Cloud patterns have four interpretable dimensions

Martin Janssens[1], Jordi Vilà-Guerau de Arellano[2], Marten Scheffer[3], Coco Antonissen[4], Pier Siebesma[4], and Franziska Glassmeier[5]

[1]Wageningen University & Research
[2]Wageningen University Research
[3]Wageningen University
[4]TU Delft
[5]TU Delft

November 22, 2022

Abstract

Shallow cloud fields over the subtropical ocean exhibit many spatial patterns. The frequency of occurrence of these patterns can change under global warming. Hence, they may influence subtropical marine clouds' climate feedback. While numerous metrics have been proposed to quantify cloud patterns, a systematic, widely accepted description is still missing. Therefore, this paper suggests one. We compute 21 metrics for 5000 satellite scenes of shallow clouds over the subtropical Atlantic Ocean and translate the resulting dataset to its principal components (PCs). This yields a unimodal, continuous distribution without distinct classes, whose first four PCs explain 82% of all 21 metrics' variance. The PCs correspond to four interpretable dimensions: *Characteristic length*, *void size*, *directional alignment* and horizontal *cloud-top height variance*. These dimensions span a space in which an effective pattern description can be given, which may be used to better understand the patterns' underlying physics and feedback on climate.

Cloud patterns have four interpretable dimensions

Martin Janssens[1], Jordi Vilà-Guerau de Arellano[1], Marten Scheffer[1], Coco Antonissen[2], A. Pier Siebesma[23], Franziska Glassmeier[2]

[1]Wageningen University & Research
[2]Delft University of Technology
[3]Royal Netherlands Meteorological Institute

Key Points:

- Shallow cloud field patterns in satellite observations are quantified by 21 metrics and follow a unimodal, continuous distribution.
- Most existing metrics are redundant; 4 principal components capture 82% of the variance of 21 metrics.
- Characteristic length, void size, directional alignment and cloud-top height variance combine to effectively describe the patterns.

Corresponding author: Martin Janssens, martin.janssens@wur.nl

Abstract

Shallow cloud fields over the subtropical ocean exhibit many spatial patterns. The frequency of occurrence of these patterns can change under global warming. Hence, they may influence subtropical marine clouds' climate feedback. While numerous metrics have been proposed to quantify cloud patterns, a systematic, widely accepted description is still missing. Therefore, this paper suggests one. We compute 21 metrics for 5000 satellite scenes of shallow clouds over the subtropical Atlantic Ocean and translate the resulting dataset to its principal components (PCs). This yields a unimodal, continuous distribution without distinct classes, whose first four PCs explain 82% of all 21 metrics' variance. The PCs correspond to four interpretable dimensions: *Characteristic length*, *void size*, *directional alignment* and horizontal *cloud-top height variance*. These dimensions span a space in which an effective pattern description can be given, which may be used to better understand the patterns' underlying physics and feedback on climate.

Plain Language Summary

Satellite images show that clouds which develop in the lowest five kilometres of the atmosphere organise into many visually distinct patterns. Because different patterns have different radiative properties, a change in the relative occurrence of a pattern may influence Earth's response to warming. To study this effect, the patterns must first be quantified; numerous metrics have been developed for this task. In this paper, we compute 21 such metrics for 5000 cloud fields observed by satellite over the Atlantic Ocean east of Barbados. We show that the information contained in the 21 metrics can already very accurately be described by only 4 derived metrics, which capture a cloud field's typical cloud size, the size of connected clear sky patches, the clouds' degree of directional alignment and variance in cloud-top height. Combinations of these 4 metrics do not identify the existence of typical patterns, as previously suggested. However, they form a new, effective and interpretable pattern description, which can be used to better understand how cloud fields organise and how this impacts the wider climate system.

1 Introduction

Shallow cumulus clouds are the most abundant cloud type over the tropical oceans (Johnson et al., 1999), but result from many interacting processes and scales. This combination makes them the most uncertain aspect of how clouds will feed back onto a warming climate (e.g. Bony & Dufresne, 2005; Schneider et al., 2017). Several mechanisms that govern this response have recently been uncovered (Rieck et al., 2012; Bretherton, 2015; Klein et al., 2017). However, the origins and sensitivity of the rich spectrum of spatial patterns exhibited by shallow cloud fields has remained rather unexplored (Nuijens & Siebesma, 2019). Such spatial patterns alter precipitation distributions in cloud resolving simulations of deep convection in warmer conditions (Muller & Held, 2012; Tobin et al., 2012); recent research indicates that spatial patterning may influence the low cloud climate feedback too (Bony et al., 2020). Establishing this contribution of shallow cloud-field patterns and its underlying physics are therefore important research objectives.

The first step of such research is to classify or quantitatively measure any characteristic of the horizontal dimension of a shallow cloud field. Two comprehensive, complementing approaches were recently proposed: Expert visual inspection, which returns subjective, but interpretable classes of patterns (Stevens et al., 2019) and unsupervised machine learning, which is challenging to interpret, but gives more objectively inferred pattern measures (Denby, 2020). A third, more traditional approach is to compute one or more human-defined metrics; these are both interpretable and objective and are therefore considered in this paper.

Quantified patterns are often associated with a quantity called "organisation". This term has consequently taken on numerous interpretations. It is often synonymous with "aggregation" in studies of deep convection (Tobin et al., 2012; White et al., 2018; Holloway et al., 2017), sometimes characterised as the regular, random or clustered structure of nearest neighbour distances of cloud objects (Weger et al., 1992; Seifert & Heus, 2013; Tompkins & Semie, 2017), or connected to cloud scale (Neggers et al., 2019; Bony et al., 2020). However, cloud field organisation has also been defined by metrics of fractal analysis (Cahalan & Joseph, 1989), directional alignment (Brune et al., 2018), subcritical percolation (Windmiller, 2017) or spatial variance (de Roode et al., 2004; Wood & Hartmann, 2006). While this makes it difficult to objectively define and discuss organisation, all these interpretations share the same aim: Quantifying cloud field patterns. Hence, this diversity can potentially also be harnessed to distinguish between different patterns.

The aim of this paper is therefore to systematically extract the independent information encapsulated by the set of metrics associated with "cloud field organisation" in literature, and to use this information to describe and interpret cloud field patterns as effectively as possible. We first compute 21 diverse metrics for 5000 satellite observations of mesoscale cloud fields in the trades and synthesise these in a multivariate distribution (section 2). Next, we show that these metrics vary primarily along 4 principal components, allowing drastic dimensionality reduction (section 3.1). Analysis of these main principal components results in a pattern description that is remarkably effective, in addition to being interpretable and objective (section 3.2). We then highlight several approaches to approximate the principal components that balance the description's complexity and accuracy (section 3.3). Finally, we demonstrate and discuss the ability of our description to characterise previously diagnosed and novel regimes of characteristic patterns (section 3.4), before concluding (section 4).

2 Constructing a cloud field pattern distribution

2.1 Data

Following Stevens et al. (2019) and Bony et al. (2020), we concentrate on shallow, subtropical clouds in the marine North Atlantic trades east of Barbados (20°-30°N, 48°-58°W), which are representative for the entire trades (Medeiros & Nuijens, 2016). Our cloud fields stem from the MODIS instrument borne by NASA's Aqua and Terra satellites. Specifically, we sample daytime overpasses during December-May 2002-2020 and directly use the level 2 Cloud Water Path (CWP), Cloud-Top Height (CTH) and cloud mask products at 1km resolution (Platnick et al., 2015) as basis for our metrics. Fig. S1 shows that the results are not overly sensitive to resolution. We only interpret pixels classified as "confidently cloudy" by the cloud mask algorithm as cloud.

Our data points are scenes of cloud fields, which are 512km$\times$512km subsets sampled within the 10°$\times$10° observation region. To boost the size of our dataset, scenes are allowed to overlap 256km. We attempt to minimise the impact of errors and biases in remotely sensed cloud products by rejecting scenes with i) high clouds such as cirrus wisps, if more than 20% of the clouds' tops lie above 5km, ii) overly large sensor zenith angle, if this angle exceeds 45°, following e.g. Wood and Field (2011) and iii) sunglint errors, manually excluding scenes where these are visually found to influence the cloud mask. A set of 5004 scenes remains.

