Estimating the Shallow Convective Mass Flux from the Subcloud-Layer Mass Budget

Raphaela Vogel, Sandrine Bony, Bjorn B Stevens

Estimating the Shallow Convective Mass Flux from the Subcloud-Layer Mass Budget

RAPHAELA VOGEL AND SANDRINE BONY

LMD/IPSL, CNRS, Sorbonne University, Paris, France

BJORN STEVENS

Max Planck Institute for Meteorology, Hamburg, Germany

ABSTRACT

This paper develops a method to estimate the shallow-convective mass flux M at the top of the subcloud layer as a residual of the subcloud-layer mass budget. The ability of the mass-budget estimate to reproduce the mass flux diagnosed directly from the cloud-core area fraction and vertical velocity is tested using real-case large-eddy simulations over the tropical Atlantic. We find that M reproduces well the magnitude, diurnal cycle, and day-to-day variability of the core-sampled mass flux, with an average root-mean-square error of less than 30% of the mean. The average M across the four winter days analyzed is $12\,\mathrm{mm\,s^{-1}}$, where the entrainment rate E contributes on average $14\,\mathrm{mm\,s^{-1}}$ and the large-scale vertical velocity W contributes $-2\,\mathrm{mm\,s^{-1}}$. We find that day-to-day variations in M are mostly explained by variations in W, whereas E is very similar among the different days analyzed. Instead E exhibits a pronounced diurnal cycle, with a minimum of about $10\,\mathrm{mm\,s^{-1}}$ around sunset and a maximum of about $18\,\mathrm{mm\,s^{-1}}$ around sunrise. Application of the method to dropsonde data from an airborne field campaign in August 2016 yields the first measurements of the mass flux derived from the mass budget, and supports the result that the variability in M is mostly due to the variability in W. Our analyses thus suggest a strong coupling between the day-to-day variability in shallow convective mixing (as measured by M) and the large-scale circulation (as measured by W). Application of the method to the EUREC[4]A field campaign will help evaluate this coupling, and assess its implications for cloud-base cloudiness.

1. Introduction

The response of shallow trade wind cumuli to global warming is a major contributor to uncertainty in climate sensitivity across climate models (Bony et al. 2004; Bony and Dufresne 2005; Vial et al. 2013). A large fraction of this uncertainty can be attributed to model differences in the strength of lower-tropospheric mixing associated with convective and large-scale circulations (Sherwood et al. 2014; Zhao 2014; Brient et al. 2016). A stronger mixing in a warmer climate leads to a stronger drying of the lower cloud layer, resulting in reduced cloudiness near the base of the shallow cumulus layer and a higher climate sensitivity (Sherwood et al. 2014). The rate at

which cloudiness reduces with increasing mixing depends both on the shallowness of the present-day cloud profile (Brient et al. 2016), and on the coupling between convective mixing, surface turbulent fluxes and low-cloud radiative effects (Vial et al. 2016). In contrast to climate models, cloud-base cloudiness in large-eddy simulations (LES) and in observations is much less sensitive to changes in the large-scale environment (Bretherton et al. 2013; Blossey et al. 2013; Vogel et al. 2016; Nuijens et al. 2014, 2015).

The *cumulus-valve mechanism* provides a conceptual framework for understanding changes in cloud-base cloudiness in response to changes in the convective mass flux—an important measure for the strength of convective mixing. This mechanism could explain a larger cloud fraction with more mass flux, if the mass flux increase needed to maintain cloud base at the lifting condensation level (LCL) was mostly accommodated by increasing the area fraction of active cumuli. However, whether such a cumulus-valve mechanism is at play has

never been critically evaluated with observations, partly because of the lack of an observational estimate of the mass flux.

To test the coupling between convective mixing, cloudiness and the large-scale environment, the Elucidating the Role of Cloud-Circulation Coupling in Climate (EUREC[4]A) field campaign will collect simultaneous measurements of the properties of the cloud field and its thermodynamic and dynamic environment (Bony et al. 2017). From these measurements, the convective mass flux at the top of the subcloud layer will be inferred. The mass flux can be computed from radar-retrieved cloud-core area fraction and vertical velocity (Kollias and Albrecht 2010; Ghate et al. 2011; Lamer et al. 2015). However, given that the large-scale vertical velocity can now be measured from dropsondes (Bony and Stevens 2019), the mass flux can also be estimated as a residual of the subcloud-layer mass budget. As a first step toward testing the cumulus-valve mechanism with observations, this study develops a method to estimate the convective mass flux at the top of the subcloud layer from observations by analyzing the mass budget, and tests this method using LES.

Analyses of the mass budget have previously been used to estimate entrainment rates in stratocumulus-topped mixed layers (Stevens et al. 2003; Albrecht et al. 2016; Ghate et al. 2019). In the case of stratocumulus, turbulence is confined to the mixed layer and there is no convective mass flux out of the layer. Ahlgrimm and Randall (2006) used a simple bulk boundary layer model to estimate the entrainment rate and the cumulus mass flux at the top of the trade cumulus layer from monthly mean reanalysis data. Because the trade wind convection remains mostly confined to the cumulus layer, the mass flux vanishes at the top of the cloud layer.

Here we estimate the shallow-convective mass flux at the top of the subcloud layer, because this level is critical for understanding low-cloud feedbacks and for testing the cumulus-valve mechanism. We address the following research questions: 1) How well does the mass flux estimated from the subcloud-layer mass budget reproduce the magnitude and variability of the core-sampled mass flux? 2) What is the contribution of the different budget terms to the magnitude and variability of the mass flux?

The analyses presented in this paper are based on 4 days of LES performed with the Icosahedral Nonhydrostatic Large-Eddy Model (ICON-LEM) during the first Next-Generation Aircraft Remote Sensing for Validation Studies (NARVAL) field campaign in December 2013 (Stevens et al. 2019). Section 2 describes the conceptual model used to infer the mass flux from LES or observations. The setup of the ICON-LEM simulation, the application of the conceptual model to the ICON-LEM output and the evaluation metrics are presented in section 3. Section 4 discusses the ability of the mass budget estimate to reproduce the core-sampled mass flux (section 4a), the robustness of the estimate (section 4b), and the sources of mass flux variability in the LES (section 4c). Section 5 shows the application of the method to dropsonde data from the second NARVAL field campaign in August 2016 (Stevens et al. 2019). The summary and conclusions are presented in section 6.

2. Conceptual model

Figure 1 shows an illustration of the subcloud-layer specific mass budget, which can be expressed as

$$\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + \mathbf{V}_h \cdot \nabla h = E + W - M. \tag{1}$$

Here, h is the height of the subcloud layer and Dh/Dt is the total derivative of h, which includes both the time evolution and the horizontal advection of h. The entrainment rate E represents deepening through small-scale mixing at h; E incorporates relatively warm and dry cloud layer air into the subcloud layer and thus constitutes a mass source. Also, W is the large-scale vertical velocity at h. In the presence of divergence within the subcloud layer, W measures the subcloud-layer descent that arises from the lateral mass sink. Finally, M is the convective updraft mass flux out of the subcloud layer, an additional mass sink.

Assuming that E, W, and Dh/Dt can be measured, M can be estimated as a residual of Eq. (1). The estimation procedure is illustrated in Fig. 2.

We define h as the level of neutral buoyancy z_{nb} of a parcel lifted from the surface layer z_{sf}, plus a fixed overshoot fraction α_{os} that accounts for the momentum of the updrafts (Stull 1988; Hourdin et al. 2002), such that

$$h = z_{nb} + \alpha_{os} \times (z_{nb} - z_{sf}). \tag{2}$$

Similar to Betts (1976), this definition of h includes both the mixed layer and the transition layer below cloud base, with the latter being represented by the overshoot.

For computing the entrainment rate at h, we assume that the entrainment dynamics of cumulus-topped boundary layers are the same as those of dry-convective boundary layers, as shown to be valid for LES (Siebesma and Cuijpers 1995; Siebesma et al. 2003) and as used in bulk boundary layer models (Betts 1976; Neggers et al. 2006; Stevens 2006). With the aid of the buoyancy flux closure, which models the vertical buoyancy flux at h, $\overline{w'\theta'_v}|_h$, as a fixed
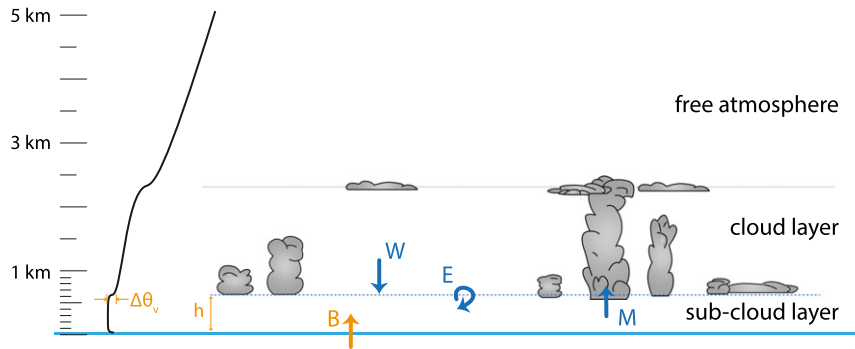
FIG. 1. Illustration of the trade wind layer structure and the terms of the subcloud-layer mass budget (adapted from Bony et al. 2017). Refer to the text for details.

fraction A of its surface value $\overline{w'\theta'_v}|_s$ (see, e.g., Tennekes 1973; Deardorff et al. 1974; Betts 1976), we can compute E from the zeroth-order flux-jump relationship (Lilly 1968) as

$$E = -\frac{\overline{w'\theta'_v}|_h}{\Delta\theta_v} = \frac{A\,\overline{w'\theta'_v}|_s}{\Delta\theta_v}, \qquad (3)$$

where $\Delta\theta_v$ denotes the jump across an assumed infinitesimally thin transition layer between the subcloud layer and the cloud layer. As the transition layer between the mixed layer and the bottom of the cloud layer in the regime analyzed here is about 100–150 m thick (as represented by α_{os}), the use of a zeroth-order jump model might only partially represent the complexity of the entrainment dynamics (vanZanten et al. 1999; Canut et al. 2012; Garcia and Mellado 2014).

