Convective cloud size distributions in idealized cloud resolving model simulations

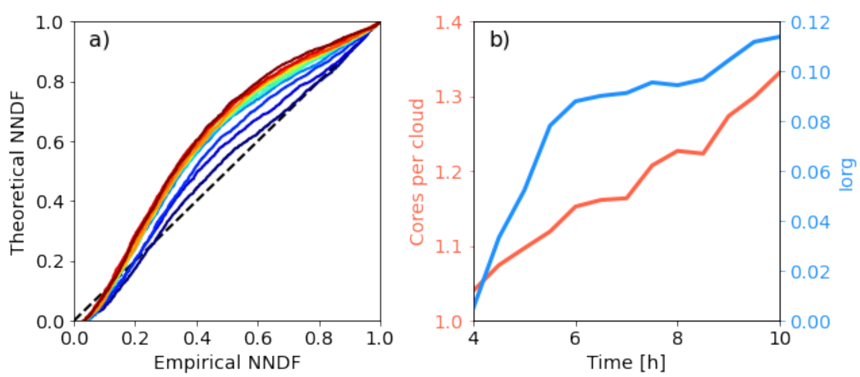
Julien Savre[1] and George C. Craig[2]

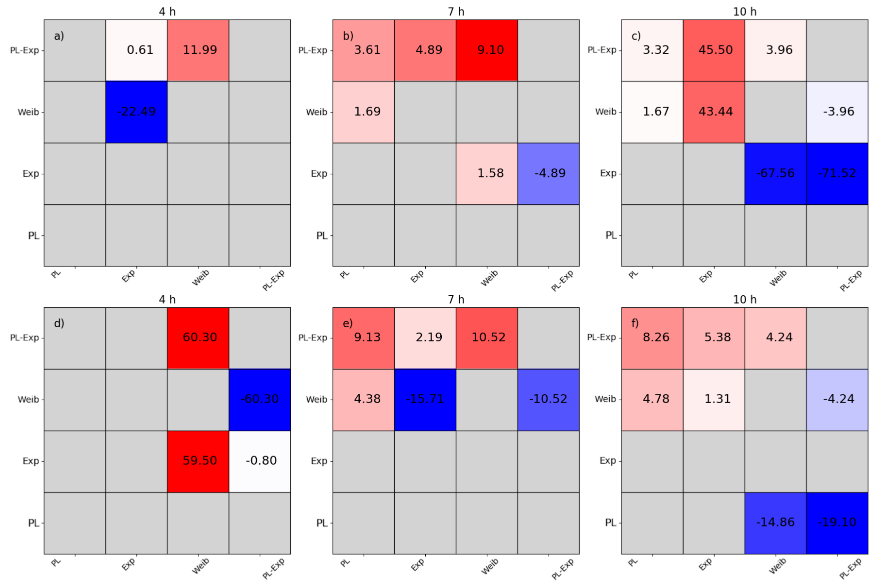
[1]Ludwig-Maximilians-University
[2]Institute of Meteorology - University of Munich

November 18, 2022

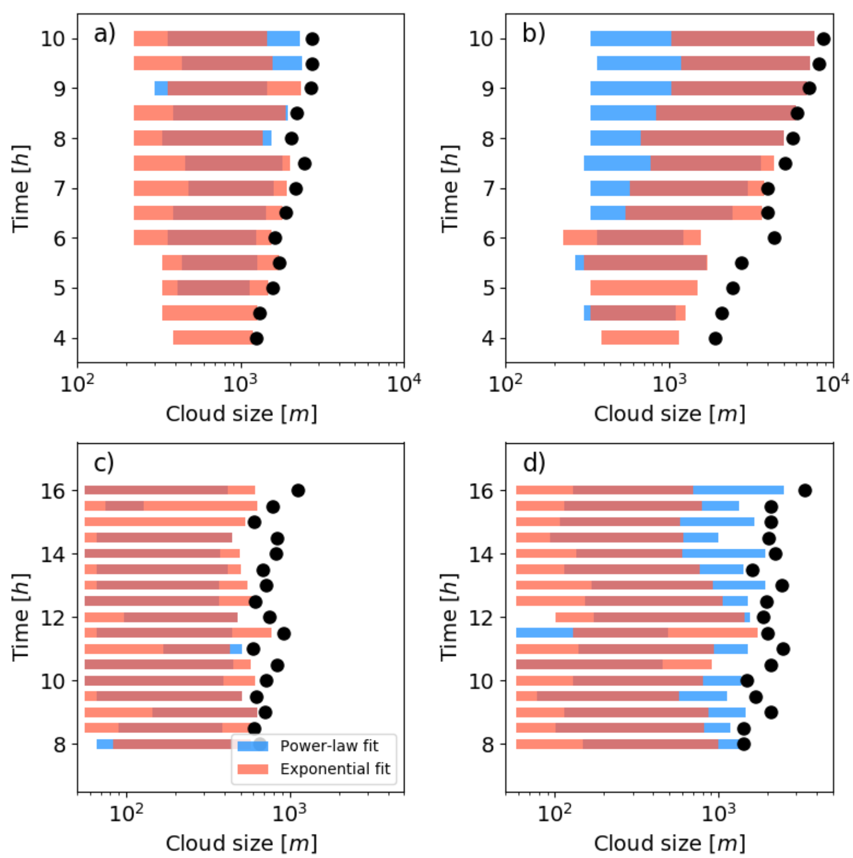
Abstract

It is now widely accepted that cumulus cloud size distributions follow power-laws, at least over part of the cloud size spectrum. Providing reliable fits to empirical size distributions is however not a simple task, and this is reflected by the large spread in power-law exponents reported in the literature. Two well-documented idealized high-resolution numerical simulations of convective situations are here performed and analyzed in order to gain a clearer understanding of cumulus size distributions. Advanced statistical methods, including maximum likelihood estimators and goodness-of-fit tests, are employed to produce the most accurate fits possible. Various candidate distributions are tested including exponentials, power-laws and other heavy-tail functions. Size distributions estimated from clouds identified just above cloud base are found to be best modeled by exponential distributions. If one considers instead clouds identified from an integrated condensed water path, robust power-law behaviors start to emerge, in particular when deep convection is involved. In general however, these empirical distributions are best represented by alternative heavy-tail distributions such as the Weibull or cutoff power-law distributions. In an attempt to explain these results, it is suggested that exponential size distributions characterize a population where clouds interact only weakly, whereas heavy-tail distributions are the manifestation of a cloud population that self-organizes towards a critical state.

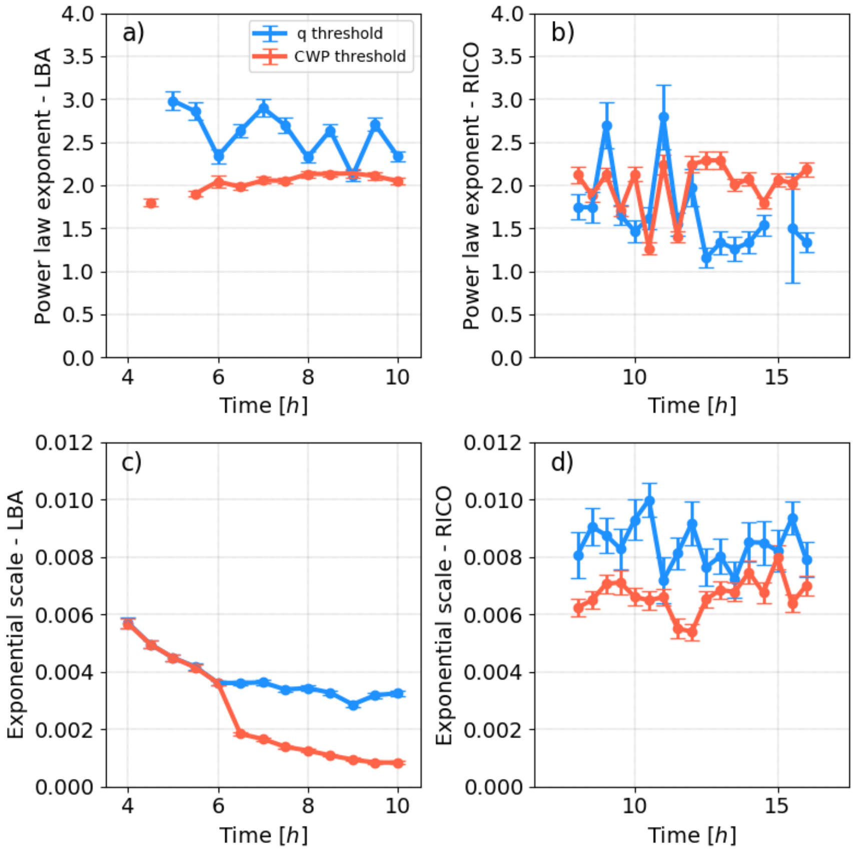


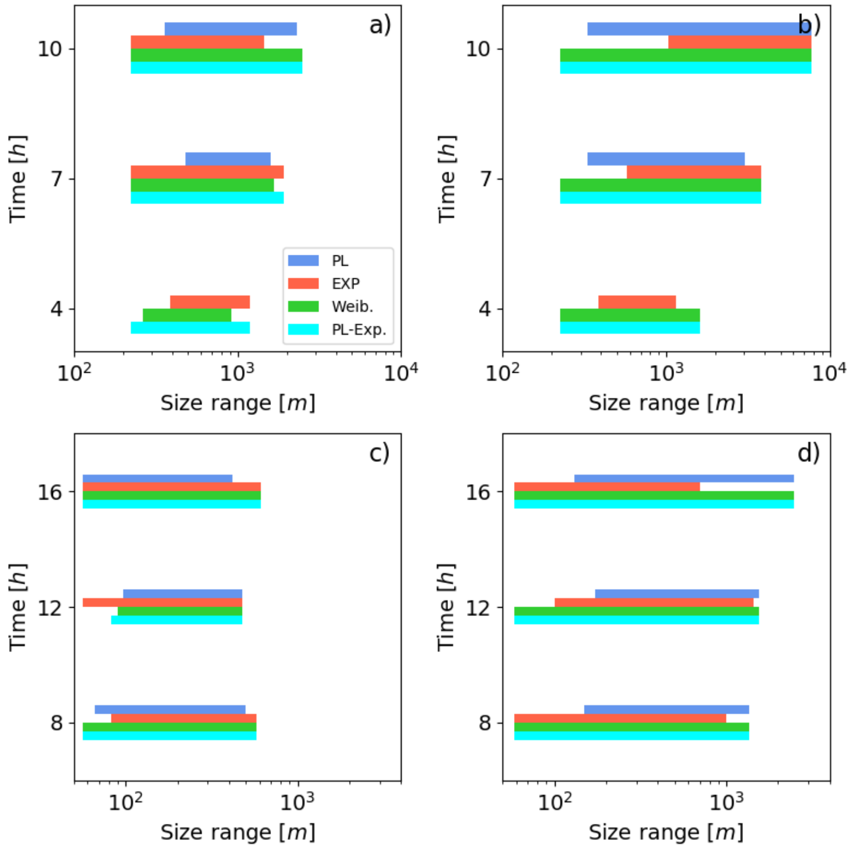


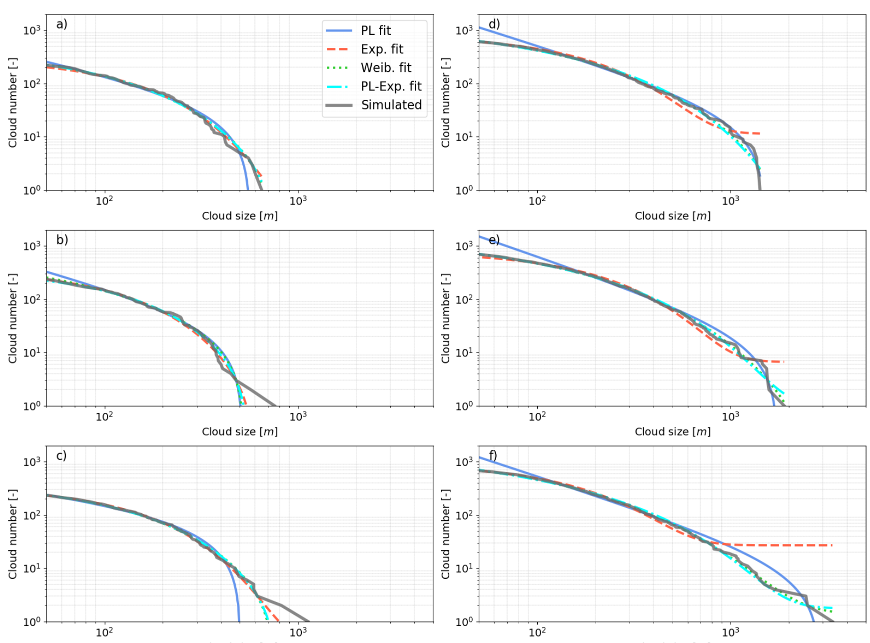
4 h
7 h
10 h
a)
b)
c)
d)
e)
f)
PL-Exp
Weib
Exp
PL
0.61
11.99
-22.49
3.61
4.89
9.10
1.69
1.58
-4.89
3.32
45.50
3.96
1.67
43.44
-3.96
-67.56
-71.52
60.30
60.30
59.50
-0.80
9.13
2.19
10.52
4.38
-15.71
-10.52
8.26
5.38
4.24
4.78
1.31
-4.24
-14.86
-19.10



a)
b)
c)
d)
Time [h]
Cloud size [m]
Power-law fit
Exponential fit


a)
b)
c)
d)
q threshold
CWP threshold
Power law exponent - LBA
Power law exponent - RICO
Exponential scale - LBA
Exponential scale - RICO
Time [h]


a)
b)
c)
d)
PL
EXP
Weib.
PL-Exp.
Time [h]
Size range [m]
10
7
4
16
12
8
10^2
10^3
10^4


a)
b)
c)
d)
e)
f)
PL fit
Exp. fit
Weib. fit
PL-Exp. fit
Simulated
Cloud number [-]
Cloud size [m]
10^0
10^1
10^2
10^3