2.2 Metrics and dimensionality reduction

To appropriately capture the body of existing organisation metrics, we require them to meet either of the following two criteria: i) Are they perceived to capture a unique aspect of the patterns? or ii) do they frequently recur or recently first appear in literature? Additionally, they must be easy to interpret. This procedure (see tab. S1 for details) diagnoses 21 metrics, which broadly divide into three categories: Statistical moments of

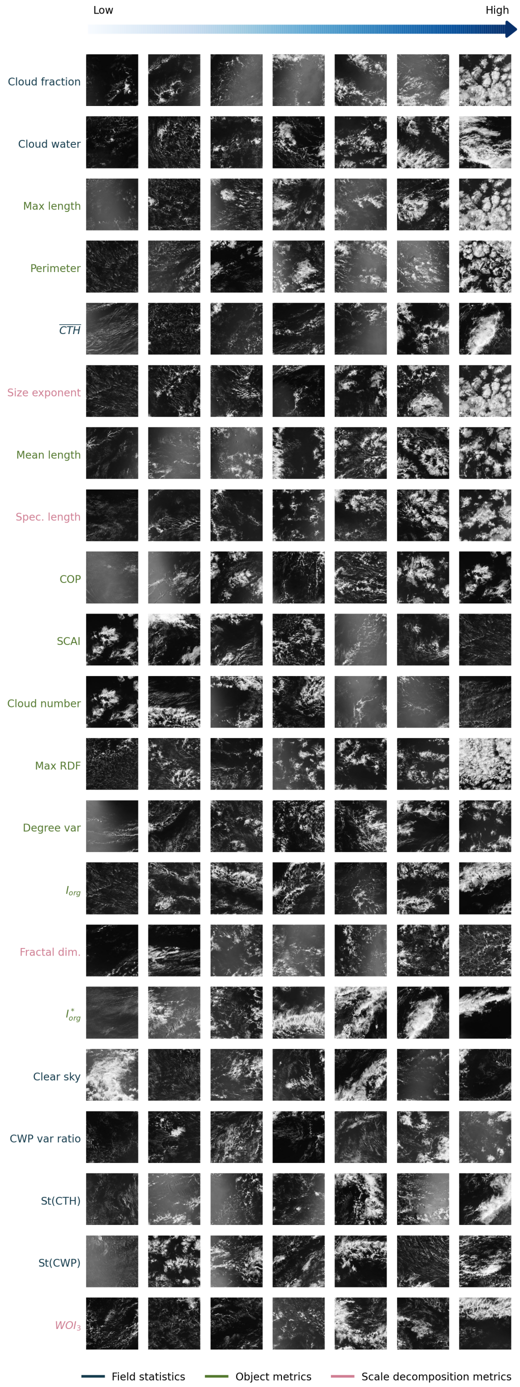

Figure 1. Visual representation of scenes ordered by metrics derived from three categories (text colour) and sampled at linear intervals. Bright backgrounds stem from sunglint, which is accounted for in metric computations.

physical cloud field properties, object-based metrics, and attributes of scale decompositions. These metrics are briefly introduced below, visually presented in fig. 1 and further detailed in Text S1.

Statistical moments of cloud field properties comprise measures of typical cloud mass and area: The cloud mask's coverage fraction (Cloud fraction), the CWP's scene integral (Cloud water) and standard deviation (St(CWP)) and the variance ratio for "mesoscale aggregation" of moisture proposed by Bretherton and Blossey (2017) (CWP var. ratio), here applied only to cloud water. Furthermore, this class contains measures of the clouds' vertical extent: The mean and standard deviation of cloud-top height ($\overline{\text{CTH}}$ and St(CTH) respectively).

Object-based metrics measure size, shape and relative positioning of individual cloud segments, which are identified from cloud mask fields using 4-connectivity labelling. To avoid artefacts at the resolution scale, objects of a smaller dimension than four times the instrument resolution are ignored. Our results are not sensitive to the chosen connectivity scheme or minimum object size (see fig. S1). The resulting metrics further divide into two categories: Scene statistics of individual object properties and measures of the spatial distribution of the objects. The first category includes the mean and maximum object length (Mean length, Max length), the number of objects (Cloud number) and the mean object perimeter (Perimeter); the second comprises the Simple Convective Aggregation Index (SCAI) (Tobin et al., 2012), Convective Organisation Potential (COP) (White et al., 2018), the peak of the average radial distribution function (Rasp et al., 2018) (Max RDF), the degree variance (Degree var) of the cloud objects' nearest-neighbour network representation (Glassmeier & Feingold, 2017) and the Organisation Index (I_{org}) (Weger et al., 1992), of which we include two versions. The first, most commonly applied form, compares the cloud field nearest-neighbour cumulative density function (NNCDF) to a Weibull distribution. The second variant (I^*_{org}) compares it to an inhibition NNCDF that accounts for object size and therefore is less likely to erroneously predict regularity in the cloud fields (Benner & Curry, 1998). This metric is similar to that introduced by Pscheidt et al. (2019).

We compute four metrics from scale decompositions: The size exponent of the cloud object size distribution modelled as a power law (Size exponent), the box-counting dimension of cloud boundaries in the cloud mask field (Fractal dim.), the Spectral length scale as defined by Jonker et al. (1999) and the deviation of variance from the mean in the horizontal, vertical or diagonal orientations of the cloud water field's stationary wavelet spectrum (WOI_3) (Brune et al., 2018). In this paper, we use these metrics as discriminators between individual cloud fields, not to measure their cumulative scaling properties. Finally, we introduce a novel metric: A scene's largest, rectangular, contiguous cloud-free area (Clear sky), as a simple measure of *lacunarity*, the degree to which continuous areas without clouds dominate a scene.

We describe patterns as a linear combination of the computed metrics, which are standardised to weight them equally. Since many metrics in fig. 1 strongly correlate (see fig. S2), they are treated to a Principal Component Analysis (PCA, e.g. Abdi and Williams (2010)). This transforms the metrics to an orthogonal basis whose components (principal components - PCs) explain the maximum variance in the dataset. If a *small* number of PCs (orthogonal dimensions) can accurately capture the metric set's variance, these form an effective pattern description.

3 Describing patterns

3.1 A four-dimensional pattern distribution

Figure 2 shows uni- and bivariate kernel density estimates on planes spanned by the first four PCs of the metric distribution, annotated with the fractional variance of the dataset explained by each PC (explained variance ratio - EVR). It reveals that multiple

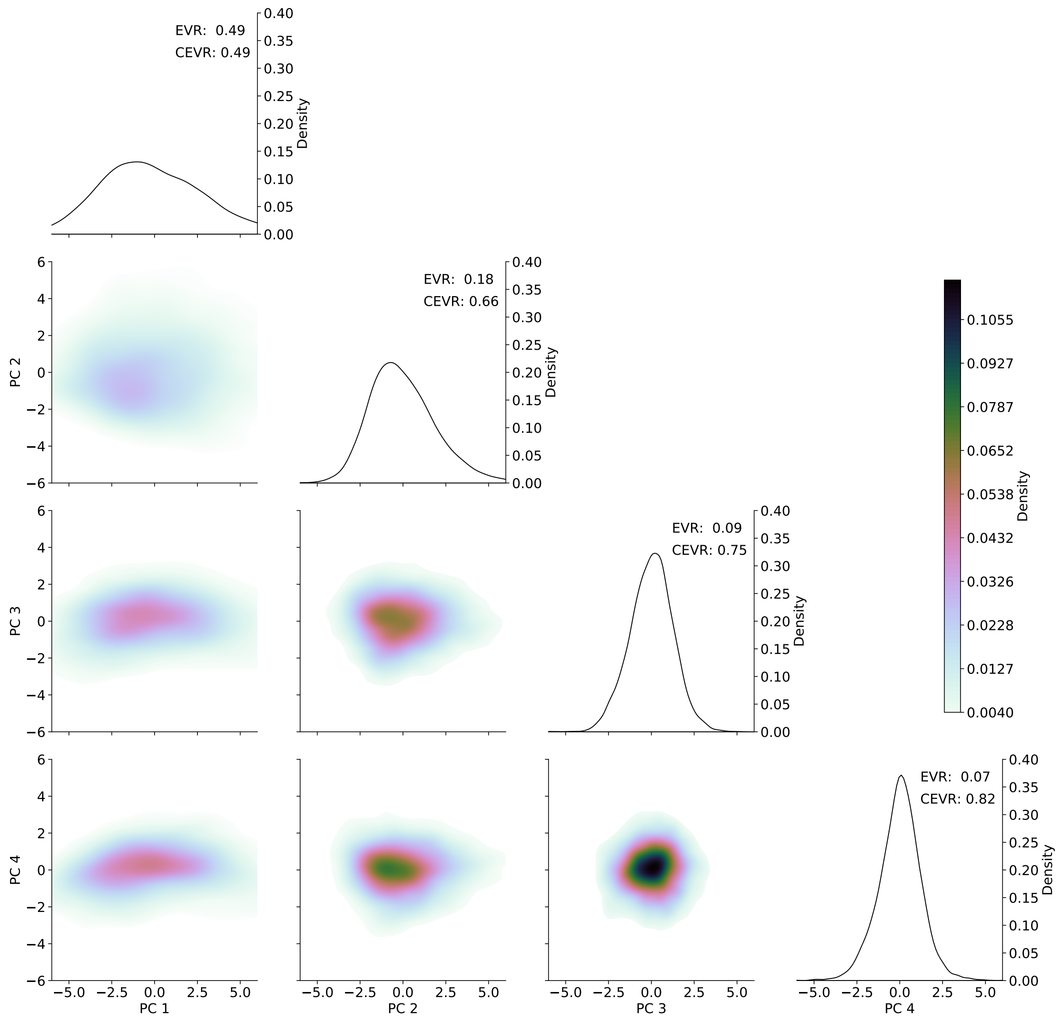

Figure 2. Univariate (diagonal, density on y-axis) and bivariate (off-diagonal, density in colour) Gaussian kernel density estimates of the first four principal components (PCs) of the pattern distribution. The annotations EVR and CEVR denote the individual and cumulative explained variance ratio of each PC, respectively. Bandwidths for the Gaussian kernels are computed using Scott's rule (Scott, 1992).

