University of Reading

The representation of the West-African Monsoon vertical cloud structure in the Met Office Unified Model: an evaluation with CloudSat

Article

Accepted Version

Open Access

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

The representation of the West-African Monsoon vertical cloud structure in the Met Office Unified Model: An evaluation with CloudSat

Thorwald H.M. Stein (t.h.m.stein@reading.ac.uk)
Douglas J. Parker,   Robin J. Hogan,   Cathryn Birch,
Christopher E. Holloway,   Grenville Lister,
John H. Marsham,   and Steven J. Woolnough

June 24, 2015

Abstract

Weather and climate model simulations of the West African Monsoon (WAM) have generally poor representation of the rainfall distribution and monsoon circulation because key processes, such as clouds and convection, are poorly characterized. The vertical distribution of cloud and precipitation during the WAM are evaluated in Met Office Unified Model simulations against CloudSat observations. Simulations were run at 40-km and 12-km horizontal grid length using a convection parameterization scheme and at 12-km, 4-km, and 1.5-km grid length with the convection scheme effectively switched off, to study the impact of model resolution and convection parameterization scheme on the organisation of tropical convection. Radar reflectivity is forward-modelled from the model cloud fields using the CloudSat simulator to present a like-with-like comparison with the CloudSat radar observations. The representation of cloud and precipitation at 12-km horizontal grid length improves dramatically when the convection parameterization is switched off, primarily because of a reduction in daytime (moist) convection. Further improvement is obtained when reducing model grid length to 4 km or 1.5 km, especially in the representation of thin anvil and mid-level cloud, but three issues remain in all model configurations. Firstly, all simulations underestimate the fraction of anvils with cloud top height above 12 km, which can be attributed to too low ice water contents in the model compared to satellite retrievals. Secondly, the model consistently detrains mid-level cloud too close to the freezing level, compared to higher altitudes in CloudSat observations. Finally, there is too much low-level cloud cover in all simulations and this bias was not improved when adjusting the rainfall parameters in the microphysics scheme. To improve model simulations of the WAM, more detailed and in-situ observations of the dynamics and microphysics targeting these non-precipitating cloud types are required.

1

1 Introduction

Global climate models (GCMs) show large uncertainties in the radiative impact of clouds (Jakob, 2002) as cloud processes are often poorly represented in parameterization schemes (Randall *et al.*, 2003; Stevens and Bony, 2013). Yet for many GCMs, little or no improvement in the representation of cloud amount with height has been made in recent years (Klein *et al.*, 2013). It is therefore vital to establish suitable measures to evaluate GCMs as well as numerical weather prediction (NWP) models in terms of the location and depth of clouds, and to test whether microphysics parameterization schemes are able to capture the variations in hydrometeor type and distribution throughout the cloud depth. With the global coverage from CloudSat and CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations), cloud vertical profiles can be obtained all around the globe and are now used to evaluate models (e.g., Bodas-Salcedo *et al.* (2008)), providing a great advance in model evaluation, especially in poorly observed regions.

The West African Monsoon (WAM) is a weather and climate phenomenon that influences the global circulation, yet low confidence persists in future rainfall projections for West Africa (Christensen and Kanikicharla, 2013). Studies and observational campaigns carried out under the AMMA programme (African Monsoon Multidisciplinary Analysis, Redelsperger *et al.* (2006)) have greatly advanced the understanding of the WAM. The AMMA-model intercomparison project (Hourdin *et al.*, 2010) indicates that although GCMs adequately simulate the main characteristics of the WAM, the models studied showed large variations in accumulated rainfall over the Sahel and the location of the African Easterly Jet (AEJ), concluding that the resulting rainfall was highly dependent on the choice of convective parameterization. Waves on the AEJ (African Easterly Waves, AEW) are associated with rainfall variability in the Sahel (e.g. Newell and Kidson (1984); Thorncroft *et al.* (2003)), and the position of the AEJ has been shown to correlate with the averaged latitudinal position of convective systems in satellite observations (Mohr and Thorncroft, 2006; Stein *et al.*, 2011a). The WAM in many ways is a "natural laboratory" for tropical clouds, illustrated by the prominence of low-level and mid-level clouds as well as deep-convective systems (Stein *et al.*, 2011a; Bouniol *et al.*, 2012). A comprehensive overview by Roehrig *et al.* (2013) of the state of the WAM in simulations from the Coupled Model Intercomparison Project phase 5 (CMIP5) highlighted three cloud features that GCMs struggle with, namely the vertical extent of deep convection, the amount and occurrence of mid-level cloud over the Sahara, and the depth and occurrence of stratus over the Gulf of Guinea.

In this study, we focus on these cloud features to analyse the vertical structure of clouds and precipitation over West Africa in the Met Office Unified Model (MetUM) using a set of simulations which were run as part of the Cascade project (Pearson *et al.*, 2010). A major focus of Cascade was to study model ability to represent clouds and convection at varying time- and length-scales as horizontal resolution was increased, including the effect of running with or without a convection parameterization scheme. Pearson *et al.* (2010) and Pearson

et al. (2014) evaluated the Cascade simulations at 12-km, 4-km, and 1.5-km horizontal grid length against observations from GERB (Geostationary Earth Radiation Budget), analysing the diurnal cycle of the size distribution of clusters of outgoing longwave radiation (OLR). Their results showed that when using a convection parameterization scheme, the cloud and precipitation occurrences peaked in the early afternoon at all cluster-size scales, whereas GERB observations indicated a gradual shift from smaller OLR clusters in the early afternoon that lasted until the evening, towards larger clusters peaking in the evening and lasting until the early morning. They showed that at 12-km grid length, simulations without a parameterization scheme simulated the diurnal cycle well and — depending on choice of sub-grid mixing scheme — could outperform simulations at 4-km grid length in this metric (Pearson *et al.*, 2014).

Using the Cascade simulations, Marsham *et al.* (2013) showed that the mean state of the WAM was influenced not only by the amount of modelled rainfall affecting net heating over the continent, but also by the timing and structure of the modelled convection. In simulations with parameterized convection, Marsham *et al.* (2013) found that the pressure gradients are modified through moist convective heating during the day, when the synoptic-scale monsoon flow is inactive because it is inhibited by dry boundary-layer convection (Parker *et al.*, 2005); in their simulations without convection parameterization, i.e. "explicit" simulations, the (more realistic) later timing of the moist convection means that latent heating modifies pressure gradients when the monsoon flow is active, thus affecting the entire monsoon circulation. Furthermore, in explicit simulations, cold pool outflows provide a significant component of the monsoon flux, while these outflows are absent from parameterized runs. This is consistent with missing cold pools being a major source of error in global forecasts for the central Sahara (Garcia-Carreras *et al.*, 2013). Birch *et al.* (2014) extended the Cascade analysis to show that the location, timing and structure of the convection affects the entire water cycle. Thus the accurate modelling of clouds in the WAM is not only important for the radiation budget and rainfall, but for the entire monsoon circulation.

The aims of this paper are to evaluate the representation of West-African Monsoon vertical cloud structure in the suite of Cascade simulations against CloudSat observations and thereby to infer the differing roles of model resolution and convective parameterization in the model errors. A brief explanation of the Cascade data is provided in section 2, followed by a description of the CloudSat observations used for evaluation in section 3. The results of the model evaluation of cloud vertical structure for different resolutions and treatment of convection are presented in section 4, including analyses of frequency of occurrence and amount when present and a focus on three cloud-type groups (deep convection and anvil, mid-level layer cloud, and low-level cloud and congestus). The vertical distribution of reflectivity and ice water content is discussed in section 5, including an additional sensitivity study that focuses on precipitating low-level cloud, followed by conclusions and outlook for further research on clouds in the West African monsoon.

Table 1: List of main MetUM simulations analysed in this study with distinguishing parameters. Convection treatment is specified in the simulation short name, i.e. "explicit" and "param" (parameterized).