Convective cloud size distributions in idealized cloud resolving model simulations

Julien Savre[1], and George Craig[1]

[1]Meteorological Institute, Physics Department, Ludwig-Maximilians-Universität, Munich, Germany

Key Points:

- Advanced statistical methods are used to fit cloud size distributions from two simulated convective cloud ensembles
- Cloud objects identified above their base, and especially cloudy updrafts, preferentially follow an exponential size distribution
- Clouds identified from an integrated condensed water criterion exhibit clear power-law scaling, but are best approximated by cutoff power-laws
- These results are consistent with the hypothesis that the cloud ensemble self-organizes towards a critical state

Corresponding author: Julien Savre, `julien.savre@lmu.de`

Abstract

It is now widely accepted that cumulus cloud size distributions follow power-laws, at least over part of the cloud size spectrum. Providing reliable fits to empirical size distributions is however not a simple task, and this is reflected by the large spread in power-law exponents reported in the literature. Two well-documented idealized high-resolution numerical simulations of convective situations are here performed and analyzed in order to gain a clearer understanding of cumulus size distributions. Advanced statistical methods, including maximum likelihood estimators and goodness-of-fit tests, are employed to produce the most accurate fits possible. Various candidate distributions are tested including exponentials, power-laws and other heavy-tail functions. Size distributions estimated from clouds identified just above cloud base are found to be best modeled by exponential distributions. If one considers instead clouds identified from an integrated condensed water path, robust power-law behaviors start to emerge, in particular when deep convection is involved. In general however, these empirical distributions are best represented by alternative heavy-tail distributions such as the Weibull or cutoff power-law distributions. In an attempt to explain these results, it is suggested that exponential size distributions characterize a population where clouds interact only weakly, whereas heavy-tail distributions are the manifestation of a cloud population that self-organizes towards a critical state.

Plain Language Summary

[Clouds constitute an important element of the climate system by reflecting incoming solar radiation back to space and emitting infra-red radiation that heats the atmosphere. The net radiative impact of clouds however depends on many factors including their size. It is thus of prime importance to characterize the size of clouds, in particular convective clouds, and understand the underlying processes shaping them. In this study, a numerical model is used to simulate two idealized convective situations at horizontal resolutions providing a fine description of all cloud processes. After defining and calculating the size of individual simulated clouds, advanced statistical methods are employed to characterize how clouds are distributed in size. It is shown that the distributions of cloud sizes measured near cloud base are close to exponential functions. In contrast, cloud size distributions as measured from above (similar to what satellite imagery would do) resemble power-law functions with a cut-off at large sizes. It is suggested that these results can be explained by a combination of basic physical principles and complex interactions organizing the cloud population.]

1 Introduction

The first Landsat satellite launched in 1972 provided high-resolution data that allowed the first large-scale analysis of cloud properties and their radiative impact on our climate. It was later recognized that spatial inhomogeneities in individual cloud scenes are an important factor determining cloud radiative properties. It quickly appeared that an essential step towards better predicting the radiative impact of broken cloud fields (cumulus and stratocumulus) consisted in understanding the large-scale structure and organization of such cloud fields, as well as the small-scale characteristics of individual clouds (including their fractal shapes and size distributions) (Wielicki & Welch, 1986; Welch et al., 1988; Kuo et al., 1988).

Earlier observational studies have suggested that cloud size distributions could be fitted using various functional forms including exponential (Wielicki & Welch, 1986) and log-normal (Lopez, 1977; Houze Jr. & Cheng, 1977) distributions. In the late 80's, Parker et al. (1986), Welch et al. (1988) and Kuo et al. (1988) have introduced the power-law as the functional form that best represents cumulus and stratocumulus cloud size distributions. Accordingly, the number of clouds n having a certain size L

obeys:

$$n(L) \propto L^{-b}, \tag{1}$$

where $\propto$ should be read "scales as", and b is a constant characterizing the power-law. Since then, power-laws (or derived forms thereof like power-laws with an exponential cutoff) have been universally recognized as the best functions to model cloud size distributions obtained from either satellite imagery (Welch et al., 1988; Sengupta et al., 1990; Kuo et al., 1993; Benner & Curry, 1998; Zhao & Di Girolamo, 2007; Koren et al., 2008; Wood & Field, 2011; Bley et al., 2017; Senf et al., 2018), aircraft measurements (Benner & Curry, 1998; Jiang et al., 2008; Wood & Field, 2011), or high-resolution simulations (Neggers et al., 2003; Xue & Feingold, 2006; Jiang et al., 2008; Dawe & Austin, 2012; Heus & Seifert, 2013; Rieck et al., 2014; Senf et al., 2018).

Being able to accurately characterize cloud size distributions and recognize them as power-laws (or any other functional form) is not just a matter of satisfying our curiosity. Doing so may indeed be of prime importance for the development of improved cumulus parameterizations, and may give crucial indications on how clouds organize.

If we first consider the parameterization issue, knowing the precise distribution of cumulus clouds may help constrain spectral schemes (Arakawa & Schubert, 1974) for which the subgrid variability associated with convection can be described using a size-resolved cloud population. By explicitly introducing information on cloud sizes, such schemes can easily be made scale-aware (Neggers, 2015) and therefore provide interesting solutions to parameterizing convection in the "grey zone" (that is at model resolutions equivalent to the characteristic convective length scale). Following this idea, Neggers (2015) and Brast et al. (2018) introduced and tested an extension of the Eddy Diffusivity Mass-Flux scheme (EDMF, (Siebesma et al., 2007)) in which the cloud population is assumed to be distributed in size following a power-law like Eq. (1). The model was found to yield promising results (Brast et al., 2018) by providing a gradual transition between fully resolved and fully parameterized convection across the "grey zone", although this transition was found to be too sharp compared to coarse grained data. This points to the necessity to further improve the representation of cloud size distributions is models.

Besides, as mentioned earlier, knowing that cloud sizes are power-law distributed may give indications on the mechanisms controlling the organization of cumulus cloud ensembles. Over the years, numerous generative mechanisms and hypotheses have indeed been put forth to explain power-law behaviors often found in social sciences, natural sciences, economics or physics (Newman, 2005; Sornette, 2006; Marković & Gros, 2015). Among these, self-organized criticality (SOC) (Bak et al., 1988; Jensen, 1998; Marković & Gros, 2015) has been frequently invoked to rationalize the emergence of power-law cumulus and rain cluster size distributions (Peters & Neelin, 2006; Peters et al., 2009, 2010; Yano et al., 2012; Teo et al., 2017; Windmiller, 2017). Systems displaying SOC properties are dynamical systems that are slowly driven towards a critical point marking the transition between two distinct states. However, SOC systems are also characterized by a quick relaxation of their internal energy when criticality is reached such that the system is effectively attracted towards the critical point. The emergence of SOC generally does not depend on the details of the system or driving mechanism, it does not require tuning (it self-organizes), and it can be described by scale-invariant properties that show up as power-law distributions. A priori, these propositions agree well with what we know from convective cloud ensembles.

Despite the large number of studies reporting power-law cloud size distributions, it should be stressed that unambiguously identifying scale-free behaviors in empirical distributions is not a simple task. This is reflected by the large spread in power-law exponents reported in the literature, varying between ~ 1 (Benner & Curry, 1998) to more than 3 (Bley et al., 2017; Senf et al., 2018) (with most reported values lying in the $1.7 - 2.4$ range). As argued by Goldstein et al. (2004) and Clauset et al. (2009), relying solely on the apparent constant slopes of empirical distributions in log-log coordinates may introduce large biases when estimating the exponent b in Eq. (1). In particular, linear regression requires the subjective determination of the range over which the power-law holds (and is sensitive thereof), and it is known to be generally too sensitivity to the noise in the distributions tails. These limitations may in part explain the range of cloud size distribution exponents mentioned previously. Consequently, Clauset et al. (2009) proposed an alternative fitting procedure dedicated to power-law distributions and making use of a combination of robust statistical methods among which maximum likelihood estimation and goodness-of-fit tests. Although the procedure has previously been applied to fit rain cluster size distributions (Peters et al., 2010; Traxl et al., 2016), we wish to emphasize here following Clauset et al. (2009) that the use of such advanced statistical methods should always be preferred over simple linear regression techniques when estimating power-law best fits to cloud size distributions.

In this work, we analyze cloud size distributions obtained from Cloud Resolving Model (CRM) simulations of two well-documented idealized convective situations: a diurnal transition from shallow to deep convection over land (Grabowski et al., 2006), and a case of maritime shallow convection in the trade-winds region (vanZanten et al., 2011). These cases were selected as they are representative of cloud populations evolving at different spatial and temporal scales. Besides, the degree of idealization adopted facilitates our analysis by suppressing variability introduced by large-scale transport or heterogeneous surface conditions. The main objective is then to adapt the methodology proposed by Clauset et al. (2009) to identify power-laws in empirical cloud size distributions, evaluate the goodness of these fits, and test possible alternatives. Thanks to the robustness of the fitting procedure, our results are later used to identify and analyze systematic behaviors (organization) in the cloud ensembles.

The paper is organized as follows. In section 2, we give details on the numerical setups used to extract cloud size distributions as well as on the statistical methods forming the basis of our fitting algorithm. Results from the procedure, testing both the power-law and exponential hypotheses, are presented in section 3. Various alternative distributions are then evaluated in section 4 and possible generative mechanisms explaining the simulated empirical cloud size distributions are discussed in section 5. Finally, we give our conclusions in section 6.

2 Methods

2.1 Numerical experiments

The first selected numerical experiment is based on the Large-scale Biosphere-Atmosphere (LBA) intercomparison study Grabowski et al. (2006), with modifications similar to Böing et al. (2012) and Savre (2021). The initial potential temperature profile was taken from Grabowski et al. (2006), while the initial moisture content was modified with a relative humidity held constant and equal to 80% between the surface and 2.5 km, and decreasing linearly to 15% up to 18 km. Horizontal winds were initially set to 0 m s^{-1} everywhere. Although a diurnal cycle is imposed through time dependent surface fluxes in the original case description, we chose here to adopt fluxes constant in time, equal to 161/343 W m^{-2} (sensible and latent fluxes respectively) following Böing et al. (2012). The numerical model is doubly periodic, uses a constant horizontal grid spacing of 100 m with 1024×1024 grid points. The domain extends vertically to 14250 m with 180 grid points.The vertical grid spacing is held constant and equal to 25 m below 2 km, and increases geometrically above. A Rayleigh damping layer with characteristic time scale of 2 h was prescribed above 12 km, and horizontal winds were nudged everywhere to their initial value with a time scale of 6 h. The simulation was run for a total of 10 h, but the first 4 h were discarded.

The second case follows the Rain In Cumulus over the Ocean (RICO) model intercomparison (vanZanten et al., 2011), as simulated for example by Dawe and Austin (2012); Heus and Seifert (2013). The overall setup follows closely the description given by vanZanten et al. (2011), with similar initial potential temperature, total water mixing ratio, horizontal wind, and large scale forcing profiles. Here again, no interactive radiation is employed. Instead, net radiative cooling is accounted for in the prescribed forcing. Surface fluxes are computed using bulk formula with a prescribed surface temperature of 299.8 K, and surface water vapor assumed to be at saturation. The numerical domain spans 20 km^2 in the horizontal, with a homogeneous grid spacing of 26 m (768×768 grid points). This is similar to the setup employed by Heus and Seifert (2013) despite a slightly smaller domain (25 km^2 in Heus and Seifert (2013)). The domain extends vertically up to 3.9 km with a constant spacing of 26 m. A Rayleigh damping layer with characteristic time scale of 2 h was also used above 3.2 km. The simulation was run for 16 h, and the results analyzed after 8 h.

The LBA simulation exhibits a fast transition from shallow to deep, precipitating convection, with clouds reaching 10 km in altitude. This transition occurs within a few hours, although a little bit slower than in the original configuration from Grabowski et al. (2006). The cloud population then continues to evolve under the influence of subcloud layer organization (Böing et al., 2012), with clouds continuously deepening and eventually forming anvils. In contrast, the RICO case is mostly driven by the large-scale forcing held constant during the course of the simulation. This results in a very slowly evolving cloud population (Heus & Seifert, 2013), with clouds remaining relatively shallow (maximum depth on the order of 2 km) and light drizzle at the surface. These two simulations, driven by mechanisms operating at different time scales, result in cloud populations presenting widely different characteristics.

Both experiments were carried out using the MISU-MIT Cloud and Aerosol (MIMICA) model solving anelastic governing equations for potential temperature, total water mass mixing ratio and momentum (Savre et al., 2014; Savre, 2021). The numerical methods employed in both cases are similar: scalar advection uses a flux-limited version of the Lax-Wendroff scheme, momentum advection uses 4th order central finite differences, and time integration is performed using a 2nd order Runge-Kutta method. Turbulent mixing was parameterized using the Smagorinsky-Lilly closure, whereas no interactive radiation was necessary in either case. Different microphysical schemes were employed for the LBA and RICO simulations. In the first case, we used the simple one-moment microphysics from Grabowski (1998) which only distinguishes between precipitating and non-precipitating cloud particles, and partitions liquid and ice based on a linear function of the temperature. In the second case, the warm part of the more advanced two-moment microphysics scheme from Seifert and Beheng (2006) was used.