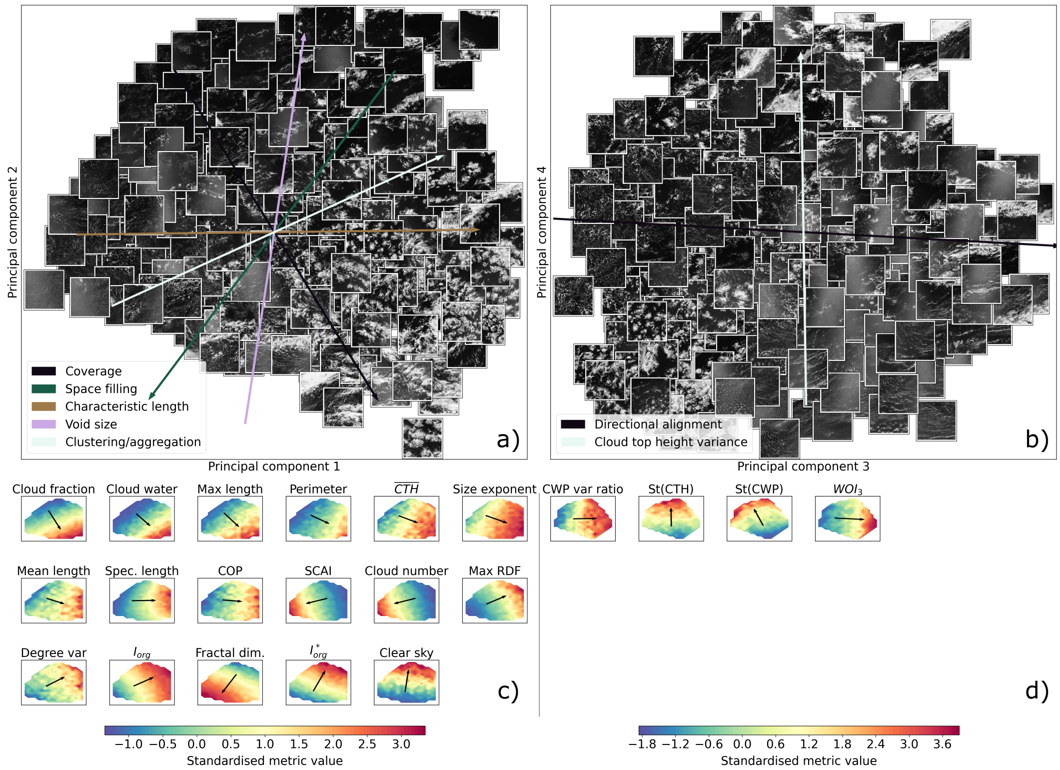

Figure 3. Top: Images of scenes projected onto planes spanned by the first and second (a) and third and fourth (b) PCs of the metric distribution, overlaid by arrows oriented along the mean gradient of several metric groups (see main text). Bottom: Filled contours of standardised metric values that have in excess of 50% of their variance explained by the first (c) and second (d) plane, constructed by piecewise linear barycentric interpolation and overlaid by an arrow pointing along the mean gradient. Subfigures b) and d) are rotated counter-clockwise by 49° in-plane to improve clarity of visualisation.

PCs (dimensions) are needed to capture the multivariate distribution's cumulative EVR (CEVR) appropriately. However, the first PC is by far the most influential (EVR=0.49 - widest distribution). Furthermore, the CEVR of the first two PCs already rises to 0.66, while including 3 and 4 of the 21 original dimensions explains 75% and 82% of the dataset's variance, respectively. After the fourth PC, EVR quickly deflates (PCs 5-9 have EVRs of 0.04, 0.03, 0.03, 0.02, 0.02), dropping below 0.01 after the tenth PC (fig. S3). These statistics show that four PCs effectively capture the information in all 21 metrics. Therefore, we reduce our 21-dimensional metric set to these four PCs.

Of course, truncating the PCA after precisely four components remains somewhat arbitrary. Yet, this choice strikes a useful balance between including enough dimensions to effectively describe patterns and sufficiently few dimensions to interpret them. This claim is visually supported by fig. 3 a) and b) (fig. S3 adds quantitative evidence): Combinations of PC1 and PC2 (fig. 3 a) consistently and coherently position visually similar (different) scenes close to (far from) each other. PC3 and PC4 (fig. 3 b) ably reveal further distinctions. Hence, linear combinations of these four PCs form an effective pattern description.

3.2 An interpretable pattern description

Our four-dimensional pattern description is not only effective; by relating the PCs to their underpinning metrics, it can also be interpreted. This interpretation is facilitated by fig. 3 c) and d), which show the in-plane gradient and mean direction of change of metrics that predominantly vary in the planes depicted in fig. 3 a) and b) respectively. By averaging the gradients of several similarly varying metrics, we identify a meaningful vocabulary that labels several directions of change in the two planes (arrows in fig. 3 a and b). Using this vocabulary, we name the principal components and relate them to several common interpretations of organisation.

Strikingly, 17/21 metrics mainly describe variations in the first two PCs (fig. 3 c, see also fig. S4). These metrics derive from all three categories (field statistics, objects and scale decompositions) and point in a rather continuous spectrum of directions, offering a remarkable number of interesting choices for interpreting fig. 3 a):

1. *Coverage* (Arrow in fig. 3 a represents the mean gradient of Cloud fraction, Max length and Cloud water)
2. *Space filling* (Fractal dim., I^*_{org})
3. *Characteristic length* (Spectral length scale, Size exponent, Mean length)
4. *Void size* (Clear sky)
5. *Aggregation* or *clustering* (I_{org}, SCAI, Cloud number, Max RDF), as commonly associated with deep convective organisation (Tompkins & Semie, 2017; Tobin et al., 2012).

We adopt the two directions that best align themselves with the PCs as names for our pattern description's first two dimensions: *Characteristic length* and *void size*. We find it both intuitive and beautiful that these two dimensions, which respectively measure the typical scale of clouds and the complementary clear sky space between them, naturally emerge from our approach.

Linear combinations of the PCs can construct different terms in the quintet above. For instance, *clustering/aggregation* differs only subtly from *characteristic length*, assigning slightly more importance to voids between cloud clusters. *Space filling* weights voids even more heavily. Finally, *coverage* distinguishes itself from *void size* by assigning marginally more importance to characteristic length. Hence, the same aspects of the patterns in fig. 3 a) can be described with different pairs of terms.

Several such pairs are already indirectly recognised as central traits of "organisation". For instance, Seifert and Heus (2013) suggest that both a spectral length scale (*characteristic length*) and I_{org} (*clustering*) may be needed to discriminate between various modes of organisation; Neggers et al. (2019) identify organisation as a combination of maximum cloud size (*coverage*) and typical nearest-neighbour distances between smaller clouds (*space filling*); chapter 5 of van Laar (2019) distinguishes "cloud field characteristics" (cloud fraction, maximum cloud size - *coverage*) from "organisation parameters" (I_{org}, SCAI, COP - *clustering*) and Bony et al. (2020) span their planar description of organisation with mean length (*characteristic length*) and I_{org} (*clustering*). The arrows in Figure 3 relate all these interpretations to each other.

However, our four-dimensional pattern description goes beyond these common, two-dimensional interpretations of organisation. Figure 3 d) shows that the third and fourth PC distinguish patterns with different directional alignment (WOI_3) of the scene's larger scales (CWP var ratio) and those with different horizontal variance of vertical cloud development (St(CTH)). Hence, variations in PC 3 and PC4 can be understood as combinations of *directional alignment* and *cloud-top height variance.*

Summarising, we propose to think of cloud field patterns, as described by organisation metrics, as a linear combination of the 4 PCs: *Characteristic length*, *void size*, *directional alignment* and *cloud-top height variance*, each term contributing a dimension that is uncorrelated to the others. However, many valid interpretations exist, especially of the first two dimensions. Therefore, the consistency with which "organisation" is understood can be considerably advanced by using the relationships between the various interpretations established in this section.

3.3 Selecting metric subsets

While four PCs describe patterns remarkably well, they still require input from *all* metrics. If added interpretability or less computation is desired, one might approximate the PCs with a subset of metrics. This approach challenges each chosen metric to do considerably more work than merely inspiring an interpretation of the PCs, as in the previous section, since no metric subset is fully orthogonal or optimally variance-capturing. Moreover, it is often not obvious that a given metric is much better suited to approximate a PC than a similarly varying one. This problem is illustrated by applying *sparse* PCA (Zou et al., 2006) to our data. Despite optimising a cost function that explicitly balances the accuracy of the approximate PCs with how many metrics contribute to them, this technique cannot robustly indicate metric subsets (see fig. S5).

One practical way to compose a subset nonetheless is choosing one metric that most closely correlates to each PC (Cadima & Jolliffe, 1995). This approach selects the Spectral length scale, Clear sky, WOI_3 and St(CTH) (CEVR=0.59) and is a reasonable approximation of the PC description (CEVR=0.82). If one's primary interest is in the first two dimensions of the pattern distribution, several roughly orthogonal metric pairs competently estimate the plane in fig. 3 a). Examples include Spectral slope and Clear sky (CEVR=0.31), Cloud fraction and Fractal dim. (CEVR=0.31) or Perimeter and I^*_{org} (CEVR=0.30). All three pairs sacrifice explained variance compared to two PCs (CEVR=0.66). Yet, they capture far more information than various metric combinations considered in literature, e.g. Cloud number and I_{org} (Bony et al., 2020, CEVR=0.18), I_{org} and Fractal dim. (Denby, 2020, CEVR=0.20), Spec. Length and I_{org} (Seifert & Heus, 2013, CEVR=0.19) or I_{org}, SCAI, COP and Max RDF (van Laar, 2019, CEVR=0.26). Therefore, we recommend to always assess the orthogonality and EVR of one's metrics with a PCA, before optionally selecting a metric subset that approximates their desirable properties appropriately.