We account for the contribution of the temporal fluctuations of h to the mass budget, but neglect the horizontal advection term, because it is difficult to robustly estimate it from the observations. h would have to be computed from individual dropsondes, whose θ_v profiles tend to be noisy. Estimation of the advection term from

dropsonde circle data from the NARVAL2 campaign (see section 5) showed that it is likely about an order of magnitude smaller than the other terms. Note that shallow trade cumuli precipitate frequently (Byers and Hall 1955; Rauber et al. 2007) and may produce downdrafts in the lower cloud layer and subcloud layer (Zuidema et al. 2012). Here we assume that such penetrative downdrafts occur too rarely to be represented explicitly, and that their effect is implicitly captured as an entrainment contribution from downward coherent structures.

The various assumptions made in the conceptual model are tested with LES. The application of the conceptual model to the LES is presented in the next section.

3. Application to LES

a. LES setup

We use ICON in LES configuration (Dipankar et al. 2015) on a domain over the tropical Atlantic, upstream of Barbados. The model setup is similar to the one used in Heinze et al. (2017) for real-case LES over Germany.
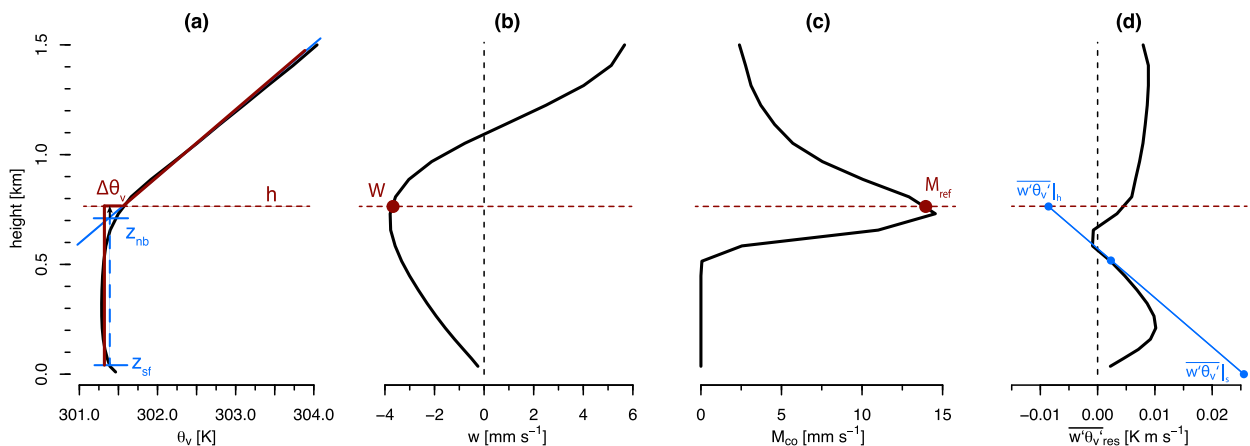

FIG. 2. Illustration of the estimation procedure. Vertical profiles of (a) θ_v, (b) large-scale vertical velocity, (c) M_{co}, and (d) resolved buoyancy flux, from the 156-m-resolution ICON simulation. Refer to the text for details.

TABLE 1. Mean properties of the 156-m-resolution simulations for the 4 days during NARVAL1 in 2013. The rows labeled day (night) represent averages from 1200 to 1600 LT (0200 to 0600 LT). Shown are total cloud cover, maximum cloud-base cloud fraction, average cloud-top height, SST, 10-m wind speed, water vapor path, surface latent and sensible heat flux, surface buoyancy flux, and surface rain rate. The last column shows the type of mesoscale cloud organization as defined by Stevens et al. (2020).

	CC (–)	CFcb (–)	CTH (m)	SST (K)	U_{10m} (m s^{-1})	WVP (kg m^{-2})	LHF (W m^{-2})	SHF (W m^{-2})	SBF (K m s^{-1})	Rsrf (mm day^{-1})	Org (–)
11 Dec											
Day	0.07	0.04	956	300.8	9.5	29.7	235.5	10.3	0.0236	0.00	Sugar
Night	0.19	0.06	1991	300.8	10.1	34.4	233.2	14.7	0.0272	0.13	Gravel
12 Dec											
Day	0.08	0.03	1441	300.8	9.6	32.4	236.3	7.6	0.0214	0.00	Sugar
Night	0.13	0.06	1273	300.8	9.3	30.3	217.5	14.3	0.0259	0.13	Sugar
15 Dec											
Day	0.07	0.03	1157	300.8	10.7	31.1	272.6	9.8	0.0255	0.02	Fish
Night	0.16	0.05	1692	300.8	11.3	29.1	302.1	17.7	0.0341	0.09	Fish
20 Dec											
Day	0.08	0.03	1493	300.6	8.2	31.2	213.3	9.7	0.0217	0.00	Gravel
Night	0.16	0.05	2151	300.6	9.6	29.9	270.4	17.7	0.0321	0.00	Gravel

ICON solves the compressible Navier–Stokes equations on an unstructured grid as detailed in Zängl et al. (2015) and Dipankar et al. (2015). The LES configuration uses the Smagorinsky turbulence scheme and the two-moment mixed-phase bulk microphysics scheme of Seifert and Beheng (2006). No subgrid cloud scheme is used, and cloud fraction is unity in a grid box if the liquid water content $q_l > 1 \times 10^{-8}$ kg kg^{-1}. Sea surface temperatures (SSTs) are prescribed from Integrated Forecast System (IFS) data of the European Centre for Medium-Range Weather Forecasts (ECMWF) and do not change during the day (see Table 1).

The LES are initialized at 0900 UTC from a storm-resolving simulation at 1.25-km grid spacing (see Klocke et al. 2017). The 1.25-km simulation is initialized at 0000 UTC from IFS data, and is nudged to the IFS data at its lateral boundaries every hour. The LES domains consist of three one-way nested grids at 625-, 313-, and 156-m horizontal grid spacing. The three LES domains are nudged at every time step to the next coarser grid. The two outer grids cover the western tropical Atlantic from about 60° to 43°W and 8° to 17°N. The innermost 156-m-resolution grid spans a smaller domain between about 59.75° and 57.25°W and 12.6° and 13.6°N. A stretched vertical grid with 150 levels is used, with a grid spacing of 20 m at the surface and about 70 m at 600-m height. The simulations are run for 27 h until 1200 UTC of the next day. We discard the first 5 h due to spinup and analyze the simulations over an almost complete diurnal cycle from 1400 to 1200 UTC on the next day, that is, from 1000 to 0800 LT in Barbados.

We analyze 4 days in December 2013 encompassing research flights from the NARVAL1 field campaign that took place upstream of Barbados (Stevens et al. 2019). We focus on the 156-m-resolution simulations on a domain spanning 59.4°–57.4°W and 12.6°–13.6°N (about 200 km × 100 km), which excludes Barbados and the borders of the domain, and is about the size of the area of the aircraft operations planned during EUREC^{4}A.

Table 1 summarizes the mean conditions of these simulations. On all days, there is a pronounced diurnal cycle of cloud cover (including clouds up to 7.5-km height), with a minimum cloud cover of 7% in the early afternoon and a maximum cloud cover of about 16% near the end of the night (see also Vial et al. 2019). The diurnal cycle is also reflected in cloud-base cloudiness and average cloud-top height (except on 12 December). All days are characterized by strong surface winds of about 10 m s^{-1}, and strong sensible and latent heat fluxes. There is negligible surface precipitation and the cloud organization is mostly characterized by small-scale and unorganized patches of shallow clouds [referred to as *Sugar* in Stevens et al. (2020)], or slightly deeper clouds with indications of arc-like structures (*Gravel*). The 15 December stands out as a day with particularly strong winds and surface fluxes, and a mesoscale organization with larger structures resembling the *Fish* category of Stevens et al. (2020).

Two additional days (14 and 16 December 2013) were simulated at a time when the method was already developed. We find similar results for these days compared to the other days and do not discuss them here.

b. From the conceptual model to the LES

In this section we discuss how the conceptual model is applied to the ICON model output. Figure 2 shows an illustration of the estimation procedure.

For computing h from the median θ_v profile following Eq. (2), the surface lifting height z_{sf} is set to 45 m, and

the surface $\theta_{v,\mathrm{sf}}$ is the average θ_v over the first 90 m (Fig. 2a). The level of neutral buoyancy z_{nb} is computed with respect to a background profile fitted to the cloud layer θ_v. The fit is done using linear regression of the cloud layer θ_v, starting from the first level lying 50 m above the height where $\theta_v > \theta_{v,\mathrm{sf}}$ (with a minimum height of 700 m), up to the level comprising two-thirds of the distance between the starting level and the trade inversion z_{inv} (maximum height: 2700 m); z_{inv} is defined as the level of the maximum relative humidity gradient above 1200 m, which tends to give a more robust estimate of the inversion than the more commonly used maximum θ_v gradient.

The fixed overshoot fraction α_{os} is set to 15%, which is about half the average overshoot attained when accounting for the subcloud-layer convective available potential energy (CAPE) in the ICON simulations (not shown), and which is slightly larger than the 9% derived in Betts (1976). α_{os} is used here as a tuning parameter to constrain h such that it gets close to the level of maximum cloud-base cloudiness, which results in the best correspondence with the core-sampled mass flux. Note that our h can lie anywhere between two vertical levels and is thus independent of the vertical grid spacing. We also tested an alternative definition of h that computes the overshoot at every time step using the subcloud-layer CAPE (see Hourdin et al. 2002; Rio and Hourdin 2008), but do not find an improved consistency with the core-sampled mass flux compared to the simpler definition with fixed α_{os}.

