

Comparing ground-based observations and a large-eddy simulation of shallow cumuli by isolating the main controlling factors of the mass flux distribution

Mirjana Sakradzija^{1,2}  | Marcus Klingebiel¹

¹The Atmosphere in the Earth System, Max Planck Institute for Meteorology, Hamburg, Germany

²Hans Ertel Centre for Weather Research, Deutscher Wetterdienst, Offenbach, Germany

Correspondence

Mirjana Sakradzija, The Atmosphere in the Earth System, Max Planck Institute for Meteorology, Hamburg, Germany.
Email: mirjana.sakradzija@mpimet.mpg.de

Funding information

German Federal Ministry of Transport and Digital Infrastructure (BMVI)

Abstract

The distribution of mass flux at the cloud base has long been thought to be independent of large-scale forcing. However, recent idealized modelling studies have revealed its dependence on some large-scale conditions. Such dependence makes it possible to isolate the observed large-scale conditions, which are similar to those in large-eddy simulations (LES), in order to compare the observed and modelled mass flux distributions. In this study, we derive for the first time the distribution of the cloud-base mass flux among individual shallow cumuli from ground-based observations at the Barbados Cloud Observatory (BCO) and compare it with the Rain In Cumulus over the Ocean (RICO) LES case study. The procedure of cloud sampling in LES mimics the pointwise measurement procedure at the BCO to provide a mass flux metric that is directly comparable with observations. We find a difference between the mass flux distribution observed during the year 2017 at the BCO and the distribution modelled by LES that is comparable to the seasonal changes in the observed distribution. This difference between the observed and modelled distributions is diminished and an extremely good match is found by subsampling the measurements under a similar horizontal wind distribution and area-averaged surface Bowen ratio to those modelled in LES. This provides confidence in our observational method and shows that LES produces realistic clouds that are comparable to those observed in nature under the same large-scale conditions. We also confirm that the stronger horizontal winds and higher Bowen ratios in our case study shift the distributions to higher mass flux values, which is coincident with clouds of larger horizontal areas and not with stronger updrafts.

KEYWORDS

cloud-base mass flux, ground-based remote sensing, large-eddy simulation (LES), mass flux distribution, shallow convection

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

1 | INTRODUCTION

In currently operational weather forecasting and climate models, the effect exerted by convective clouds on large-scale atmospheric flow is most commonly represented using a mass flux approach (Arakawa, 2004). By this approach, the convective transport of conserved atmospheric quantities is modelled by independent buoyant plumes, with the intensity of the vertical transport set by the value of the mass flux at the cloud base. The effects of a cloud ensemble can be represented using a single bulk buoyant plume (Yanai *et al.*, 1973) or a spectrum of buoyant plumes that group clouds of similar properties together (Arakawa and Schubert, 1974). In a spectral approach, the probability distribution of the cloud-base mass flux is estimated and multiple buoyant plumes are modelled to calculate the vertical structure of the convective layer. To understand which processes control the mass flux probability distribution and what sets its parameters, theoretical convection models based on mass flux distributions have to be supported by evidence from observations. The early studies of Ogura and Cho (1973) and Nitta (1975) diagnosed the statistics of cumulus cloud populations from observed large-scale meteorological conditions using the theoretical models of Yanai *et al.* (1973) and Arakawa and Schubert (1974). However, it remained unclear from these studies whether the mass flux distribution was determined by the large-scale conditions and processes and if so, what the main control parameters were. To answer this question, direct observational evidence of the probability distributions of the mass flux in cumulus clouds is needed.

With the use of the cloud resolving models and large-eddy simulations (LES), greater insight has been gained on the probability distributions of cloud populations and the mass flux distribution (Cohen and Craig, 2006; Sakradzija *et al.*, 2015; Sakradzija and Hohenegger, 2017). These studies arose from the need to understand what controls the probability distribution of the mass flux for use in the stochastic parameterization of convective fluctuations at high model resolutions (Plant and Craig, 2008; Sakradzija *et al.*, 2015, 2016; Sakradzija and Klocke, 2018). The distribution of the convective mass flux was suspected to be invariant to the changes in large-scale forcing (Craig and Cohen, 2006; Plant and Craig, 2008). It was thought that these changes affected the total convective mass flux in a particular region by changing the number of cloud elements, rather than by increasing the mass flux of individual clouds. Thus, in models the average mass flux per cloud is treated as a constant (Plant and Craig, 2008; Sakradzija *et al.*, 2015; 2016).

However, more recent studies based on LES and convection-permitting simulations have revealed that the mass flux distributions do change with certain environmental conditions. In an LES study that involved a dozen simulations based on one case over the tropical Atlantic Ocean and

one over central North America, Sakradzija and Hohenegger (2017) isolated the partitioning of the surface heating into sensible and latent heat fluxes, expressed by the Bowen ratio B , as the main controlling factor of the mass flux distribution of the shallow convection. In their study, it was also suspected that wind speed might play a role in setting the slope of the mass flux distribution; however, this was not further investigated. Wind speed is an important factor that controls shallow convection in the trade winds region, where stronger winds lead to deeper boundary layers (Nuijens and Stevens, 2012). Deeper boundary layers are associated with wider eddies in the subcloud layer, which might also form wider clouds. Another hypothesis – that the mass flux distribution might vary with geographical location – was posed by Rasp *et al.* (2018) in their study of the applicability of the theory of Craig and Cohen (2006) to realistic situations using a convection-permitting ensemble.

Although there is a lack of observational studies on the large-scale control of the mass flux distribution, the cloud size distribution has been studied extensively. In a recent study by Mieslinger *et al.* (2019), ASTER satellite imagery and the ERA-Interim reanalysis are used to test the controls of the large-scale meteorological conditions on the shallow cloud properties over the tropical oceans. Wind speed is recognized as a major factor controlling the cloud size distribution, among other cloud-based statistics. They further show that the Bowen ratio has an influence on the estimated slopes of the cloud size distribution, cloud cover and cloud top height, despite the small range of the Bowen ratio tested in their study. Based on these findings and hypotheses about the cloud size distribution, in our study we isolate the wind speed next to the surface Bowen ratio as a major factor that might control and determine the mass flux distribution of shallow cumuli. We note here that the dependence of the mass flux distribution on wind speed has not yet been established.

The goal of this study is to derive for the first time the distribution of the mass flux at the cloud base from ground-based observations and to compare the observed distributions of the mass flux with an idealized LES study. Our observation method represents a first attempt to measure the mass flux of individual clouds; as such it is not an established routine and it involves several assumptions (see section 2). It is thus necessary to compare the results of our measurements with a well-established method such as LES. Although LES models are not without their limitations, they are widely used to study shallow convective clouds and their evaluation and intercomparison studies have been well documented (see, e.g., Siebesma *et al.*, 2003; van Zanten *et al.*, 2011). On the other hand, a good match between our observations and LES would increase our confidence in the LES estimate of the cloud-base mass flux. Thus, we see a comparison between observations and LES not as a one-way validation procedure

but as a two-way street that will provide greater confidence in both methods.

The LES of atmospheric shallow convection are commonly based on observations over a given time period at a single observation site and are forced by the average steady conditions representative of that site and a given synoptic situation (see, e.g., Siebesma *et al.*, 2003). As a result of such an idealized simulation set-up, one possible realization of the cloud field is produced assuming uniformity of the meteorological conditions and the large-scale forcing. Furthermore, the forcing imposed on the convection that develops in LES is carefully selected to isolate certain processes of interest and to eliminate other processes that might affect the simulated convection. On the other hand, when a realistic case is simulated in LES using the instantaneous observed structure of the planetary boundary layer (PBL) and time-varying large-scale forcing, it is challenging to measure individual cloud properties and to collect enough samples of clouds under the exact conditions isolated by the LES, because the time span of the LES is usually only a single day. Therefore, in order to compare the mass flux distribution observed at the BCO during the year 2017 with the modelled mass flux distribution, it is necessary to subsample the observations in order to isolate similar large-scale conditions to those imposed by the LES. Such a comparison will also assess the selected large-scale conditions as the controlling factors of the mass flux distribution in cases of high similarity between the observed and modelled distributions.

