in a scene and not from an increase in the mean area of the cells. This increase in the number of cells may be accompanied by a small decrease in the degree of clustering, if anything. Reciprocally, the number of cells is drastically reduced in dry scenes as compared to moister ones, but the mean area of the cells can remain unaffected. Dry scenes can contain both small and large

cells, but only few cells of each category. Hence, keeping precipitation constant while moving into drier scenes in our phase-space representation is accompanied by an increase in the mean area of the cells, which compensates for the reduction in the number of cells in drier scenes. An increase in the mean area of the cells also comes with an increase in the degree of clustering. In this sense, clustering may be seen as important to maintain precipitation amounts in drier atmospheres. Moreover, taking two scenes with the same WVP, the scene with the largest mean area of its cells, and hence the higher degree of clustering, exhibits the largest scene-averaged precipitation. This seems to be the only positive effects of clustering on the precipitation amounts from the different investigated aspects.

The above results were derived for the tropical belt comprised between $5°N \pm 20°$ and for oceanic regions only. Repeating the analysis for the land region confirmed the poor correlation between the degree of clustering and precipitation. In that case, I_{org} and precipitation are even anti-correlated, with an increase in precipitation accompanied by more randomly distributed cells in a scene.

Our analysis benefitted from the high spatial and temporal resolution of the simulation, as compared to the typical resolution of observational datasets. This allowed us to build a robust dataset and to consider organization even on very small scales. As a next step, it would be interesting to repeat the analysis using observations. The similarity of the simulated precipitation distribution in the $\overline{A}-N$ phase space to a similar type of distribution derived from radar observations taken over Darwin, as well as the poor correlation found between organization and precipitation using radar and satellite observations over Germany in a previous study, give us confidence that similar relationships should also be evident in observations spanning the full Tropics.

ORCID

Cathy Hohenegger https://orcid.org/0000-0002-7478-6275
Bjorn Stevens https://orcid.org/0000-0003-3795-0475

REFERENCES

Arakawa, A. and Wu, C.M. (2013) A unified representation of deep moist convection in numerical modeling of the atmosphere. Part 1. *Journal of the Atmospheric Sciences*, 70, 1977–1992.

Becker, T., Bretherton, C.S., Hohenegger, C. and Stevens, B. (2018) Estimating bulk entrainment with unaggregated and aggregated convection. *Geophysical Research Letters*, 45, 455–462.

Bony, S., Stevens, B., Frierson, D.M.W., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T.G., Sherwood, S.C., Siebesma, A.P., Sobel, A.H., Watanabe, M. and Webb, M.J. (2015) Clouds, circulation and climate sensitivity. *Nature Geoscience*, 8, 261–268.

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.

Bretherton, C.S., Peters, M.E. and Back, L.E. (2004) Relationships between water vapor path and precipitation over the tropical oceans. *Journal of Climate*, 17, 1517–1528.

Brune, S., Kapp, F. and Friederichs, P. (2018) A wavelet-based analysis of convective organization in ICON large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, 144, 2812–2829.

Davies, L., Jakob, C., May, P., Kumar, V.V. and Xie, S. (2013) Relationships between the large-scale atmosphere and the small-scale convective state for Darwin, Australia. *Journal of Geophysical Research*, 118, 11534–11545.

Doneaud, A., Ionescu-Niscov, S., Priegnitz, D.L. and Smith, P.L. (1984) The area–time integral as an indicator for convective rain volumes. *Journal of Climate and Applied Meteorology*, 23, 555–561.

Fiolleau, T. and Rocca, R. (2013) An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 4302–4315.

Hohenegger, C., Kornblueh, L., Klocke, D., Becker, T., Cioni, G., Engels, J.F., Schulzweida, U. and Stevens, B. (2020) Climate statistics in global simulations, from 80- to 2.5-km grid spacing. *Journal of the Meteorological Society of Japan*. https://doi.org/10.2151/jmsj.2020-005

Hohenegger, C. and Stevens, B. (2016) Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *Journal of Advances in Modeling Earth Systems*, 8, 1468–1482.

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

Houze, R.A. (2004) Mesoscale convective systems. *Reviews of Geophysics*, 42, 6885–6904.

Houze, R.A. and Betts, A.K. (1981) Convection in GATE. *Reviews of Geophysics*, 19, 541–576.

Kadoya, T. and Masunaga, H. (2018) New observational metrics of convective self-aggregation: methodology and a case study. *Journal of the Meteorological Society of Japan*, 96, 535–548.

Lebsock, M.D., L'Ecuyer, T.S. and Pincus, R. (2017) An observational view of relationships between moisture aggregation, cloud, and radiative heating profiles. *Surveys in Geophysics*, 38, 1237–1254.