2.2 Cloud identification

The first step towards estimating cloud size distributions is to define cloud objects and extract their properties. To do so, horizontal slices are first extracted at 2000 m and 1200 m in the LBA and RICO cases respectively. Two different criteria are then used to identify cloud objects: a cloud water mixing ratio q exceeding 0.01 g kg^{-1}, or a Condensed Water Path (CWP, including both liquid and ice particles) exceeding 50 g m^{-2} in the LBA case, or 5 g m^{-2} in the RICO case (different thresholds were chosen to accommodate the different cloud depths). In both situations, clouds are simply identified as clusters of connected grid boxes respecting either criterion above. Only four point connectivity is considered here, that is connected grid boxes must at least share a face. Clusters consisting of a single grid points were discarded. Once cloud objects are identified, their equivalent sizes are calculated as $L = \sqrt{\Delta x \Delta y N_{pts}}$, with N_{pts} the number of grid points covered by a cloud.

The distinction between clouds identified from q and CWP criteria is justified by the fact that both definitions correspond to distinct situations commonly met in atmospheric sciences. The CWP criterion allows the identification of clouds as seen from above, and in particular as retrieved from satellite imagery. In contrast, the q criterion is relevant in the context of convection parameterizations since these latter are mostly concerned with cloud properties near their base and as they develop in the lower free troposphere.

2.3 Fitting size distributions

2.3.1 Theoretical distributions

Four kinds of theoretical distributions are evaluated: power-law, exponential, Weibull and power-law distributions with an exponential cutoff (cutoff power-laws). These can be written:

$$\begin{aligned} p_{PL}(L) &= C_{PL} L^{-\alpha} && (2) \\ p_E(L) &= C_E \exp(-\lambda L) && (3) \\ p_W(L) &= C_W L^{\beta-1} \exp\left(-\eta L^{\beta}\right) && (4) \\ p_{PE}(L) &= C_{PE} L^{-\nu} \exp(-\mu L), && (5) \end{aligned}$$

where L is the cloud size, p_X are the theoretical distribution functions and C_X are appropriate normalizing factors. α, λ, β, η, μ and ν are the parameters characterizing each distribution that we seek to estimate. Log-normal distributions have also been considered as possible alternative distributions, but they generally yielded worse fits than any other function and were therefore ruled out as viable choices.

The corresponding complementary cumulative distribution functions (CCDFs) are defined from $p_X(L)$ as:

$$P_X(L) = \int_L^{L_{max}} p_X(L^*)\, dL^*, \tag{6}$$

where L varies from L_{min} to L_{max} (the size bounds over which the fits are performed), and $P_X(L)$ represents the probability that a given cloud has a size larger than L. This yields, for the four tested distributions:

$$\begin{aligned} P_{PL}(L) &= \frac{L^{1+\alpha} - L_{max}^{1+\alpha}}{L_{min}^{1+\alpha} - L_{max}^{1+\alpha}} && (7) \\ P_E(L) &= \frac{\exp(-\lambda L) - \exp(-\lambda L_{max})}{\exp(-\lambda L_{min}) - \exp(-\lambda L_{max})} && (8) \\ P_W(L) &= \frac{\exp\left(-\eta L^{\beta}\right) - \exp\left(-\eta L_{max}^{\beta}\right)}{\exp\left(-\eta L_{min}^{\beta}\right) - \exp\left(-\eta L_{max}^{\beta}\right)} && (9) \\ P_{PE}(L) &= \frac{\Gamma(1+\nu, \mu L) - \Gamma(1+\nu, \mu L_{max})}{\Gamma(1+\nu, \mu L_{min}) - \Gamma(1+\nu, \mu L_{max})} && (10) \end{aligned}$$

with Γ the upper incomplete gamma function defined by:

$$\Gamma(a, b) = \int_b^{+\infty} x^{a-1} e^{-x} dx. \tag{11}$$

Theoretical distributions are here defined over a finite size range $L_{min} - L_{max}$ following Deluca and Corral (2013) (what is called "truncated distributions"). The use of truncated functions appears as a necessity since in most cases the maximum cloud sizes remain small and the power-law scalings are relatively narrow (at most an order of magnitude).

Let's stress that the advantage of using and plotting CCDFs instead of standard frequency distributions is twofold. First, calculating cumulative distributions does not require binning the data. Second, when plotted, heavy-tailed distributions often appear noisy, especially in the tail, while their cumulative counterparts are smoother. As a result, linear regression in log-log coordinates generally introduces important biases when fitting frequency distributions (a 40% error can be expected for a power-law with $\alpha = 2.5$) (Goldstein et al., 2004; Clauset et al., 2009).

2.3.2 The Clauset et al. method

Clauset et al. (2009) proposed a detailed procedure to identify power-law distributions in empirical data and estimate their exponents with great precision. The suitability and precision of the whole procedure were tested thoroughly by Clauset et al. (2009) for many classical empirical distributions exhibiting power-law behaviors.

In the following, a modified version of the method proposed by Deluca and Corral (2013); Peters et al. (2010) is described and used. The modifications introduced allow the application of the procedure to truncated distributions for which both the lower and upper bounds of validity must be simultaneously estimated. The procedure can be summarized as follows (each method is described in detail in section 2.3.3):

1. Best fit power-law exponents $\widehat{\alpha}$ are estimated using Maximum Likelihood Estimation (MLE) for all possible values of L_{min} and L_{max};
2. The goodness-of-fit between the empirical and theoretical distributions is calculated for each triplet $\{\widehat{\alpha}, L_{min}, L_{max}\}$ using a Kolmogorov-Smirnov (KS) goodness-of-fit test;
3. The best fit triplet is the one that maximizes the ratio $r = L_{max}/L_{min}$, while maintaining the computed KS statistics (D) below an arbitrary threshold set to 0.05. The p-value associated with the KS statistics for the retained triplet is then computed using Monte-Carlo sampling. The fit is accepted only if the p-value is below an arbitrarily chosen confidence level of 5%;
4. The power-law fit can finally be compared to alternative distributions over the same range $L_{min} - L_{max}$ using the likelihood ratio (LR) test.

Although the method was originally designed to estimate power-law fits only, we will also employ it to determine best fits to the alternative distributions introduced in section 2.3.1. The mathematical foundation behind the application of the technique to distributions other than power-laws may be weaker, we are considering it for the sake of comparison.

2.3.3 Statistical methods

The procedure described above makes use of several statistical methods briefly summarized below, starting with MLE. Let's assume a set of empirical data $\mathbf{x} = (x_{min}, ..., x_{max})$ that we wish to approximate by a known distribution $p_{X|\Theta}$ described by N parameters $\Theta = (\theta_1, ..., \theta_N)$. Defining the log-likelihood function as:

$$\ell(\Theta, \mathbf{x}) = \ln \prod_{x=x_{min}}^{x_{max}} p_{X|\Theta}(x), \tag{12}$$

the set of parameters $\widehat{\Theta}$ yielding the best fit is the one that maximizes $\ell(\Theta, \mathbf{x})$. For numerous standard distributions such as non-truncated power-law and exponential distributions, the optimal parameters maximizing $\ell(\Theta, \mathbf{x})$ can be found analytically by letting $\partial\ell(\Theta, \mathbf{x})/\partial\Theta = 0$. In the more general case however, and in particular for the truncated distributions introduced in section 2.3.1 for which $\{L_{min}, L_{max}\} \in \Theta$, $\widehat{\Theta}$ cannot be reduced to a simple analytical formula. In this situation, $\widehat{\Theta}$ is obtained from the numerical maximization of the log-likelihood function ℓ.

Once optimal parameters $\widehat{\Theta}$ have been found, the KS statistics D is computed to give an estimate of how good the fit it. D is simply defined by the maximum absolute distance between the empirical and theoretical CCDFs between x_{min} and x_{max}:

$$D = \sup_{x_{min}<x<x_{max}} \left| P_e(x) - P_{X|\widehat{\Theta}}(x) \right|, \tag{13}$$

where P_e and $P_{X|\widehat{\Theta}}$ denote the empirical and theoretical cumulative distributions.

Whereas lower D values intuitively indicate better fits, the statistical significance of this quantity depends strongly on the number of data points considered. For this reason, the statistics must be complemented by a $p-$value computed from the probability density function (PDF) of D, indicating the probability that the underlying hypothesis should be rejected. In the situation where D is calculated for a theoretical distribution whose parameters are estimated, Monte-Carlo sampling must be used (Clauset et al., 2009; Deluca & Corral, 2013). First, n points (n being the number of points in the empirical data set) are drawn randomly from the estimated theoretical distribution $p_{X|\widehat{\Theta}}$. New values for the best fit parameters are then obtained for the simulated data, and the associated KS statistics is computed. Repeating the procedure a sufficiently large number of times (500 in our case), a D PDF can be constructed, and if the probability of occurrence of the original D value is equal to or lower than a predefined threshold (here set to 0.05), the test is rejected.

Finally, to compare power-law distributions to other plausible hypotheses an alternative goodness-of-fit test is employed, the LR test. More generally, best fits obtained for any distribution p_X can be compared to any other alternative hypothesis by means of the LR test. For any best fit distribution $p_{X|\widehat{\Theta}}$ determined over the range $x_{min} - x_{max}$, and any alternative distribution $p_{X|\widehat{\Phi}}$ described by a set of parameter $\widehat{\Phi}$ fitted over the same x range, the log-likelihood ratio is defined by:

$$LR = -2\ln\frac{\mathcal{L}\left(\widehat{\Theta}, \mathbf{x}\right)}{\mathcal{L}\left(\widehat{\Phi}, \mathbf{x}\right)} = -2\left[\ell\left(\widehat{\Theta}, \mathbf{x}\right) - \ell\left(\widehat{\Phi}, \mathbf{x}\right)\right] \tag{14}$$

where $\mathcal{L}$ is the standard likelihood function. Because larger ℓ values indicate a higher probability for the data to be drawn from the hypothesized distribution, a negative likelihood ratio $LR < 0$ means that the null hypothesis is more likely than the alternative hypothesis. In contrast, a positive LR value means that the alternative hypothesis is more likely. In practice, LR must be sufficiently negative (respectively positive) for the null hypothesis (respectively the alternative hypothesis) to be unambiguously identified as the better hypothesis.

3 Power-law and exponential fits

We are initially interested in fitting simulated cloud size distributions to power-law and exponential functions only. These two functions are the simplest among those presented in section 2.3 as they only depend on a single degree of freedom. Besides, power-laws are the most common functions used to fit empirical cloud size distributions, while exponentials are ubiquitous in natural sciences and physics, including atmospheric sciences (Craig & Cohen, 2006).

3.1 Visual inspection

Cumulative Cloud Size Distributions (or CCSDs) extracted at three different times into the LBA and RICO simulations are shown in figures 1 and 2 respectively. At this point, we are only concerned with the empirical distributions depicted as solid grey lines. In the LBA simulation, starting at 4 h, clouds identified using a CWP threshold are generally bigger and twice as numerous as those determined from a q threshold ($\sim$ 6000 with CWP as opposed to $\sim$ 3200 with q). At this time, most clouds remain indeed relatively shallow and do not reach an altitude of 2000 m where they can be identified using the q criterion. At later times, the largest clouds now reach about 2 km and 4 km when determined based on q and CWP thresholds respectively. The number of clouds identified using CWP is also largely reduced to $\sim$ 3500. The emergence of much bigger clouds indicates the development of convective outflows and the possible merging of cloud objects as convection becomes deeper. After 10 h, a clear linear scaling in log-log coordinates between 200 m and 3000 m is visible, and the largest clouds reach up to 9 km in size. Despite the narrower overall cloud size range, a power-law scaling between 250 m and 1500 m is also visible on the CCSD obtained from a q threshold.

As expected, CCSDs extracted from the RICO simulation (Figure 2) do not show the same evident temporal evolution as the ones obtained from the LBA case. The total number of clouds identified based on the q criterion is relatively constant at $\sim$ 220, but the biggest clouds in the domain get bigger over time (from 700 m to 1200 m). No power-law scaling can be identified here. Again, many more clouds can be identified using a CWP threshold (between 600 and 700), with cloud objects reaching up to 3.3 km in size after 16 h. Again, no clear power-law behavior is evident at first sight. Note however that the visual determination of power-law behaviors is here biased by the fact that truncated power-law CDFs do not necessarily appear as straight lines in log-log coordinates. This is a direct consequence of the way the truncated function $P_{PL}(L)$ is defined in equation 7.

3.2 Estimated power-law and exponential best fits

The procedure introduced in section 2.3.2 is now applied to the CCSDs described in the previous section. At the moment, we only focus on finding best fits to both truncated power-law and exponential distributions.

Figure 3 shows the temporal evolution of the best fit power-law exponents α and exponential rate parameters λ in the LBA and RICO simulations. In the LBA case, our algorithm was not able to find reliable fits to power-law distributions early in the analysis period. From 5.5 h, considering only clouds defined based on CWP, α is relatively constant in time, taking values between 2.05 and 2.13, with an average of 2.07 between 6 h and 10 h. To evaluate the robustness of these estimates, Root Mean Square Errors (RMSE) can be computed for each fit using a standard bootstrapping procedure. The method involves resampling the empirical data with replacement, computing the corresponding best fit parameters for the resampled data, and repeating the operation a sufficient number of times (here 5000 times) to calculate reliable statistics. The average RMSE (represented by error bars on Figures 3) computed between 6 h and 10 h for the considered case remains low at about 0.04, indicating that the identified power-law scalings are a robust feature of these distributions. More variability is found for α values obtained for clouds identified from q, with estimates ranging between 2.1 and 2.9, and associated RMSE averaging to 0.075 between 6 h and 10 h.

Best fit λ values in the LBA case are seen to decrease continuously with time. λ is indeed directly related to the inverse of the mean cloud size and is thus expected to decrease as clouds get bigger. Early in the simulation, both CWP and q thresholds yield precisely the same λ estimates which suggests that clouds interact only weakly in the shallow cumulus phase. After 6 h, λ values obtained for clouds computed based on CWP decrease abruptly below 0.002 as the exponential best fit range shifts from the bulk to the tail of the distributions (see section 3.3). At this time, clouds suddenly get wider and deeper, and deep convective outflows develop. In contrast, cloud sizes calculated from a q criterion yield λ values relatively constant in time. Again, RMSE calculated for all λ values remain small, giving us confidence into our λ estimates.