The observations were conducted using a cloud radar and a Doppler lidar at the observation site in Barbados. We chose the LES case based on observations near Antigua and Barbuda in winter 2004/2005 – the Rain in Cumulus over the Ocean (RICO) case. Thus observations recorded at a single site are compared with one realization of an idealized convective case in a nearby region, which is based on an independent dataset. Since it is challenging to isolate different processes in observations that might control the cloud statistics, we use findings based on LES as a guide for the choices we make in analysing the observational data. In order to make the comparison possible, the observational data is subsampled to include only those samples that were observed under similar wind conditions as modelled in LES. In the next step, the large-scale conditions estimated from the ERA-Interim reanalysis during the observation period of 2017 are used to determine which parameters in addition to wind speed could be responsible for the similarities between the observed and modelled distributions.

2 | METHODS

2.1 | Barbados Cloud Observatory (BCO)

The island of Barbados is located to the east of the Caribbean Sea and sits in the trade winds region of the Atlantic Ocean.

The Barbados Cloud Observatory (BCO) is located on the east coast of the island at Deebles Point, 13.16°N, 59.43°W (Stevens *et al.*, 2016). It contains several active and passive remote sensing instruments to profile the aerosol and cloud properties inside the atmosphere. In addition to standard meteorological instruments (measuring temperature, humidity, pressure, wind, rain rate, solar radiation, etc.), the main instruments used for this study are a polarized Doppler cloud radar and a Doppler lidar. The radar has been operating since June 2015 and has a frequency of 35.5 GHz (K_a -band) with a sensitivity of 57 dBZ at an altitude of 5 km. This makes it ideal for detecting cloud particles as well as hygroscopically grown sea salt particles in the subcloud layer (Klingebiel *et al.*, 2019). To differentiate between different particle types, such as liquid or ice particles, a Linear Depolarization Ratio is also available. With a Doppler velocity precision of less than $0.02 \text{ m}\cdot\text{s}^{-1}$, a vertical resolution of 31.18 m and a temporal resolution of 10 s, the radar detects clouds in an altitude range between 150 m and 18 km and delivers information about the clouds' vertical structure and turbulence.

The Doppler lidar is a HALO Photonics Stream Line Pro instrument. It uses a 1,500 nm laser to measure vertical velocities up to $\pm 20 \text{ m}\cdot\text{s}^{-1}$ in an altitude range between 50 m and around 1 km (depending on atmospheric conditions) with a vertical resolution of 30 m. The instrument is located as close as possible to the radar dish (about 2 m), and with a temporal resolution of 1.3 s it is able to provide information about the turbulence in the boundary layer.

2.2 | Large-eddy simulation (LES) and case study

The LES simulation and the postprocessing analysis used in this study are adopted in full from the earlier study of Sakradzija and Hohenegger (2017). Thus, in the following, we only briefly describe the LES model, the simulation set-up and the case specification.

LES are run using the UCLA-LES (University of California, Los Angeles Large-Eddy Simulation) model (Stevens *et al.*, 1999; Stevens, 2010). The model solves the Ogura–Phillips anelastic equations for the prognostic variables of wind, liquid water potential temperature, total water mixing ratio, rain mass mixing ratio and rain number mixing ratio. The equations are discretized on a doubly periodic uniform Arakawa C grid. A third-order Runge–Kutta scheme is used for the time integration, a directionally split monotone upwind scheme is used for the advection of scalars, and a directionally split fourth-order centred scheme is used for the momentum advection. The double-moment warm rain scheme of Seifert and Beheng (2001) is used to compute the cloud microphysics. The subgrid turbulent fluxes are computed using the Smagorinsky scheme (as described in Stevens *et al.*, 1999). The effects of radiation are prescribed

as net forcing tendencies. A more detailed description of the UCLA-LES model equations, numerical methods and sub-grid physics is provided in Stevens (2010) and the references therein.

The RICO field measurement campaign that took place during winter 2004/2005 upwind of the islands of Antigua and Barbuda (Rauber, 2007) is used as a basis for the set-up of the RICO LES case. The initial profiles of the potential temperature, specific humidity and horizontal wind components are constructed by fitting the averaged profiles from the radiosonde measurements taken over Barbuda (van Zanten *et al.*, 2011, fig. 2 and table 2). The subsidence profiles are prescribed as horizontally homogeneous and time-invariant and act on the temperature and moisture fields only. The large-scale horizontal advection tendencies of moisture and temperature are prescribed as horizontally homogeneous and time-invariant (van Zanten *et al.*, 2011, table 2). The radiative and advective cooling rates are prescribed jointly as a constant cooling rate of $2.5 \text{ K}\cdot\text{day}^{-1}$ at all vertical levels. The geostrophic wind profiles are prescribed as time-invariant and are equal to the initial wind profiles (van Zanten *et al.*, 2011), while the background wind is set to $u = -5 \text{ m}\cdot\text{s}^{-1}$ and $v = -4 \text{ m}\cdot\text{s}^{-1}$. The sea surface temperature is set to 299.8 K, and the surface turbulent fluxes are computed interactively using a surface-layer bulk aerodynamic parameterization (see van Zanten *et al.*, 2011). The LES simulation is performed over a domain of $51.2 \times 51.2 \text{ km}$ with a horizontal grid spacing of 25 m and a vertical resolution of 25 m up to the top of the domain at 5 km. Five vertical grid levels are used as damping layers at the top of the domain. The duration of the simulation is 60 hr.

2.3 | Definition of cloud properties in observations and LES

The vertical flux of mass through the cloud base of the i th cloud is defined as

$$m_i = \rho a_i w_i. \quad (1)$$

To estimate the shallow convective mass flux for a single cloud with remote sensing measurements, it is necessary to quantify the parameters as follows: density ρ ($\text{kg}\cdot\text{m}^{-3}$); cloud area a_i (m^2); in-cloud vertical velocity w_i ($\text{m}\cdot\text{s}^{-1}$). Similarly to Ghate *et al.* (2011), we assume the air density to be constant and equal to $1.2 \text{ kg}\cdot\text{m}^{-3}$.

The cloud area a_i is not directly measurable with our current remote sensing instruments, because the upward-looking Doppler cloud radar and the Doppler lidar deliver only two-dimensional measurements (see Figure 1). For this reason, we calculate a_i for every single cloud based on its cloud chord length. We assume that every cloud base has a circular shape and that the cloud chord length represents the diameter of the cloud base. The cloud chord length (see

Figure 1) is estimated by the product of the radar overpass time for each cloud and the horizontal wind speed at the associated height of the cloud base. Fragmented clouds are considered as one cloud when the time gap of the recorded data between the fragments is less than 30 s and the horizontal wind speed at the cloud base $v_{cb,i}$ is estimated by following Hellman (1916), who extrapolates the horizontal windspeed v_0 from $h_0 = 2 \text{ m}$ above ground to the cloud-base height h_{cb} :

$$v_{cb} = v_0 \left(\frac{h_{cb}}{h_0} \right)^\alpha. \quad (2)$$

The shear exponent, $\alpha = 0.1$, represents an ocean terrain type and the cloud-base height is defined by the first radar range gate with a cloud signal ($\geq 50 \text{ dBZ}$). Using this method, we found good agreement between the extrapolated horizontal wind speed and the wind speed measured with radiosondes at the cloud base.

Because we want to focus on the upward shallow convective mass flux in this study, we only consider upward motion and neglect downward motion, which shortens the cloud chord length (presented in Figure 1) to $l_{up,i}$. Based on these assumptions, the cloud area a_i is calculated as

$$a_i = \left(\frac{l_{up,i}}{2} \right)^2 \pi. \quad (3)$$

Most of the time, however, the detected shallow cumulus clouds do not pass with their centre over the BCO. This leads to an underestimation of $l_{up,i}$. On the other hand, the model simulation described below uses a similar method that makes the measured and simulated convective mass fluxes comparable. For the comparison shown in this study, we are using remote sensing measurements at the BCO from the year 2017.