Simulation short name	Horizontal grid spacing (km)	Number of vertical levels
40-km param	40	38
12-km param	12	38
12-km explicit	12	38
4-km explicit	4	70
1.5-km explicit	1.5	70

2 Cascade simulations

In the Cascade project, using the Met Office Unified Model (MetUM) a nested suite of limited-area model simulations were performed for the 10-day period from 25 July to 3 August 2006 over the region of West Africa. The simulations were run at horizontal grid lengths ranging from 40 km down to 4 km (see Table 1), with a 1.5-km simulation run from 25 July to 30 July 2006. At 40-km and 12-km horizontal grid length with 38 vertical levels, simulations were performed with the Gregory and Rowntree (1990) convection scheme turned on to parameterize deep convection on the sub-grid scale. Both these simulations took their initialization and lateral boundary conditions from analyses provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The convection-permitting grid lengths of 4 km and 1.5 km were one-way nested inside the 12-km simulation, with sub-grid turbulence in the horizontal parameterized by the 2D Smagorinsky mixing scheme and vertical mixing treated using the Lock *et al.* (2000) non-local boundary-layer scheme. In the explicit simulations, the CAPE-closure time scale is made CAPE-dependent, increasing rapidly with CAPE, which results in a negligible rainfall contribution from the convection scheme (Pearson *et al.*, 2014); we thus consider the convection scheme effectively switched off. A further 12-km simulation was performed with explicit treatment of convection and sub-grid mixing as in the 4-km and 1.5-km simulations (Birch *et al.*, 2014). At 12-km grid spacing, the model is not expected to faithfully represent the physics of explicit convection and Smagorinsky turbulence mixing, but this 12-km explicit simulation allows us to study separately the impact of convection parameterization and the impact of model resolution between the parameterized 12-km and the explicit 4-km simulations, respectively. For further details regarding the model configurations and choices of parameterization, we refer the reader to Pearson *et al.* (2014).

The MetUM uses a single-moment microphysics scheme for prognostic cloud ice and cloud liquid (Wilson and Ballard, 1999), which has been developed to include prognostic rain for simulations without convection parameterization. The rain particle size distribution (PSD) is based on Marshall and Palmer (1948) and the ice-aggregate PSD is modelled following Cox (1988). The microphysics scheme contains a diagnostic split between ice crystals and aggregates, the for-

mer modelled with the same PSD shape but its prefactor multiplied by a factor 20; both ice-particle habits use the Brown and Francis (1995) mass-diameter relationship and their different fall-speed relationships are based on Mitchell (1996).

In order to provide a like-with-like comparison with observations, the model fields were converted into 94-GHz radar reflectivities using the CloudSat simulator (Haynes *et al.*, 2007), which includes treatment of attenuation of the radar signal through hydrometeors. This simulator was run as part of the cloud-observation simulator package (COSP) (Bodas-Salcedo *et al.*, 2011). For the 40-km, 12-km, and 4-km simulations, sub-grid sampling of cloud overlap in COSP, based on Klein and Jakob (1999), was set to provide 40, 20, and 10 individual columns respectively for each grid box with model hydrometeor distributions spread across each sub-grid column. The number of sub-grid columns was chosen to balance between accuracy and speed of calculation, following Di Michele *et al.* (2012). Rainfall was distributed among the sub-columns via an additional step, matching convective and stratiform rain with convective and stratiform cloud, respectively; a thorough explanation of these methods is provided by Zhang *et al.* (2010). Finally, the CloudSat simulator was run for each of the columns separately. No sub-grid sampling was applied to the 1.5-km simulation, as this resolution is comparable to the CloudSat footprint. For the 1.5-km simulation, we tested an extension to the CloudSat simulator to model multiple scattering based on Hogan and Battaglia (2008), but found that it had little impact on our results. For instance, although the total fraction of cumulonimbus profiles increases from 1.3% to 1.8% and the fraction of congestus increases from 2.5% to 3.2% when multiple scattering is modelled, these changes are small compared to the broad differences between the model simulations and the observations that we discuss in this paper.

The MetUM convection scheme produces ice and liquid mixing ratios and fluxes, but has no microphysical parameterization. However, the radiation scheme treats ice and liquid in convective cloud similar to the large-scale scheme, so for the CloudSat simulator, the convective cloud and precipitation microphysical parameters were assumed identical to those from the large-scale scheme.

The simulations were analysed for the 5-day period of 26 July to 30 July 2006 to include the 1.5-km simulation; data from the 25 July 2006 were ignored to avoid the model spin-up period. Only the model outputs at 0100 UTC and 1300 UTC were considered for comparison with CloudSat observations at 0130 LT and 1330 LT. Results from Pearson *et al.* (2014) indicate that apart from the 12-km param simulation, the different configurations have a comparable diurnal cycle of convection — as identified by OLR clusters — to GERB observations (including the 40-km param simulation, which was not shown), although all simulations tend to produce low OLR too widely. Similarly, Birch *et al.* (2014) found too much rainfall and too strong a diurnal cycle in all model configurations, with the time of peak rainfall shifted to 1300 LT in the parameterized simulations compared to 1800 LT in the observations and explicit simulations. The analyses of OLR and rainfall miss several diurnally varying cloud-types that impact the monsoon circulation, including congestus and mid-

level cloud, as well as nocturnal stratus (which peak after 0200 LT, Schrage and Fink (2012)), which are all observed by CloudSat (Stein *et al.*, 2011a). A like-with-like comparison of the state of the vertical structure of cloud and precipitation in models at 0100 UTC and 1300 UTC with CloudSat observations at 0130 LT and 1330 LT should therefore be appropriate to establish model differences in cloud locations (vertical and latitudinal) and amounts and to relate these to the previously diagnosed model differences in the diurnal cycles of OLR and rainfall. Since each CloudSat overpass over West Africa can be considered a north-south transect, the following preliminary analysis was performed at each fixed model longitude, as if it were a CloudSat transect, using a reflectivity threshold of -27.5 dBZ to identify cloud and precipitation in the simulated reflectivities:

1. For every $1°$ latitude and 50 hPa pressure interval (at 0100 UTC and 1300 UTC), the radar hydrometeor fraction was calculated, combining all sub-grid reflectivity columns at all qualifying latitudes.

2. If the radar hydrometeor fraction was above 5%, a "cloud event" was recorded to calculate the frequency of occurrence.

The same preliminary analysis was performed on the CloudSat observations. Here, we use the term "radar hydrometeor fraction" (RHF) (Marchand *et al.*, 2009) to indicate that both cloud and precipitation are considered, without explicitly distinguishing between the two. While RHF is a measure of the presence of cloud averaged in time and space, the "frequency of occurrence" (FOC) is used to specifically evaluate how frequently any significant amount of cloud is present. When the RHF is averaged only over the cloud events, we obtain the "mean amount when present" (MAP), which informs of the spatial coherence of hydrometeor layers. Thus, if most of the RHF is due to sporadic mesoscale systems, the FOC will be low but the MAP will be high; if the RHF is due to regular broken stratocumulus, the FOC will be high but the MAP will be low.

RHF, FOC, and MAP are related as follows, for a given RHF threshold t (values range between 0 and 1):

$$\begin{aligned} \text{RHF} \quad = \quad & \text{FOC}_{\text{RHF}>t} \times \text{MAP}_{\text{RHF}>t} \qquad\qquad\qquad (1) \\ + \quad & (1 - \text{FOC}_{\text{RHF}>t}) \times \text{MAP}_{\text{RHF}\leq t} \ , \end{aligned}$$

where

$$\text{MAP}_{\text{RHF}>t} = \overline{\text{RHF}_{\text{RHF}>t}} \ . \qquad\qquad\qquad (2)$$

Since $\text{MAP}_{\text{RHF}\leq t} \leq t$ by definition, $\text{FOC} \times \text{MAP}$ is a useful but crude approximation of RHF, provided that t is small. Taking into account the U-shaped distribution of the frequency of cloud fraction (Hogan *et al.*, 2001) we tested several values for t. A value of $t = 5\%$ led to results and conclusions broadly consistent with greater values, but showed additional cloud events from the 40-km param simulation.
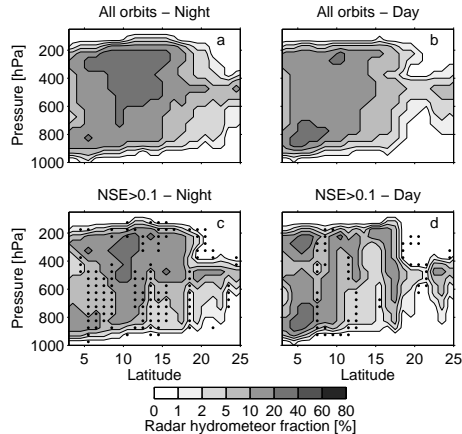
Figure 1: Radar hydrometeor fraction (RHF) for July and August 2006–2009 observed by CloudSat for (a) all nighttime overpasses and (b) all daytime overpasses; (c) and (d) are as (a) and (b), but showing the mean for overpasses for which the Nash-Sutcliffe efficiency (NSE) of the AEJ position compared with the AEJ position during 25th–31st July 2006 was greater than 0.1; dots indicate where the all-orbit RHF is outside the bootstrapped 95% confidence interval of the sub-sampled RHF.