In the RICO case, best fit α values obtained for both cloud definitions show stronger temporal variability compared to the LBA case. Only after 12 h do the estimated α seem to stabilize, with mean values of 2.1 and 1.35 for clouds identified based on CWP and q thresholds respectively. The corresponding mean RMSE is also larger than in the LBA case with values of 0.085 and 0.13 respectively. That the power-law best fit estimates are here less reliable than in the LBA case is likely a consequence of the smaller cloud samples available in each RICO scene. Similar remarks can be made regarding best fit exponential parameters for RICO. Compared to the LBA case, λ remains on average constant in time, but display larger temporal variability and larger RMSE.

The mean α estimates discussed previously are in relatively good agreement with power-law exponents reported in the literature where values ranging between ~ 1 to ~ 3.3 can be found depending on the case. Considering deep convection, exponents between $\sim 1.7-1.9$ (Kuo et al., 1993; Rieck et al., 2014) to $\sim 2.6-3.3$ (Bley et al., 2017; Senf et al., 2018) were reported from high-resolution simulations and satellite imagery. Considering maritime shallow convection, our average α estimate is somewhat smaller than the value of 2.42 given by Heus and Seifert (2013) for simulations performed under similar conditions, but generally larger than values reported for other similar cases between 1.7 (Neggers et al., 2003) and 1.9 (Dawe & Austin, 2012). Satellite retrieval yields a broader range of exponents, from 1.6 to 2.2 (Benner & Curry, 1998; Zhao & Di Girolamo, 2007; Koren et al., 2008; Wood & Field, 2011).

3.3 Power-law and exponential best fit ranges

Figure 4 displays the calculated best fit size ranges obtained for the LBA and RICO simulations. In the LBA case, when clouds are identified based on a q threshold, the power-law range remains very narrow (less than a decade), and is always narrower than the exponential range. This latter frequently extends up to the biggest clouds in the domain, especially before 8 h. In contrast, the power-law range for clouds determined from CWP increases over time as clouds get deeper and wider, with a cloud size ratio r reaching 23.5 at 10 h. A clear transition can be seen at 6.5 h, with the power-law fits becoming valid over an increasingly broader size range, and the exponential fits shifting to the distributions tails. Interestingly, the power-law fits are also found to extend up to the largest clouds identified, something made possible by the use of truncated functions. Judging goodness-of-fit based on the fit range only, CCSDs computed based on a q criterion are best represented by exponential distributions, whereas those obtained from CWP are best modeled by power-law distributions (especially in the deep convection regime after 6 h).

Considering the RICO simulation, exponential distributions generally produce best fits valid over a broader range than power-laws for clouds computed from q. That best fits rarely extend to the biggest clouds identified may be a consequence of the relatively small cloud samples available. Considering clouds identified from a CWP threshold, both power-law and exponential best fits extend over about one order of magnitude at almost all times (r ranging between 9 and 20). Contrary to the LBA case, exponential best fits are generally valid over broader size ranges than power-laws, and preferentially cover the bulk of the empirical CCSDs. This suggests that exponential distributions may be the best choice to represent all CCSDs from RICO.

3.4 Comparison to linear regression

To evaluate biases introduced when fitting power-laws using linear regression in log-log coordinates, power-law exponents were recomputed with this technique for all empirical CCSDs analyzed previously. The results are compared to our best fit estimates in Figure 5. Linear regression was here applied over size ranges determined visually from the distributions plotted in Figures 1 and 2.

Linear regression generally tends to overestimate power-law exponents obtained by MLE in both LBA and RICO cases. In the LBA case, linearly regressed exponents differ only moderately from the MLE ones with α values averaged between 6 h and 10 h of 2.24 and 2.85 (to be compared to the MLE estimates of 2.07 and 2.52). The corresponding relative errors computed for clouds identified from CWP and q thresholds amount to 23% and 35% respectively. In contrast, estimates from linear regression computed for the RICO simulation show a large spread around the MLE ones, with mean errors reaching almost 100% with both cloud identification criteria.

The somewhat erratic power-law exponents obtained by linear regression can be directly attributed to the difficulty of clearly identifying the range over which the fits must be performed. As an illustration, best fit exponents for CWP clouds at time 10 h in the LBA case were recomputed using lower and upper size bounds varying from 200 m to 500 m, and from 1000 m to 3000 m respectively (Figure 6). α is found to increase systematically with increasing L_{min} and L_{max} such that the smallest estimated exponent of 2.05 is found in the range $200-1000$ m, and the largest exponent of 2.32 in the range $500 - 3000$ m. This sensitivity can be explained by the fact that linear regression tends to be overly sensitive to the few, noisy points located in the distributions tails where the slope is steeper.

Overall, these results indicate that the use of linear regression may explain some of the largest exponent estimates reported in the literature. Lower estimates may in turn be explained by the subjective choice of fitting ranges biased towards smaller cloud sizes where distributions are generally flatter.

4 Alternative distributions

4.1 Visual impression and direct goodness-of-fit

Figures 1 and 2 allow us to visually compare best fits obtained for all distributions proposed in section 2.3.1. Fits were obtained applying the procedure described in section 2.3.2 to all distributions. Corresponding best fit parameters are reported in tables 6 and 6. In both the LBA and RICO cases, the exponential and power-law functions only irregularly extend to the largest cloud sizes, whereas the Weibull and cutoff power-law distributions consistently provide excellent fits to the distributions tails with minimal errors. Large differences are also obtained for smaller clouds where both exponential and power-law distributions generally provide poorer fits.

A more objective way to compare the different models is given by the range $L_{min} - L_{max}$ over which each fit is valid (see Figures 7). Note that since D depends implicitly on the size range, it cannot be used here to draw a fair comparison between fits obtained over different ranges. First considering the LBA case (top panels), the size ranges indicate that exponential distributions provide better fits than power-laws for CCSDs calculated from a q threshold, while the opposite is true for CCSDs computed from a CWP threshold. The Weibull and cutoff power-law distributions are generally valid over comparable size ranges, these being systematically broader than those obtained with both the exponential and pure power-law. In the RICO case (bottom panels), conclusions consistent with our previous analysis can be drawn, with the exponential distribution being a better model for CCSDs calculated from q (broader

$L_{min} - L_{max}$ range). The results are more contrasted for CWP clouds, with the exponential distribution being perhaps again the better choice. As in the LBA case, both alternative distributions are valid over similar size ranges being broader than those obtained with the exponential and pure power-law functions.

Consistent with our visual inspection, this analysis suggests that in all situations, both the Weibull and cutoff power-law distributions provide the best representations of our empirical CCSDs. However, the two hypotheses can't be easily differentiated without further analyses.

4.2 Likelihood ratio tests

As mentioned in sections 2.3.2 and 2.3.3, the LR test constitutes a more robust tool to assess the goodness of various hypotheses. The strength of the LR test resides in the fact that different hypotheses can be compared over the same ranges. Matrices of LR values for each possible combination of null and alternative hypotheses are shown on Figures 8 and 9. Null hypotheses (the reference) are shown on the horizontal axes, while alternative hypotheses can be read on the vertical axes. If a matrix element corresponding to given null and alternative distributions is blue (respectively red), corresponding to a negative (resp. positive) LR, the reference (resp. alternative) distribution can be regarded as the better fit. Grey squares indicate that no reliable fit could be found with the corresponding alternative distribution.

Focusing first on the LBA simulation (Figure 8), the LR test confirms that the cutoff power-law is the distribution yielding the best fits in all cases. This can be deduced from the positive LR values obtained when the cutoff power-law is used as the alternative distribution (red squares in the top rows), and from the negative LR values obtained when it is used as the reference distribution (blue squares in the rightmost columns). Both exponential (at earlier times) and Weibull (at later times) distributions are also found to yield reasonable fits in certain situations. In contrast, the power-law appears to be constantly outperformed by other alternatives (red squares in the leftmost columns, and grey or blue squares in the bottom rows).

In the RICO case (Figure 9), both exponential and power-law fits are outperformed by the two alternative distributions at any time and with any cloud definition. For clouds identified from a q threshold, the Weibull and cutoff power-law distributions generally result in equally good fits (low LR values), with a slight advantage to the cutoff power-law. For clouds based on a CWP threshold, the two alternative distributions again perform similarly well, although the Weibull distribution now emerges as the better choice.

In summary, the cutoff power-law distribution generally constitutes the best option to approximate CCSDs estimated from the LBA simulation, while the cutoff power-law, Weibull and exponential distributions may all be regarded as the better model in the RICO case depending on the situation. None of the estimators used here suggest that power-laws should be the preferred option to model empirical CCSDs.

5 Physical interpretation and generative mechanisms

5.1 Interpreting exponential fits

Using arguments from statistical physics, Craig and Cohen (2006) predicted that the mass flux distribution of a convective cloud ensemble should follow an exponential distribution. This result was obtained assuming that the system is in equilibrium, that clouds do not interact, and that the mass flux associated with individual clouds is an independent and identically distributed (i.i.d.) random variable. This result is consistent with the maximum entropy principle, according to which the exponential distribution is the distribution with specified mean, supported on the interval $[0, +\infty)$ that maximizes the information entropy. In other words, the exponential distribution is the most likely model to represent cloud mass fluxes in a non-interacting cloud ensemble.

Following the same line of thought, the maximum entropy principle applied to cloud sizes (having a finite mean and being supported on the open interval $[0, +\infty)$) predicts that the most likely model describing CCSDs is the exponential distribution. This is again valid only as long as clouds do not interact strongly. Deviations from this principle may explain why empirical distributions are only loosely represented by pure exponential distributions.

It is interesting to note that only clouds identified from a q threshold appear to preferentially follow exponential distributions. Clouds determined from a CWP threshold are indeed more likely to be subject to complex interactions, merging and possibly overlapping at various altitudes. This may also be true for clouds developing in presence of strong self-organization (for example in the LBA case at 10 h). In this situation, the probability that a cloud forms at a particular location is not uniform but depends strongly on the underlying dynamics (clouds are for example known to develop preferentially at the intersection of propagating cold pool fronts (Haerter et al., 2019)). The emergence of power-law scalings in CCSDs may therefore be a manifestation of strong self-organization, cloud clustering and cloud merging. Conversely, exponential CCSDs may indicate an absence of self-organization.

5.2 Power-laws and heavy-tailed distributions

In general, cloud size distributions were found to be better approximated by Weibull or cutoff power-law distributions, a characteristic already noted by Windmiller (2017) and van Laar et al. (2019). More than just a coincidence, this general feature might give us clues regarding the mechanisms organizing the cloud ensemble.

A distinctive property of the Weibull and cutoff power-law distributions is that they are expressed as the product of a power-law and an exponential. This is reminiscent of the characteristic distributions found for sub-critical percolating (Stauffer & Ahaorny, 1994) and SOC (Bak et al., 1988; Jensen, 1998) systems:

$$p\left(s\right) = s^{-\tau} \mathcal{G}\left(s/s_0\right) \tag{15}$$

with s a characteristic size, s_0 a correlation length, τ a critical exponent and $\mathcal{G}$ a scaling function. In a system of finite size, the correlation length is related to the system size L via $s_0 \propto L^d$, with d a critical exponent. Note that the pure power-law behavior is only recovered asymptotically as the correlation length diverges. For s_0 sufficiently small, finite-size scaling will affect the power-law pre-factor and critical exponents.

By analogy, several authors have proposed SOC as a possible explanation to describe convective cloud and rain cluster ensembles (Peters & Neelin, 2006; Peters et al., 2009, 2010; Teo et al., 2017; Yano et al., 2012). For a system to exhibit SOC, it is generally believed that certain conditions have to be fulfilled (Bak et al., 1988; Jensen, 1998): 1) an external force drives the system slowly towards an unstable state, 2) a threshold exists beyond which avalanches are triggered, 3) there exists a strong scale separation between the slow external force and the fast relaxation following each avalanche (relaxation is key to ensure that energy is conserved). All three criteria are generally met in our simulations and, more generally, in convective situations.

Following these ideas, a simple model of the convective atmosphere can be designed. Here, atmospheric moisture plays the role of the control parameter, as suggested by Peters and Neelin (2006); Peters et al. (2009, 2010), and the following principles apply:

1. The atmosphere is slowly driven towards instability by surface fluxes of energy and moisture. Convective thermals originating from the surface induce local bursts of moisture similar to grains being added one by one in the sandpile model from Bak et al. (1988).
2. When one of these thermals becomes supersaturated, that is when the system's critical threshold is exceeded, a cloud forms, that is an avalanche is triggered.
3. Entrainment and detrainment mix the cloud with its environment leading to moisture being diffused to nearest neighbors. The avalanche grows as long as these neighbors become supersaturated. An avalanche may thus spread to other critical sites nearby, just like clouds may merge with their neighbors.
4. The cloud eventually dissipates, having moistened its environment. A new cycle may then start again.

Note that, as explained by Marković and Gros (2015), systems apparently showing critical behaviors may in fact not reach criticality at all. This applies notably to non-conservative systems as well as when small variations are introduced in classical SOC models. In such situations, the resulting characteristic distributions may differ from equation 15. In non-conservative SOC systems, energy is dissipated each time an avalanche is triggered which provides a stabilizing mechanism that constrains avalanches (clouds) to be of finite sizes even in virtually infinite systems. This state was termed "self-organized quasi-criticality" (SOqC) (Bonachela & Muñoz, 2009) as the system seems to approach criticality without ever reaching it. Going back to the previous analogy between a convective cloud ensemble and SOC, large-scale drying through compensating subsidence in the environment may play the role of the stabilizing factor by removing moisture. It is thus suggested here that SOqC might be a more appropriate model than SOC to describe and explain convective cloud ensembles.