The cloud-base mass flux just above the cloud-base level is estimated using the LES output following a method that provides a similar metric of the mass flux as in the observations. For this purpose, we developed a routine for cloud identification along a horizontal line in the LES output to mimic the pointwise measurements at the observation site. The horizontal cross-sections at the level that lies 100 m above the lifting condensation level are selected every 15 min from the 24th to the 60th hour of the RICO case simulation. The horizontal lines along one spatial dimension with a distance of 3 km are extracted from the horizontal snapshots. The difference between the statistics based on samples collected along the x or y dimensions is minimal and can be disregarded, so we choose the y direction. The distance of 3 km was chosen as an estimate of the maximum cloud size in the RICO simulation. We define the cloud chord length as the number of grid points along these lines that contain liquid water and have a vertical velocity greater than zero times the grid resolution

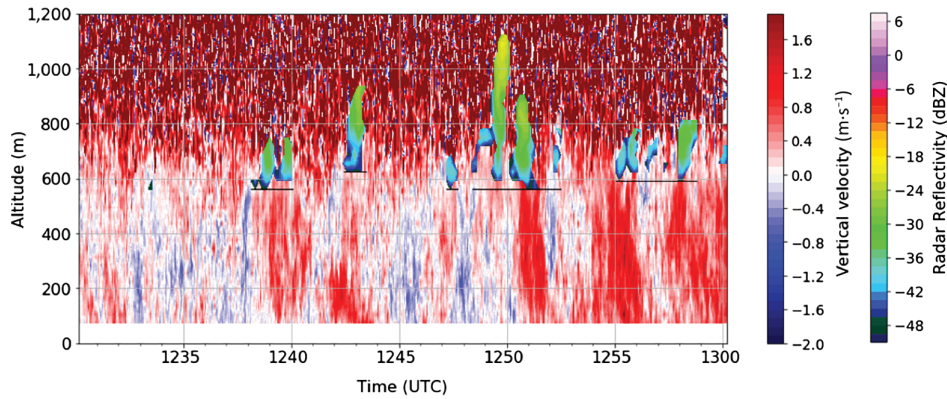


FIGURE 1 Radar reflectivity of some shallow cumulus clouds from April 13, 2017 in combination with Doppler lidar measurements of the vertical air motion in the boundary layer. The horizontal black lines represent the cloud chord length. The detection limit of the Doppler lidar is at an altitude of around 750 m, which is indicated by the noise

$\Delta x = 25$ m. This length is assumed to be the diameter of the circular cloud area defined as

$$a_i = \left(\frac{N_i \Delta x}{2} \right)^2 \pi, \quad (4)$$

where i is the index of a single cloud and N is the total number of points along the line that crosses a single cloud. Similarly to the procedure for the observations, only those points with positive vertical velocity are taken into account. The vertical velocity of a single cloud is taken as the average along the cloud chord length

$$w_i = \sum_{n=1}^{N_i} \frac{w_n}{N_i}, \quad (5)$$

where n is the index of a single point along the cloud chord length. The statistics derived in this way are directly comparable with the observed cloud statistics.

3 | CUMULUS CLOUD STATISTICS, RESULTS AND DISCUSSION

3.1 | Comparison between the observed and modelled mass flux distributions

Cloud samples are collected from the observations and from the LES output and the probability density distributions of the cloud-base mass flux are calculated based on these samples. The sample sizes of the distributions are 6,125 clouds in the RICO LES case during the 24–60 hr period and 16,989 clouds from the BCO observations during 2017. The probability density distribution is computed using the generic R function *hist* (R Core Team, 2015). The width of the bins used to compute the probability density of the cloud-base mass fluxes is exponentially increasing for higher values of the mass flux. Due to

the finite resolution of the LES model, mass flux values lower than $600 \text{ kg} \cdot \text{s}^{-1}$ are removed from the model and, for consistency, from the observed cloud samples. The two distributions are compared in Figure 2.

The observed cloud mass flux distribution at the BCO from the year 2017 has a shallower slope than the distribution in the RICO LES case (Figure 2a). Clouds observed at the BCO have a higher frequency of large mass flux values. This is also shown on the cumulative distribution plot (Figure 2b), where the distance between the two distributions demonstrates a higher probability of high mass flux values in observations. As a quantitative measure of the distance between the modelled and the observed distributions, the Kolmogorov–Smirnov (KS) statistic D is calculated. D is defined as $D = \max_m |P_{BCO}(m) - P_{LES}(m)|$, where $P_{BCO}(m)$ and $P_{LES}(m)$ represent the cumulative distributions of the mass flux in the BCO and LES cloud samples, respectively. D can be interpreted as a maximum absolute distance between the empirical distribution of the BCO cloud sample and that of the LES sample. In the present case, the KS statistic is equal to $D = 0.135$ (calculated using the library “stats”; R Core Team, 2015). At the 95% level, the critical value for the KS statistic can be approximated as $D_{crit,0.05} = 1.36 \sqrt{\left(\frac{1}{n_{BCO}} + \frac{1}{n_{LES}} \right)}$ (see Conover, 1999, table A20), where n_{BCO} and n_{LES} are the sample sizes of the two cloud samples. Since $D > D_{crit,0.05}$, the null hypothesis that the two distributions are similar is not supported and must be rejected based on the KS test. The quantile–quantile (Q–Q) plot also shows the discrepancy between the two distributions, as the sample pairs fall far below the $x = y$ line (Figure 2c).

Discrepancies between the observed and modelled mass flux distributions are to be expected because the RICO LES case is based on a different set of observations in a different region and at different time intervals from our BCO dataset. Furthermore, the large-scale forcing and the initial conditions

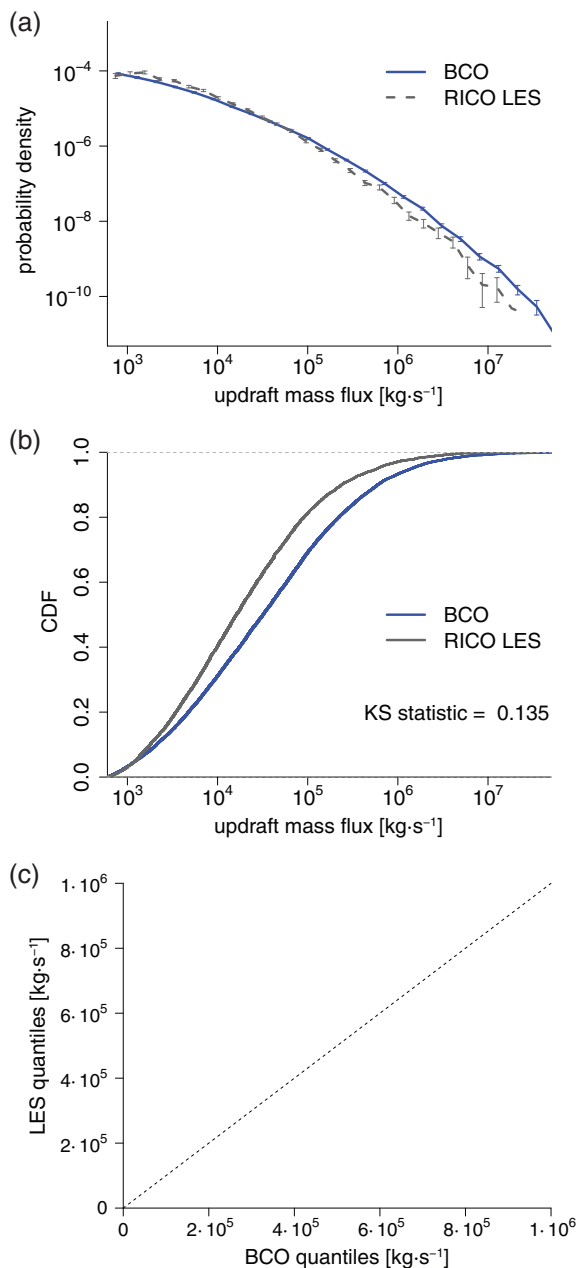


FIGURE 2 Comparison of the mass flux distribution at the cloud base observed at the BCO in 2017 with the RICO LES case mass flux distribution: (a) probability density, (b) cumulative distribution, and (c) quantile–quantile (Q–Q) plot. Vertical bars represent the sampling uncertainty based on the 95% confidence intervals

of the RICO LES case are constructed based on averaged conditions over several weeks and are further simplified and smoothed, which makes the LES case idealized. The observations at the BCO span variable conditions and include seasonal changes over a single year, so it is to be expected that the mass flux distribution as observed at the BCO shows a distinct slope from the one that originates from LES.