For the entrainment rate E [Eq. (3)], $\Delta\theta_v$ is computed as the difference between the average subcloud-layer θ_v and the fitted background θ_v at h (see Fig. 2a). The ratio of the entrainment to the surface buoyancy flux A is set to the typical value of 0.2 (Stull 1976; Betts 1976; Stevens 2006). In the absence of total buoyancy flux profile output from the LES, a rough estimate of A can be obtained from the resolved buoyancy flux (Fig. 2d), which at the given resolution is only a small fraction of the total flux near the surface and near h (Dipankar et al. 2015). Similar to the procedure of Deardorff et al. (1974), the entrainment buoyancy flux at h is approximated by fitting a line from the surface buoyancy flux, through the value of the first level below the inflection point near the mixed layer top, up to h (see the illustration in Fig. 2d). With this method, we find the mean A to be between about 0.29 (on 15 December) and 0.45 (on 12 December), with no systematic diurnal cycle. As this test only gives a rough estimate of A due to the missing subgrid contribution, we keep $A = 0.2$ fixed here, as this renders the mass budget estimate closest to the core-sampled mass flux. A is therefore used as a tuning parameter to match the mean magnitude of M for the ICON-LEM simulations. Because

the buoyancy flux profile is influenced by the depth of the transition layer, which the simulations poorly resolve, the optimal value of A may be different for the dropsonde observations or other LES [as discussed in section 4b for a simulation with the University of California, Los Angeles (UCLA)-LES (Stevens et al. 2005)]. Note that our h is close to the level of maximum cloud-base cloud fraction and therefore already in the region of positive buoyancy flux.

We define W as the domain-averaged vertical velocity evaluated at h (Fig. 2b), and $\partial h/\partial t$ is computed as the centered difference of h between the previous and the subsequent time step, divided by the time interval. We apply the estimation procedure to hourly averaged LES output, as this is approximately the time required to fly one circle during EUREC⁴A (Bony et al. 2017).

The magnitude and variability of the horizontal advection term is assessed by computing h from the median θ_v profiles of blocks of $0.125° \times 0.125°$. The gradients of h in the x and y directions are computed with the linear regression method of Bony and Stevens (2019) and multiplied by the wind speed at the average h of the different blocks. We find the advection of h in the LES to be similar in magnitude and variability compared to the temporal fluctuation of h, with a mean and standard deviation of about $-1.5 \pm 8.2\,\mathrm{mm\,s^{-1}}$. The estimates show little sensitivity to the block size. The advection term tends to be larger on smaller domains, as variability in h can be noisy and local gradients in h therefore quite large. As mentioned in section 2, the advection term is neglected in the following analyses, as it is difficult to estimate it robustly (especially from the observations).

The sensitivity of the mass flux to the various parameter choices will be discussed in section 4b.

c. Evaluation metrics

In the following analyses, we consider two mass flux estimates from the mass budget. An equilibrium estimate,

$$M = E + W,\qquad(4)$$

which neglects the contribution of temporal fluctuations of h. And a temporally varying estimate,

$$M' = E + W - \frac{\partial h}{\partial t},\qquad(5)$$

which accounts for the contribution of temporal fluctuations of h.

For the LES, the two mass flux estimates are compared with the core-sampled mass flux M_{ref} diagnosed directly from the cloud-core area fraction and cloud-core vertical velocity (see Fig. 2c) as

$$M_{\mathrm{ref}} = M_{\mathrm{co}}|_h = (a_{\mathrm{co}}w_{\mathrm{co}})|_h.\qquad(6)$$
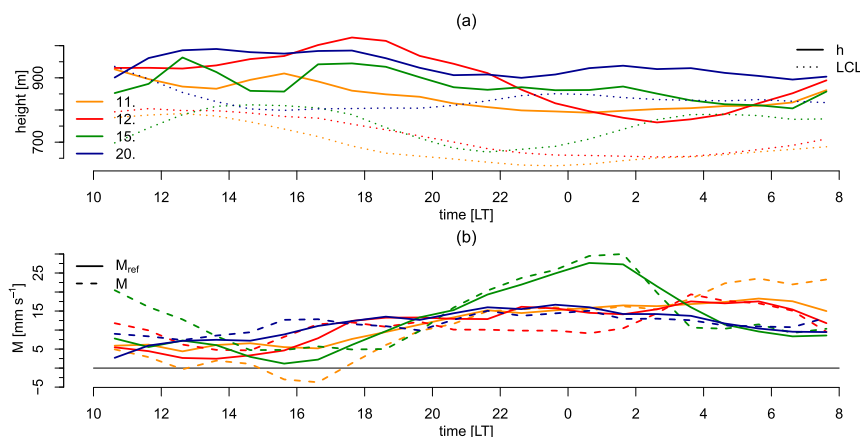
FIG. 3. Time series of (a) h and LCL and (b) M and M_{ref}. Shown are the 4 days of December 2013 during NARVAL1 at 156-m resolution.

We define the cloud-core area fraction a_{co} as the fraction of all cloudy grid points with positive vertical velocity ($w > 0\,\mathrm{m\,s^{-1}}$ and $q_l > 1 \times 10^{-8}\,\mathrm{kg\,kg^{-1}}$). This core definition facilitates comparison with observations, as both the vertical velocity and the area fraction of cloud cores can be retrieved from Doppler radars (see, e.g., Kollias and Albrecht 2010; Lamer et al. 2015). There are other possible ways to define M_{co}, for example, by using the buoyancy to define the cloud core or by considering also the cloudy areas with downward motion (Siebesma and Cuijpers 1995). The appendix shows that different definitions lead to a similar variability, but a different magnitude of M_{co}.

During EUREC^{4}A, we ultimately want to examine the relationship between the mass flux and cloud-base cloudiness. Our M_{ref} should thus be close to the maximum M_{co} near cloud base, where the cloud fraction is also maximum. The appendix shows that this is generally the case.

Note that our mass fluxes are in units of millimeters per second and are thus specific mass fluxes or volume fluxes. They should be multiplied by the air density to be in units of kilograms per square meter per second.

4. LES results

a. Ability of M to reproduce M_{ref}

This section shows how the mass flux estimated from the mass budget is able to reproduce the magnitude and variability of the core-sampled mass flux. Figure 3a shows the subcloud-layer depth h, diagnosed as explained in section 3b, for all four days: it ranges from 800 to 1000 m, which is about 50 to 200 m above the LCL [computed with air properties at z_{sf} following Bolton (1980)]. h shows a slight diurnal cycle on all days, with a peak in the late afternoon and a decrease during the night. The average M_{ref} of about $12\,\mathrm{mm\,s^{-1}}$ has the same order of magnitude as estimates from previous LES intercomparison cases of shallow convection over tropical oceans (Siebesma et al. 2003; VanZanten et al. 2011) and observations (Lamer et al. 2015; Ghate et al. 2011). The equilibrium estimate M appears to capture the magnitude, the diurnal cycle, and also much of the day-to-day variability of M_{ref} (Fig. 3b). For example, on 15 December, M captures well the minimum in M_{ref} in the afternoon and the very pronounced maximum around midnight. The absolute difference between M and M_{ref} is generally smaller than $5\,\mathrm{mm\,s^{-1}}$. Inclusion of the $\partial h/\partial t$ term in M' leads to noisier estimates (not shown).

Figure 3b shows that in a few cases, M becomes negative. An explanation for the negative values may be the neglection of the horizontal advection and temporal fluctuation of h in M. Because we interpret M as the updraft mass flux carried by the cloud core, and because M_{ref} is by definition positive, we set all negative values of M (and M') to 0 for the following analyses. We find differences between M and M_{ref} to be anticorrelated with the subcloud-layer CAPE (correlation coefficient $R = -0.67$ for all simulations considered), which itself is anticorrelated with the precipitation rate (not shown). CAPE reduces when evaporating precipitation cools the lower portion of the subcloud layer, which leads to lower values of h and often a larger difference between M and M_{ref}. All instances when CAPE is lower than $0.75\,\mathrm{m^2\,s^{-2}}$ are excluded in the following.

The diagnostics in Table 2 and the scatterplot in Fig. 4 include the modifications discussed in the previous paragraph and allow for a more detailed analysis of estimation skill. The daily mean M is with 11–$14\,\mathrm{mm\,s^{-1}}$ very similar among the different days, and agrees well with the mean M_{ref}. Differences between

TABLE 2. Mean properties and skill of the mass budget estimates for the 156-m-resolution simulations using the default setting. Shown are the daily mean and standard deviation of the mass flux ($\overline{M}$ and σ_M) and the difference in 4-h means between the nighttime maximum and the daytime minimum ($\Delta M_{\mathrm{diurn}}$) for both mass budget estimates M and M', and the core-sampled M_{ref}; and the root-mean-square error (RMSE), correlation coefficient R, and linear-regression coefficient β_{LM} for M and M' with respect to M_{ref}.

	11 Dec			12 Dec			15 Dec			20 Dec		
	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}
$\overline{M}$ (mm s^{-1})	11.0	11.4	11.8	11.1	11.7	11.4	13.8	15.1	12.7	11.7	11.8	11.4
σ_M (mm s^{-1})	8.4	7.1	5.0	3.8	7.3	5.2	8.2	9.9	8.3	2.2	4.9	3.8
$\Delta M_{\mathrm{diurn}}$ (mm s^{-1})	17.0	12.8	11.3	12.0	11.7	13.9	19.8	14.3	21.2	6.3	7.5	9.2
RMSE (−)	3.9	3.1	—	3.5	5.2	—	3.6	5.3	—	2.3	3.3	—
R (−)	0.96	0.92	—	0.73	0.68	—	0.91	0.87	—	0.81	0.73	—
β_{LM} (−)	1.63	1.29	—	0.54	0.92	—	0.90	1.02	—	0.47	0.93	—

hourly values of M and M_{ref} are quantified by the root-mean-square error (RMSE) and range between 2.3 mm s^{-1} on 20 December and 3.9 mm s^{-1} on 11 December. The RMSE is on average less than 30% of the mean. The correlation between M and M_{ref} ranges from 0.74 to 0.96.