3.4 Regimes of patterns

Asking how many dimensions cloud field patterns possess is not equal to asking how many fundamental types of cloud patterns exist. Dividing clouds into distinct classes (e.g. cumulus or cirrus) is a classical approach, which recently inspired efforts to also classify shallow cloud field patterns, using both the human eye (Stevens et al., 2019) and metrics (Bony et al., 2020). We compare our pattern description to these classes ("sugar", "gravel", "fish" and "flowers") by identifying seven k-means clusters in the four-dimensional PC distribution (fig. 4).

Scenes arguably dominated by "sugar" and "gravel" reside in clusters 5 (brown) and 3 (maroon). These patterns should, in the terminology from section 3.2, be understood as small-scale with rather small voids (or disaggregated/unclustered); "gravel" distinguishes itself through its higher cloud-top height variance and low directional alignment (see also left side of fig. 3 b). Cluster 1 (navy) comprises i.a. "fish", which shares gravel's void size, cloud-top height variance and low degree of directional alignment, only at larger scales. Finally, one may see "flowers" in cluster 7 (blue), as large-scale, aggregated structures with little directional alignment and low cloud-top height variance.

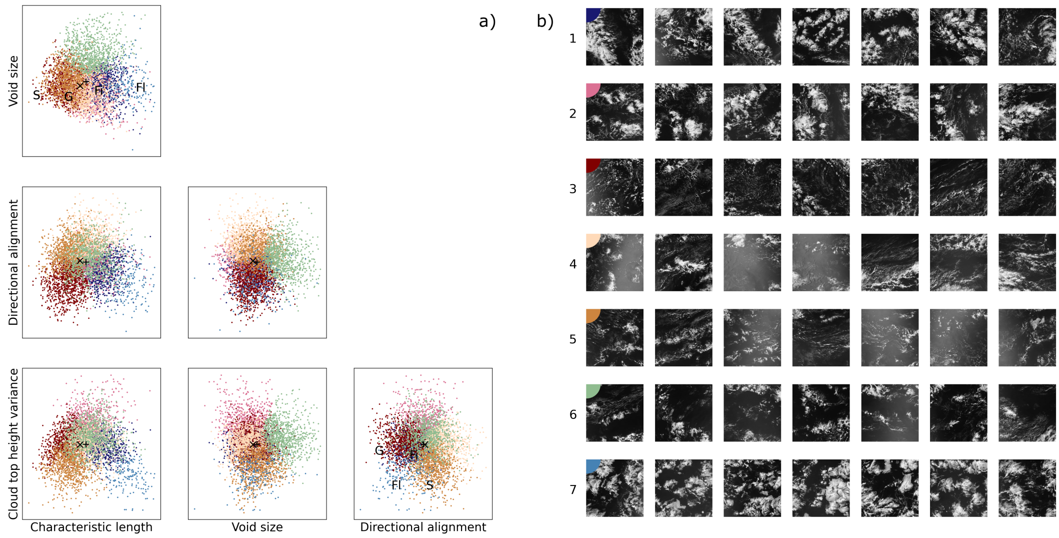

Figure 4. Seven regimes of the 4D pattern description, identified as k-means clusters of different colour: a) Scenes scattered over planes defined by the first four PCs, each normalised to unit variance, named using the convention from section 3.2; b) seven examples of scenes in each regime. Pluses and crosses indicate the distribution's mean and mode, respectively. S, G, Fi and Fl suggest typical locations for the "sugar", "gravel", "fish" and "flowers" patterns diagnosed by Stevens et al. (2019), in the two planes shown in fig. 3, determined by eye.

The natural emergence of these regimes from our systematic metric analysis is encouragingly consistent with human pattern identification (Stevens et al., 2019) and solidifies Bony et al. (2020)'s conclusion that these patterns can be objectively identified. However, even in an unrealistic scenario where all scenes in these four regimes could unambiguously be labelled sugar, gravel, fish or flowers, they would contain only 52% of the observations in our dataset. Figure 4 indicates several other regimes that differ in important regards. For instance, many scenes possess vast voids (cluster 6, sea green). In this regime, clouds likely affect the region's climate much less than sugar, gravel, fish or flowers, which all have higher cloud cover. Analyses of the patterns' climate sensitivity must probably consider this and other different regimes explicitly.

In fact, pattern classification is itself an approximation. The pattern distribution is *unimodal* and *continuous* (fig. 2), and therefore does not inherently possess multiple "classes", "clusters" or "modes". Breaking the continuum into clusters neglects subtly different patterns within a cluster. For instance, the band-like sub-regime at high directional alignment in fig. 3 b) falls within cluster 4 (peach) in fig. 4, even if this sub-regime is visually distinct from all displayed scenes in cluster 4. To capture such subtleties, we recommend shifting focus from regimes, classes or clusters of patterns to a more fitting, continuous representation.

Finally, many of the human-identified patterns (sugar, flowers) appear on our distribution's extremes (see also fig. 3 a) and b)). While this may explain why they are most easily distinguished by humans, they lie far from the distribution's statistical mean and mode (indicated by pluses and crosses respectively in fig. 4 a) and are thus not typical. Instead, the modal pattern is partial to smaller scales and voids, which characterise scenes with shallow, cold-pool dominated convection (clusters 3, 5) or processes on a wide range of scales (cluster 4); this space may be most relevant to the climatology of patterns.

4 Conclusion and outlook

Research on the climate feedback of patterns in shallow trade-wind cloud fields requires a consistently understood description of those patterns. In this paper, we have systematically developed such a description for square, 500 km^2 satellite-observed cloud fields east of Barbados. By projecting one new and 20 previously developed organisation metrics onto a set of PCs, we show that cloud patterns can be effectively described as a 4-dimensional, linear combination of *characteristic length*, *void size*, *directional alignment* and *cloud-top height variance*. This description is objective and interpretable, in contrast to direct unsupervised machine learning (objective, not usually interpretable) or human pattern identification (interpretable, not objective). It also demonstrates that patterns follow a continuous, unimodal distribution without distinct classes and that visually striking patterns are extreme, rather than typical. Future studies of the physics behind and climate impact of shallow cloud field patterns can therefore rely either on our PCs or, if accuracy is less important, on metrics that correlate closely to them.

The effectiveness of our approach may well extend to descriptions of deep convective organisation. Many relationships between our metrics are consistent with those found for deep convective cloud fields (Rempel et al., 2017; Brueck et al., 2020), suggesting that an effective, low-dimensional description of deep convective organisation is attainable. Our pattern description could also be used for forecast verification (Jolliffe & Stephenson, 2012), using the pattern distribution's dimensions as matching criteria between model and observation in similar fashion to e.g. the criteria developed by Wernli et al. (2008). In turn, the forecast verification community may offer useful insights to descriptions of cloud field patterns.

Finally, our approach can itself be refined in several regards. First, using predefined metrics to describe patterns leaves potentially undiscovered information from the description. Therefore, it may be fruitful to compare our approach to more unsupervised machine learning (e.g. Denby, 2020). However, the completeness of a pattern description should ideally be assessed in terms of how fully the underlying processes are separated. This requires process-resolving numerical simulations and/or temporally evolving observations, which link the evolution of the pattern continuum to that of the atmospheric state. Next, our conclusions are tied to our observation scales (1-500 km), meaning that we may inadequately capture this scale window's extremes. Furthermore, we treat this scale window in an integral sense and ignore patterns that appear on one scale, but may be cancelled by another (Nair et al., 1998). Hence, a further refinement could be to consider pattern distributions on a per-scale basis. Lastly, some subjectivity will likely remain in how different researchers interpret "organisation". This attests the richness of the underlying patterns, which we hope remains appreciated.

Acknowledgments

MJ and FG acknowledge support by The Branco Weiss Fellowship - Society in Science, administered by ETH Zürich; FG also acknowledges a Veni grant from the Dutch Research Council (NWO). A.P.S acknowledges funding by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 820829 (CONSTRAIN project). The Aqua and Terra MODIS Clouds 5-Min L2 Swath 1km products, which contain the cloud mask, cloud water path and cloud-top height products used in this study, were extracted from NASA's Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) (`http://dx.doi.org/10.5067/MODIS/MYD06_L2.061`; `http://dx.doi.org/10.5067/MODIS/MOD06_L2.061`). All download, preprocessing, metric computation and analysis was done using Python and its Numpy (van der Walt et al., 2011), Pandas (McKinney, 2010) and Scipy (Virtanen et al., 2020) libraries. Field processing and segmentation is done with Scikit Image (van der Walt et al., 2014), PCA and clustering with Scikit Learn (Pedregosa et al., 2011), sparse PCA with Ristretto (Erichson et al., 2020) and plots with Matplotlib (Hunter, 2007) and Seaborn (Waskom & the seaborn de-

velopment team, 2020). The integrated code, along with detailed instructions on how to run it, are available in a living GitHub repository (`https://github.com/martinjanssens/cloudmetrics`) and its frozen image at the time of submission (`https://doi.org/10.6084/m9.figshare.12687302.v1`).

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, *2*(4), 433–459.