The sampling variability of the mass flux distributions is very low in both the LES and BCO cases, except towards the ends of the right tails of the distributions (Figure 2a). This

means that the differences between the distributions cannot be attributed to sampling issues. The higher variability by high mass flux values is indicative of a limited sample size of the largest possible cloud mass flux. The LES case has higher sampling variability due to its smaller sample size in comparison with the BCO case. The sampling variability is based on 95% confidence intervals computed for each distribution bin (vertical bars in Figure 2a). The confidence intervals are calculated by a bootstrapping method with replacements using 1,000 random samples.

3.2 | Seasonal changes of the cloud statistics as observed at the BCO

The distribution of the cloud-base mass flux of individual clouds at the BCO changes throughout the year. To show these changes, the distributions that correspond to the four trimesters in 2017 (FMA, MJJ, ASO and NDJ) are plotted and compared with the LES distribution (Figure 3). The greatest similarity between the modelled and observed distributions is found in the third trimester (ASO; see Figure 3a–c). The plot of the cumulative distribution shows greater similarity between the observed and the modelled distributions with a lower KS distance of $D = 0.071$, compared to 0.135 obtained previously. The sample size in the ASO trimester is 3,881, which gives the critical value of $D_{crit, 0.05} = 0.0205$. The KS distance is closer to the critical value; however, the null hypothesis of the similarity between the two distributions still has to be rejected based on the 95% confidence interval. The Q–Q plot also shows in greater detail that the discrepancy between the observed and the modelled samples is still present (Figure 3d).

By examining the distribution plots of the cloud-base area (Figure 4a) and the distribution of the horizontal wind speed at the level of the cloud base (Figure 4c), it is evident that stronger winds are coincident with larger clouds, as the cumulative distribution function of the cloud area in the BCO samples shows a similar behaviour to the wind distributions. Namely, in the ASO trimester, the distribution of the cloud-base area is shifted towards smaller clouds and the distribution of the wind speed is shifted towards lower wind speed values. During the other three trimesters, the distributions of the cloud-base area are very similar to each other, although a trend coincident with that of the distribution of the wind speed can be recognized. The vertical velocity distributions do not show any significant differences between the trimesters in the summer or winter seasons and do not seem to resemble the changes in the distribution of the horizontal wind speed (Figure 4b). Since the mass flux is a product of the cloud area and vertical velocity, we can confirm based on these results that the stronger winds are related to larger cloud areas at cloud bases that result in higher mass fluxes, and are not necessarily related to stronger updrafts.

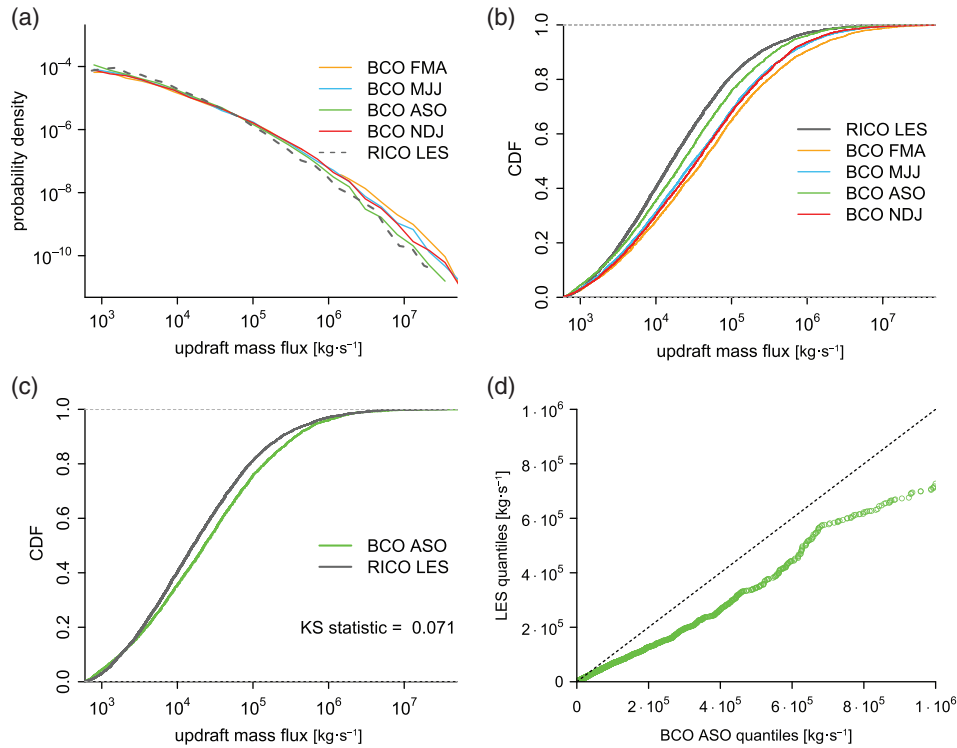


FIGURE 3 As Figure 2, but with the BCO data split into four trimesters. For greater transparency, plots (c) and (d) show only the most similar observed and modelled distributions and their Q–Q plot

3.3 | Isolating the possible controlling factors

In the subsequent analysis, we examine the effect of the horizontal wind speed on the mass flux distribution and attempt to uncover additional factors that might control it.

To examine the effect of the horizontal wind speed on the mass flux distributions, samples of the horizontal wind speed are collected at 2 m above ground and extrapolated to the cloud-base height (see section 2.3) for each cloud that passes over the Doppler lidar at the BCO. To obtain a sample of wind statistics that is comparable to observations, we collect data from clouds that are cross-cut by a sampling line in the LES domain at a height level of 100 m above the liquid condensation level (LCL). The wind speed is then averaged over the time a cloud is passing above the BCO and over the length of a line that crosses through a cloud in LES. Collected samples of the wind speeds of individual clouds are then used to calculate the distributions of the wind speed at the BCO for 2017 and in the RICO LES case for the simulation time period 24–60 hr (Figure 5). Distributions of the wind speed at the cloud base at the BCO are also plotted for different trimesters in Figure 4c.

The distributions of the wind speed at the BCO and in LES are similar in shape, but they differ in the mean and particularly in the variance (Figure 5). The distribution calculated from the yearly cloud samples at the BCO has a higher mean value of $9.9 \text{ m}\cdot\text{s}^{-1}$ and a larger standard deviation of $3.54 \text{ m}\cdot\text{s}^{-1}$ compared to the LES mean of $8.99 \text{ m}\cdot\text{s}^{-1}$

and standard deviation of $0.58 \text{ m}\cdot\text{s}^{-1}$ (see Table 1). The distribution shape is symmetric and can be modelled by a Gaussian distribution function. We will use the LES wind distribution to subsample the BCO cloud sample in the following, so we estimate a fit to the LES distribution using the Gaussian function (Figure 5a, red line):

$$f(u_{cb}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u_{cb}-\mu)^2}{2\sigma^2}}, \quad (6)$$

where the mean equals $\mu = 8.992 \text{ m}\cdot\text{s}^{-1}$ and the standard deviation is $\sigma = 0.5789 \text{ m}\cdot\text{s}^{-1}$. The mean and the variance of the Gaussian fit are estimated using the method of moments.