As Figs. 3b and 4 show, the strength of the diurnal cycle of the mass flux differs among days. The difference in 4-h means between the nighttime maximum and the daytime minimum M_{ref} ($\Delta M_{\mathrm{diurn}}$) ranges from 9 mm s^{-1} (20 December) to 21 mm s^{-1} (15 December). M reproduces the magnitude of the core-sampled $\Delta M_{\mathrm{diurn}}$ well,
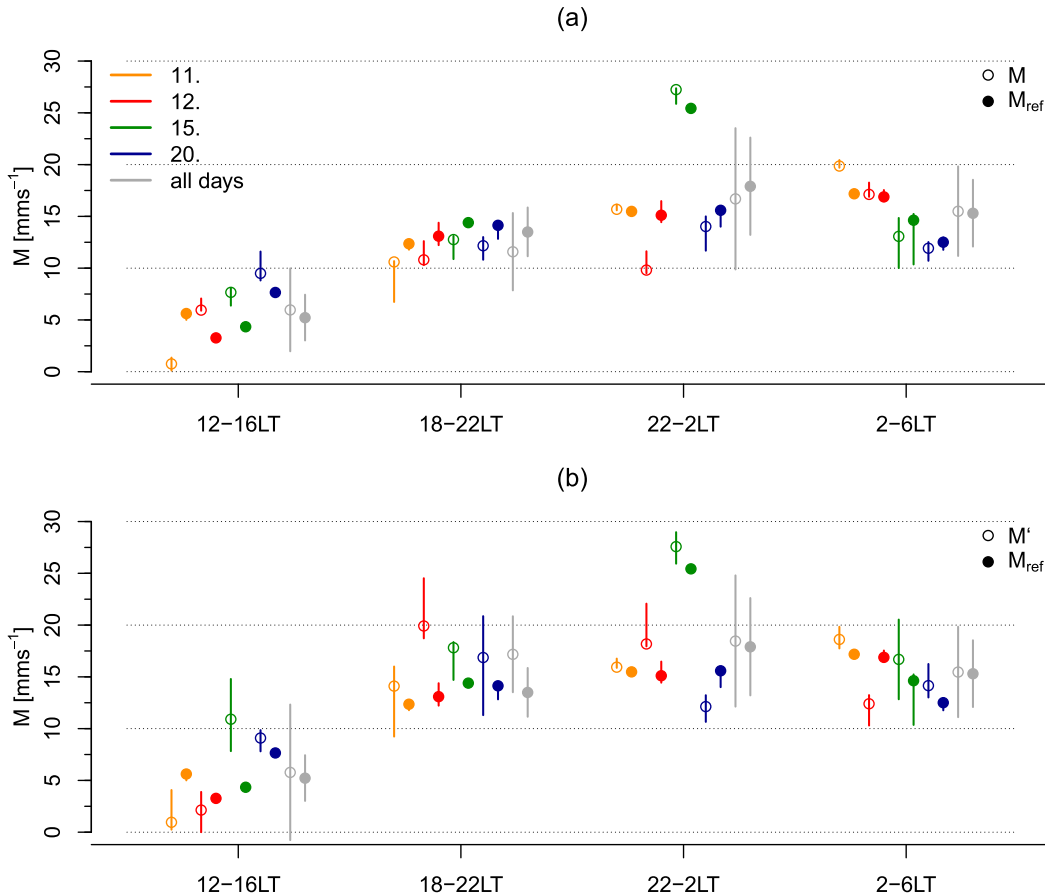

FIG. 4. Estimates of 4 h-mean (a) M and (b) M' at four different times of the day, together with the corresponding 4-h mean of M_{ref}. Shown are values for the 4 days individually, and for all days together representing the day-to-day mean and standard deviation (vertical lines). For the individual days, the vertical line represents the estimation uncertainty, defined as the minimum and maximum mass flux of a set of plausible parameter choices (see section 4b).
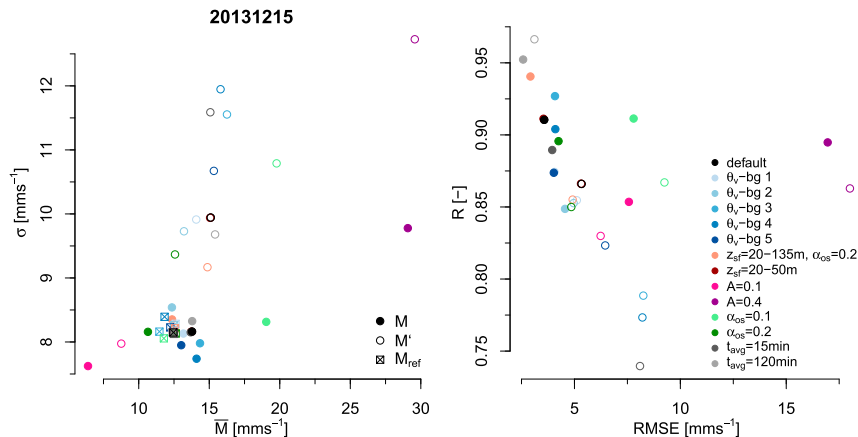
FIG. 5. (left) The daily mean and standard deviation of M, M', and M_{ref} and (right) the RMSE and the correlation of M and M' with respect to M_{ref}, for the different parameter choices explained in the text. Values are shown for 15 Dec.

except for an overestimation of ΔM_{diurn} on 11 December. The standard deviation of hourly values (σ) also indicates that the diurnal variability on 11 December is overestimated in M compared to M_{ref}, whereas σ is slightly underestimated on 12 and 20 December.

The day-to-day variability in M_{ref} is also well reproduced by M. Figure 4a shows that this is particularly true between 2200 and 0600 LT, when day-to-day differences are up to $10\,\mathrm{mm\,s^{-1}}$. The day-to-day σ shown in gray is slightly overestimated at all times.

The value of M generally agrees better with M_{ref} than M' concerning the daily mean and the day-to-day variability. However, the inclusion of the $\partial h/\partial t$ term in M' tends to improve the diurnal variability. This is evident in the linear-regression coefficients β_{LM} between M' and M_{ref} being closer to 1 compared to M on all days (see Table 2).

b. Robustness of the estimate and sensitivity to LES setup

Relating the conceptual model to ICON model output requires to make some hypotheses and subjective choices. Here we test the sensitivity of the results to these choices, so as to assess the robustness of the estimate. We consider in particular the specification of the background θ_v profile, the surface lifting height, the entrainment-to-surface buoyancy flux ratio A, the overshoot fraction α_{os}, and the temporal averaging.

Figure 5 shows that the mean, standard deviation σ, and RMSE of M on 15 December are largely insensitive to the number and the height of the levels used to fit the background profile to the cloud-layer θ_v (sensitivity tests referred to as "θ_v-bg 1–5" in Fig. 5). Only at times when the cloud-layer lapse rate is extremely variable can it impact the estimation. Also, the choice of the surface

lifting height z_{sf} and the averaging period hardly influence the skill of M.

Changing the entrainment ratio A to 0.4 doubles the entrainment rate E, which strongly increases the mean M and thus also the RMSE. Commensurate changes accompany a halving of A. The factor α_{os} directly influences h and for a constant entrainment ratio A, a higher h leads to a smaller M, because $\Delta\theta_v$ increases and E decreases. Depending on the narrowness of the M_{co} peak near cloud base, this can lead to large discrepancies with M_{ref} and thus a large RMSE, as is the case for $\alpha_{os} = 0.1$. Allowing for variability in α_{os} by computing it every time step using the subcloud-layer CAPE did not lead to improvements in M (not shown). Overall, both the changes in A and α_{os} therefore mostly influence the magnitude of M, but not its variability.

The sensitivity to the specification of the background θ_v-profile is larger for M' than for M, because changes in the background profile lead to changes in h and thus different magnitudes of $\partial h/\partial t$. A longer averaging period strongly improves both RMSE and correlation of M', as short-term fluctuations in h are smoothed. Similar results are found on the other days.

The different parameter choices can be used to calculate the estimation uncertainty of M and M'. In Fig. 4 we show this uncertainty by displaying the minimum and maximum M from the set of plausible parameter choices (excluding the cases with different A and α_{os}). As expected, the estimation uncertainty is slightly larger for M' compared to M. However, the uncertainty is not necessarily the largest when the difference between M and M_{ref} is largest: M_{ref} often lies outside the estimation uncertainty of M. The apparent underestimation of the uncertainty might be due to assuming homogeneous

TABLE 3. As in Table 2, but for the 313-m-resolution simulations for the same domain as the 156-m-resolution simulation (SD) and for a fourfold-larger domain (LD), and the minimum and maximum value of the four 313-m-resolution simulations with the small domain shifted to the east and north (ΔSD). Note that the vertical resolution of the simulations is the same here. Only 11 and 15 December 2013 are shown.

| | 11 Dec | | | | | | | | | 15 Dec | | | | | | | | |
| | SD | | | LD | | | ΔSD | | | SD | | | LD | | | ΔSD | | |
	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}	M	M'	M_{ref}
$\overline{M}$ (mm s^{-1})	9.9	10.6	11.8	10.4	9.6	11.6	11.3–16.0	9.9–15.9	11.2–14.7	13.3	14.6	12.1	10.5	11.6	11.5	7.5–9.8	9.7–13.1	7.6–11.0
σ (mm s^{-1})	8.1	6.8	5.9	6.7	6.0	5.5	6.7–7.9	6.9–8.7	5.3–6.9	7.8	10.4	8.8	3.8	6.0	6.3	4.1–4.8	8.9–10.3	3.8–7.2
ΔM_{diurn} (mm s^{-1})	12.6	8.5	12.4	16.5	15.6	13.7	7.1–21.1	14.1–19.6	13.3–18.2	18.5	13.3	21.7	8.6	9.4	16.7	−5.4–11.6	−9.0–19.2	4.4–17.8
RMSE (mm s^{-1})	3.7	3.6	—	2.8	2.5	—	1.5–4.8	4.0–5.1	—	3.9	5.8	—	3.6	4.1	—	4.5–6.5	8.7–10.7	—
R (−)	0.94	0.85	—	0.93	0.95	—	0.79–0.97	0.73–0.91	—	0.90	0.84	—	0.87	0.77	—	−0.04–0.72	0.17–0.38	—
β_{LM} (−)	1.30	0.97	—	1.14	1.05	—	0.85–1.19	0.82–1.27	—	0.81	0.98	—	0.52	0.74	—	−0.05–0.52	0.23–0.70	—

conditions and not representing the potential influence of mesoscale organization and penetrative downdrafts, due to neglecting the temporal fluctuation and horizontal advection of h, or due to keeping parameters like α_{os} fixed. Note that also M_{ref} varies slightly with the different choices due to the changing h.