Note that whereas Weibull distributions are generally not considered to be an emerging feature of SOC systems, it has been shown that they often provide excellent fits to empirical distributions exhibiting power-law scalings with quickly decaying tails (Laherrere & Sornette, 1998). As mentioned by Laherrere and Sornette (1998), the Weibull distribution could stem from the superposition of finite-size scaling (pure exponential decay), and deviations from a pure power-law. This is particularly relevant in the RICO case for which the sample size and overall cloud size range are small. In addition, deviations in the tail of the simulated empirical distributions may also be accentuated by dynamical feedbacks constraining the size of the biggest clouds, and possibly being affected by the size of the numerical domain (Heus & Seifert, 2013). From this perspective, the Weibull distribution is not necessarily inconsistent with the SOC hypothesis, and is perhaps a manifestation of natural deviations from the general form represented by equation 15.

5.3 Transition from exponential to power-law scaling

In the quickly organizing LBA simulation, a clear transition from an exponential behavior at earlier times to heavy-tailed cloud size distributions at later times could be identified (see Figure 4). As suggested in sections 5.1 and 5.2, exponential cloud size distributions characterize cloud ensembles where individual clouds only weakly interact, whereas a power-law scaling is a manifestation of self-organized states where short-range interactions prevail. Following this idea, we can expect cloud clustering and cloud merging to become more frequent during the transition.

Cloud clustering can be visualized on Figure 10a where, following Nair et al. (1998), the simulated nearest-neighbour cumulative distribution functions (NNCDF) for CWP clouds in the LBA case are plotted against theoretical NNCDF obtained for randomly distributed clouds. The simulated NNCDFs constantly lie above the diagonal line suggesting that clustering happens at all times (Nair et al., 1998). The earlier NNCDFs (blue lines) indicate however that the cloud ensemble is still close to spatial randomness before 5 h. As time progresses, clustering becomes stronger and peaks at the end of the simulation.

Clustering can alternatively be quantified by means of the I_{org} parameter (Tompkins & Semie, 2017) which is proportional to the NNCDF integral with respect to the diagonal line in Figure 10a. Consistent with the displayed NNCDFs, I_{org} increases with time (Figure 10b), from a value slightly above 0 (clouds being randomly distributed in space) at 4 h, to 0.12 at 10 h. A break in the slope of the I_{org} time evolution at 6 h reveals a change in the behavior of the cloud ensemble. This change coincides with the emergence of a clear power-law scaling in the empirical CCDFs. Although this is insufficient to conclude on a causal relationship between clustering and power-law CCDFs, it is here hypothesized that the two are intimately connected.

Figure 10b displays the time evolution of the mean number of active cores within each identified cloud. To determine active cores, local w maxima were first identified within each cloud object using windows of 5×5 pixels. Only distinct maxima larger than 3 m s^{-1} were then counted as independent, and clouds with no such local maximum were counted as having a single core. Overall, the number of cores per cloud increases throughout the simulation indicating that individual clouds are more likely to contain several cores, or that individual clouds are made up of an increasingly large number of cores. This suggests that as clouds organize and cluster (increasing I_{org}), they are also more likely to merge. As a result, the very large clouds constituting the tail of the empirical CCSDs (in particular after 6 h) are in fact formed by a collection of several individual cores that have merged in regions of strong clustering.

6 Conclusion

In order to provide more robust fits to cumulus cloud size distributions, and therefore permit unbiased interpretations of the properties and organization of a cloud ensemble, an advanced fitting algorithm, inspired by the works of Clauset et al. (2009), was described and applied to simulated convective situations. The method is based on the following principles: 1) smooth cumulative distributions are used instead of the more noisy frequency distributions; 2) a robust maximum likelihood estimator is employed to determine best fits to predefined theoretical distributions; 3) a goodness-of-fit test is employed to find the optimal size range over which these fits hold. In addition, the described algorithm also permits direct comparisons between best fits obtained with various distributions including exponential, power-law, Weibull and cutoff power-law functions. Overall, the method directly addresses some of the main issues generally associated with fitting techniques based on linear regression.

The algorithm was demonstrated using two cloud resolving model simulations representative of the diurnal shallow-to-deep convection transition over land, and maritime shallow cumuli in the trade-winds region. In addition, two criteria were tested to identify cloud objects: the first is based on a condensed water content (q) threshold, and the second on condensed water path (CWP). For clouds identified based on q, empirical size distributions were generally reasonably well approximated by exponential distributions, although alternative distributions (Weibull and cutoff power-law) frequently yielded better fits. Clear power-law scalings were identified for CWP based cloud size distributions. However, despite these robust estimates, the alternative heavy-tail functions tested were found to unequivocally be the best models representing the simulated distributions.

Two main mechanisms were invoked to explain the emergence of both exponential and heavy-tail cloud size distributions in our simulations. Exponential distributions can be derived from the maximum entropy principle: the exponential is the most

Table 1. Best fit parameters estimated from the CCSDs obtained in the LBA simulation. Cloud objects are identified based either on a q threshold at $z = 2000\ m$ (top rows), or a CWP threshold (bottom rows).

Time	$\widehat{\lambda}$	r	$\widehat{\alpha}$	r	$\widehat{\eta}$-$\widehat{\beta}$	r	$\widehat{\mu}$-$\widehat{\nu}$	r
10 h	3.26×10^{-3}	6.5	-2.34	6.4	0.15-0.47	11.1	1.27×10^{-3}-1.16	11.1
7 h	3.64×10^{-3}	8.6	-2.90	3.4	0.042-0.66	7.5	3.0×10^{-3}-0.36	8.6
4 h	5.71×10^{-3}	3.1	-	-	0.26-0.44	3.5	6.71×10^{-3}--0.92	5.3
10 h	0.84×10^{-3}	7.5	-2.05	23.5	1.39-0.20	34.8	2.1×10^{-4}-1.71	34.8
7 h	1.66×10^{-3}	6.7	-2.06	9.1	0.21-0.42	17.1	7.1×10^{-4}-1.37	17.1
4 h	5.68×10^{-3}	3.0	-	-	0.04-0.66	7.2	3.37×10^{-3}-0.11	7.2

probable size distribution model as long as clouds interact only weakly and do not organize (random occurrence). In contrast, the emergence of power-law scalings was suggested to be the manifestation of self-organized criticality (SOC) or, to be more precise, self-organized quasi-criticality (SOqC). A direct analogy between convective cloud ensembles and a typical (non-conservative) SOC model could indeed be drawn where water vapor plays the role of the control parameter, and clouds correspond to avalanches triggered when water vapor locally exceeds a critical threshold (saturation). Note however that as attractive as the concept may be, a more careful evaluation of convective cloud ensembles and their characteristics should be conducted before concluding on the relevance of SOC.

The two mechanisms mentioned here were shown to be consistent with the transition observed in the continental convection case from an exponentially distributed cloud population at earlier times (when shallow clouds prevail), to heavy-tail cloud size distributions at later times (deep convective regime). As the transition takes place, it was indeed shown that both cloud clustering and cloud merging increase, thus supporting the fact that exponentials characterize weakly interacting clouds, while heavy-tail functions are a manifestation of self-organization.

The two cases analyzed in this study yielded comparable power-law exponents of about 2.1 (when applicable), a value consistent with most estimates reported in the literature, and with the theoretical critical exponent of 2 expected from relevant SOC models. Note however that previously published best fit exponents take values ranging between ~ 1 to > 3, a variability that can only in part be explained by the use of unsuitable fitting techniques. While we can expect the proposed method to increase the confidence associated with best fit exponent estimates, other factors are likely responsible for much of the variability. For example, it has been shown previously that exponent values may vary as clouds become more mature. It was also suggested that boundary layer properties influence the size of the biggest clouds in a cloud scene, and possibly the power-law scaling. Overall, this suggests that more efforts should be put into trying to understand the factors influencing scaling exponents of cumulus size distributions.

Finally, it should be stressed that our analysis of empirical cloud size distributions may be affected by sub-sampling biases. If the cloud sample available is too small, we may indeed expect errors in the distribution tails as the largest clouds become under-represented. Ultimately, the best solution to minimize these biases and yield reliable fits remains to use very large samples obtained from long-term, large domain simulations.

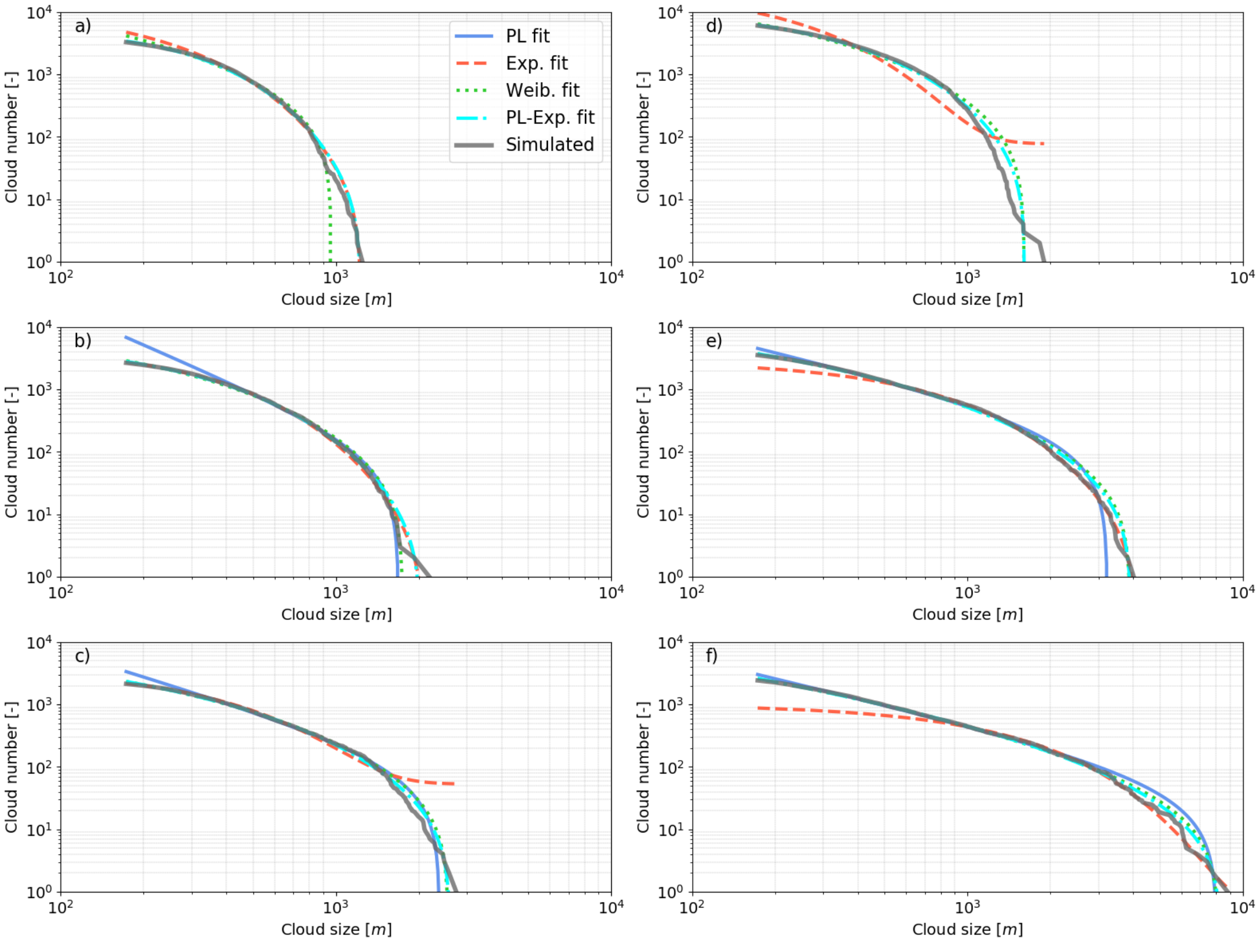


Figure 1. CCSDs calculated for the LBA diurnal cycle case, at three different times into the simulation. From top to bottom: at 4 h, 7 h and 10 h. Left column: clouds are identified based on a condensed water content q threshold of 0.01 $g.m^{-3}$ at 2000 m. Right column: clouds are identified based on a CWP threshold of 50 g m^{-2}. Best fits obtained with power-law, exponential, Weibull and cutoff power-law functions are also plotted. The corresponding best fit parameters are collected in Table 6).

Table 2. Same as table 6 but for the RICO simulation.