The distributions of the wind speed during the four trimesters of 2017 at the BCO have similar symmetric shapes and similar standard deviations, but different mean values (Figure 4c). The lowest mean wind is observed in the ASO trimester, while the highest mean value is observed during FMA (Table 1). The mean wind speed during the ASO trimester is the closest to the calculated mean wind speed in the LES. From this we hypothesize that the wind speed might be the reason for the close match in the mass flux distributions between the samples observed in the ASO trimester and LES, as well as the reason why the FMA trimester is the farthest away from the LES (see Figure 3b). The standard deviation is similar among the wind distributions observed over the four trimesters at the BCO, but the difference in the standard

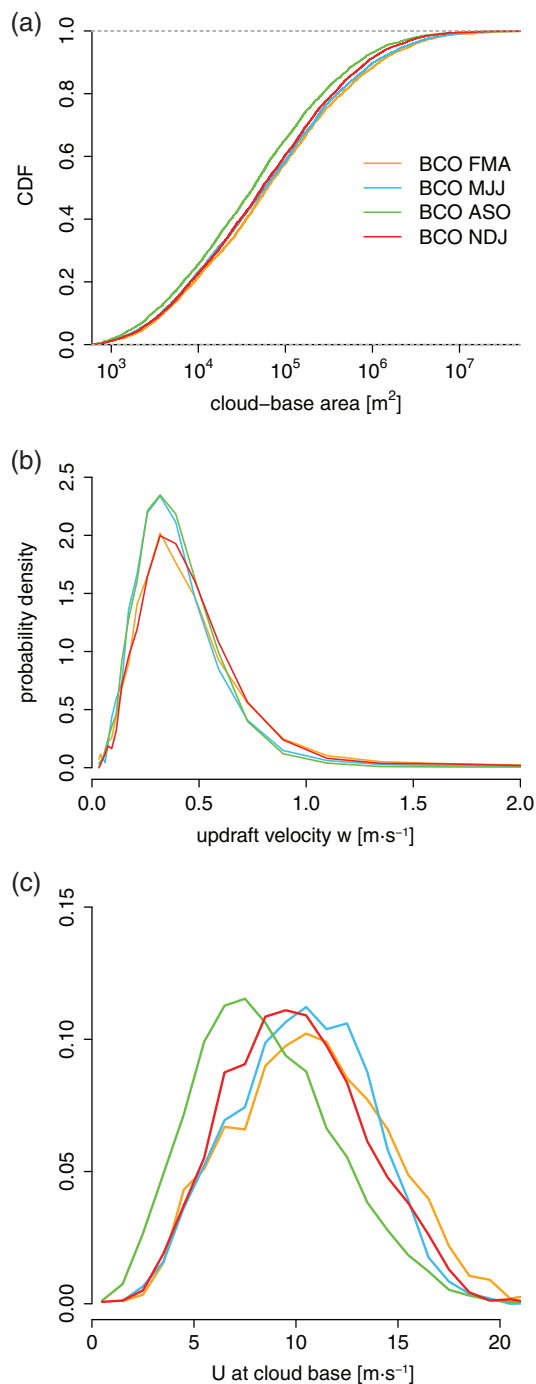


FIGURE 4 The observed (a) cumulative distribution of the individual cloud-base areas, (b) probability density distribution of the updraft vertical velocity, and (c) probability density distribution of the horizontal wind speed at the BCO in four trimesters of 2017

deviation between observations and LES is substantial. So, it is most likely the case that both the mean value and the variability of the wind determine the difference between the observed and the modelled distributions. The limitation of the sample size in our study prevents us from deriving more certain conclusions.

To assess the similarity between the observed and the modelled mass flux distributions under the same wind

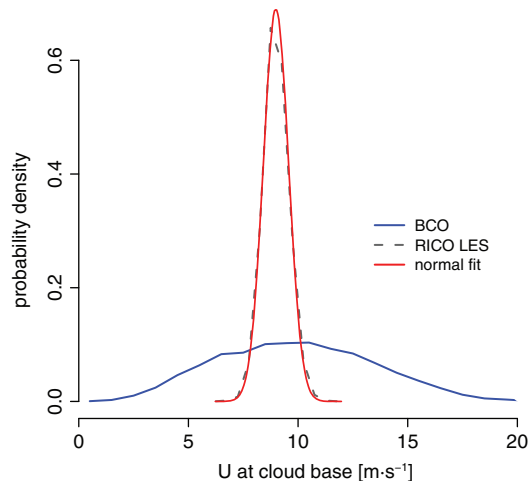


FIGURE 5 The distribution of the wind speed at the cloud base: comparison between the LES distribution and the BCO distribution, and the normal fit to the LES distribution

TABLE 1 Average wind speeds and standard deviations of the wind speed in the LES and at the BCO

Time	Mean [$\text{m}\cdot\text{s}^{-1}$]	Std [$\text{m}\cdot\text{s}^{-1}$]
RICO LES	8.99	0.58
BCO 2017	9.9	3.54
BCO FMA	10.74	3.74
BCO MJJ	10.26	3.27
BCO ASO	8.49	3.43
BCO NDJ	10.01	3.38

conditions, we select a subsample of clouds such that the wind distribution observed at the BCO exactly matches the wind distribution of the LES (Figure 5). This is achieved by randomly generating samples of the Gaussian distribution with the mean and standard deviation equal to those of the LES wind distribution, and then browsing the observed samples and collecting the same number of counts for each of the 20 bins of the Gaussian distribution. In this way we construct the distribution of the observed wind speed based on only those samples that constitute a similar distribution to the wind distribution of the LES (see Figure 5). This procedure is repeated for each trimester. The corresponding mass flux samples are paired with the drawn wind samples and the mass flux distributions are plotted in Figure 6.

The subsamples of the mass flux distributions that are based on a similar wind distribution to that of the LES result in a significantly closer match between the observed and the modelled mass flux distributions only in the MJJ case (Figure 6a,b and Table 2). We demonstrate the similarity between the MJJ case and the LES by using a cumulative distribution plot and a corresponding Q–Q plot (Figure 6c,d). The KS statistic has a remarkably low value in the MJJ case with $D = 0.029$, which is lower than the critical value for

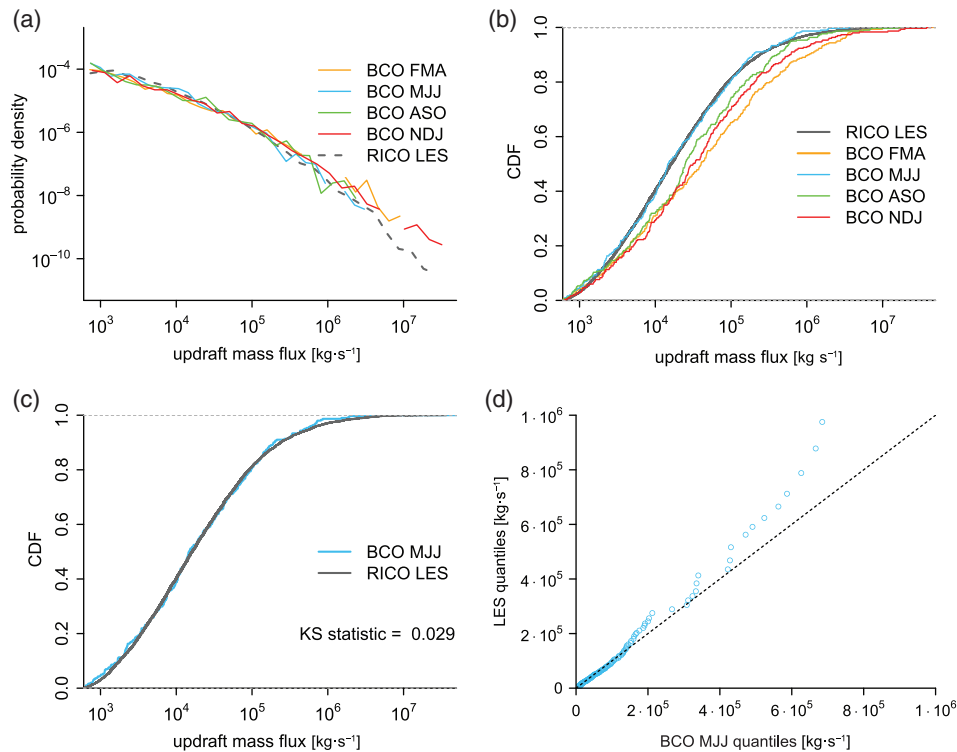


FIGURE 6 As Figure 3, but now samples consist of only those clouds for which the cloud-base horizontal velocity forms an identical distribution to the LES distribution of winds