To test the sensitivity of the estimation to the horizontal resolution and domain size of the LES, Table 3 shows results for the 313-m-resolution simulations both on the same domain as the 156-m-resolution simulation, and on a fourfold-larger domain spanning 59.4°–55.4°W and 12.1°–14.1°N. Comparison with Table 2 shows that the estimates are largely insensitive to halving the horizontal resolution from 156 to 313 m (at the same vertical resolution). The quadrupling of the domain size strongly improves the skill of M' as fluctuations in h are smoothed. Also, the differences between M and M_{ref} tend to reduce on a larger domain because of the smoother input data.

The influence of variability on the meso-β (20–200 km) and meso-α (200–2000 km) scales (Orlanski 1975) for the 313-m-resolution simulation is analyzed by shifting the small domain by (i) 1° north, (ii) 2° east, (iii) 6° east, and (iv) 6° east and 1° east. The columns ΔSD in Table 3 show that mesoscale variability between nearby domains is considerable. The mean and diurnal cycle of both M and M_{ref} vary strongly among the different domains, especially on 15 December. The mesoscale variability also affects the estimation skill, with RMSE ranging from 1.5 to 6.5 mm s^{-1} for M.

We also tested the estimation procedure with model output from the UCLA-LES for an idealized trade cumulus case (see Vogel et al. 2020). The UCLA-LES differs from the ICON-LEM simulations in that it has higher resolution (50-m horizontal grid spacing), different numerics and microphysics, and different boundary conditions. Despite these differences, we find that the estimation also works well for this simulation (not shown). M' shows even more skill than M in reproducing the variability of M_{ref} for the UCLA-LES, because temporal fluctuations in h are smaller due to the homogeneous boundary conditions. However, a ratio of the entrainment to surface buoyancy flux A of 0.4 is needed to get the correct magnitude of M, showing again that the choice of A can be critical for the mean magnitude of M.

c. Sources of mass flux variability

The simple estimation framework makes it possible to study the contribution of the different terms of the mass budget to the magnitude and variability of M. Figure 6b shows that on average across all days, $E = 14$ mm s^{-1} and $W = -2$ mm s^{-1}. The day-to-day variability in M is mostly explained by the variability in W, whereas E is
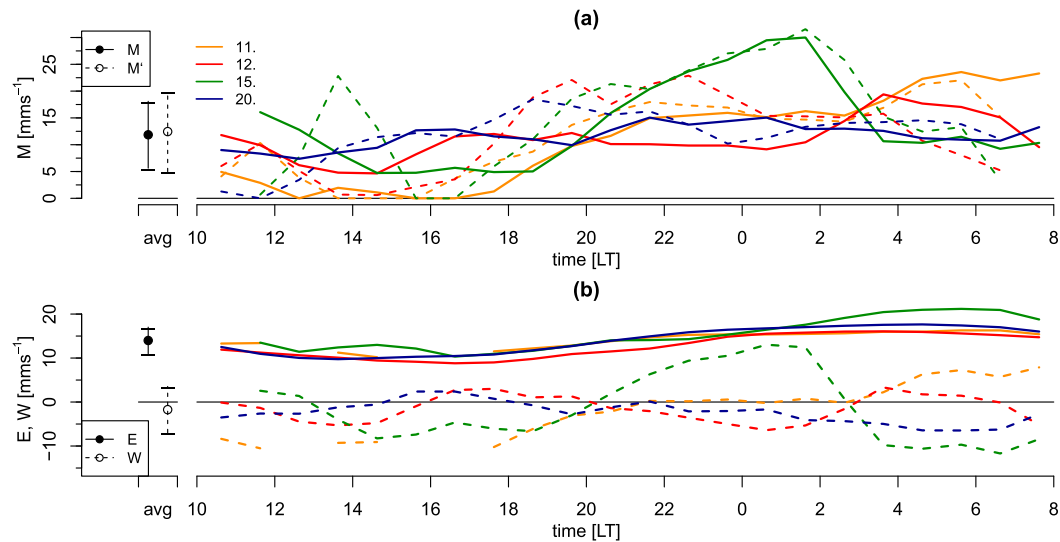
FIG. 6. Diurnal cycles of (a) M an M' and (b) E and W. The mean and standard deviation for all days and times are shown on the left side of each panel. The times when M and M' were negative and set to zero on 11 Dec are left out in (b).

very similar among the different days analyzed. E instead exhibits a pronounced diurnal variability. On all days, E has a minimum of about $10\,\mathrm{mm\,s^{-1}}$ in the afternoon and a maximum of about $18\,\mathrm{mm\,s^{-1}}$ at night. Averaging all days together, also W exhibits diurnal variability comparable to E, but with an $\sim2\,\mathrm{h}$ earlier minimum in the afternoon and an $\sim4\,\mathrm{h}$ earlier maximum at night. On the individual days, the diurnal variability in W tends to either strengthen the diurnal cycle in M (on 11 and 15 December) or weaken it (on 12 and 20 December).

Both the surface buoyancy flux and the inversion θ_v jump show a diurnal cycle, with enhanced buoyancy fluxes (mostly along with larger wind speeds) and reduced θ_v jumps during nighttime (not shown). The relative contributions of the two terms to the variability of E can be assessed by averaging one of the terms and only retain the variability of the other. The two terms are $E'_{\overline{\mathrm{SBF}}} = (A\,\overline{w'\theta'_v}|_s/\Delta\theta_v) - \overline{E}$ and $E'_{\overline{\Delta\theta_v}} = (A\,\overline{w'\theta'_v}|_s/\overline{\Delta\theta_v}) - \overline{E}$, where the prime denotes fluctuations from the daily mean. We find that both terms contribute about equally to the diurnal cycle in E on all days (not shown). The physical interpretation of this diurnal cycle is not yet settled. The $\Delta\theta_v$, for example, will depend on whether trade cumuli are mostly detraining near h, decreasing $\Delta\theta_v$, or farther aloft, thereby mixing down drier air and increasing $\Delta\theta_v$ (Neggers 2015). Vial et al. (2019) showed that the shallow and deep trade cumulus populations have distinct diurnal cycles, which likely influence the diurnal cycle of $\Delta\theta_v$. The diurnal cycle in $\Delta\theta_v$ is also strongly correlated to the diurnal cycle in h (see Fig. 3a). As the SST shows no diurnal cycle in the trade cumulus regime due to the

strong surface wind (Vial et al. 2019), and is fixed in the LES used here, it is likely that radiative mechanisms drive the diurnal cycle of E and trade wind cloudiness. The detailed physical mechanisms underlying this diurnal cycle, including the coupling of convection with surface winds, will be a major focus of the EUREC^{4}A campaign.

Hourly values of $\partial h/\partial t$ can reach $\pm15\,\mathrm{mm\,s^{-1}}$, and can change by the same amount between subsequent hours. No diurnal cycle of $\partial h/\partial t$ is evident, but both the magnitude and variability of $\partial h/\partial t$ tend to be smaller at night (not shown).

The diurnal cycle in E is insensitive to halving the resolution and quadrupling the domain size. Also, the variability in W continues to explain most of the (somewhat lower) day-to-day variability of M on a fourfold-larger domain.

5. Application to NARVAL2 observations

This section presents the first estimates of M from observations made during the NARVAL2 field campaign, which took place in the same area as NARVAL1 near Barbados, but in summer (Stevens et al. 2019; Konow et al. 2019). The observations used are derived from dropsondes released at high frequency during research flights RF03 and RF06 on 12 and 19 August 2016. During these flights, the German High Altitude and Long Range Research Aircraft (HALO) flew pairs of horizontal circles with a diameter of about $170\,\mathrm{km}$ at 9-km height and released 12 dropsondes per circle, one sonde every 4 min. Bony and Stevens (2019)
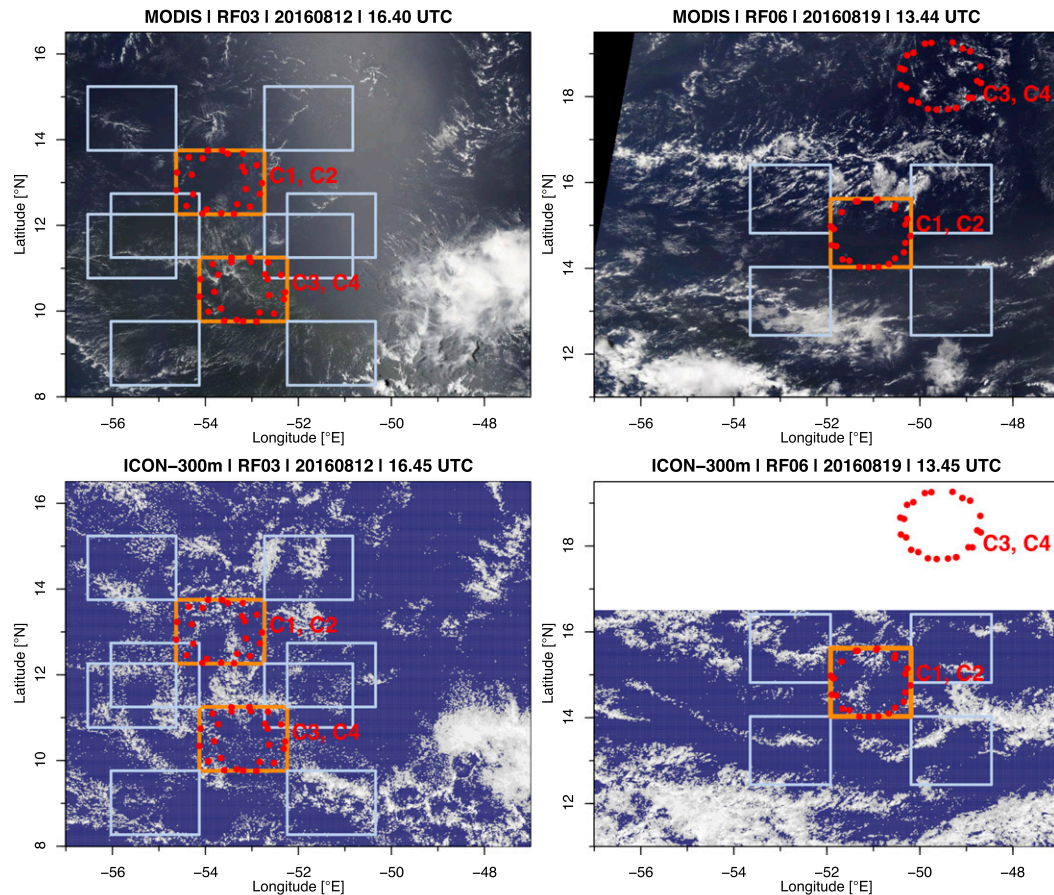
FIG. 7. (top) MODIS Worldview visible images for RF03 and RF06 of NARVAL2 and (bottom) the corresponding liquid water path fields of the 313-m-resolution ICON simulation. Also indicated are the dropsonde circles flown (red dots and labels indicating the circle numbers) and the ICON domains used (orange rectangles for the centered domain and light blue rectangles for the four additional domains at the corners). Note that the ICON-LEM simulations only cover the region up to 16.5°N.

demonstrated that the large-scale vertical velocity can be robustly computed by vertically integrating the divergence of the horizontal wind field inferred from the dropsondes.