Time	$\widehat{\lambda}$	r	$\widehat{\alpha}$	r	$\widehat{\eta}$-$\widehat{\beta}$	r	$\widehat{\mu}$-$\widehat{\nu}$	r
12 h	7.93×10^{-3}	11.0	-1.34	7.5	0.029-0.78	11.0	5.61×10^{-3}-0.47	11.0
8 h	9.18×10^{-3}	8.6	-1.97	4.9	0.17-0.51	5.3	7.72×10^{-3}-0.14	5.8
4 h	8.08×10^{-3}	7.0	-1.75	7.5	0.037-0.75	10.4	5.75×10^{-3}-0.59	10.4
12 h	7.01×10^{-3}	12.0	-2.18	19.2	0.10-0.56	43.0	2.2×10^{-3}-1.05	43.0
8 h	5.40×10^{-3}	14.3	-2.24	9.0	0.078-0.59	26.7	2.55×10^{-3}-0.93	26.7
4 h	6.24×10^{-3}	17.1	-2.12	9.1	0.073-0.6	23.5	2.63×10^{-3}-0.9	23.5

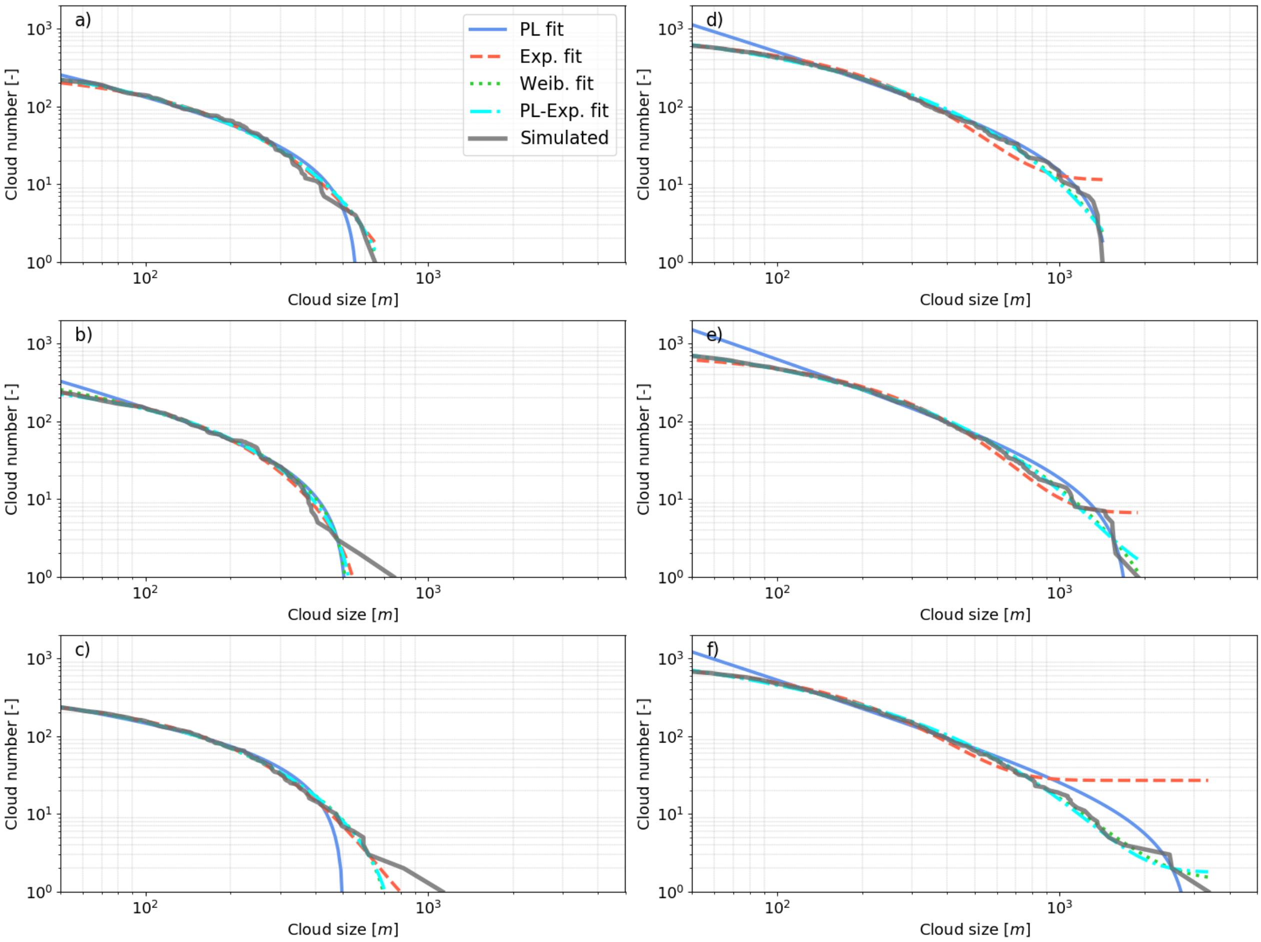


Figure 2. Same as Figure 1 for the RICO case. From top to bottom: at 8 h, 12 h and 16 h. The corresponding best fit parameters are collected in Table 6).

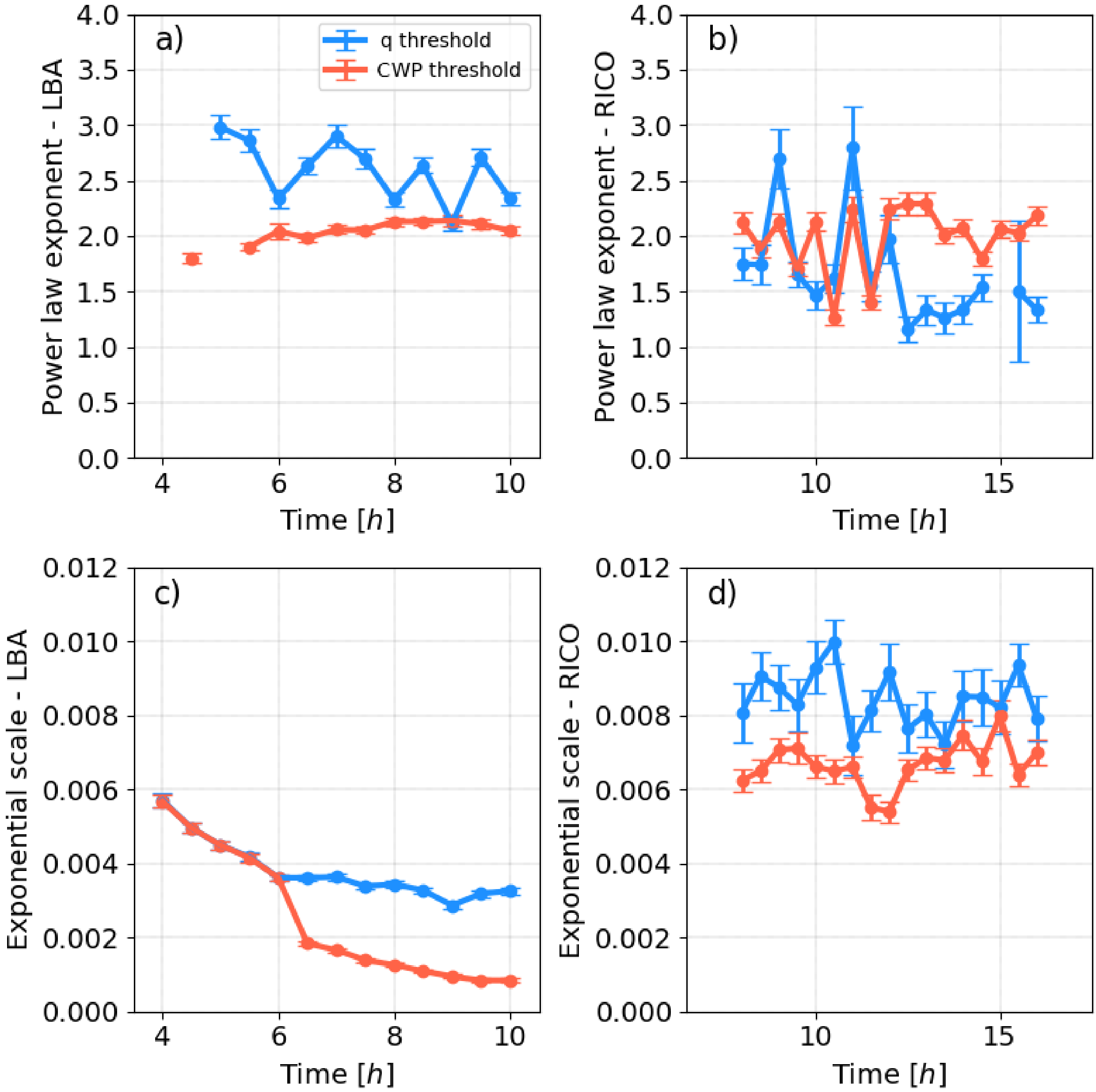


Figure 3. Time series of estimated power-law exponents α, and exponential scales λ estimated in the LBA (left, panels a and c) and RICO (right, panels b and d) cases. Uncertainty bars corresponding to the root mean square errors computed directly from bootstrapping (see text) are also depicted. Estimates that do not satisfy the $p-$value condition are omitted.

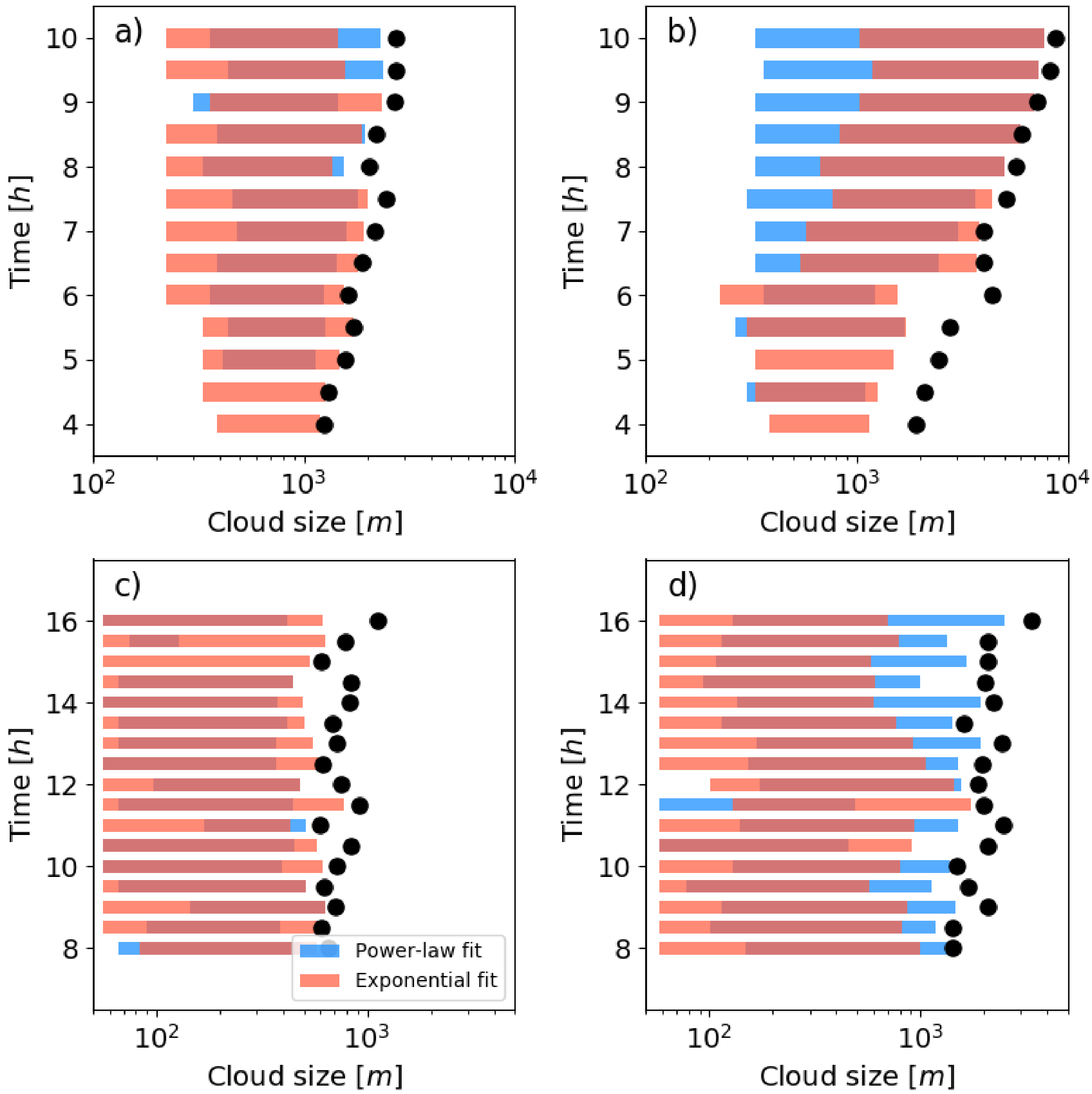


Figure 4. Time evolution of power-law (blue) and exponential (red) best fit size ranges computed for the CWP and q CCSDs (left and right column respectively). Top row: LBA case; bottom row: RICO case.

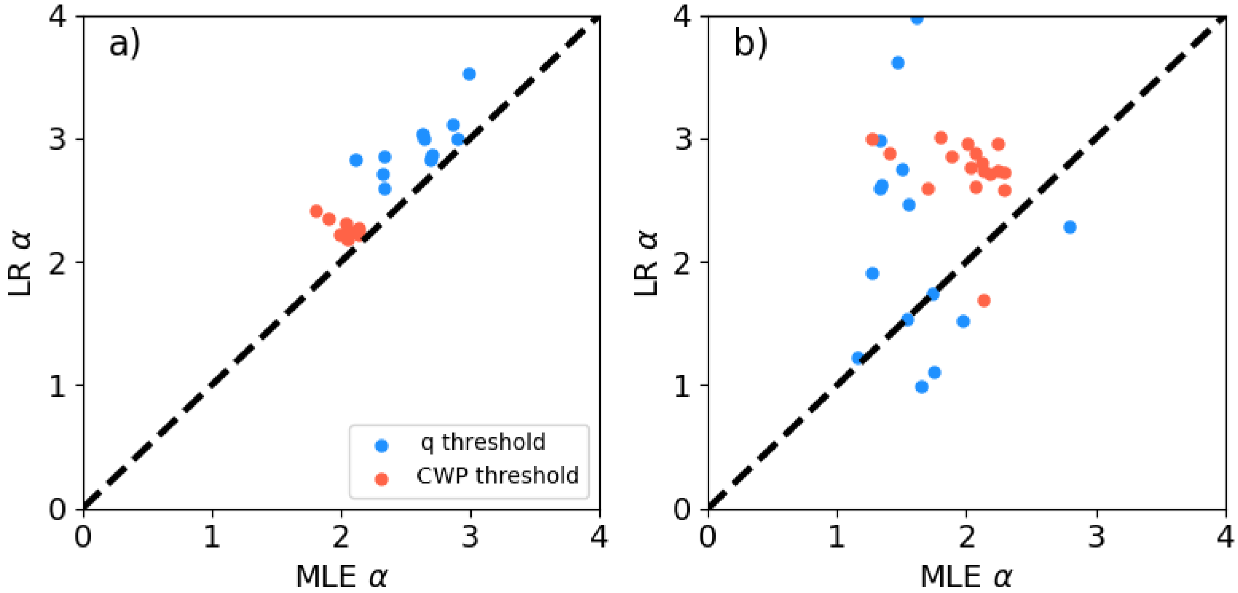


Figure 5. Estimated power-law exponents as plotted on figures 3 for the LBA and RICO simulations, against equivalent best-fit estimates obtained using linear regression.