3   CloudSat observations

Launched in 2006, CloudSat provides high-resolution vertical cloud profiles across the globe, with a return period of approximately 16 days (Stephens *et al.*, 2002). The 94-GHz cloud-profiling radar (CPR) measures profiles of equivalent radar reflectivity factor at approximately 1.5 km horizontal and 240 m vertical resolution; the radar has a sensitivity of about -30 dBZ (Marchand *et al.*, 2008), which allows it to detect at least 50% of all ice cloud downward from the $-51°$C level (Stein *et al.*, 2011b), but will miss thin cirrus. The CloudSat "2B-GEOPROF" product (Marchand *et al.*, 2008) was used to determine hydrometeor occurrences when its value was above 20 (arbitrary units), indicating confidence in hydrometeor detection. Following Marchand *et al.* (2009), we used an additional threshold of -27.5 dBZ to identify presence of hydrometeors.

We limit our analysis to two variables that are readily derived from the model and observations, namely CloudSat reflectivity, which is directly measured and is simulated from the model hydrometeor fields; and ice water content (Section 5.1), which is a prognostic model variable and is retrieved from CloudSat-radar and CALIPSO-lidar observations based on the Delanoë and Hogan (2010) optimal estimation algorithm. A comprehensive analysis of model microphysics using A-Train observations should incorporate rainfall identification products (e.g. "2C-PRECIP-COLUMN" (Haynes *et al.*, 2009)) or lidar cloud detection (e.g. "2B-GEOPROF-LIDAR" (Mace and Zhang, 2014)), but such analysis is

beyond the purpose of our study.

During the period of the Cascade simulations, CloudSat had ten overpasses through the region. Because of this small sample, the CloudSat data were instead considered for all of July and August for the years 2006–2009. However, this larger sample covers a wide range of synoptic conditions — for instance due to wave disturbances of the AEJ — which are not necessarily representative of the synoptic situation during the period of the simulations. Since the AEJ has been shown to correlate with the position of convective systems (Mohr and Thorncroft, 2006; Stein $et\ al.$, 2011a), we use its position as a proxy for synoptic conditions over West Africa and to select suitable CloudSat overpasses. We aim to sample the observations in such a way that the AEJ variability is very similar to that during the simulation period. Therefore, the AEJ position across the region was considered for a 5-day period centered in time on each orbit overpass and compared with the 5-day period centered on 28 July 2006. The jet position was calculated from ERA-Interim re-analyses, using the same methodology as in Stein $et\ al.$ (2011a).

For our sub-sampling of CloudSat overpasses we consider the mean AEJ position as well as its variance, which could signify African Easterly Wave activity. We consider the "overpass" values of the AEJ latitude, $\phi_o(t, \lambda)$, for a 5-day window around the time of the CloudSat overpass, with AEJ latitudes specified for every 1° longitude (λ) between 10°W and 10°E and for every 6 hours (t). "Cascade period" values, $\phi_c(t, \lambda)$ are then the AEJ latitudes for a 5-day window centered on any time between 0000 UTC on 26 July 2006 and 0000 UTC on 31 July 2006. The Nash-Sutcliffe efficiency (NSE) was defined by Nash and Sutcliffe (1970) as:

$$\mathrm{NSE} = 1 - \frac{\sum_t \sum_\lambda \left(\phi_o(t, \lambda) - \phi_c(t, \lambda) \right)^2}{\sum_t \sum_\lambda \left(\phi_c(t, \lambda) - \overline{\phi_c} \right)^2} \ , \tag{3}$$

which can be interpreted as the variance of the AEJ during the Cascade period minus the mean squared error in AEJ position between the current overpass and the Cascade period, divided by the variance during the Cascade period. If a CloudSat overpass had its jet latitude everywhere equal to the mean during the Cascade period, $\overline{\phi_c}$, then NSE = 0, so any value of NSE > 0 would be an improved estimate, with a perfect match achieving an efficiency of NSE = 1. Negative values of NSE can occur when the ϕ_o have an incorrect mean and/or variance. For this study, a threshold efficiency of 0.1 was chosen to distinguish those CloudSat overpasses with ϕ_a that were similar to the ϕ_c for any 5-day window during the simulation period. This value led to 128 overpasses (29%) included in the sample and 318 (71%) excluded. When considering nighttime and daytime observations separately, we are left with 73 and 55 overpasses, respectively. Because of these small samples, following Liu $et\ al.$ (2010), we applied the bootstrap method to estimate the 95% confidence interval for the mean RHF; that is, we resampled the population (number of overpasses) with replacement 1,000 times, resulting in 1,000 estimates of the RHF. From these estimates we derive the 95% confidence interval and we will indicate where the

mean fraction over all orbits is outside the 95% confidence interval for the mean estimated from the sub-sample.

In Figure 1, we compare the mean vertical cloud structure over all July and August 2006–2009 overpasses to the structure from only those overpasses with NSE greater than 0.1. The vertical distribution of RHF shows common features for the subset and the full set of overpasses, which are 1) a maximum fraction of 20–40% around 250 hPa at nighttime, associated with RHF above 10% down to 1 km above the surface, 2) a mid-level local maximum around 500 hPa that is especially noticeable over the Sahara where fractions up to 20% are observed, and 3) a maximum around 800 hPa at daytime dominated by low-level hydrometeors south of 10°N with average fractions up to 40%. During the period of the simulations, the mean AEJ location was 14°N and the main effects of sub-sampling the data to similar synoptic conditions are observed in the latitudinal distribution of these common cloud features, as indicated by the dotted regions in Figure 1c and d where the all-orbit mean (Figure 1a and b) is outside the bootstrapped 95% confidence interval. When only overpasses with NSE > 0.1 are included in the analysis, the nighttime 250-hPa maximum and its associated deep cloud are restricted to 8–12°N, at least 2° south of the mean AEJ location. For daytime overpasses with NSE > 0.1, the low-level maximum extends north to 6°N compared to 9°N when all overpasses are included. By conditioning the CloudSat observations on the synoptic state of the AEJ during the simulation period, we can qualitatively evaluate the simulations in terms of latitudinal variability of cloud types as well as vertical cloud structure.

4 The vertical cloud structure of the West African monsoon

This study focuses on the region between 10°W and 10°E, typically used for analysis of the north-south structure of cloud and precipitation during the WAM (Nicholson, 2009; Stein *et al.*, 2011a; Bouniol *et al.*, 2012). Analyses of the WAM historically highlight the east-west homogeneity (e.g. Hamilton and Archbold (1945)) and different cloud-types are known to be associated with different regions, for instance low-level cloud and congestus over the coastal Guinean region are associated with the moist monsoon layer and mid-level cloud over Sahara are associated with the deep boundary layer (Stein *et al.*, 2011a). Latitudes are considered between 3°N and 25°N, encompassing the Guinea coastal region and a large part of the Sahara; this range was limited by the extent of the domain of the 1.5-km simulation. The period of interest runs from 26–31 July 2006, which was a period of considerable wave activity along the AEJ (e.g. Bain *et al.* (2011)) and substantial rain in the Sahara observed at Tamanrasset (22°N, 5°E) (Cuesta *et al.*, 2010).