TABLE 2 Kolmogorov–Smirnov statistics as a measure for the distribution distance between observations and LES

Trimester	D of the full sample	D of the wind-based sample
ERA FMA	0.181	0.179
ERA MJJ	0.140	0.029
ERA ASO	0.071	0.131
ERA NDJ	0.148	0.149

this sample of 350 clouds, $D < D_{crit,0.05} = 0.075$, based on the 95% confidence interval. So the null hypothesis of the similarity between the MJJ and LES distributions is not rejected in this case. The p -value of the KS test is $p = .97$, which means that there is a 97% chance of the two tested samples coming from the same distribution. The Q–Q plot shows very good agreement between the two samples. From these results we infer that the wind speed acts as a controlling factor for the mass flux distribution, because in at least one of the trimesters we find an excellent match between the observed and the modelled distributions. However, it is not the only factor that can shape the mass flux distribution, because the distributions in the other three trimesters do not closely match the LES distribution.

Based on the previous LES study of Sakradzija and Hohenegger (2017), the Bowen ratio B is likely another important factor that, in addition to the wind speed, controls

the mass flux distribution. The Bowen ratio in the region upwind of Barbados (12–17°N and 59–44°W) is calculated from the surface turbulent fluxes extracted from the ERA-Interim reanalysis with a time frequency of 12 hr (Dee *et al.*, 2011; Figure 7). The calculated Bowen ratio shows high daily variability. To assess the trend on a monthly basis, we also plot the running average (blue line in Figure 7a). The lowest values of the Bowen ratio are observed consistently during the months MJJ, on average 0.052 (Table 3). In the RICO LES case, the Bowen ratio does not change much over time (Figure 7b) and its value is 0.051 as averaged over the 24–60 hr simulation period. Thus, the value of the Bowen ratio calculated for the MJJ trimester is the closest to the Bowen ratio calculated in the LES (Figure 8, Table 3). We recognize this as a possible reason for the excellent match between the MJJ and LES mass flux distributions derived in the previous section when conditioned on similar wind conditions (Figure 6). In the ASO and NDJ trimesters, the observed Bowen ratio is around 0.07, which is also consistent with the larger difference in the mass flux distributions for these two trimesters (Figure 6). The trimester FMA is an outlier and although it has a Bowen ratio closer to the LES value compared to ASO and NDJ, the mass flux distribution is the furthest away from the LES distribution.

We further investigate whether other variables could explain the similarity between the MJJ and LES mass flux distributions. The surface buoyancy flux, sensible and latent heat fluxes, and the boundary layer height do not produce a

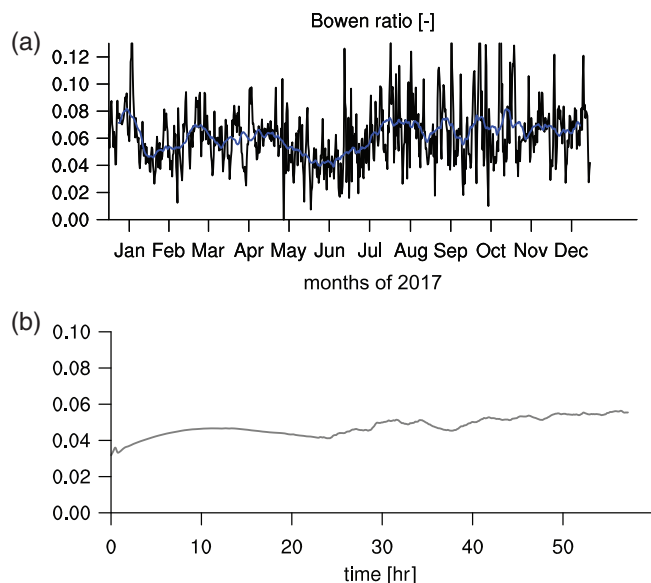


FIGURE 7 (a) Daily values of the Bowen ratio calculated from the ERA-Interim reanalysis data for 2017. The domain that represents the upwind conditions at the BCO site covers the region from 12°N to 17°N, and from 59°W to 44°W. A symmetric running average over 31 time steps is applied (blue line) to visually emphasize the trend over the year. (b) Time series of the Bowen ratio in the RICO LES case

good match between the reanalysis for the MJJ trimester and the LES and cannot explain the similarity between the mass flux distributions (Table 3 and Figure 8). It is intriguing that the buoyancy flux in the ASO trimester has the closest value to that of the buoyancy flux in the LES; however, we did not find a good match between the ASO and LES mass flux distributions. The mass flux distribution of the ASO trimester has the most similarity to the LES distribution before the subsampling, which is most likely because its wind distribution is the closest to the LES wind distribution (Figures 3 and 5), but it is not the closest match to the LES under the same wind conditions (Figure 6).

Based on these results, the observed distribution of the cloud-base mass flux is most similar to the distribution calculated from the LES output in the case of a subsample that has the closest Bowen ratio to the LES value of 0.05, but only in the case when the wind distribution closely resembles the wind distribution simulated in LES.

3.4 | Discussion

The influence of wind speed on the mass flux distribution does not come as a surprise. In the study of Nuijens and Stevens (2012), wind speed was found to influence the deepening of the boundary layer and cause deeper cloud layers. The effect of the wind on the boundary layer clouds is also confirmed by observations (Brueck *et al.*, 2015; Mieslinger *et al.*, 2019). Nuijens and Stevens (2012) further emphasize that shallow cumuli get deeper with stronger winds, but not more numerous or more energetic, while the wind speed does not change the buoyancy flux in the mixed layer in equilibrium. This finding is not in contradiction with observations that show the strong effect of the wind speed on the cloud size distribution and hence on the mass flux distribution. The mass flux of individual clouds is strongly correlated to the cloud area, and shows a very low correlation with the updraft velocity (Sakradzija *et al.*, 2015). Furthermore, the probability distributions of the mass flux strongly resemble the cloud size distributions and are not controlled by the buoyancy flux in the subcloud layer (Sakradzija and Hohenegger, 2017).