The dropsonde data are compared to 313-m-resolution ICON-LEM simulations performed for the respective days and averaged over a domain centered on the circle pairs. The comparison allows to test whether the ICON-LEM simulations reproduce the observed conditions, and to study the mesoscale variability by performing the analyses for four additional domains at the corners of the centered domain. Figure 7 shows visible MODIS images and corresponding liquid-water path fields from the LES for RF03 and RF06. The structure of the simulated cloud fields agrees well with the observed cloud fields. For the first circle pair of RF06 (C1 + 2), the profiles in Fig. 8 show that the trade wind layer in the centered ICON domain (in blue) is about 0.5–1 K colder compared to the dropsondes, with a slightly deeper and

drier subcloud layer, and a weaker subsidence in the lowest kilometer. For the two circle pairs of RF03, the simulated trade wind layer is also about 1 K colder and $1 \, \text{g} \, \text{kg}^{-1}$ drier, but subsidence is stronger compared to the dropsondes (not shown). Because the ICON simulations are only nudged to IFS at the boundaries of the outermost grid, the simulations diverge from the observations and we do not expect exact correspondence between the centered ICON domain and the dropsondes.

Both the θ_v and the large-scale vertical velocity profile measured by the dropsondes at 30 m vertical resolution are used for the mass budget estimation. Because of the lack of surface flux observations during NARVAL2, we compute E using the surface buoyancy flux from the ECMWF ERA5 reanalysis at hourly temporal resolution, averaged over a domain encompassing the circle pair. The surface buoyancy fluxes from ERA5 differ from the centered ICON domain by less than ±10% (not shown).
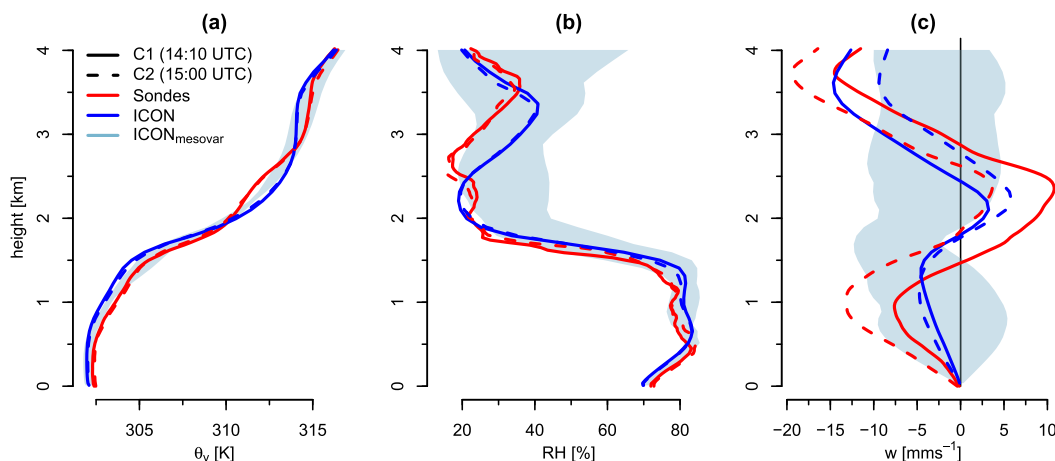
FIG. 8. Vertical profiles of (a) θ_v, (b) relative humidity, and (c) large-scale vertical velocity for circles C1 and C2 of RF06. Shown are averages for the 12 dropsondes released per circle (red) and for the 313-m-resolution ICON simulation averaged over the flight hours for the domain encompassing the circle pair (blue). Also shown are the minima and maxima of the four additional domains at the corners of the centered domain (light blue shading).

Figure 9a shows that the sonde-derived h for the four circle pairs lies between 670 and 850 m, which tends to be within ± 60 m of the respective LCL. M estimated from the sondes ranges from 4 to 17 mm s^{-1}, whereby the largest M for RF06 C3 + 4 is caused by the large positive W of this circle pair (Figs. 9b,c). Again, most of the variability[1] in M among the circle pairs is due to variability in W, whereas E is less variable.

The vertical lines in Fig. 9 represent the estimation uncertainty. RF03 C1 + 2 has a particularly large estimation uncertainty of the sonde-derived h, E, and thus M, which is mainly due to the very shallow cloud layer of about 300 m depth that complicates the estimation (not shown). For RF03 C1 + 2, we also find larger discrepancies between the M derived from the sondes and the centered ICON domain. However, the estimates of the additional ICON domains at the corners of the centered domain (light blue; see Fig. 7 for their location) also indicate that there is substantial variability on the meso-β scale (Orlanski 1975), with the sonde-derived estimates lying within the range of the ICON estimates. The mesoscale variability is mostly reflected in variability in W among the different domains.

Overall, this section shows that—without any adjustments or tuning of the method—we get reasonable mass flux estimates when applying our mass budget estimation method to the dropsonde data of NARVAL2. The somewhat larger discrepancies between M and M_{ref} for the ICON simulations are likely due to the more variable conditions of the summer trades, and also motivate

an analysis of the influence of the mesoscale organization pattern (Stevens et al. 2020) on the estimation skill.

During EUREC^{4}A, the robustness of the mass flux estimation will be facilitated by a number of factors: namely, (i) the steadier conditions of the winter trades, (ii) the coincident measurement of the buoyancy flux at the surface and in the subcloud layer, (iii) the increase in the number of circles flown, and (iv) additional measurements from a suite of other platforms and instruments (see Bony et al. 2017). The larger number of circles flown will allow for a longer averaging time and a more robust estimation of W, E, and particularly h, which may also allow to account for the temporal fluctuations and the horizontal advection of h. Coincident lidar observations of the humidity jump near cloud base might further help mapping the spatial variability of h and thus allow a robust estimation of the advection term. Analysis of the budgets of moisture and heat, as well as other tracers like isotopes, will yield independent estimates of E that will help constrain the magnitude of A. Also, the measurements of the buoyancy flux in the subcloud layer will help evaluate the optimal value of A, as well as the extent to which A should be allowed to vary to balance the budgets of mass, heat, and moisture.

6. Summary and conclusions

This paper develops a method to estimate the shallow convective mass flux M at the top of the subcloud layer as a residual of the subcloud-layer mass budget. The goal of the method is to be applied to observations, but a prerequisite is that the ability of M to reproduce the mass flux diagnosed directly from the cloud-core area

---

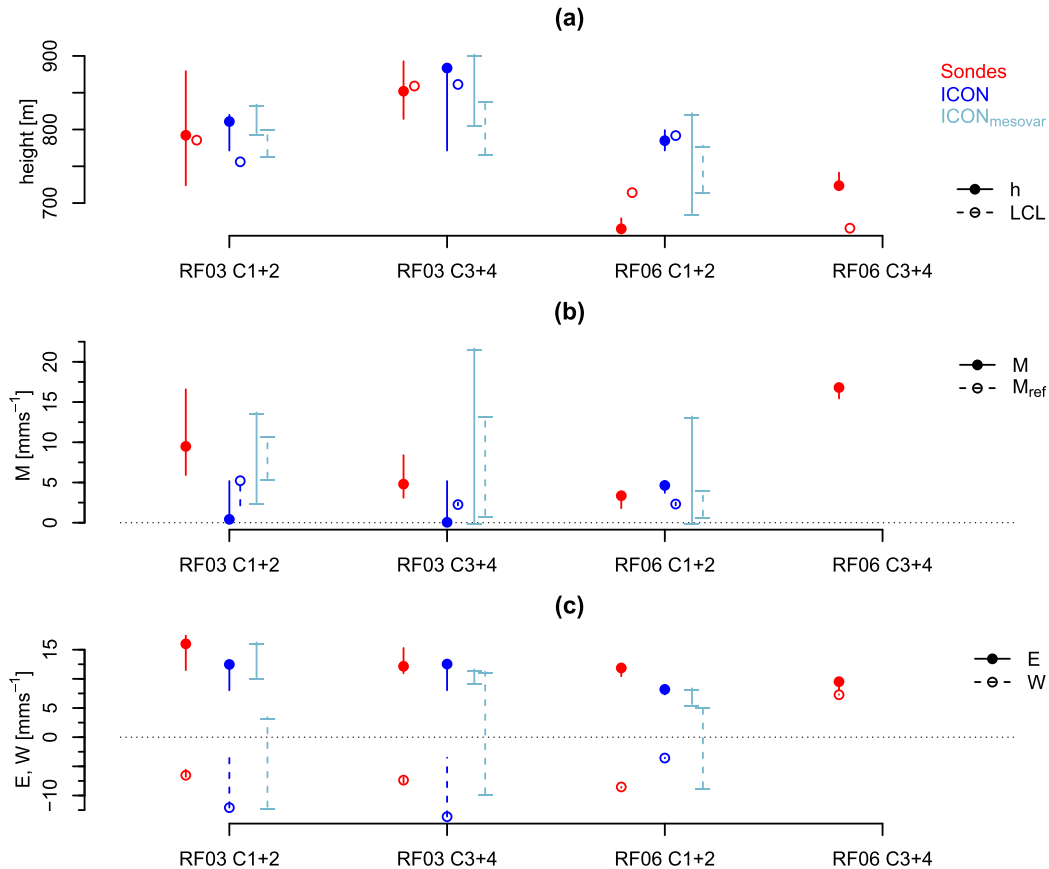[1] Containing spatial, day-to-day, and also some diurnal variability.