Benner, T. C., & Curry, J. A. (1998). Characteristics of small tropical cumulus clouds and their impact on the environment. *Journal of Geophysical Research: Atmospheres*, *103*(D22), 28753–28767.

Bony, S., & Dufresne, J.-L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, *32*(20).

Bony, S., Schulz, H., Vial, J., & Stevens, B. (2020). Sugar, gravel, fish and flowers: Dependence of mesoscale patterns of trade-wind clouds on environmental conditions. *Geophysical Research Letters*.

Bretherton, C. (2015). Insights into low-latitude cloud feedbacks from high-resolution models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2054), 20140415.

Bretherton, C., & Blossey, P. (2017). Understanding mesoscale aggregation of shallow cumulus convection using large-eddy simulation. *Journal of Advances in Modeling Earth Systems*, *9*(8), 2798–2821.

Brueck, M., Hohenegger, C., & Stevens, B. (2020). Mesoscale marine tropical precipitation varies independently from the spatial arrangement of its convective cells. *Quarterly Journal of the Royal Meteorological Society*, *146*(728), 1391–1402.

Brune, S., Kapp, F., & Friederichs, P. (2018). A wavelet-based analysis of convective organization in ICON large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2812–2829.

Cadima, J., & Jolliffe, I. T. (1995). Loading and correlations in the interpretation of principle compenents. *Journal of applied Statistics*, *22*(2), 203–214.

Cahalan, R. F., & Joseph, J. H. (1989). Fractal statistics of cloud fields. *Monthly weather review*, *117*(2), 261–272.

Denby, L. (2020). Discovering the importance of mesoscale cloud organization through unsupervised classification. *Geophysical Research Letters*, *47*(1), e2019GL085190.

de Roode, S. R., Duynkerke, P. G., & Jonker, H. J. (2004). Large-eddy simulation: How large is large enough? *Journal of the atmospheric sciences*, *61*(4), 403–421.

Erichson, N. B., Zheng, P., Manohar, K., Brunton, S. L., Kutz, J. N., & Aravkin, A. Y. (2020). Sparse principal component analysis via variable projection. *SIAM Journal on Applied Mathematics*, *80*(2), 977–1002.

Glassmeier, F., & Feingold, G. (2017). Network approach to patterns in stratocumulus clouds. *Proceedings of the National Academy of Sciences*, *114*(40), 10578–10583.

Holloway, C. E., Wing, A. A., Bony, S., Muller, C., Masunaga, H., L'Ecuyer, T. S., ... Zuidema, P. (2017). Observing convective aggregation. *Surveys in Geophysics*, *38*(6), 1199–1236.

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. doi: 10.1109/MCSE.2007.55

Johnson, R. H., Rickenbach, T. M., Rutledge, S. A., Ciesielski, P. E., & Schubert, W. H. (1999). Trimodal characteristics of tropical convection. *Journal of climate*, *12*(8), 2397–2418.

Jolliffe, I. T., & Stephenson, D. B. (2012). *Forecast verification: a practitioner's guide in atmospheric science.* John Wiley & Sons.

Jonker, H. J., Duynkerke, P. G., & Cuijpers, J. W. (1999). Mesoscale fluctuations in scalars generated by boundary layer convection. *Journal of the atmospheric sciences*, *56*(5), 801–808.

Klein, S. A., Hall, A., Norris, J. R., & Pincus, R. (2017). Low-cloud feedbacks from cloud-controlling factors: a review. In *Shallow clouds, water vapor, circulation, and climate sensitivity* (pp. 135–157). Springer.

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In Stéfan van der Walt & Jarrod Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (p. 56 - 61). doi: 10.25080/Majora-92bf1922-00a

Medeiros, B., & Nuijens, L. (2016). Clouds at Barbados are representative of clouds across the trade wind regions in observations and climate models. *Proceedings of the National Academy of Sciences*, *113*(22), E3062–E3070.

Muller, C. J., & Held, I. M. (2012). Detailed investigation of the self-aggregation of convection in cloud-resolving simulations. *Journal of the Atmospheric Sciences*, *69*(8), 2551–2565.

Nair, U., Weger, R., Kuo, K., & Welch, R. (1998). Clustering, randomness, and regularity in cloud fields: 5. the nature of regular cumulus cloud fields. *Journal of Geophysical Research: Atmospheres*, *103*(D10), 11363–11380.

Neggers, R., Griewank, P., & Heus, T. (2019). Power-law scaling in the internal variability of cumulus cloud size distributions due to subsampling and spatial organization. *Journal of the Atmospheric Sciences*, *76*(6), 1489–1503.

Nuijens, L., & Siebesma, A. P. (2019). Boundary layer clouds and convection over subtropical oceans in our current and in a warmer climate. *Current Climate Change Reports*, *5*(2), 80–94.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Platnick, S., Ackerman, S., King, M., Menzel, P., Wind, G., & Frey, R. (2015). *MODIS Atmosphere L2 Cloud Product (06_L2).* Goddard Space Flight Center, USA: NASA MODIS Adaptive Processing System. doi: http://dx.doi.org/10.5067/MODIS/MYD06_L2.061

Pscheidt, I., Senf, F., Heinze, R., Deneke, H., Trömel, S., & Hohenegger, C. (2019). How organized is deep convection over germany? *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2366–2384.

Rasp, S., Selz, T., & Craig, G. C. (2018). Variability and clustering of midlatitude summertime convection: Testing the Craig and Cohen theory in a convection-permitting ensemble with stochastic boundary layer perturbations. *Journal of the Atmospheric Sciences*, *75*(2), 691–706.

Rempel, M., Senf, F., & Deneke, H. (2017). Object-based metrics for forecast verification of convective development with geostationary satellite data. *Monthly Weather Review*, *145*(8), 3161–3178.

Rieck, M., Nuijens, L., & Stevens, B. (2012). Marine boundary layer cloud feedbacks in a constant relative humidity atmosphere. *Journal of the Atmospheric Sciences*, *69*(8), 2538–2550.

Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., & Siebesma, A. P. (2017). Climate goals and computing the future of clouds. *Nature Climate Change*, *7*(1), 3–5.

Scott, D. W. (1992). *Multivariate density estimation: theory, practice, and visualization.* John Wiley & Sons.

Seifert, A., & Heus, T. (2013). Large-eddy simulation of organized precipitating trade wind cumulus clouds. *Atmos. Chem. Phys*, *13*(11), 5631–5645.

Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., … others (2019). Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds. *Quarterly Journal of the Royal Meteorological Society*, 1–12.

Tobin, I., Bony, S., & 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*(20), 6885–6904.

Tompkins, A. M., & 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*(2), 1046–1068.

van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, *13*(2), 22–30.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., . . . the scikit-image contributors (2014, 6). scikit-image: image processing in Python. *PeerJ*, *2*, e453. Retrieved from `https://doi.org/10.7717/peerj.453` doi: 10.7717/peerj.453

van Laar, T. W. (2019). *Spatial patterns in shallow cumulus cloud populations over a heterogeneous surface* (Doctoral dissertation, University of Cologne). Retrieved from `http://kups.ub.uni-koeln.de/id/eprint/10221`

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: https://doi.org/10.1038/s41592-019-0686-2

Waskom, M., & the seaborn development team. (2020, September). *mwaskom/seaborn.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.592845` doi: 10.5281/zenodo.592845

Weger, R., Lee, J., Zhu, T., & Welch, R. (1992). Clustering, randomness and regularity in cloud fields: 1. theoretical considerations. *Journal of Geophysical Research: Atmospheres*, *97*(D18), 20519–20536.

Wernli, H., Paulat, M., Hagen, M., & Frei, C. (2008). SAL — a novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, *136*(11), 4470–4487.

White, B., Buchanan, A., Birch, C., Stier, P., & Pearson, K. (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*(2), 425–450.

Windmiller, J. M. (2017). *Organization of tropical convection* (Doctoral dissertation, Ludwig-Maximilian University of Munich). Retrieved from `https://edoc.ub.uni-muenchen.de/21245/`

Wood, R., & Field, P. R. (2011). The distribution of cloud horizontal sizes. *Journal of Climate*, *24*(18), 4800–4816.

Wood, R., & Hartmann, D. L. (2006). Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection. *Journal of Climate*, *19*(9), 1748–1764.

Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of computational and graphical statistics*, *15*(2), 265–286.

Supporting Information for "Cloud patterns have four interpretable dimensions"

DOI: 10.1002/

Martin Janssens[1], Jordi Vilà-Guerau de Arellano[1], Marten Scheffer[1], Coco Antonissen[2], A. Pier Siebesma[23], Franziska Glassmeier[2]

[1]Wageningen University & Research

[2]Delft University of Technology

[3]Royal Netherlands Meteorological Institute

Contents of this file

Corresponding author: M. Janssens, Departments of Meteorology & Air Quality and Aquatic Ecology & Water Quality Management, Wageningen University & Research, Wageningen, Lumen Building, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands, (martin.janssens@wur.nl)

Introduction

This supplement contains further descriptions of the metrics that we use to characterise our cloud field pattern distribution (Text S1). Specifically, we elaborate upon details of and justify choices made in their computation. Code that evaluates these metrics given input scenes of cloud mask, cloud water path and cloud top height can be found in our accompanying GitHub repository (`https://github.com/martinjanssens/cloudmetrics`) and Figshare copy of this repository at the time of publication `https://figshare.com/projects/Cloud_field_organisation_description_with_metrics/86303`. The supplement also contains five figures, that quantify i) the sensitivity of our metric distribution to field resolution, object segmentation strategy and minimum cloud size, ii) the absolute Pearson correlation between all metrics, iii) the fraction of variance in each metric explained by every PC, iv) an estimate of the quality of our metric-based approach to approximating cloud field patterns and v) the sensitivity to free parameters of approximating principal components with a subset of metrics through sparse principal component analysis.