Liu, C. and Zipser, E. (2013) Regional variation of morphology of organized convection in the tropics and subtropics. *Journal of Geophysical Research*, 118, 453–466.

Liu, C., Zipser, E.J., Cecil, D.J., Nesbitt, S.W. and Sherwood, S. (2008) A cloud and precipitation feature database from nine years of

TRMM observations. *Journal of Applied Meteorology and Climatology*, 47, 2712–2728.

Lopez, R.E. (1978) Internal structure and development processes of C-scale aggregates of cumulus clouds. *Monthly Weather Review*, 106, 1488–1494.

Louf, V., Jakob, C., Protat, A., Bergemann, M. and Narsey, S. (2019) The relationship of cloud number and size with their large-scale environment in deep tropical convection. *Geophysical Research Letters*, 46, 9203–9212.

Mapes, B., Tulich, S., Lin, J. and Zuidema, P. (2006) The mesoscale convection life cycle: building block or prototype for large-scale tropical waves?. *Dynamics of Atmospheres and Oceans*, 42, 3–29.

Moncrieff, M.W., Waliser, D.E., Miller, M.J., Shapiro, M.A., Asrar, G.R. and Caughey, J. (2012) Multiscale convective organization and the YOTC virtual global field campaign. *Bulletin of the American Meteorological Society*, 93, 1171–1187.

Moseley, C., Henneberg, O. and Haerter, J.O. (2019) A statistical model for isolated convective precipitation events. *Journal of Advances in Modeling Earth Systems*, 11, 360–375.

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.

Nesbitt, S.W., Cifelli, R. and Rutledge, S.A. (2006) Storm morphology and rainfall characteristics of TRMM precipitation features. *Monthly Weather Review*, 134, 2702–2721.

Nuijens, L., Stevens, B. and Siebesma, A.P. (2009) The environment of precipitating shallow cumulus convection. *Journal of the Atmospheric Sciences*, 66, 1962–1979.

Peters, K., Crueger, T., Jakob, C., and Möbis, B. (2017) Improved MJO-simulation in ECHAM6.3 by coupling a Stochastic Multi-cloud Model to the convection scheme. *Journal of Advances in Modeling Earth Systems*, 9, 193–219.

Pscheidt, I., Senf, F., Heinze, R., Deneke, H., Trömel, S. and Hohenegger, C. (2019) How organized is deep convection over Germany?. *Quarterly Journal of the Royal Meteorological Society*, 145, 2366–2384. https://doi.org/10.1002/qj.3552

Roca, R., Aublanc, J., Chambon, P., Fiolleau, T. and Viltard, N. (2014) Robust observational quantification of the contribution of mesoscale convective systems to rainfall in the tropics. *Journal of Climate*, 27, 4952–4958.

Senf, F., Klocke, D. and Brueck, M. (2018) Size-resolved evaluation of simulated deep tropical convection. *Monthly Weather Review*, 146, 2161–2182.

Stanfield, R.E., Jiang, J.H., Dong, X., Xi, B., Su, H., Donner, L., Rotstayn, L., Wu, T., Cole, J. and Shindo, E. (2016) A quantitative assessment of precipitation associated with the ITCZ in the CMIP5 GCM simulations. *Climate Dynamics*, 47, 1863–1880.

Stein, T.H., Holloway, C.E., Tobin, I. and Bony, S. (2017) Observed relationships between cloud vertical structure and convective aggregation over tropical ocean. *Journal of Climate*, 30, 2187–2207.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C.S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S.J., Putman, W.M., Shibuya, R., Neumann, P., Röber, N., Vanniere, B., Vidale, P.L., Wedi, N. and Zhou, L. (2019) DYAMOND: the Dynamics of the Atmospheric general circulation Modeled on Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, 6(61). https://doi.org/10.1186/s40645-019-0304-z

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.

Tompkins, A.M. and Craig, G.C. (1998) Radiative-convective equilibrium in a three-dimensional cloud-ensemble model. *Quarterly Journal of the Royal Meteorological Society*, 124, 2073–2097.

Tompkins, A.M. and Semie, A.G. (2017) Organization of tropical convection in low vertical wind shears: role of updraft entrainment. *Journal of Advances in Modeling Earth Systems*, 9, 1046–1068. https://doi.org/10.1002/2016MS000802

Weger, R.C., Lee, J., Zhu, T. and Welch, R.M. (1992) Clustering, randomness and regularity in cloud fields: 1. Theoretical considerations. *Journal of Geophysical Research*, 97, 20519–20536.

White, B.A., Buchanan, A.M., Birch, C.E., Stier, P. and Pearson, K.J. (2018) Quantifying the effects of horizontal grid length and parameterized convection on the degree of convective organization using a metric of the potential for convective interaction. *Journal of the Atmospheric Sciences*, 75, 425–450.

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*, 6, 59–74.