FIG. 9. Estimates of (a) h and LCL, (b) M and M_{ref}, and (c) E and W for the 4 circle pairs of RF03 and RF06 during NARVAL2 in August 2016 (see Fig. 7 for the location of the different circle pairs). Shown are the dropsondes (red) and the 313-m-resolution ICON simulations (except for RF06 C3 + 4), with both the domain centered on the flight area (blue) and the range of mean values from the four additional domains at the corners of the centered domain (light blue). For the sondes and the centered ICON domain, the estimation uncertainty is represented by the vertical line showing the minimum and maximum values of the set of plausible parameter choices.

fraction and vertical velocity is tested using large-eddy simulation (LES).

We find that the magnitude, diurnal cycle, and day-to-day variability of the core-sampled mass flux are very well reproduced by the mass budget estimate. Omitting both temporal fluctuations of the subcloud-layer depth h and periods of low subcloud-layer convective available potential energy (due to substantial precipitation), the average root-mean-square error is less than 30% of the mean. This is about 3–6 times smaller than the diurnal cycle in M, and 1.5–3 times smaller than the day-to-day variability in M at different times of the day. When the contributions of the temporal fluctuations of h are accounted for, the diurnal cycle tends to be further improved. However, short-term variability in h can be large in the absence of sufficiently long temporal (longer than ~2 h) and spatial (larger than ~800 km^2) averaging, which limits the reliability of the estimate. The mass budget estimation also shows promising results when

applied to dropsonde observations from the NARVAL2 field campaign.

Our equilibrium estimate M in the LES is largely insensitive to halving the resolution and quadrupling the domain size. The various parameter choices mostly affect the magnitude of M, but not its variability. As the magnitude of the core-sampled mass flux anyway depends on the cloud-core definition, and because the focus of the EUREC^{4}A campaign lies mostly in relating the variability of M to the variability in cloudiness, some uncertainty in the mean magnitude of M is acceptable. The simple estimation framework still underestimates some of the variability in M, which is likely due to omitting the potential influence of mesoscale organization and downdrafts, neglecting the horizontal advection and temporal fluctuation of h, or due to keeping parameters like the overshoot fraction fixed. The robustness of the mass budget estimate for the EUREC^{4}A campaign will benefit from the expected diversity of conditions to be

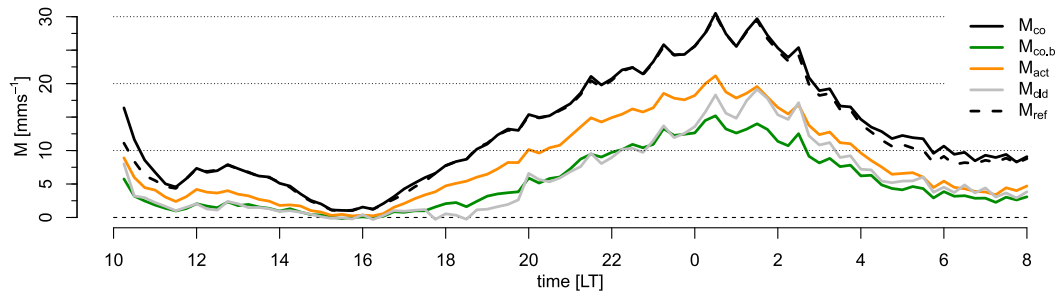FIG. A1. Time series of the different core-sampled mass flux definitions at the level of maximum cloud-base cloudiness for 15 Dec. Also shown is M_{ref}, which is the value of M_{co} at h.

encountered over a month with intense sampling, and the suite of additional measurements planned, which will help close the heat and moisture budgets, and yield independent estimates of the entrainment rate E.

The simple estimation framework allows us to analyze the sources of variability in M. Across the four winter days analyzed, the average entrainment rate E is $14\,\mathrm{mm\,s^{-1}}$ and the average large-scale vertical velocity W is $-2\,\mathrm{mm\,s^{-1}}$, yielding an average M of $12\,\mathrm{mm\,s^{-1}}$. Day-to-day variations in M are mostly due to variations in W; E is very similar among the different days analyzed, albeit with a pronounced diurnal cycle that is due to a diurnal cycle in both the surface buoyancy flux and the inversion θ_v jump. Also, W exhibits substantial variability on the diurnal time scale. The observations support the result that most of the variability in M is due to the variability in W.

The physical mechanisms underlying the diurnal cycle in E, as well as the seemingly strong coupling between the variability in shallow convective mixing (as represented by M) and the large-scale circulation (represented by W) merit further investigations. From the analysis presented here, it is difficult to establish causality between M and W. However, Bony and Stevens (2019) show that the autocorrelation time scale of the horizontal divergence (and thus W) on the scale of 200 km is more than 8 h in the winter trades. This is much larger than the typical cumulus time scale of 15–30 min and thus representative of the large-scale environment, suggesting that M most likely responds to W. However, M can also influence W to some degree, for example, by affecting the stability and the spatial and vertical distribution of radiative cooling.

Assuming that changes in the core-sampled mass flux are mostly associated with changes in the cloud-core area fraction rather than the vertical velocity within clouds (Sakradzija and Hohenegger 2017), then the observational estimates of M suggest a threefold to fourfold difference in cloudiness depending on M. Such a large difference seems plausible based on satellite imagery, and is testable given the increased emphasis on simultaneously

measuring the large-scale vertical velocity and cloud amount in the upcoming EUREC[4]A field campaign.

APPENDIX

Definition of M_{co}

Different definitions for the core part of a cloud have been used in the past (see, e.g., Stull 1985; Siebesma and Cuijpers 1995). Here we compare our M_{co}, defined as the mass flux of all cloudy grid points with positive vertical velocity, to the following other definitions:

- $M_{co.b}$: mass flux of all cloudy grid points that are positively buoyant with respect to the slab average
- M_{act}: mass flux of all cloudy grid points that have positive vertical velocity and are positively buoyant with respect to the slab average (i.e., active clouds)
- M_{cld}: mass flux of all cloudy grid points

Because the θ_v field has a lot of spatial variability due to the realistic LES setup, we also computed the positive

buoyancy with respect to the mean θ_v in a 10 km radius. We only find little influence of the θ_v averaging scale on the resulting $M_{\text{co.b}}$. Similarly, we find no difference if we condition on the vertical velocity being greater than 0 or greater than the (domain mean) large-scale vertical velocity.

Figure A1 shows that the core definition strongly influences the magnitude of the mass flux. Our M_{co} has the largest mass flux, and is more than twice as large as $M_{\text{co.b}}$. This is both because of vertical velocities being twice as large and also the area fraction being slightly larger for M_{co} compared to $M_{\text{co.b}}$ (not shown). Regarding M_{act}, the ~50% larger vertical velocity due to the more restrictive core definition compared to M_{co} is over-compensated by the more than 50% decrease in area fraction. On the contrary, the total cloudy mass flux M_{cld} including the downdraft parts of the cloud is less than half of M_{co} because of an 80% reduction in the vertical velocity despite of a doubling of the area fraction.

Despite the differences in the magnitude, the diurnal cycle of the mass flux is very similar among the different definitions used. Similarly, the core definition hardly influences the day-to-day variability of the mass flux (not shown). Figure A1 also shows that M_{ref}, which is the value of M_{co} at h, is generally very close to the maximum cloud-base M_{co}. The same is true for the other days analyzed.

REFERENCES

Ahlgrimm, M., and D. A. Randall, 2006: Diagnosing monthly mean boundary layer properties from reanalysis data using a bulk boundary layer model. *J. Atmos. Sci.*, **63**, 998–1012, https://doi.org/10.1175/JAS3669.1.

Albrecht, B., M. Fang, and V. Ghate, 2016: Exploring stratocumulus cloud-top entrainment processes and parameterizations by using Doppler cloud radar observations. *J. Atmos. Sci.*, **73**, 729–742, https://doi.org/10.1175/JAS-D-15-0147.1.

Betts, A. K., 1976: Modeling subcloud layer structure and interaction with a shallow cumulus layer. *J. Atmos. Sci.*, **33**, 2363–2382, https://doi.org/10.1175/1520-0469(1976)033<2363:MSLSAI>2.0.CO;2.

Blossey, P. N., and Coauthors, 2013: Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison. *J. Adv. Model. Earth Syst.*, **5**, 234–258, https://doi.org/10.1002/JAME.20025.

Bolton, D., 1980: The computation of equivalent potential temperature. *Mon. Wea. Rev.*, **108**, 1046–1053, https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**, L20806, https://doi.org/10.1029/2005GL023851.

——, and B. Stevens, 2019: Measuring area-averaged vertical motions with dropsondes. *J. Atmos. Sci.*, **76**, 767–783, https://doi.org/10.1175/JAS-D-18-0141.1.

——, J.-L. Dufresne, H. Le Treut, J.-J. Morcrette, and C. Senior, 2004: On dynamic and thermodynamic components of cloud changes. *Climate Dyn.*, **22**, 71–86, https://doi.org/10.1007/s00382-003-0369-6.

——, and Coauthors, 2017: EUREC[4]A: A field campaign to elucidate the couplings between clouds, convection and circulation. *Surv. Geophys.*, **38**, 1529–1568, https://doi.org/10.1007/s10712-017-9428-0.

Bretherton, C. S., P. N. Blossey, and C. R. Jones, 2013: Mechanisms of marine low cloud sensitivity to idealized climate perturbations: A single-LES exploration extending the CGILS cases. *J. Adv. Model. Earth Syst.*, **5**, 316–337, https://doi.org/10.1002/jame.20019.

Brient, F., T. Schneider, Z. Tan, S. Bony, X. Qu, and A. Hall, 2016: Shallowness of tropical low clouds as a predictor of climate models' response to warming. *Climate Dyn.*, **47**, 433–449, https://doi.org/10.1007/s00382-015-2846-0.

Byers, H. R., and R. K. Hall, 1955: A census of cumulus-cloud height versus precipitation in the vicinity of Puerto Rico during the winter and spring of 1953–1954. *J. Atmos. Sci.*, **12**, 176–178, https://doi.org/10.1175/1520-0469(1955)012<0176:ACOCCH>2.0.CO;2.