Text S1. - Details of metrics

Statistical moments of cloud field properties

We quantify several statistics of the extracted cloud field products. Some of these are straightforward computations that do not feature design choices (cloud fraction, total cloud water, standard deviation of cloud water over cloudy pixels). The other metrics require further qualification.

Mean and standard deviation of cloud top height ($\overline{\mathrm{CTH}}$ and St(CTH) respectively) i) explicitly ignore clouds higher than 5km, as cirrus wisps were found to disproportionately affect the results otherwise and ii) only consider cloudy pixels. Higher-order moments of these fields were small and are therefore not included.

Cloud water variance ratio R (CWP var. ratio) is directly adopted from Bretherton and Blossey (2017), but instead of being applied to the total, vertically integrated moisture field, it is here only applied to the cloud water:

$$R = \frac{Std\left(\overline{CWP_b} - \overline{CWP}\right)}{Std(CWP)} \tag{1}$$

In this relation, $\bar{\cdot}$ denotes a domain average and CWP_b indicates the cloud water contained in blocks of 16x16 pixels.

Object-based metrics

Object-based metrics follow from segmenting the cloud mask field into N_o objects according to their 4-connectivity. To avoid artefacts at the grid scale, we only consider objects with areas larger than 4 pixels. Each extracted object covers an area A_i, such that a typical length scale for that object is $l_i = \sqrt{A_i}$.

Mean object size is defined as $\frac{1}{N_o}\sum_i l_i$

Max object size is defined as $\max l_i$.

Mean perimeter is derived by extracting the perimeter of each object P_i and defining the mean perimeter $\overline{P} = \frac{1}{N_o}\sum_i P_i$.

The *Simple Convective Aggregation Index (SCAI)* (Tobin et al., 2012) is defined as:

$$\text{SCAI} = \frac{N_o D_0}{N_{max}} \tag{2}$$

Where N_{max} is the number of pixels in a scene, $D_0 = \sqrt[N_p]{\prod_i^{N_p} d_i}$ is the geometric mean of Euclidian pairwise distance between all object centroids d_i and $N_p = N_o(N_o - 1)/2$.

The Convective Organisation Potential (COP) (White et al., 2018) is:

$$\text{COP} = \frac{1}{N_p}\sum_{i=0}^{N_o}\sum_{j=i+1}^{N_o}\frac{l_i + l_j}{\sqrt{\pi} d_{ij}} \tag{3}$$

Where d_{ij} now explicitly represents the distance between two object centroids.

Max RDF is the maximum value of the radial distribution function $\text{RDF}(r)$ as proposed in Rasp, Selz, and Craig (2018):

$$\text{RDF}(r) = \frac{1}{N_i}\sum_i \frac{\sum_{r \leq r_i < r+dr} 1}{L\left(\pi\left(r+dr\right)^2 - r^2\right)} \tag{4}$$

Where r_i are pairwise distances from the i$^{\text{th}}$ centroid to all other centroids, dr denotes the width of a radial annulus over which we sum such distances, L is the length of the scene's side, and N_i are the number of centroids that lie within a distance r_{max} from the domain edges. We only consider coordinates within a radius r_{max} from any original centroid. We set $r_{max} = 20$ pixels, as in practice $\arg\max \mathrm{RDF}(r) < 20$ always, and use $dr = 1$ (the results are not sensitive to these parameters).

Degree variance of nearest-neighbour network representations of the scenes are quantified by constructing a Voronoi tessellation from the computed object centroids and measuring the variance in the degree (number of neighbours) distribution of the identified Voronoi cells.

I_{org} (Weger et al., 1992) is included in two flavours. The first is the original metric, which integrates the area under the curve defined by the NNCDF, the cumulative density function of nearest neighbour distances d_N between object centroids (y axis) and the corresponding Weibull distribution (x axis):

$$W = 1 - \exp\left(\frac{N_o}{L^2}\pi d_N^2\right) \tag{5}$$

If the object centroids are scattered as a Poisson point process, they should follow W exactly, resulting in $I_{org} = 0.5$. $I_{org} < 0.5$ if they are regularly spaced; if they appear in clusters, $I_{org} > 0.5$. As pointed out by Benner and Curry (1998), this overestimates the regularity of the cloud field, because in reality separate cloud objects are inhibited from forming within the area covered by another object. To account for this, we also include a second version of I_{org}, which we name I^*_{org}. This metric compares the cloud field NNCDF to an inhibition NNCDF, which is constructed by randomly scattering N_o objects throughout the scene, provided that they do not fall within the circular area of an object that has already been placed. The computer-generated random positions of this approach are less robust than the Weibull distribution (Weger et al., 1992), but we find that repeating the computations 3 times does not impact the resulting I^*_{org} below the third significant digit.

Scale decomposition metrics

Size exponent b is computed by counting all cloud sizes N_c in bins of exponentially increasing width, and fitting the resulting cloud size distribution with a power law:

$$\log N_c \propto b \log l \tag{6}$$

The average coefficient of determination R^2 of fitting this relation to all scenes is good: 0.923. We also investigated a fit according to subcritical percolation theory that incorporates an exponential term. However, undersampling of large cloud structures make such fits quite unrealistic on a per-scene basis, even though the fit converges when sampling a large number of scenes at similar cloud fraction (not shown). It is therefore likely that these cloud fields obey the rules of subcritical percolation. Yet, the parameters of the corresponding fit cannot reliably be identified on a per scene basis.

The *box-counting dimension* D_f (fractal dim.) of each cloud mask field is derived by counting the number of square boxes N_c of dimension l_b that are neither fully cloud-free nor fully cloudy (i.e. boxes that contain cloud borders). D_f is then computed by least-squares fitting the following relation over a range of l_b:

$$\log N_c \propto D_f \log l_b \tag{7}$$

The average R^2 of this fit is 0.997, indicating an excellent goodness of fit.

The Spectral Length Scale (Spectral length) Λ is derived from the field's Fourier transform. Computing this value requires several design choices. First, the scenes are tilt-compensated by subtracting a scene's best-fit plane. Next, one would normally apply a radially symmetric window function to account for the scenes' aperiodicity. However, we find that the application of such a function occludes so much spatial information that our scenes are ordered much less coherently. Hence, we refrain from applying window functions. Next, we Fourier transform the scenes and construct their 1D PSD $S(k)$ by averaging the transform's power signals over shells of radial wavenumber k. The validity of this approach rests on the assumption that the satellite scenes are spatially isotropic, which they are often not. Yet, we find that on a scale from 0-1 (0 representing a 2D PSD where the power is equally distributed over the azimuthal direction and 1 representing the case where all power is concentrated in a single direction), the average anisotropy of all scenes is 0.104. We judge that this justifies the use of the 1D PSD. Finally, Λ is computed from the distribution's first moment, as suggested in Jonker, Duynkerke, and Cuijpers (1999):

$$\Lambda^{-a} = \frac{\int_0^{k_{Ny}} k^a S(k) dk}{\int_0^{k_{Ny}} S(k) dk}; \quad a \neq 0 \tag{8}$$

Where k_{Ny} is the Nyquist wavenumber and we choose to set $a = 1$.

We compute *Wavelet-based Organisation Indices (WOIs)* following Brune, Kapp, and Friederichs (2018). These metrics are based on the domain-averaged, squared coefficients of the 2D stationary wavelet transform (SWT) of each scene's cloud water path (CWP) field, E_{CWP}. We use the Haar wavelet as our basis. E_{CWP} contains a scale decomposition over three (horizontal, vertical, diagonal) directions, with each scale representing a power of 2 that exactly fits the 512 pixel field. Using E_{CWP}, we derive the metrics proposed by (Brune et al., 2018):

$$WOI_1 = \frac{\overline{E^l_{CWP}}}{\overline{E_{CWP}}} \tag{9}$$

$$WOI_2 = \frac{\overline{E_{CWP}}}{N_c} \tag{10}$$

$$WOI_3 = \frac{1}{3}\sqrt{\sum_d \left(\frac{E^l_{CWP_d} - \overline{E^l_{CWP}}}{\overline{E^l_{CWP}}}\right)^2 + \left(\frac{E^s_{CWP_d} - \overline{E^s_{CWP}}}{\overline{E^s_{CWP}}}\right)^2} \tag{11}$$

Where $\cdot^l$ and $\cdot^s$ indicate total energy contained in the large scales (resolution $2^1 - 2^5$) and small scales (resolution $2^6 - 2^9$) respectively, $\bar{\cdot}$ indicates averaging over all three directions and N_c is the number of cloudy pixels in a scene. These metrics measure the fraction of cloud water contained in the scene's large scales (WOI_1), the average cloud water in cloudy pixels (WOI_2) and the anisotropy in the spectrum's three directions (WOI_3). Since WOI_1 and WOI_2 are almost exact mirrors of R (eq. 1) and cloud water variance in cloudy pixels respectively, respectively, we choose to only include WOI_3 in our analysis.