In Figure 2 the RHF with pressure is shown averaged between 10°W and 10°E for simulations between the period 26–31 July 2006 and for the subset of CloudSat observations with NSE > 0.1, for nighttime and daytime overpasses,

respectively. The third and fourth columns of Figure 2 show the fractional cloud-type cover (FCC), which is the fraction of profiles with a given cloud type identified (at any height). Following Stein *et al.* (2011a), these cloud types are defined as:

1. Low-level clouds, which have cloud-top height below the 700-hPa level (approximately 3.5 km above mean sea level);

2. Congestus, which have cloud-top height between the 350-hPa (about 9 km) and 700-hPa levels and extend to within 1 km of the surface;

3. Mid-level clouds, which have cloud-top height between the 350-hPa and 700-hPa levels as well as base above 1 km above the surface;

4. Deep convective clouds, which extend from above the 350-hPa level down to within 1 km of the surface;

5. Anvil, which has cloud-top height above the 350-hPa level and base above 1 km above the surface.

In Figure 3, the FOC and MAP are shown for the same set of observations, assuming a RHF threshold $t = 5\%$ to identify events.

In the following subsections, we focus on three cloud-type groups to further evaluate model performance.

4.1 High clouds: deep convection and anvils

The observed FCC from anvils, shown in Figure 2c and d, is about 0.3 at night and 0.2 during the day south of 15°N, while deep convection FCC only peaks between 0.05–0.1 at 10°N (night and day) and 17°N (night only). The 4-km and 1.5-km simulations typically have anvil FCC a factor 1.5 lower than observed, while the underlying north-south distribution is comparable to observations, including the peaks of deep convection at 10°N and at 17°N. The RHF in the upper troposphere in these simulations in Figure 2 compares well with observations, though both simulations lack cloud around 200 hPa especially during the day. However, the model could be generating widespread ice cloud but at low mixing ratios so that the simulated reflectivities remain below the -27.5 dBZ threshold; we will revisit this sensitivity issue when evaluating model cloud ice in Section 5. The 4-km and 1.5-km simulations compare well in terms of night-time FOC in Figure 3, while daytime FOC are generally lower than observed. The nighttime MAP for high cloud in the 4-km and 1.5-km simulations between 5°–10°N is generally lower than observed, suggesting that the simulations lack some convective organisation in this region.

The 40-km param simulation also has anvil FCC a factor 1.5 lower than observed, but the anvil does not spread as far north as in the observations and in the 4-km and 1.5-km simulations. The peak of deep convection is too far south at 8°N and the RHF, FOC, and MAP in the upper troposphere show a lack of cloud during the day, while the nighttime values are comparable to observations.

10

The difference between the nighttime and daytime values in the 40-km param simulation suggests that the diurnal cycle for clouds may be comparable to that observed. This would imply a delay between the rainfall and cloud cycles, as Birch *et al.* (2014) showed that the rainfall peak was too early in the 40-km simulation — indeed, the FOC shows a low-level maximum between 8°–13°N indicating frequent congestus but with MAP less than 20%. Closer inspection of the model 3D cloud fields indeed shows that the daytime rainfall in the 40-km param simulation is produced by the convection scheme while the scheme does not produce much anvil, whereas the large-scale cloud scheme produces the majority of cloud ice and its slower response leads to increased anvil FCC at night.

The 12-km param simulation, however, has more anvil FCC during the day than at night, which suggests that the cloud and rainfall cycles are more synchronised in this simulation. The large anvil FCC in the 12-km param simulation can be seen to be due to too high FOC as well as a high MAP. In contrast with the 40-km param simulation, the 12-km param simulation was configured to allow the convection scheme to produce anvil.

The 12-km explicit performs similarly to the 4-km and 1.5-km simulations, apart from the extensive RHF and anvil FCC during the day. From Figure 3 we can tell that the RHF is due to a high FOC but with low MAP, which implies that the convection scheme is active in many grid boxes, but produces low convective RHF per grid box. Indeed, individual transects of the 12-km explicit simulation (not shown) show convective plumes in many adjacent grid boxes. Thus, despite the negligible rain amounts from the convection scheme, the CAPE-dependent time scale still allows the convection scheme to generate noticeable amounts of cloud in the 12-km explicit simulation.

4.2 Mid-level clouds

The observed mid-level cloud FCC ranges across all latitudes between 0.06–0.2 at night and 0.03–0.12 during the day, with a daytime local maximum near 17°N and nighttime maxima around 10°N and 16°N. Both the 4-km and the 1.5-km simulation overestimate mid-level cloud FCC slightly at night and during the day, while the 1.5-km simulation has too much mid-level cloud FCC north of 20°N. The vertical distribution of RHF shows observed mid-level maxima around the 500-hPa level, while in all simulations the maxima are situated around the 0°C level, between 550–600 hPa. This suggests that the model too readily produces cloud and precipitation around the freezing level — a mixed dynamical and microphysical issue — although this signal may be due to the microphysics scheme used. In particular, the MetUM large-scale cloud scheme includes a diagnostic split between ice crystals and aggregates, based on the distance from cloud top, so that ice and mixed-phase clouds with tops near the freezing level will have most ice modelled as crystals. In forward-modelled reflectivities, the melting of crystals will lead to a sharp increase across the 0°C level, as the numerous small crystals will have a relatively low reflectivity given their total mass compared to if they were aggregates (Stein *et al.*, 2014).

Thus we see for instance in the 40-km param that the RHF increases downward across the melting layer, as crystals below the reflectivity threshold melt to form precipitation above the threshold.

The 40-km param simulation overestimates nighttime mid-level cloud FCC between 5–10°N by a factor 2.5. This region coincides with the location of peak deep convection and anvil and the mid-level cloud is a result of extensive detrainment at the 0°C level, which is both frequent and at large amounts, as evident in Figure 3e and g. The daytime mid-level cloud FCC is simulated reasonably well in the 40-km param. The 12-km param simulation generally underestimates mid-level cloud FCC, especially during the day when deep convection and anvil dominate throughout the region. In the 12-km explicit simulation, mid-level cloud FCC is similar to the 4-km and 1.5-km explicit simulations and compares well with observations, apart from the detrainment near the 0°C level, which is common in all simulations.

4.3   Low cloud: low-level cloud and congestus

Fractional cover from low-level cloud and congestus gradually decreases in observations from 0.2 at 4°N to less than 0.05 at 10°N at night, while during the day it stays around 0.2 between 3–7°N and is absent north of 13°N where the boundary layer is too deep (Cuesta *et al.*, 2009). Congestus FCC is generally low, but peaks around 0.05 near 8°N at night and is approximately 0.05 between 3–11°N. The 4-km and 1.5-km simulations have slightly too little congestus during the day, though at night the 4-km simulation compares well with observations while the 1.5-km simulation overestimates congestus FCC by a factor 2. Low-level cloud appears too extensive in all simulations, with FCC too high by a factor 2 or more between 3–7°N at night and day and with FCC as far north as 15°N where it is not observed. The 12-km explicit and the 4-km and 1.5-km simulations do get the correct signal of low cloud FCC during the day extending farther north than at night.

All simulations have low MAP for the low-level cloud during the day combined with high FOC, suggesting broken cloud, similar to observations in Figure 3b and d. At night, the simulations have relatively high MAP for low-level cloud with values above 0.4, while in observations the mean amount is below 0.3. Knippertz *et al.* (2011) used ground-based synoptic reports to show the extensive nocturnal stratus over land south of 10°N, which is not regularly observed by CloudSat. Although the simulated reflectivities are not considered below 1 km above the surface, if the model simulates low-level cloud at higher altitudes, it will show up in the analysis presented here. Indeed, much of the low-level RHF occurs at 800 hPa, well above the surface where CloudSat should be able to observe such cloud. Secondly, the model may produce too high liquid water contents and precipitation from low-level clouds, which could generate reflectivities above the −27.5-dBZ threshold more frequently than in observations; we will revisit the model sensitivity to rainfall parameterization in section 5.2.

5 The vertical distribution of hydrometeors

In this section, the CloudSat profiles of reflectivity are used to study the vertical distribution of hydrometeors throughout the region. Contoured frequency-by-altitude diagrams (CFAD, Yuter and Houze Jr (1995)) are shown in Figure 4 to illustrate the variation of the reflectivity distribution with height in cumulonimbus and congestus profiles; daytime and nighttime observations are combined. An increase in reflectivity as altitude decreases (pressure increases) suggests aggregation of ice into larger particles; as reflectivity relates to mass squared, a factor 2 increase in mass leads to a factor 4 increase in reflectivity, or approximately 6 dB. The decrease of the quantiles towards the $0°C$ level and the subsequent increase are also partly associated with attenuation due to snow and supercooled water in deep convection and the increase in dielectric factor as snow melts into rain (Sassen *et al.*, 2007). Below the $0°C$ level, reflectivity deciles decrease primarily due to strong attenuation of the radar signal in precipitating profiles.