Based on the earlier study of Sakradzija and Hohenegger (2017) and on what we have just demonstrated, the buoyancy flux does not appear to control the mass flux of individual clouds and their distributions. So, we cannot explain the important role that the Bowen ratio has in shaping the statistics of shallow cumuli based on its influence on the surface buoyancy flux. For small Bowen ratios, the latent heat flux is dominant over the sensible heat flux, and only about 10% of the total surface flux contributes to the buoyancy flux. For large Bowen ratios, almost all of the heat flux at surface contributes to the buoyancy flux (e.g., Thomas *et al.*, 2018). In addition to the control on the buoyancy flux, the surface Bowen ratio also controls the limit up to which the heat flux at the surface can be converted into mechanical work of the convective circulations in the boundary layer (Shutts and Gray, 1999; Kleidon, 2016, sect. 10.2). This was recognized as a reason for the control that the Bowen ratio has on the mass flux distribution in the LES study of Sakradzija and Hohenegger (2017). The fact that the Bowen ratio sets the

TABLE 3 Comparison of the large-scale conditions estimated from ERA-Interim with the LES values of the Bowen ratio B , the surface buoyancy flux b_0 , the surface sensible heat flux SH , the latent heat flux LH and the boundary layer height H

Case/trimester	B	b_0 ($\text{cm}^2\cdot\text{s}^{-3}$)	SH ($\text{W}\cdot\text{m}^{-2}$)	LH ($\text{W}\cdot\text{m}^{-2}$)	H (m)
LES 24–60 hr	0.051	6.8	8.54	167.70	985.40
ERA FMA	0.058	7.6	10.38	175.70	910.87
ERA MJJ	0.052	6.3	7.97	155.44	751.72
ERA ASO	0.069	6.9	10.09	150.54	688.31
ERA NDJ	0.068	8.2	12.03	177.70	839.93

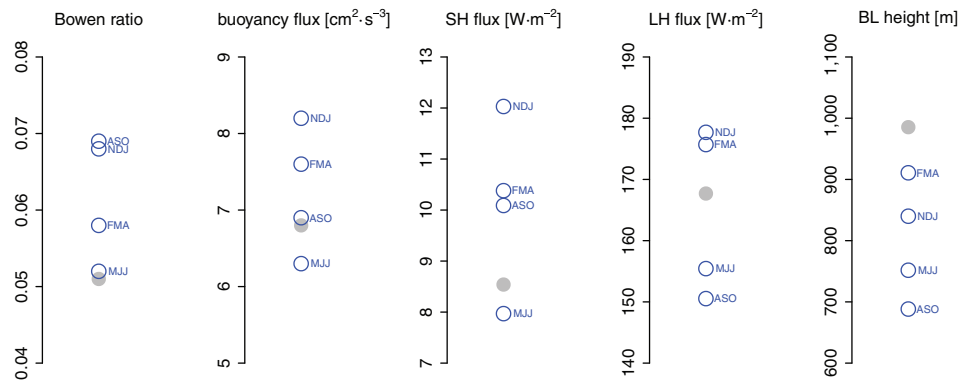


FIGURE 8 Comparison of the large-scale conditions estimated from ERA-Interim with the LES values of the Bowen ratio, surface buoyancy flux, surface sensible and latent heat fluxes, and boundary layer height. The grey dots correspond to the LES case

limit of the energy transformations in the convective boundary layer also explains why the magnitude of the buoyancy flux cannot explain the best match between the observed and the modelled distributions as in the MJJ trimester.

4 | SUMMARY AND CONCLUSIONS

We have presented a first attempt to measure the cloud-base mass flux of shallow cumuli and the corresponding probability distributions that are directly comparable to the cloud-base mass flux distributions simulated using idealized large-eddy simulation (LES). The observations were collected at the Barbados Cloud Observatory (BCO) during 2017. We used the idealized Rain In Cumulus over the Ocean (RICO) LES case study as defined in van Zanten *et al.* (2011), which is based on average conditions in winter 2004/2005 upwind of Antigua and Barbuda. We compared these two mass flux distributions and isolated the main factors that could explain the discrepancies between the two distributions. As suggested previously in the literature, the horizontal wind speed and the surface Bowen ratio were selected as the main controlling factors of the mass flux distribution. The horizontal wind speed was measured directly at the BCO, while for the other large-scale properties, including the Bowen ratio, we used the ERA-Interim reanalysis. Additional large-scale properties, such as the surface buoyancy flux, surface sensible and latent heat fluxes, and the boundary layer height were calculated using the ERA-Interim reanalysis for the region upwind of Barbados. These large-scale properties were also tested to understand the similarities between modelled and observed mass flux distributions.

We found that the mass flux distributions measured at the BCO and modelled by LES differed to some extent, but this difference was not much larger than the variability in the distribution measured at the BCO over the year 2017. The distribution measured at the BCO has a shallower slope and a longer tail compared to the LES distribution, which in general corresponds to larger clouds and stronger mass flux

values per cloud. The closest match between the observed and modelled distributions was found in the ASO trimester, which is most likely due to the horizontal wind speeds being most similar to those in the LES simulation. This match is, however, not statistically significant, and similarity between the two distributions still cannot be claimed. The measured and modelled distributions become similar and pass the statistical test only if the observed clouds are subsampled to take into account only those clouds that exist under the same wind conditions as in the LES. An excellent match is then found for the MJJ trimester, with a relatively high p -value of 97%. Since there is a good match in only one of the trimesters, we concluded that there must be another factor that exerts a control on the mass flux distribution. By analyzing the ERA-Interim fields upwind of Barbados, we found that the Bowen ratio has a value of 0.05 in both the LES and in the MJJ trimester. This led us to conclude that the wind speed and the surface Bowen ratio are the two factors that can explain the excellent match between the observed and modelled distributions in this case study. Such strong similarity between the observed and modelled mass flux distributions under the same large-scale conditions, selected based on previous research findings, provides us with greater confidence in our measurement methods. On the other hand, it shows us that the LES are capable of simulating realistic clouds that can be observed in nature within a narrow range of idealized large-scale forcing.

Here we have compared the cloud statistics based on a single LES case and a single year of observations. To investigate the robustness of our results and conclusions, more study cases and a longer observational period are necessary. The importance of our study lies in the first direct measurements of the cloud-base mass flux in individual clouds, in the first comparison of the mass flux distributions between observations and the LES and in the fact that we found an excellent match between the two. Furthermore, we have hinted at a physical background of the processes that could explain such excellent agreement between the observed and modelled distributions, which poses a basis for further investigation.

ACKNOWLEDGEMENTS

This research was carried out at the Hans Ertel Centre for Weather Research, a German research network consisting of universities, research institutions and the Deutscher Wetterdienst (German Meteorological Service) and which is funded by the Bundesministerium für Verkehr und digitale Infrastruktur (Federal Ministry of Transport and Digital Infrastructure). Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the author's work are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de. The Barbados Cloud Observatory (BCO) data are open source and are accessible to the broader research community, as detailed in Stevens *et al.* (2016). Primary BCO data can be accessed at <https://www.mpimet.mpg.de/en/science/the-atmosphere-in-the-earth-system/working-groups/tropical-cloud-observation/barbadosstation1/instrumentation-and-data/>. Computation was performed using the resources of the Deutsches Klimarechenzentrum (German Climate Computing Centre). Statistical analyses and plotting are done using the R statistical software. The authors thank Cathy Hohenegger and Bjorn Stevens for their guidance and discussions during the course of this study and the two anonymous reviewers for their constructive and helpful comments. Special thanks also go to Lutz Hirsch, Friedhelm Jansen, Ilya Serikov, Ludwig Worbes and Björn Brüggemann, who have kept the BCO running since 2010.