Our simple *Clear Sky* metric extracts the scene's largest rectangular area spanned by the horizontal and vertical lines drawn through any cloud-free pixel whose ends are the first cloudy pixel encountered along those lines. This rectangle is normalised by the domain size, to arrive at a fraction that represents the largest, contiguous, clear sky area.

References

Benner, T. C., & Curry, J. A. (1998). Characteristics of small tropical cumulus clouds and their impact on the environment. *Journal of Geophysical Research: Atmospheres*, *103*(D22), 28753–28767.

Bretherton, C., & Blossey, P. (2017). Understanding mesoscale aggregation of shallow cumulus convection using large-eddy simulation. *Journal of Advances in Modeling Earth Systems*, *9*(8), 2798–2821.

Brueck, M., Hohenegger, C., & Stevens, B. (2020). Mesoscale marine tropical precipitation varies independently from the spatial arrangement of its convective cells. *Quarterly Journal of the Royal Meteorological Society*, *146*(728), 1391–1402.

Brune, S., Kapp, F., & Friederichs, P. (2018). A wavelet-based analysis of convective organization in ICON large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2812–2829.

Denby, L. (2020). Discovering the importance of mesoscale cloud organization through unsupervised classification. *Geophysical Research Letters*, *47*(1), e2019GL085190.

Erichson, N. B., Zheng, P., Manohar, K., Brunton, S. L., Kutz, J. N., & Aravkin, A. Y. (2020). Sparse principal component analysis via variable projection. *SIAM Journal on Applied Mathematics*, *80*(2), 977–1002.

Glassmeier, F., & Feingold, G. (2017). Network approach to patterns in stratocumulus clouds. *Proceedings of the National Academy of Sciences*, *114*(40), 10578–10583.

Jonker, H. J., Duynkerke, P. G., & Cuijpers, J. W. (1999). Mesoscale fluctuations in scalars generated by boundary layer convection. *Journal of the atmospheric sciences*, *56*(5), 801–808.

Neggers, R., Griewank, P., & Heus, T. (2019). Power-law scaling in the internal variability of cumulus cloud size distributions due to subsampling and spatial organization.

Journal of the Atmospheric Sciences, *76*(6), 1489–1503.

Rasp, S., Selz, T., & Craig, G. C. (2018). Variability and clustering of midlatitude summertime convection: Testing the Craig and Cohen theory in a convection-permitting ensemble with stochastic boundary layer perturbations. *Journal of the Atmospheric Sciences*, *75*(2), 691–706.

Tobin, I., Bony, S., & 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*(20), 6885–6904.

van Laar, T. W. (2019). *Spatial patterns in shallow cumulus cloud populations over a heterogeneous surface* (Doctoral dissertation, University of Cologne). Retrieved from `http://kups.ub.uni-koeln.de/id/eprint/10221`

Weger, R., Lee, J., Zhu, T., & Welch, R. (1992). Clustering, randomness and regularity in cloud fields: 1. theoretical considerations. *Journal of Geophysical Research: Atmospheres*, *97*(D18), 20519–20536.

White, B., Buchanan, A., Birch, C., Stier, P., & Pearson, K. (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*(2), 425–450.

Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of computational and graphical statistics*, *15*(2), 265–286.

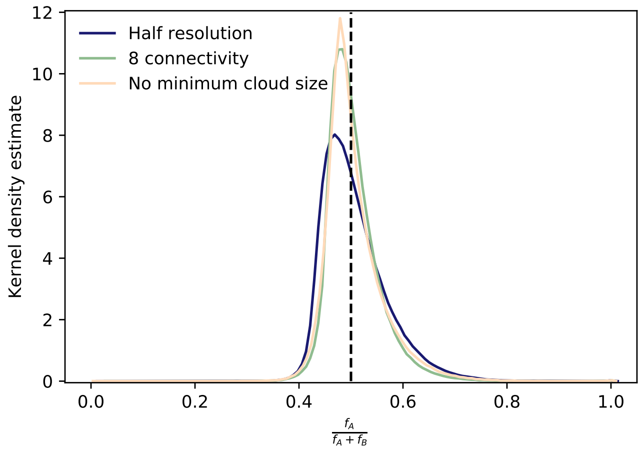

Figure S1. Gaussian kernel density estimates of the ratio $D = \frac{f_A}{f_A - f_B}$, which is constructed from high-dimensional kernel density estimates of the reference metric distribution used in the main text (f_A) and three separately perturbed metric distributions (f_B). An identical distribution to the original would yield a Dirac pulse centred at 0.5 (dashed line); deviations from this line quantify the contrast between the original and perturbed distributions. Sensitivities are quantified with respect to i) scenes that are downsampled to half the original resolution (most sensitive), ii) object segmentation based on 8-connectivity rather than 4-connectivity and iii) not including a lower bound to the minimum cloud size that is considered an object (least sensitive). All perturbed distributions are narrow and have an expected value around 0.5, indicating the robustness of the distribution presented in the main text. Furthermore, the visual relation between metrics is largely unaffected (not shown).

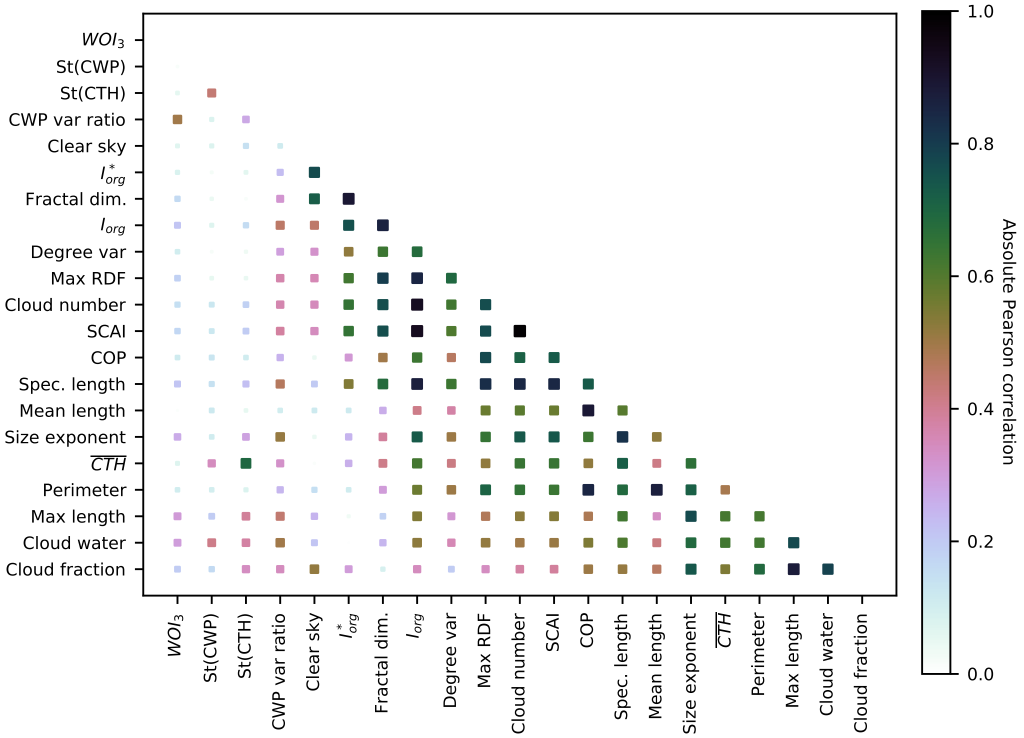

Figure S2. Standardised metric correlation matrix, with squares sized and coloured according to absolute Pearson correlation between a metric pair. Many metrics closely correlate, indicating that their cumulative information can be captured by a smaller number of effective indicators. Several closely correlating metrics follow well-known relationships, e.g. perimeter and mean length (any combination yields approximately constant fractal dim.), or cloud number and numerous aggregation metrics (this relation is similar for deep convective organisation (Brueck et al., 2020)). Others follow rather trivial ones, e.g. max length and cloud fraction, or the Spectral Length and size exponent. Several strong correlations are at first sight not trivial. For instance, I_{org} (both versions) and Fractal dim. are highly similar (up to a factor -1). Hence, highly concentrated shallow cloud clusters in rather empty scenes (high I^*_{org}) tend towards “lines” (low fractal dimension, approaching 1 from above); $I^*_{org} = 0.5$ and fractal dim.=2 both indicate random scattering of points. Finally, while some effort has been invested in contrasting and improving aggregation/clustering measures (e.g. SCAI, I_{org} and max RDF (van Laar, 2019)), these are extremely similar. Instead, shifting focus to metrics that are comparatively *uncorrelated* might be more more fruitful to further develop our understanding of shallow cloud field organisation.