The 40-km param and 12-km param do not adequately represent the hydrometeor distribution in the cumulonimbus and congestus profiles. Ice reflectivities in cumulonimbus profiles are typically 10 dB weaker than observed, which suggests that ice water contents may be a factor 3 or more too low in these simulations. Below the $0°C$ level, the 25th percentile is 10 dB lower than observed, indicating too frequent light precipitation from cumulonimbus profiles. The congestus profiles in the 40-km param and 12-km param simulations also show median reflectivity typically 10 dB lower than observed, suggesting a decrease in rainfall rate by a factor of about 4 using the Marshall and Palmer (1948) relationship of $Z[\mathrm{mm^6 m^{-3}}] = 200R[\mathrm{mm\,hr^{-1}}]^{1.6}$. These results indicate that precipitation from individual cumulonimbus and congestus profiles (or sub-grid columns in the model) is typically weaker than observed and that these simulations require more cover from congestus and cumulonimbus, or more precipitation from cumulus, to achieve the correct domain-averaged rainfall rate. Similar to the mid-level cloud, the congestus in all simulations typically have their cloud top near 500–550 hPa or the $0°C$ level, while in observations the congestus profiles appear deeper.

For the low-level, mid-level, and anvil profiles, the cloud-top height distribution in the third column of Figure 4 confirms that all simulations underestimate the fraction of anvil at high altitudes, i.e. with cloud tops above 12 km, and that the mid-level cloud is generally at altitudes that are too low, with tops between 4–6 km rather than the observed 6–8 km. In the fourth column of Figure 4 we show the distribution of cloud-base height conditioned on cloud-top height for low-level, mid-level, and anvil clouds. All simulations overestimate the frequency of anvils with cloud base below 4 km — the base below the freezing level implies these clouds are precipitating. This overestimate may be due to mis-classification of cumulonimbus as anvil when the radar reflectivity is attenuated before reaching the surface due to heavy precipitation; such a mis-classification is less likely in the observations, as multiple scattering counteracts some of the effects of attenuation (e.g., Hogan and Battaglia (2008)). In the 1.5-km sim-

13

ulation, for instance, grid boxes with surface rainfall rates above 10 mm hr^{-1} were most likely to have the lowest cloud layer classified as anvil, although these occurrences account for less than 1% of all anvil profiles.

For mid-level cloud, all simulations underestimate the frequency of cloud base above 4 km and overestimate the frequency of cloud base below 4 km. Low-level clouds tend to have lower tops than observed and appear to be thicker. The 4-km and 1.5-km explicit generally match the observed cloud-base distribution better than the 40-km param and 12-km param, especially in terms of representing thin anvil and mid-level cloud, which are on the diagonal in the fourth column of Figure 4.

5.1 Ice water content distribution

We expand the analysis of cloud vertical distribution using ice-water-content (IWC) retrievals from CloudSat and CALIPSO observations. An optimal estimation algorithm was developed by Delanoë and Hogan (2010) to retrieve cloud-ice properties from CloudSat, CALIPSO, and MODIS observations, and was used by Delanoë *et al.* (2011) to evaluate cloud-ice in the Met Office and ECMWF global models; this product will be referred to as DARDAR. Due to limited availability of the DARDAR ice product, all overpasses for July–August 2006–2008 were included, regardless of the synoptic situation. Since model ice-water contents are provided as grid-box means, we only consider the 1.5-km explicit simulation for this analysis. In Figure 5, cloud-ice retrievals using the DARDAR algorithm are compared with grid-box mean ice water content in the 1.5-km simulation, using the in-cloud cumulative distribution calculated for IWC $> 10^{-6}$ kg m^{-3}.

The radar-lidar retrieval is well-constrained between 200–400 hPa, where joint observations from both the CloudSat radar and the CALIPSO lidar occur most frequently (Stein *et al.*, 2011b). In this range, the 1.5-km explicit simulation clearly has too low IWC compared to the retrievals, as the quartiles for DARDAR are about a factor 3 higher than the quartiles from the model, which agrees with the lower reflectivity quantiles for cumulonimbus profiles in Figure 4u. The IWC quartiles from the DARDAR retrievals have a local minimum at 500 hPa, similar to that in the CloudSat reflectivity distribution (not shown); this is likely due to mid-level cloud, which may have lower IWC than anvil and deep convection at these levels. Towards the melting layer, the median increases by a factor 10, though the retrieval is subject to uncertainty as it cannot distinguish between ice and rain (Delanoë *et al.*, 2011), which may coexist at this level, and it is sensitive to radar scattering assumptions for high reflectivities (Stein *et al.* (2011b), note that DARDAR was called VarCloud in this earlier paper). In the 1.5-km explicit simulation, the quartiles decrease towards the melting layer, possibly due to mid-level cloud which dominates at a lower level than in the observations and may have lower IWC than deep convection and anvils.

Since DARDAR incorporates CALIPSO observations, more cloud samples are included in the IWC comparison than the CloudSat-only comparison, par-

ticularly observations of thin cirrus and tops of ice clouds missed by CloudSat. The 1.5-km explicit simulation clearly lacks an appropriate representation of IWC at upper levels (above the 200-hPa level), as the all-sky IWC frequencies (not shown) are too low compared to DARDAR, whereas the in-cloud quartiles in Figure 5 are too high. When lowering the IWC threshold used to calculate the cumulative percentiles, the 1.5-km explicit simulation has the peak of the IWC distribution at 200 hPa around 10^{-8} kg m^{-3}, at least a factor 10 below the lowest value retrieved with DARDAR, whereas less than 10% of in-cloud IWC reach above 10^{-6} kg m^{-3} at this level in the model. This suggests that the IWC at upper levels is too low in the simulations, which agrees with previous results comparing cloud-ice in the MetUM with in-situ and radar observations (Delanoë *et al.*, 2011; Baran *et al.*, 2011). Recent experiments using the MetUM by Furtado *et al.* (2014) indicate that changes to the ice-particle fall speed or the use of the Field *et al.* (2007) moment-estimation ice-microphysics scheme may lead to greater ice water contents at high altitudes and a better cloud-ice representation overall.

5.2 Sensitivity of results to rainfall parameterization

The high fraction of low-level cloud in the simulations and the possibility that it is precipitating suggests that our results could be sensitive to rainfall parameterization. In the Cascade simulations, the MetUM used the Marshall and Palmer (1948) rainfall particle size distribution (PSD) parameters. Abel and Boutle (2012) have shown that these parameters are only suitable for rain rates above 10 mm day^{-1} as the fixed intercept parameter, N_0, does not capture the change in rainfall PSD to more numerous, small droplets at lower rainfall rates. Using the rainfall PSD parameters proposed by Abel and Boutle (2012), for a given rain rate, the rainfall PSD will be skewed towards more numerous small droplets compared to the standard configuration used in the Cascade simulations, which will lead to lower forward-modelled reflectivities. In addition, Abel and Boutle (2012) show that the evaporation rates for low rainfall rates more closely resemble observations when using their rainfall PSD parameters compared to the Marshall and Palmer (1948) values. Therefore, we may expect a difference in the RHF as well as the in-cloud reflectivity distribution when changing to the Abel and Boutle (2012) rainfall PSD parameters. The additional model run was performed at 4-km horizontal grid length and was identical to the 4-km explicit simulation, apart from the use of the Abel and Boutle (2012) parameters; this run will be referred to as the 4-km AB2012 simulation. The two simulations are compared for the period 26–29 July 2006, as the AB2012 simulation was run for this smaller period only.