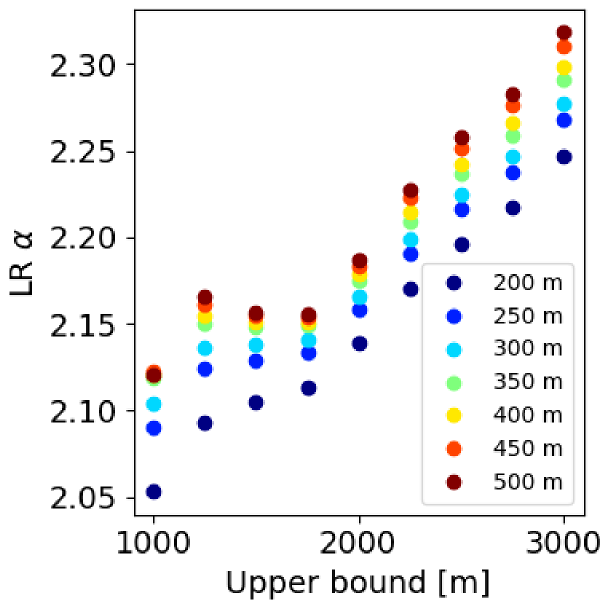


Figure 6. Sensitivity of linearly regressed exponents to the lower and upper size bounds in the LBA case for CWP clouds at 10 h.

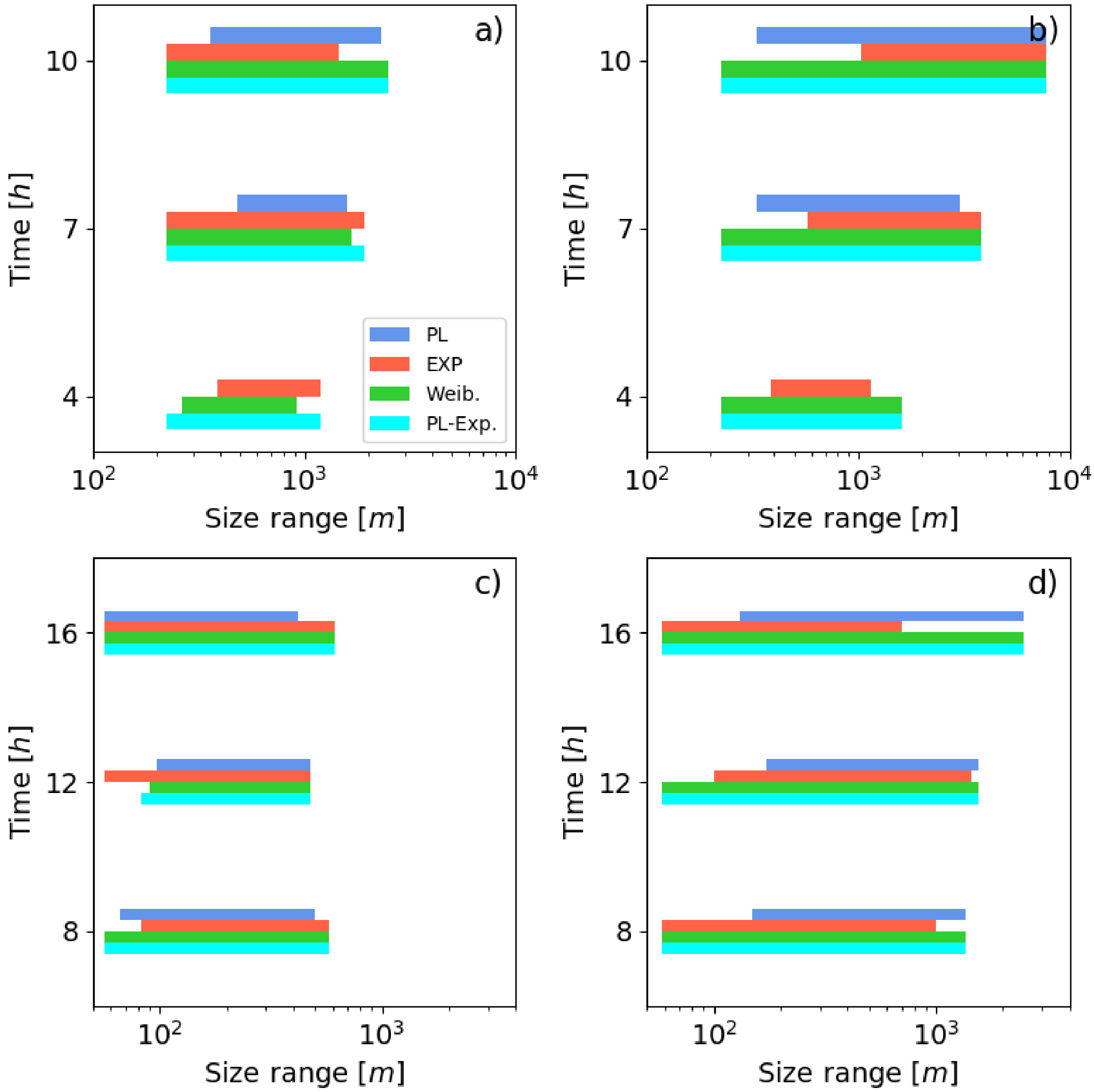


Figure 7. Size ranges ($\Delta L = L_{max} - L_{min}$) over which best-fits with power-law (blue), exponential (red), Weibull (green) and cutoff power-law (cyan) distributions were obtained. Only three different times for both the LBA (top row) and RICO (bottom row) simulations are shown. Size ranges for CWP and q based CCSDs are on the left and right respectively.

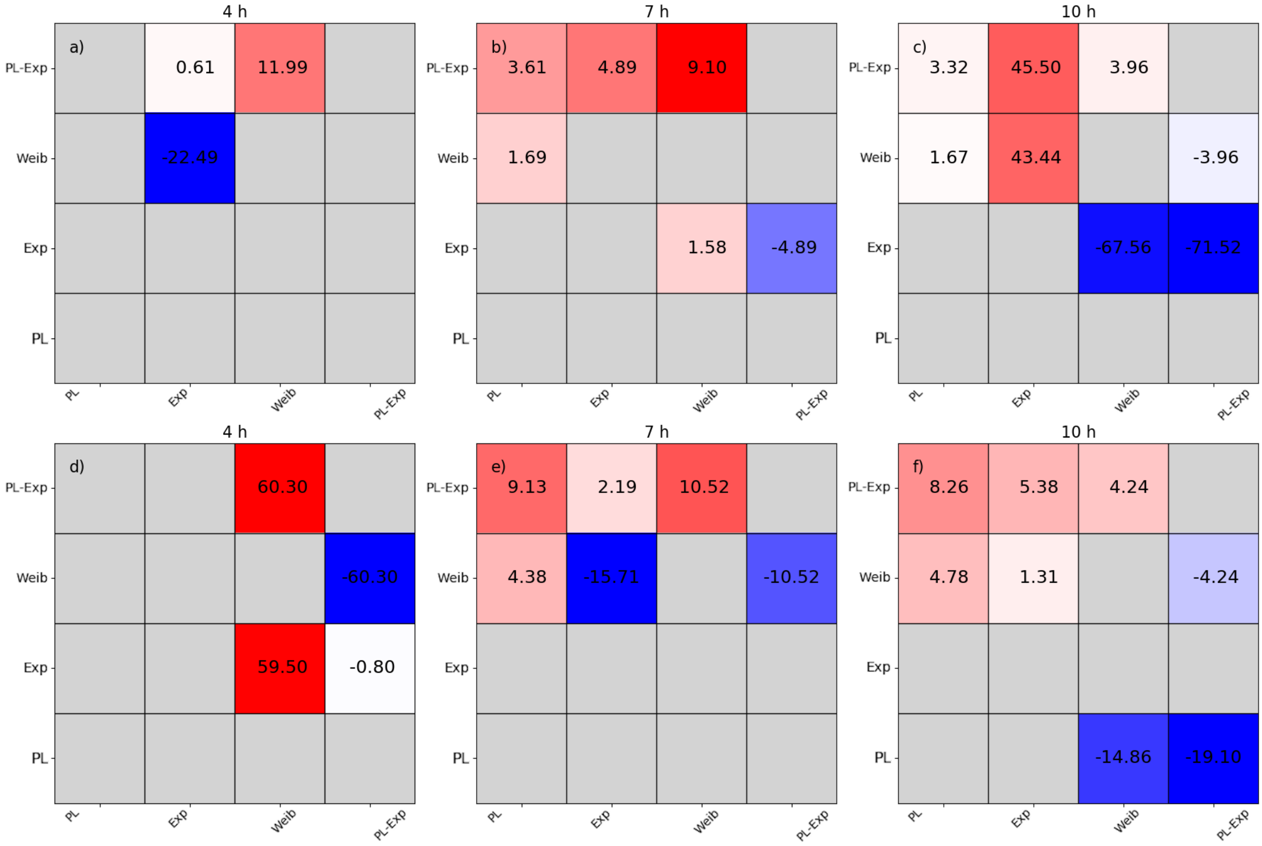


Figure 8. Results from likelihood-ratio tests applied to clouds identified based on q (top row) and CWP (bottom row) thresholds in the LBA simulation. The LR test is applied to all possible combinations of reference and alternative distributions. In each subplot, the reference distribution is read on the horizontal axis, and it is tested against an alternative distribution read on the vertical axis. The color code should be understood as follows: negative values in blue indicate that the reference distribution provides a better fit than the alternative distribution over its optimal size range; positive values in red indicate that the alternative distribution is a better fit than the reference one. Grey cells indicate that no acceptable fit could be found with the alternative distribution.

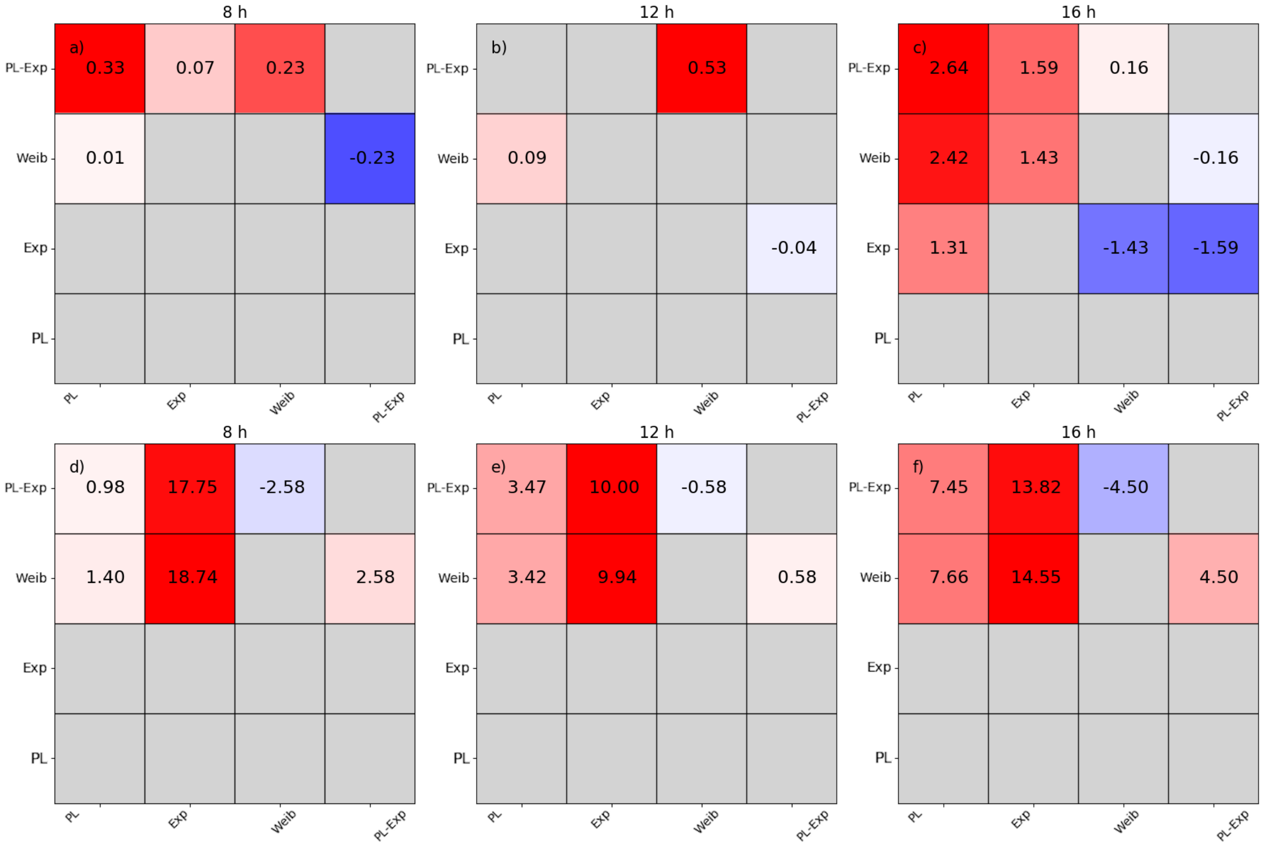


Figure 9. Same as figure 8 but for cloud distributions from the RICO simulation.