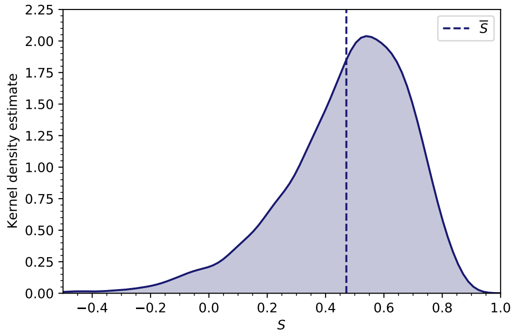

Figure S3. Gaussian kernel density estimate of $S = 1 - \frac{\|x_{a_i} - x_{n_i}\|_2}{\|x_{a_i} - x_{r_i}\|_2}$ compiled from $0 \leq i < 3951$ scenes, where x_{a_i} is the vector of the metrics for an "anchor scene", x_{n_i} are the metrics of a "neighbour scene" that overlaps with half the area of the anchor scene and x_{r_i} are the average metrics of 100 randomly sampled scenes. S measures how much the metrics minimise the Euclidian distance between an anchor and its half-overlapping scene, relative to the average Euclidian distance to a randomly sampled scene. If $S = 0$, the metrics estimate that a half-overlapping scene is equally similarly organised as a randomly sampled scene; if $S = 1$, the anchor and half-overlapping scene are estimated to be identically organised. Since half-overlapping scenes share numerous spatial features, they should usually be more similarly organised than random scenes ($S > 0$) - a feature we expect the metrics to capture. As 96% of the distribution exceeds $S = 0$, this inspires confidence in this ability. The dashed line indicates the mean, $\overline{S} = 0.47$. While this lies significantly below 1, we expect the desired upper bound of S to also lie below 1, since half-overlapping scenes are (by visual inspection) rarely *identically* organised. Estimating this bound requires knowing how far a typical pattern extends beyond a scene's boundaries; this demands a better characterisation of the relation between the measurement scale ("scene") and the true scale of a pattern. However, even without an explicit upper bound on $S < 1$, this distribution shows that our metrics on average come closer to that bound than to being random. Proficiency of a cloud field description can also be assessed by comparing S across approaches. A version of S already served as cost function for a machine-learned pattern description (Denby, 2020). One could also compile statistics on how similar humans find half-overlapping scenes compared to random scene pairs. Comparing both resulting S to our metrics could more objectively assess which approach to pattern description (human, metrics or machine) is best.

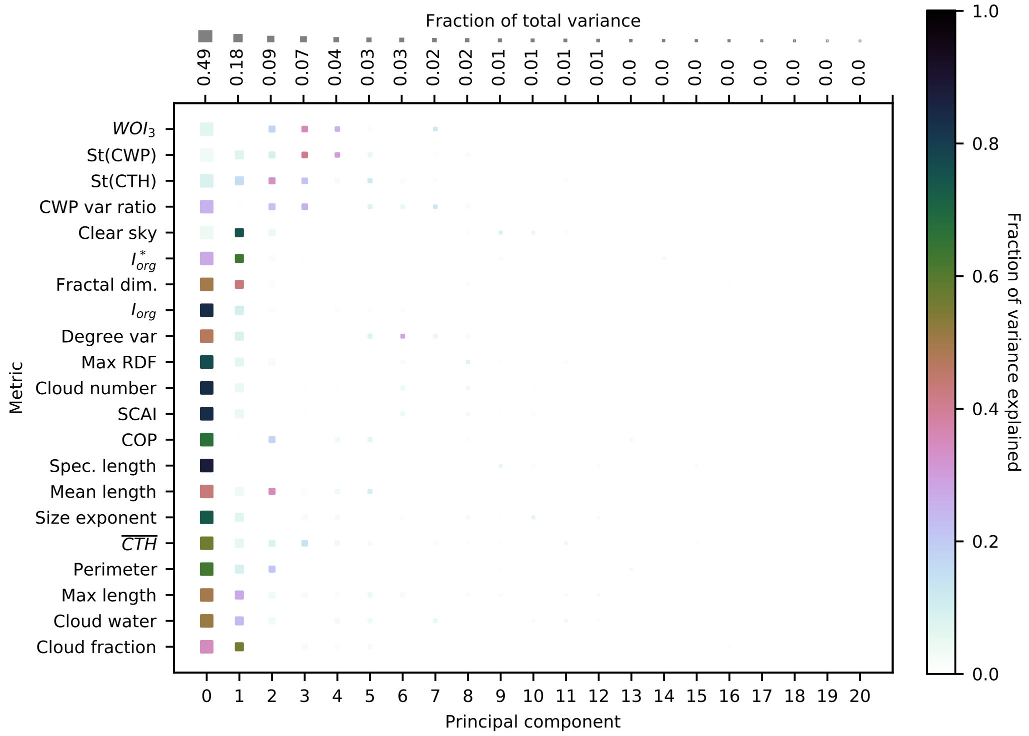

Figure S4. Fraction of variance (colour) in each metric (vertical axis) explained by each PC (horizontal axis). Sizes of squares are scaled by the total dataset's explained variance fraction in each PC (top horizontal axis). 17/21 metrics have more than 70% of their variance captured by the first two PCs; the remaining 4/21 metrics reach this threshold after four PCs.

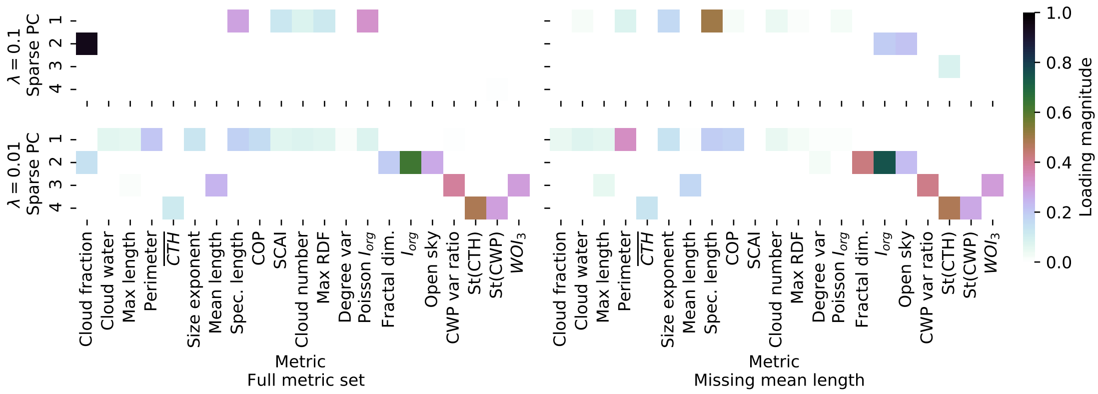

Figure S5. Sensitivity of Sparse Principal Component Analysis (SPCA, Zou et al., 2006). SPCA encourages sparsity in the weighting of metrics that form each of the four main, approximate PCs ("loadings") by casting the PCA as a regression problem, whose cost function contains at least i) a least squares error term of the PCA fit and ii) a penalty (in the L_0 or L_1 norm) on the magnitude of the regression coefficients (the loadings). This penalty is weighted by a regularisation parameter λ. We solve the resulting non-convex optimisation problem using the approach developed by Erichson et al. (2020) and refer to that paper for further details. This figure shows the optimal sparsity structure in the loadings identified by SPCA under four combinations of two free parameters: The magnitude of the sparsity penalty λ (top row vs bottom row) and the omission of a single, seemingly redundant, metric (SCAI, left column vs right column). Unfortunately, the optimal sparsity structure i) is rather sensitive and ii) reacts relatively unpredictably to changes in these free parameters. This is true also when other metrics are excluded, when a different sparsity-inducing algorithm is used or when the sparsity penalty is in the L_0 norm, rather than the L_1 norm as displayed here. These considerations curb SPCA's utility for metric selection and prevent us from recommending its use.

Table S1. Metrics quantified for initial analysis. Selection for paper is guaranteed by meeting either criteria 1 or 2, and separately meeting criterion 3, as presented in section 2.2. This excludes the lower portion of the table. Two metrics in the table's middle section meet the criteria, but are still excluded: WOI_1, WOI_2 (see Text S1). Metrics annotated with (*) are not included in the coded library.

Metric	**Criterion 1** *Unique*	**Criterion 2** *Recurrent/recent*	**Criterion 3** *Interpretable*
Cloud fraction	No	Yes	Yes
Cloud water	Yes	Yes	Yes
Max length	No	Yes	Yes
Perimeter	No	Yes	Yes
$\overline{CTH}$	Yes	Yes	Yes
Size exponent	Yes	Yes	Yes
Mean length	No	Yes	Yes
Spectral length scale	No	Yes	Yes
COP	No	Yes	Yes
SCAI	No	Yes	Yes
Cloud number	No	Yes	Yes
Max RDF	No	Yes	Yes
Degree var.	Yes	Yes	Yes
I_{org}	No	Yes	Yes
Fractal dimension	Yes	Yes	Yes
I^*_{org}	Yes	No	Yes
Open sky	Yes	No	Yes
CWP var ratio	Yes	Yes	Yes
St(CTH)	Yes	Yes	Yes
St(CWP)	Yes	Yes	Yes
WOI_3	Yes	Yes	Yes
WOI_1	No	Yes	Yes
WOI_2	No	Yes	Yes
Multifractality index (*)	Yes	Yes	No
Multifractal intermittency (*)	Yes	Yes	No
Object eccentricity	No	No	Yes
Covariance-based orientation	No	No	Yes
Raw moment-based orientation	No	No	Yes
b_{org} in small clouds (Neggers et al., 2019)	Yes	Yes	No
Skewness/kurtosis of CTH, CWP	Yes	Yes	No
Geometric mean nearest neighbour distance	No	No	Yes
Variance of CTH, CWP in largest cloud	No	No	Yes
1D PSD slope	No	No	Yes
Variance in azimuthal PSD	No	No	Yes
Aboav-Wearie fit (Glassmeier & Feingold, 2017)	Yes	Yes	No