Figure 6 shows the FCC of low-level cloud, congestus, and cumulonimbus from the 4-km AB2012 simulation and the 4-km explicit simulation. There is on average a 50% reduction in FCC from congestus when switching to AB2012, as well as a reduction in cumulonimbus, both at night and during the day, which is also apparent in the vertical distribution of RHF (not shown), while we noted little change in the mean precipitation. Changes in low-level cloud FCC

are generally within 10% of the 4-km explicit simulation, which suggests that this basic change does not affect low-level cloud significantly. The reflectivity distributions for congestus and cumulonimbus profiles in Figure 6c and d are comparable for the two 4-km simulations, which can be expected as the AB2012 parameters lead to a PSD similar to Marshall and Palmer (1948) for rainfall rates above 10 mm day^{-1}, which should dominate in these cloud types. From this analysis, the use of Abel and Boutle (2012) rainfall parameters is no obvious improvement over the standard configuration of the Cascade simulations.

6   Discussion and conclusions

We have evaluated the vertical cloud structure of the West African monsoon (WAM) in simulations of the Met Office Unified Model (MetUM) against Cloud-Sat observations, highlighting model errors in cloud-top height, cloud-type cover, and vertical distribution of radar reflectivity and ice water content. The Cloud-Sat observations used for evaluation were restricted to synoptic conditions for July and August 2006–2010 that were similar to the simulation period of 26–31 July 2006, assuming that the AEJ position and variability control the locations of deep convection. Reflectivities were obtained from the model cloud fields using the CloudSat simulator (Haynes *et al.*, 2007) and a model sensitivity study changing the rainfall parameters for the particle size distribution showed little dependence of the results on such a change. The vertical profiles of reflectivity were studied statistically to understand model performance in terms of hydrometeor distribution, as well as in-cloud ice water contents using the DARDAR ice retrievals of Delanoë and Hogan (2010). In line with previous Cascade studies of MetUM performance over West Africa, we have found improved model performance at 12-km grid length with explicit convection compared to parameterized convection, with additional improvement when resolution is increased, though we note little difference between the 4-km and 1.5-km simulations.

Despite concerns about the physics of a 12-km simulation without a convection parameterization, compared to the 12-km param simulation, the 12-km explicit has a better vertical structure of RHF with latitude, including a distribution of cumulonimbus profiles comparable to observations and a higher FCC from mid-level cloud. In terms of reflectivity profiles, the 12-km explicit also compares well to CloudSat observations, having increased ice reflectivities in cumulonimbus profiles and increased reflectivities in the precipitating part of congestus profiles compared to the 12-km param simulation. In the explicit simulations, the CAPE-closure time scale is made CAPE-dependent, increasing rapidly with CAPE, which results in a negligible rainfall contribution from the convection scheme. While the convection parameterization scheme is effectively switched off, it still generates narrow convective plumes in the 12-km explicit simulation, which lead to a high frequency of occurrence of RHF above 5% during the day. Although these convective plumes produce little precipitation, it will be important to quantify the radiative importance of the anvil cloud in comparison to the 4-km and 1.5-km explicit simulations, where this excessive

daytime cloud is absent.

The main improvement when reducing horizontal grid length to 4-km or 1.5-km is found in the distribution of mid-level and anvil cloud base. The 4-km and 1.5-km simulations both have cloud-base height distributions comparable to CloudSat observations, showing more thin anvil and mid-level cloud. Differences in performance between the 4-km and 1.5-km simulations were marginal, in line with analyses of the water cycle in Cascade simulations (Marsham *et al.*, 2013; Birch *et al.*, 2014) and the diurnal cycle of organisation of OLR clusters (Pearson *et al.*, 2014). Although these Cascade studies show little improvement at 1.5-km grid length compared to 4 km, further analysis of individual convective features could show improvements in the representation of convection when grid length is further reduced (Miyamoto *et al.*, 2013; Stein *et al.*, 2014).

When compared with DARDAR ice water content retrievals, the 1.5-km simulation had generally too low in-cloud ice water contents between the 0°C level and 300 hPa, suggesting that the ice water contents in deep convective clouds are too low, as apparent from the reflectivity distribution in Figure 4u. The DARDAR retrievals include thin cirrus observations, which indicate that IWC at upper levels is too low in the 1.5-km simulation, in line with previous MetUM evaluation studies (Delanoë *et al.*, 2011; Baran *et al.*, 2011).

Three issues are consistent among all simulations, namely a lack of high-level anvil, a detrainment of mid-level cloud too close to the 0°C level, and an abundance of low-level cloud and precipitation. Firstly, the lack of anvil with cloud-top height above 12 km is likely related to the low IWC at high altitudes. The MetUM split of ice into aggregates and crystals leads to lower reflectivities near cloud top than if only aggregates were used and thus lower cloud-top heights when the -27.5 dBZ level is considered. It will be important to establish the radiative impact of these biases in cloud-top height and low IWC, as well as the lack of thin anvil in the 40-km and 12-km simulations, possibly using groud-based radar observations from Niamey following Bouniol *et al.* (2012).

Secondly, in all simulations, mid-level cloud does not extend high enough above the 0°C level and its base is on average too far below the same level, suggesting that the model detrains cloud too readily and allows mid-level cloud to precipitate too frequently. Updated MetUM microphysics schemes have removed the diagnostic split between ice crystals and aggregates, which would likely affect cloud and precipitation from mid-level clouds (Stein *et al.*, 2014), however these updated versions were not available when the Cascade simulations were run. Previous studies have highlighted that these mid-level clouds are often observed with supercooled liquid near cloud tops (Stein *et al.*, 2011a) and are important for the radiative budget (Bouniol *et al.*, 2012), whilst their microphysical processes, including ice and rain formation, are likely strongly affected by dust and aerosol (Rosenfeld *et al.*, 2001). However, in order to identify the relevant microphysical and dynamical processes that influence their development, and to properly evaluate these clouds in models, targeted observational and modelling studies are essential.

Finally, the model overestimates low-level cloud FCC by at least a factor 2

in all simulations. A sensitivity study at 4-km grid length using the Abel and Boutle (2012) rainfall parameters was expected to reduce the low-level RHF, but only a reduction in congestus and cumulonimbus was found. The model overestimate could be due to an instrument sensitivity issue, as CloudSat underestimates low-level cloud FCC and will miss cloud that is too close to the surface (e.g. Schrage and Fink (2012)), while the simulated reflectivities may be artificially high due to choices in the model microphysics scheme. However, the model low-level cloud FCC is also much too high when compared to radar-lidar obervations (not shown). Further sensitivity studies are therefore advised to understand the effect of low-level cloud on the monsoon circulation, including improvements in droplet autoconversion. We also recommend targeted observational studies, such as the field and aircraft campaign planned during the DACCIWA (Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa, 2013–2018) project (Knippertz *et al.*, 2015), since these low-level clouds are generally poorly observed by the current satellite observational systems.

We have shown that the West African monsoon is a suitable laboratory for evaluating model cloud and that with the availability of the CloudSat simulator, CloudSat data can highlight model deficiencies in the vertical distribution of clouds. In future work, it will be advised to evaluate simulations over a longer period, such as the 40-day Cascade simulations studied by Birch *et al.* (2014), to increase statistics as well as to study synoptic controls on cloud types and vice versa. Additionally, the radiative impact of the different cloud types in models and observations should be quantified, for instance following Bouniol *et al.* (2012), as Marsham *et al.* (2013) have already highlighted the role of deep convection on the monsoon circulation. Finally, the importance of congestus in the monsoon water cycle is still unclear: the CloudSat overpass times do not enable us to study the life cycle of these cloud types or their success rate of developing into cumulonimbus, which is necessary to quantify their contribution to the precipitation budget.

7   Acknowledgments

and the CloudSat simulator.

References

Abel SJ, Boutle IA. 2012. An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Quarterly Journal of the Royal Meteorological Society* **138**(669): 2151–2162.

Bain CL, Parker DJ, Dixon N, Fink AH, Taylor CM, Brooks B, Milton SF. 2011. Anatomy of an observed African easterly wave in july 2006. *Quarterly Journal of the Royal Meteorological Society* **137**(657): 923–933.

Baran AJ, Bodas-Salcedo A, Cotton R, Lee C. 2011. Simulating the equivalent radar reflectivity of cirrus at 94 GHz using an ensemble model of cirrus ice crystals: a test of the Met Office global numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society* **137**(659): 1547–1560.