Canut, G., F. Couvreux, M. Lothon, D. Pino, and F. Saïd, 2012: Observations and large-eddy simulations of entrainment in the sheared Sahelian boundary layer. *Bound.-Layer Meteor.*, **142**, 79–101, https://doi.org/10.1007/s10546-011-9661-x.

Deardorff, J. W., G. E. Willis, and D. K. Lilly, 1974: Comment on the paper by A. K. Betts 'Non-precipitating cumulus convection and its parameterization.' *Quart. J. Roy. Meteor. Soc.*, **100**, 122–123, https://doi.org/10.1002/QJ.49710042311.

Dipankar, A., B. Stevens, R. Heinze, C. Moseley, G. Zängl, M. Giorgetta, and S. Brdar, 2015: Large eddy simulation using the general circulation model icon. *J. Adv. Model. Earth Syst.*, **7**, 963–986, https://doi.org/10.1002/2015MS000431.

Garcia, J. R., and J. P. Mellado, 2014: The two-layer structure of the entrainment zone in the convective boundary layer. *J. Atmos. Sci.*, **71**, 1935–1955, https://doi.org/10.1175/JAS-D-13-0148.1.

Ghate, V. P., M. A. Miller, and L. DiPretore, 2011: Vertical velocity structure of marine boundary layer trade wind cumulus clouds. *J. Geophys. Res.*, **116**, D16206, https://doi.org/10.1029/2010JD015344.

——, D. B. Mechem, M. P. Cadeddu, E. W. Eloranta, M. P. Jensen, M. L. Nordeen, and W. L. Smith Jr., 2019: Estimates of entrainment in closed cellular marine stratocumulus clouds from the MAGIC field campaign. *Quart. J. Roy. Meteor. Soc.*, **145**, 1589–1602, https://doi.org/10.1002/QJ.3514.

Heinze, R., and Coauthors, 2017: Large-eddy simulations over Germany using ICON: A comprehensive evaluation. *Quart. J. Roy. Meteor. Soc.*, **143**, 69–100, https://doi.org/10.1002/QJ.2947.

Hourdin, F., F. Couvreux, and L. Menut, 2002: Parameterization of the dry convective boundary layer based on a mass flux representation of thermals. *J. Atmos. Sci.*, **59**, 1105–1123, https://doi.org/10.1175/1520-0469(2002)059<1105:POTDCB>2.0.CO;2.

Klocke, D., M. Brueck, C. Hohenegger, and B. Stevens, 2017: Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nat. Geosci.*, **10**, 891–896, https://doi.org/10.1038/s41561-017-0005-4.

Kollias, P., and B. Albrecht, 2010: Vertical velocity statistics in fair-weather cumuli at the ARM TWP Nauru climate research facility. *J. Climate*, **23**, 6590–6604, https://doi.org/10.1175/2010JCLI3449.1.

Konow, H., and Coauthors, 2019: A unified data set of airborne cloud remote sensing using the HALO Microwave Package (HAMP). *Earth Syst. Sci. Data*, **11**, 921–934, https://doi.org/10.5194/ESSD-11-921-2019.

Lamer, K., P. Kollias, and L. Nuijens, 2015: Observations of the variability of shallow trade wind cumulus cloudiness and mass flux. *J. Geophys. Res. Atmos.*, **120**, 6161–6178, https://doi.org/10.1002/2014JD022950.

Lilly, D. K., 1968: Models of cloud-topped mixed layers under a strong inversion. *Quart. J. Roy. Meteor. Soc.*, **94**, 292–309, https://doi.org/10.1002/qj.49709440106.

Neggers, R. A. J., 2015: Exploring bin-macrophysics models for moist convective transport and clouds. *J. Adv. Model. Earth Syst.*, **7**, 2079–2104, https://doi.org/10.1002/2015MS000502.

——, B. Stevens, and J. D. Neelin, 2006: A simple equilibrium model for shallow-cumulus-topped mixed layers. *Theor. Comput. Fluid Dyn.*, **20**, 305–322, https://doi.org/10.1007/s00162-006-0030-1.

Nuijens, L., I. Serikov, L. Hirsch, K. Lonitz, and B. Stevens, 2014: The distribution and variability of low-level cloud in the North Atlantic trades. *Quart. J. Roy. Meteor. Soc.*, **140**, 2364–2374, https://doi.org/10.1002/QJ.2307.

——, B. Medeiros, I. Sandu, and M. Ahlgrimm, 2015: Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer. *J. Adv. Model. Earth Syst.*, **7**, 1741–1764, https://doi.org/10.1002/2015MS000483.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, **56**, 527–530, https://doi.org/10.1175/1520-0477-56.5.527.

Rauber, R. M., and Coauthors, 2007: Rain in shallow cumulus over the ocean: The RICO campaign. *Bull. Amer. Meteor. Soc.*, **88**, 1912–1928, https://doi.org/10.1175/BAMS-88-12-1912.

Rio, C., and F. Hourdin, 2008: A thermal plume model for the convective boundary layer: Representation of cumulus clouds. *J. Atmos. Sci.*, **65**, 407–425, https://doi.org/10.1175/2007JAS2256.1.

Sakradzija, M., and C. Hohenegger, 2017: What determines the distribution of shallow convective mass flux through a cloud base? *J. Atmos. Sci.*, **74**, 2615–2632, https://doi.org/10.1175/JAS-D-16-0326.1.

Seifert, A., and K. D. Beheng, 2006: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteor. Atmos. Phys.*, **92**, 45–66, https://doi.org/10.1007/S00703-005-0112-4.

Sherwood, S. C., S. Bony, and J.-L. Dufresne, 2014: Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, **505**, 37–42, https://doi.org/10.1038/nature12829.

Siebesma, A. P., and J. W. M. Cuijpers, 1995: Evaluation of parametric assumptions for shallow cumulus convection. *J. Atmos. Sci.*, **52**, 650–666, https://doi.org/10.1175/1520-0469(1995)052<0650:EOPAFS>2.0.CO;2.

——, and Coauthors, 2003: A large eddy simulation intercomparison study of shallow cumulus convection. *J. Atmos. Sci.*, **60**, 1201–1219, https://doi.org/10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2.

Stevens, B., 2006: Bulk boundary-layer concepts for simplified models of tropical dynamics. *Theor. Comput. Fluid Dyn.*, **20**, 279–304, https://doi.org/10.1007/s00162-006-0032-z.

——, and Coauthors, 2003: On entrainment rates in nocturnal marine stratocumulus. *Quart. J. Roy. Meteor. Soc.*, **129**, 3469–3493, https://doi.org/10.1256/qj.02.202.

——, and Coauthors, 2005: Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Mon. Wea. Rev.*, **133**, 1443–1462, https://doi.org/10.1175/MWR2930.1.

——, and Coauthors, 2019: A high-altitude long-range aircraft configured as a cloud observatory: The NARVAL expeditions. *Bull. Amer. Meteor. Soc.*, **100**, 1061–1077, https://doi.org/10.1175/BAMS-D-18-0198.1.

——, and Coauthors, 2020: Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds. *Quart. J. Roy. Meteor. Soc.*, **146**, 141–152, https://doi.org/10.1002/QJ.3662.

Stull, R. B., 1976: The energetics of entrainment across a density interface. *J. Atmos. Sci.*, **33**, 1260–1267, https://doi.org/10.1175/1520-0469(1976)033<1260:TEOEAD>2.0.CO;2.

——, 1985: A fair-weather cumulus cloud classification scheme for mixed-layer studies. *J. Climate Appl. Meteor.*, **24**, 49–56, https://doi.org/10.1175/1520-0450(1985)024<0049:AFWCCC>2.0.CO;2.

——, 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic, 666 pp.

Tennekes, H., 1973: A model for the dynamics of the inversion above a convective boundary layer. *J. Atmos. Sci.*, **30**, 558–567, https://doi.org/10.1175/1520-0469(1973)030<0558:AMFTDO>2.0.CO;2.

vanZanten, M. C., P. G. Duynkerke, and J. W. M. Cuijpers, 1999: Entrainment parameterization in convective boundary layers. *J. Atmos. Sci.*, **56**, 813–828, https://doi.org/10.1175/1520-0469(1999)056<0813:EPICBL>2.0.CO;2.

——, and Coauthors, 2011: Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *J. Adv. Model. Earth Syst.*, **3**, M06001, https://doi.org/10.1029/2011MS000056.

Vial, J., J. L. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dyn.*, **41**, 3339–3362, https://doi.org/10.1007/s00382-013-1725-9.

——, S. Bony, J.-L. Dufresne, and R. Roehrig, 2016: Coupling between lower-tropospheric convective mixing and low-level clouds: Physical mechanisms and dependence on convection scheme. *J. Adv. Model. Earth Syst.*, **8**, 1892–1911, https://doi.org/10.1002/2016MS000740.

——, and Coauthors, 2019: A new look at the daily cycle of trade wind cumuli. *J. Adv. Model. Earth Syst.*, **11**, 3148–3166, https://doi.org/10.1029/2019MS001746.

Vogel, R., L. Nuijens, and B. Stevens, 2016: The role of precipitation and spatial organization in the response of trade-wind clouds to warming. *J. Adv. Model. Earth Syst.*, **8**, 843–862, https://doi.org/10.1002/2015MS000568.

——, ——, and ——, 2020: Influence of deepening and mesoscale organization of shallow convection on stratiform cloudiness in the downstream trades. *Quart. J. Roy. Meteor. Soc.*, **146**, 174–185, https://doi.org/10.1002/qj.3664.

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf, 2015: The ICON (Icosahedral Non-Hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quart. J. Roy. Meteor. Soc.*, **141**, 563–579, https://doi.org/10.1002/qj.2378.

Zhao, M., 2014: An investigation of the connections among convection, clouds, and climate sensitivity in a global climate model. *J. Climate*, **27**, 1845–1862, https://doi.org/10.1175/JCLI-D-13-00145.1.

Zuidema, P., and Coauthors, 2012: On trade wind cumulus cold pools. *J. Atmos. Sci.*, **69**, 258–280, https://doi.org/10.1175/JAS-D-11-0143.1.