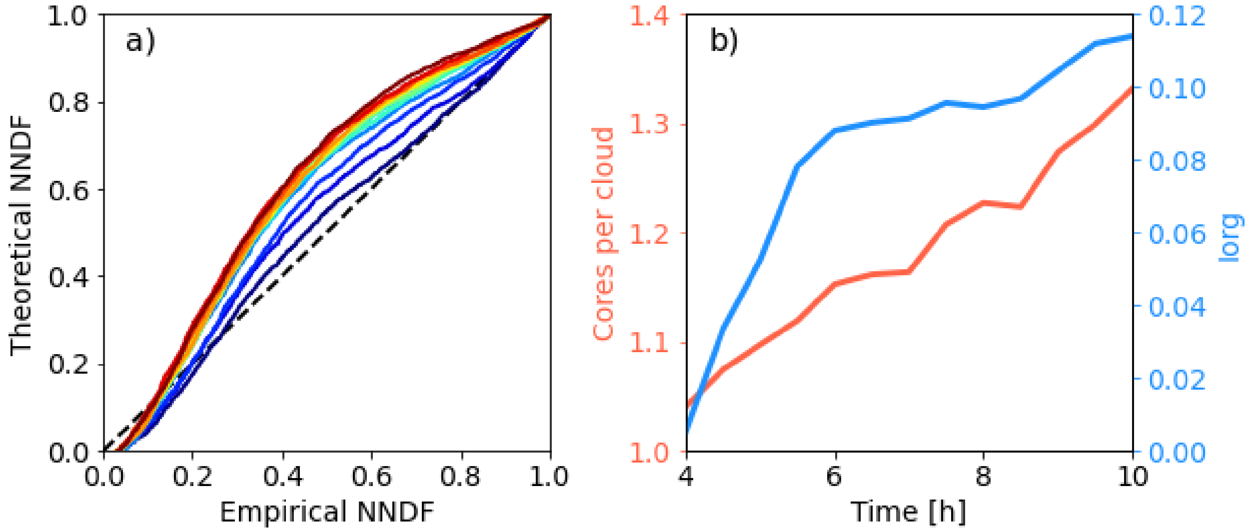


Figure 10. Clustering indices plotted as a function of time for the LBA case: a) NNCDF distributions (see text for further information) with lines colored according to time, from 4 h (blue) to 10 h (dark red), b) I_{org} and mean number of cores ($w > 4$ m s^{-1}) per cloud object.

Acknowledgments

The authors would like the acknowledge the Leibniz-Rechenzentrum (LRZ) for providing the computing resources necessary to successfully complete this project.

References

Arakawa, A., & Schubert, W. (1974). Interaction of a cumulus cloudensemble with the large-scale environment, part i. *J. Atmos. Sci.*, *31*, 674–701. doi: 10.1175/1520-0469(1974)031⟨0674:IOACCE⟩2.0.CO;2

Bak, P., Tang, C., & Wiesenfeld, K. (1988). Self-organized criticality. *Phys. Rev. A*, *38*, 364–374.

Benner, T., & Curry, J. (1998). Characteristics of small tropical cumulus clouds and their impact on the environment. *J. Geophys. Res.*, *27*, 28753–28767.

Bley, S., Deneke, H., Senf, F., & Scheck, L. (2017). Metrics for the evaluation of warm convective cloud fields in a large-eddy simulation with meteosat images. *Q. J. R. Meteorol. Soc.*, *143*, 2050–2060.

Böing, S. J., Jonker, H. J. J., Siebesma, A. P., & Grabowski, W. W. (2012). Influence of the subcloud layer on the development of a deep convective ensemble. *J. Atm. Sci.*, *69*(9), 2682–2698. doi: 10.1175/JAS-D-11-0317.1

Bonachela, J., & Muñoz, M. (2009). Self-organization without conservation: true or just apparent scale-invariance? *J. Stat. Mech.*, *2009*, P09009. doi: 10.1088/1742-5468/2009/09/p09009

Brast, M., Schemann, V., & Neggers, R. (2018). Investigating the scale adaptivity of a size-filtered mass flux parameterization in the gray zone of shallow cumulus convection. *J. Atmos. Sci.*, *75*, 1195–1214. doi: 10.1175/JAS-D-17-0231.1

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. doi: 10.1175/1520-0469(1998)055⟨3440:EADINS⟩2.0.CO;2

Craig, G. C., & Cohen, B. G. (2006). Fluctuations in an equilibrium convective ensemble. part i: Theoretical formulation. *J. Atmos. Sci.*, *63*, 1996–2004. doi: 10.1175/JAS3709.1

Dawe, J. T., & Austin, P. H. (2012). Statistical analysis of an les shallow cumulus cloud ensemble using a cloud tracking algorithm. *Atmos. Chem. Phys.*, *12*, 1101–1119.

Deluca, A., & Corral, A. (2013). Fitting and goodness-of-fit of non-truncated and truncated power-law distributions. *Acta Geophys.*, *61*(6), 1351–1394. doi: 10.2478/s11600-013-0154-9

Goldstein, M., Morris, S., & Yen, G. (2004). Problems with fitting to the power-law distribution. *Eur. Phys. J. B*, *41*, 255–258. doi: 10.1140/epjb/e2004-00316-5

Grabowski, W. W. (1998). Toward cloud resolving modeling of large-scale tropical circulations: A simple cloud microphysics parameterization. *J. Atmos. Sci.*, *55*(21), 3283–3298. doi: 10.1175/1520-0469(1998)055⟨3283:TCRMOL⟩2.0.CO;2

Grabowski, W. W., Bechtold, P., Cheng, A., Forbes, R., Halliwell, C., Khairoutdinov, M., ... Xu, K.-M. (2006). Daytime convective development over land: A model intercomparison based on lba observations. *Quart. J. Roy. Meteteor. Soc.*, *132*(615), 317–344. doi: 10.1256/qj.04.147

Haerter, J., Böing, S., Henneberg, O., & Nissen, S. (2019). Circling in on convective organization. *Geophys. Res. Lett.*, *46*, 7024–7034. doi: 10.1029/2019GL082092

Heus, T., & Seifert, A. (2013). Automated tracking of shallow cumulus clouds in large domain, long duration large-eddy simulations. *Geosci. Mod. Dev.*, *6*, 1261–1273. doi: 10.5194/gmd-6-1261-2013

Houze Jr., R., & Cheng, C.-P. (1977). Radar characteristics of tropical convection observed during gate: mean properties and trends over the summer season.

Mon. Wea. Rev., *105*, 964–980.

Jensen, H. (1998). *Self-organized criticality: Emergent complex behavior in physical and biological systems.* Cambridge University Press.

Jiang, H., Feingold, G., Jonsson, H., Lu, M.-L., Chuang, P., Flagan, R., & Seinfeld, J. (2008). Statistical comparison of properties of simulated and observed cumulus clouds in the vicinity of houston during the gulf of mexico atmospheric composition and climate study (gomaccs). *J. Geophys. Res.*, *113*, D13205. doi: 10.1029/2007JD009304

Koren, I., Oreopoulos, L., Feingold, G., Remer, L., & Altaratz, O. (2008). How small is a small cloud? *Atmos. Chem. Phys.*, *8*, 3855–3864.

Kuo, K.-S., Welch, R., & Sengupta, S. (1988). Structural and textural characteristics of cirrus clouds observed using high spatial resolution landsat imagery. *J. App. Met.*, *27*, 1242–1260.

Kuo, K.-S., Welch, R., Weger, R., Engelstad, M., & Sengupta, S. (1993). The three-dimensional structure of cumulus clouds over the ocean: 1. structural analysis. *J. Geophys. Res.*, *11*, 20685–20711.

Laherrere, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales. *Eur. Phys. J. B*, *2*, 525–539. doi: 10.1007/s100510050276

Lopez, R. (1977). The lognormal distribution and cumulus cloud populations. *Mon. Wea. Rev.*, *105*, 865–872.

Marković, D., & Gros, C. (2015). Power laws and self-organized criticality in theory and nature. *Physics Reports*, *536*, 41–74.

Nair, U., Weger, R., Kuo, K., & Welch, R. (1998). Clustering, randomness and regularity in cloud fields 5. the nature of regular cumulus cloud fields. *J. Geophys. Res.*, *103*, 11363–11380.

Neggers, R. (2015). Exploring bin-macrophysics models for moist convective transport and clouds. *Journal of Advances in Modeling Earth Systems*, *7*(4), 2079-2104. doi: https://doi.org/10.1002/2015MS000502

Neggers, R., Jonker, H., & Siebesma, A. (2003). Size statistics of cumulus cloud populations in large-eddy simulations. *J. Atmos. Sci.*, *60*, 1060–1074. doi: 10.1175/1520-0469(2003)60

Newman, M. (2005). Power laws, pareto distributions and zipf's law. *Contemporary Phys.*, *46*, 323–351.

Parker, L., Welch, R., & Musil, D. (1986). Analysis of spatial inhomogeneities in cumulus clouds using high spatial resolution landsat data. *J. App. Meteor. and Clim.*, *25*, 1301–1314.

Peters, O., Deluca, A., Corral, A., Neelin, J. D., & Holloway, C. E. (2010). Universality of rain event size distributions. *J. Stat. Mech.*, 11030. doi: 10.1088/1742-5468/2010/11/P11030

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

Peters, O., Neelin, J. D., & Nesbitt, S. (2009). Mesoscale convective systems and critical clusters. *J. Atmos. Sci.*, *66*, 2913–2924. doi: 10.1175/2008JAS2761.1

Rieck, M., Hohenegger, C., & van Heerwaarden, C. (2014). The influence of land surface heterogeneities on cloud size development. *Mon. Wea. Rev.*, *142*, 3830–3846.

Savre, J. (2021). Formation and maintenance of subsiding shells around non-precipitating and precipitating cumulus clouds. *Quart. J. Roy. Meteor. Soc.*, *147*, 728—745. doi: 10.1002/qj.3942

Savre, J., Ekman, A. M. L., & Svensson, G. (2014). Technical note: Introduction to mimica, a large-eddy simulation solver for cloudy planetary boundary layers. *J. Adv. Mod. Earth Sys.*, *6*(3), 630-649. doi: 10.1002/2013MS000292

Seifert, A., & Beheng, K. D. (2006). A two-moment cloud microphysics parameterization for mixed-phase clouds. part i: Model descroption. *Meteorology and At-*

mospheric Physics, *92*, 45-66.

Senf, F., Klocke, D., & Brueck, M. (2018). Size-resolved evaluation of simulated deep tropical convection. *Mon. Wea. Rev.*, *146*, 2161–2182.

Sengupta, S., Welch, R., Navar, M., Berendes, T., & Chen, D. (1990). Cumulus cloud field morphology and spatial patterns derived from high spatial resolution landsat imagery. *J. App. Met.*, *29*, 1245–1267.

Siebesma, A., Soares, P., & Teixeira, J. (2007). A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *J. Atmos. Sci*, *64*, 1230–1248. doi: 10.1175/JAS3888.1

Sornette, D. (2006). *Critical phenomena in natural sciences* (2nd ed.). Springer.

Stauffer, D., & Ahaorny, A. (1994). *Introduction to percolation theory* (2nd ed.). London: Taylor & Francis.

Teo, C.-K., Huynh, H.-N., Koh, T.-Y., Cheung, K., Legras, B., Chew, L., & Norford, L. (2017). The universal scaling characteristics of tropical oceanic rain clusters. *J. Geophys. Res.*, *122*, 5582–5599.

Tompkins, A., & Semie, A. (2017). Organization of tropical convection in low vertical wind shears:role of updraft entrainment. *J. Adv. Model. Earth Sys.*, *9*, 1046–1068. doi: 10.1002/2016MS000802.

Traxl, D., Boers, N., Rheinwalt, A., Goswami, B., & Kurths, J. (2016). The size distribution of spatiotemporal extreme rainfall clusters around the globe. *Geophys. Res. Lett.*, *43*, 9939–9947.

van Laar, T. W., Schemann, V., & Neggers, R. A. J. (2019). Investigating the diurnal evolution of the cloud size distribution of continental cumulus convection using multiday les. *J. Atmos. Sci.*, *767*, 729–747. doi: 10.1175/JAS-D-18-0084.1

vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., ... Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during rico. *J. Adv. Mod. Earth Sys.*, *3*, 1–19. doi: 10.1029/2011MS000056

Welch, R., Kuo, K., Wielicki, B., Sengupta, S., & Parker, L. (1988). Marine stratocumulus cloud fields off the coast of southern california observed using landsat imagery. part i: structural characteristics. *J. App. Met.*, *27*, 341–362.

Wielicki, B., & Welch, R. (1986). Cumulus cloud properties derived using landsat satellite data. *J. Clim. and App. Met.*, *25*, 261–276.

Windmiller, J. (2017). *Organization of tropical convection* (Unpublished doctoral dissertation). Ludwig-Maximilians-Universität München.

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

Xue, H., & Feingold, G. (2006). Large-eddy simulations of trade wind cumuli: investigation of aerosol indirect effects. *J. Atmos. Sci.*, *63*, 1605–1622.

Yano, J.-I., Liu, C., & Moncrieff, M. (2012). Self-organized criticality and homeostasis in atmospheric convective organization. *J. Atmos. Sci.*, *69*(12), 3449–3462. doi: 10.1175/JAS-D-12-069.1

Zhao, G., & Di Girolamo, L. (2007). statistics on the macrophysical properties of trade wind cumuli over the tropical wester atlantic. *J. Geophys. Res.*, *112*, D10204.