Birch CE, Parker DJ, Marsham JH, Copsey D, Garcia-Carreras L. 2014. A seamless assessment of the role of convection in the water cycle of the West African Monsoon. *Journal of Geophysical Research: Atmospheres* **119**(6): 2890–2912.

Bodas-Salcedo A, Webb MJ, Bony S, Chepfer H, Dufresne JL, Klein SA, Zhang Y, Marchand R, Haynes JM, Pincus R, John VO. 2011. COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society* **92**(8): 1023–1043.

Bodas-Salcedo A, Webb MJ, Brooks ME, Ringer MA, William KD, Milton SF, Wilson DR. 2008. Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities. *J. Geophys. Res.* **113**: D00A13.

Bouniol D, Couvreux F, Kamsu-Tamo PH, Leplay M, Guichard F, Favot F, O'Connor EJ. 2012. Diurnal and seasonal cycles of cloud occurrences, types, and radiative impact over West Africa. *Journal of Applied Meteorology and Climatology* **51**(3): 534–553.

Brown PRA, Francis PN. 1995. Improved measurements of the ice water content in cirrus using a total-water probe. *J. Atmos. Ocean. Tech.* **12**: 410–414.

Christensen JH, Kanikicharla KK. 2013. Climate phenomena and their relevance for future regional climate change. In: *Climate Change 2013 - The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker T, Qin D, Plattner GK, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds), ch. 14, Cambridge University Press: Cambridge, UK and New York, NY, USA, pp. 1217–1308.

Cox GP. 1988. Modelling precipitation in frontal rainbands. *Quarterly Journal of the Royal Meteorological Society* **114**(479): 115–127.

Cuesta J, Lavaysse C, Flamant C, Mimouni M, Knippertz P. 2010. Northward bursts of the West African monsoon leading to rainfall over the Hoggar Massif, Algeria. *Quarterly Journal of the Royal Meteorological Society* **136**(S1): 174–189.

Cuesta J, Marsham JH, Parker DJ, Flamant C. 2009. Dynamical mechanisms controlling the vertical redistribution of dust and the thermodynamic structure of the West Saharan atmospheric boundary layer during summer. *Atmos. Sci. Lett.* **10**: 34–42.

Delanoë J, Hogan RJ. 2010. Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *J. Geophys. Res.* **115**: D00H29.

Delanoë J, Hogan RJ, Forbes RM, Bodas-Salcedo A, Stein THM. 2011. Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data. *Quarterly Journal of the Royal Meteorological Society* **137**(661): 2064–2078.

Di Michele S, Ahlgrimm M, Forbes R, Kulie M, Bennartz R, Janisková M, Bauer P. 2012. Interpreting an evaluation of the ECMWF global model with CloudSat observations: ambiguities due to radar reflectivity forward operator uncertainties. *Quarterly Journal of the Royal Meteorological Society* **138**(669): 2047–2065.

Field PR, Heymsfield AJ, Bansemer A. 2007. Snow size distribution parameterization for midlatitude and tropical ice clouds. *Journal of the Atmospheric Sciences* **64**(12): 4346–4365.

Furtado K, Field PR, Cotton R, Baran AJ. 2014. The sensitivity of simulated high clouds to ice crystal fall speed, shape and size distribution. *Q. J. R. Meteorol. Soc.* doi:10.1002/qj.2457.

Garcia-Carreras L, Marsham JH, Parker DJ, Bain CL, Milton S, Saci A, Salah-Ferroudj M, Ouchene B, Washington R. 2013. The impact of convective cold pool outflows on model biases in the Sahara. *Geophysical Research Letters* **40**(8): 1647–1652.

Gregory D, Rowntree PR. 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Monthly Weather Review* **118**(7): 1483–1506.

Hamilton RA, Archbold JW. 1945. Meteorology of Nigeria and adjacent territory. *Q. J. R. Meteorol. Soc.* **71**: 231–264.

Haynes JM, L'Ecuyer TS, Stephens GL, Miller SD, Mitrescu C, Wood NB, Tanelli S. 2009. Rainfall retrieval over the ocean with spaceborne W-band radar. *Journal of Geophysical Research: Atmospheres* **114**(D8): D00A22.

Haynes JM, Luo Z, Stephens GL, Marchand RT, Bodas-Salcedo A. 2007. A multipurpose radar simulation package: Quickbeam. *Bulletin of the American Meteorological Society* **88**(11): 1723–1727.

Hogan RJ, Battaglia A. 2008. Fast lidar and radar multiple-scattering models. Part II: Wide-angle scattering using the time-dependent two-stream approximation. *Journal of the Atmospheric Sciences* **65**(12): 3636–3651.

Hogan RJ, Jakob C, Illingworth AJ. 2001. Comparison of ECMWF winter-season cloud fraction with radar-derived values. *J. Appl. Met.* **40**: 513–525.

Hourdin F, Musat I, Grandpeix JY, Polcher J, Guichard F, Favot F, Marquet P, Boone A, Lafore JP, Redelsperger JL, *et al.* 2010. AMMA-model intercomparison project. *Bulletin of the American Meteorological Society* **91**(1): 95–104.

Jakob C. 2002. Ice clouds in numerical weather prediction models. In: *Cirrus*, Lynch DK (ed), ch. 16, Oxford University Press, New York, pp. 327–345.

Klein SA, Jakob C. 1999. Validation and sensitivies of frontal clouds simulated by the ECMWF model. *Monthly Weather Review* **127**(10): 2514–2531.

Klein SA, Zhang Y, Zelinka MD, Pincus R, Boyle J, Gleckler PJ. 2013. Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator. *Journal of Geophysical Research: Atmospheres* **118**(3): 1329–1342.

Knippertz P, Coe H, Chiu JC, Evans MJ, Fink AH, Kalthoff N, Liousse C, Mari C, Allan RP, Brooks B, *et al.* 2015. The DACCIWA project: Dynamics-aerosol-chemistry-cloud interactions in West Africa. *Bulletin of the American Meteorological Society* doi:10.1175/BAMS-D-14-00108.1.

Knippertz P, Fink AH, Schuster R, Trentmann J, Schrage JM, Yorke C. 2011. Ultra-low clouds over the southern West African monsoon region. *Geophysical Research Letters* **38**(21): L21 808.

Liu Z, Marchand R, Ackerman T. 2010. A comparison of observations in the tropical western Pacific from ground-based and satellite millimeter-wavelength cloud radars. *J. Geophys. Res.* **115**: D24 206.

Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB. 2000. A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Monthly Weather Review* **128**(9): 3187–3199.

Mace GG, Zhang Q. 2014. The CloudSat radar-lidar geometrical profile product (RL-GeoProf): Updates, improvements, and selected results. *Journal of Geophysical Research: Atmospheres* **119**(15): 9441–9462.

Marchand R, Haynes J, Mace GG, Ackerman T, Stephens G. 2009. A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations. *Journal of Geophysical Research: Atmospheres* **114**(D8): D00A20.

Marchand R, Mace GG, Ackerman T, Stephens G. 2008. Hydrometeor detection using CloudSat - An earth orbiting 94-GHz cloud radar. *J. Atmos. Ocean. Tech.* **25**: 519–533.

Marshall JS, Palmer WMK. 1948. The distribution of raindrops with size. *Journal of Meteorology* **5**(4): 165–166.

Marsham JH, Dixon NS, Garcia-Carreras L, Lister GMS, Parker DJ, Knippertz P, Birch CE. 2013. The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophysical Research Letters* **40**(9): 1843–1849.

Mitchell DL. 1996. Use of mass- and area-dimensional power laws for determining precipitation particle terminal velocities. *Journal of the Atmospheric Sciences* **53**(12): 1710–1723.

Miyamoto Y, Kajikawa Y, Yoshida R, Yamaura T, Yashiro H, Tomita H. 2013. Deep moist atmospheric convection in a subkilometer global simulation. *Geophysical Research Letters* **40**(18): 4922–4926.

Mohr KI, Thorncroft CD. 2006. Intense convective systems in West Africa and their relationship to the African easterly jet. *Q. J. R. Meteorol. Soc.* **132**: 163–176.

Nash J, Sutcliffe J. 1970. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* **10**(3): 282 – 290.

Newell RE, Kidson JW. 1984. African mean wind changes between Sahelian wet and dry periods. *J. Climatol.* **4**: 27–33.