ORCID

Mirjana Sakradzija  <https://orcid.org/0000-0002-1544-8268>

REFERENCES

- Arakawa, A. (2004) The cumulus parameterization problem: past, present, and future. *Journal of Climate*, 17(13), 2493–2525. [https://doi.org/10.1175/1520-0442\(2004\)017<2493:RATCPP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2).
- Arakawa, A. and Schubert, W.H. (1974) Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *Journal of the Atmospheric Sciences*, 31, 674–701.
- Brueck, M., Nuijens, L. and Stevens, B. (2015) On the seasonal and synoptic time-scale variability of the North Atlantic trade wind region and its low-level clouds. *Journal of the Atmospheric Sciences*, 72(4), 1428–1446. <https://doi.org/10.1175/JAS-D-14-0054.1>.
- Cohen, B.G. and Craig, G.C. (2006) Fluctuations in an equilibrium convective ensemble. Part II: numerical experiments. *Journal of the Atmospheric Sciences*, 63(8), 2005–2015. <https://doi.org/10.1175/JAS3710.1>.
- Conover, W.J. (1999) *Practical Nonparametric Statistics*, 3rd edition. New York, NY: John Wiley & Sons.
- Craig, G.C. and Cohen, B.G. (2006) Fluctuations in an equilibrium convective ensemble. Part I: theoretical formulation. *Journal of the Atmospheric Sciences*, 63(8), 1996–2004. <https://doi.org/10.1175/JAS3709.1>.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>.
- Ghate, V.P., Miller, M.A. and DiPretore, L. (2011) Vertical velocity structure of marine boundary layer trade wind cumulus clouds. *Journal of Geophysical Research: Atmospheres*, 116, D16206. <https://doi.org/10.1029/2010JD015344>.
- Hellman, G. (1916) Über die Bewegung der Luft in den untersten Schichten der Atmosphäre. *Meteorologische Zeitschrift*, 34, 273–285.
- Kleidon, A. (2016) *Thermodynamic Foundations of the Earth System*, 1st edition. Cambridge, England: Cambridge University Press.
- Klingebiel, M., Ghate, V.P., Naumann, A.K., Ditas, F., Pöhlker, M.L., Pöhlker, C., Kandler, K., Konow, H. and Stevens, B. (2019) Remote sensing of sea salt aerosol below trade wind clouds. *Journal of the Atmospheric Sciences*, 76(5), 1189–1202. <https://doi.org/10.1175/JAS-D-18-0139.1>.
- Mieslinger, T., Horváth, A., Buehler, S.A. and Sakradzija, M. (2019) The dependence of shallow cumulus macrophysical properties on large-scale meteorology as observed in ASTER imagery. *Journal of Geophysical Research: Atmospheres*, 124. <https://doi.org/10.1029/2019JD030768>.
- Nitta, T. (1975) Observational determination of cloud mass flux distributions. *Journal of the Atmospheric Sciences*, 32(1), 73–91. [https://doi.org/10.1175/1520-0469\(2003\)60<1201:ALESIS>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2).
- Nuijens, L. and Stevens, B. (2012) The influence of wind speed on shallow marine cumulus convection. *Journal of the Atmospheric Sciences*, 69(1), 168–184. <https://doi.org/10.1175/JAS-D-11-02.1>.
- Ogura, Y. and Cho, H.-R. (1973) Diagnostic determination of cumulus cloud populations from observed large-scale variables. *Journal of the Atmospheric Sciences*, 30(7), 1276–1286. [https://doi.org/10.1175/1520-0469\(1973\)030<1276:DDOCCP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<1276:DDOCCP>2.0.CO;2).
- Plant, R.S. and Craig, G.C. (2008) A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences*, 65(1), 87–105. <https://doi.org/10.1175/2007JAS2263.1>.
- R Core Team (2015) *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/> [Accessed 22 October 2019].
- Rasp, S., Selz, T. and 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. <https://doi.org/10.1175/JAS-D-17-0258.1>.
- Rauber, R.M., Stevens, B., Ochs, H.T., Knight, C., Albrecht, B.A., Blyth, A.M., Fairall, C.W., Jensen, J.B., Lasher-Trapp, S.G., Mayol-Bracero, O.L., Vali, G., Anderson, J.R., Baker, B.A., Bandy, A.R., Burnet, E., Brenguier, J., Brewer, W.A., Brown, P.R., Chuang, R., Cotton, W.R., Di Girolamo, L., Geerts, B., Gerber, H., Göke, S., Gomes, L., Heikes, B.G., Hudson, J.G., Kollias, P., Lawson,

- R.R., Krueger, S.K., Lenschow, D.H., Nuijens, L., DO'Sullivan, D.W., Rilling, R.A., Rogers, D.C., Siebesma, A.P., Snodgrass, E., JStith, J.L., Thornton, D.C., Tucker, S., Twohy, C.H., and Zuidema, P. (2007) Rain in (shallow) cumulus over the ocean: the RICO campaign. *Bulletin of the American Meteorological Society*, 88, 1912–1928.
- Sakradzija, M. and Hohenegger, C. (2017) What determines the distribution of shallow convective mass flux through a cloud base? *Journal of the Atmospheric Sciences*, 74(8), 2615–2632. <https://doi.org/10.1175/JAS-D-16-0326.1>.
- Sakradzija, M. and Klocke, D. (2018) Physically constrained stochastic shallow convection in realistic kilometer-scale simulations. *Journal of Advances in Modeling Earth Systems*, 10(11), 2755–2776. <https://doi.org/10.1029/2018MS001358>.
- Sakradzija, M., Seifert, A. and Dipankar, A. (2016) A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone. *Journal of Advances in Modeling Earth Systems*, 8(2), 786–812. <https://doi.org/10.1002/2016MS000634>.
- Sakradzija, M., Seifert, A. and Heus, T. (2015) Fluctuations in a quasi-stationary shallow cumulus cloud ensemble. *Nonlinear Processes in Geophysics*, 22, 65–85. <https://doi.org/10.5194/npg-22-65-2015>.
- Seifert, A. and Beheng, K.D. (2001) A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmospheric Research*, 59–60, 265–281. [https://doi.org/10.1016/S0169-8095\(01\)00126-0](https://doi.org/10.1016/S0169-8095(01)00126-0).
- Shutts, G.J. and Gray, M.E.B. (1999) Numerical simulations of convective equilibrium under prescribed forcing. *Quarterly Journal of the Royal Meteorological Society*, 125(559), 2767–2787. <https://doi.org/10.1002/qj.49712555921>.
- Siebesma, A.P., Bretherton, C.S., Brown, A., Chlond, A., Cuxart, J., Duijkerke, P.G., Jiang, H., Khairoutdinov, M., Lewellen, D., Moeng, C.-H., Sanchez, E., Stevens, B. and Stevens, D.E. (2003) A large eddy simulation intercomparison study of shallow cumulus convection. *Journal of the Atmospheric Sciences*, 60(10), 1201–1219. [https://doi.org/10.1175/1520-0469\(2003\)60<1201:ALESIS>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2).
- Stevens, B. (2010) *Introduction to UCLA-LES, version 3.2.1*. Available at: https://github.com/uclaes/uclaes/blob/master/doc/les_doc.pdf [Accessed 11th April 2019].
- Stevens, B., Farrell, D., Hirsch, L., Jansen, F., Nuijens, L., Serikov, I., Brüggmann, B., Forde, M., Linne, H., Lonitz, K. and Prospero, J.M. (2016) The Barbados Cloud Observatory: anchoring investigations of clouds and circulation on the edge of the ITCZ. *Bulletin of the American Meteorological Society*, 97(5), 787–801. <https://doi.org/10.1175/BAMS-D-14-00247.1>.
- Stevens, B., Moeng, C.-H. and Sullivan, P.P. (1999) Large-eddy simulations of radiatively driven convection: sensitivities to the representation of small scales. *Journal of the Atmospheric Sciences*, 56, 3963–3984.
- Thomas, L., Malap, N., Grabowski, W.W., Dani, K. and Prabha, T.V. (2018) Convective environment in pre-monsoon and monsoon conditions over the Indian subcontinent: the impact of surface forcing. *Atmospheric Chemistry and Physics*, 18(10), 7473–7488.
- van Zanten, M.C., Stevens, B., Nuijens, L., Siebesma, A.P., Ackerman, A.S., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D.C., Mechem, D., Nakamura, K., Noda, A., Shipway, B.J., Slawinska, J., Wang, S. and Wyszogrodzki, A. (2011) Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *Journal of Advances in Modeling Earth Systems*, 3, M06001. <https://doi.org/10.1029/2011MS000056>.
- Yanai, M., Esbensen, S. and Chu, J.-H. (1973) Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *Journal of the Atmospheric Sciences*, 30(4), 611–627. [https://doi.org/10.1175/1520-0469\(1973\)030<0611:DOBPOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2).

How to cite this article: Sakradzija M, Klingebiel M. Comparing ground-based observations and a large-eddy simulation of shallow cumuli by isolating the main controlling factors of the mass flux distribution. *Q J R Meteorol Soc.* 2020;146:254–266. <https://doi.org/10.1002/qj.3671>