Nicholson SE. 2009. A revised picture of the structure of the "monsoon" and land ITCZ over West Africa. *Clim. Dyn.* **32**: 1155–1171.

Parker DJ, Thorncroft CD, Burton RR, Diongue-Niang A. 2005. Analysis of the African easterly jet, using aircraft observations from the JET2000 experiment. *Q. J. R. Meteorol. Soc.* **131**: 1461–1482.

Pearson K, Hogan R, Allan R, Lister G, Holloway C. 2010. Evaluation of the model representation of the evolution of convective systems using satellite observations of outgoing longwave radiation. *Journal of Geophysical Research* **115**(D20): D20 206.

Pearson KJ, Lister GMS, Birch CE, Allan RP, Hogan RJ, Woolnough SJ. 2014. Modelling the diurnal cycle of tropical convection across the 'grey zone'. *Quarterly Journal of the Royal Meteorological Society* **140**: 491–499.

Randall D, Khairoutdinov M, Arakawa A, Grabowski W. 2003. Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society* **84**(11): 1547–1564.

Redelsperger JL, Thorncroft CD, Diedhiou A, Lebel T, Parker DJ, Polcher J. 2006. African monsoon multidisciplinary analysis: An international research project and field campaign. *Bulletin of the American Meteorological Society* **87**(12): 1739–1746.

Roehrig R, Bouniol D, Guichard F, Hourdin F, Redelsperger JL. 2013. The present and future of the West African monsoon: A process-oriented assessment of CMIP5 simulations along the AMMA transect. *Journal of Climate* **26**(17): 6471–6505.

Rosenfeld D, Rudich Y, Lahav R. 2001. Desert dust suppressing precipitation: A possible desertification feedback loop. *Proc. Natl. Acad. Sci.* **98**: 5975–5980.

Sassen K, Matrosov S, Campbell J. 2007. CloudSat spaceborne 94 GHz radar bright bands in the melting layer: An attenuation-driven upside-down lidar analog. *Geophysical Research Letters* **34**(16): L16 818.

Schrage JM, Fink AH. 2012. Nocturnal continental low-level stratus over tropical West Africa: Observations and possible mechanisms controlling its onset. *Monthly Weather Review* **140**(6): 1794–1809.

Stein T, Parker D, Delanoë J, Dixon N, Hogan R, Knippertz P, Maidment R, Marsham J. 2011a. The vertical cloud structure of the West African monsoon: A 4 year climatology using CloudSat and CALIPSO. *Journal of Geophysical Research: Atmospheres (1984–2012)* **116**(D22).

Stein THM, Delanoë J, Hogan RJ. 2011b. A comparison among four different retrieval methods for ice-cloud properties using data from CloudSat, CALIPSO, and MODIS. *Journal of Applied Meteorology and Climatology* **50**(9): 1952–1969.

Stein THM, Hogan RJ, Hanley KE, Nicol JC, Lean HW, Plant RS, Clark PA, Halliwell CE. 2014. The three-dimensional morphology of simulated and observed convective storms over southern england. *Monthly Weather Review* **142**: 3264–3283.

Stephens GL, Vane DG, Boain RJ, Mace GG, Sassen K, Wang Z, Illingworth AJ, O'Connor EJ, Rossow WB, Durden SL, Miller SD, Austin RT, Benedetti A, Mitrescu C. 2002. The CloudSat mission and the A-Train. *Bull. Amer. Meteor. Soc.* **83**: 1771–1790.

Stevens B, Bony S. 2013. What are climate models missing? *Science* **340**(6136): 1053–1054.

Thorncroft CD, Parker DJ, Burton RR, Diop M, Ayers JH, Barjat H, Devereau S, Diongue A, Dumelow R, Kindred DR, Price NM, Saloum M, Taylor CM, Tompkins AM. 2003. The JET2000 project: Aircraft observations of the African easterly jet and African easterly waves. *Bull. Amer. Meteor. Soc.* **84**: 337–351.

Wilson DR, Ballard SP. 1999. A microphysically based precipitation scheme for the UK meteorological office unified model. *Quarterly Journal of the Royal Meteorological Society* **125**(557): 1607–1636.

Yuter SE, Houze Jr RA. 1995. Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Monthly Weather Review* **123**(7): 1941–1963.

Zhang Y, Klein SA, Boyle J, Mace GG. 2010. Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. *Journal of Geophysical Research: Atmospheres* **115**(D12): n/a–n/a.
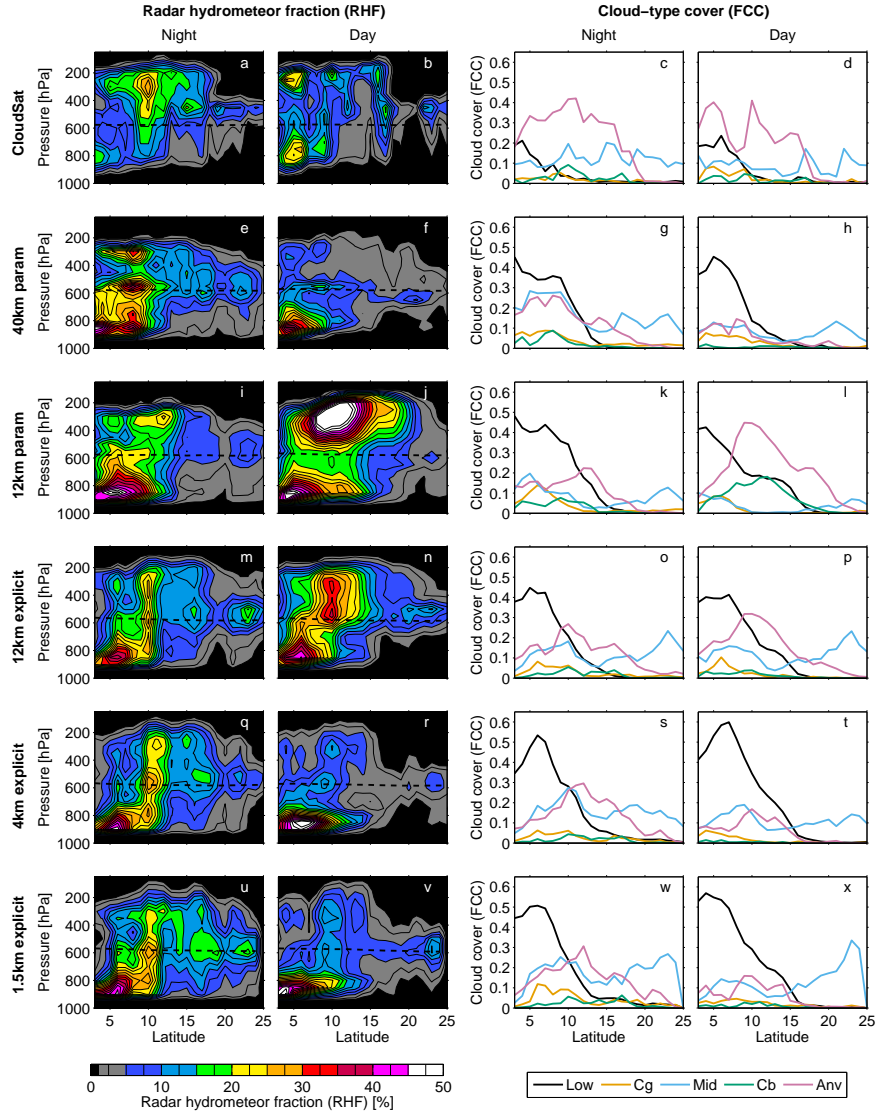
Figure 2: Radar hydrometeor fraction (RHF) and fractional cloud-type cover (FCC) for nighttime overpasses and 0100 UTC model data (first and third column) and for daytime overpasses and 1300 UTC model data (second and fourth column). Results are shown for CloudSat observations (a–d), 40-km param (e–h), 12-km param (i–l), 12-km explicit (m–p), 4-km explicit (q–t), and 1.5-km explicit (u–x). The dashed line in the left two columns indicates the mean pressure of the $0°C$ level. Cloud types in the right two columns are low-level cloud (black), congestus (orange), mid-level cloud (blue), anvil (red), and deep convection (green), as defined in the text.
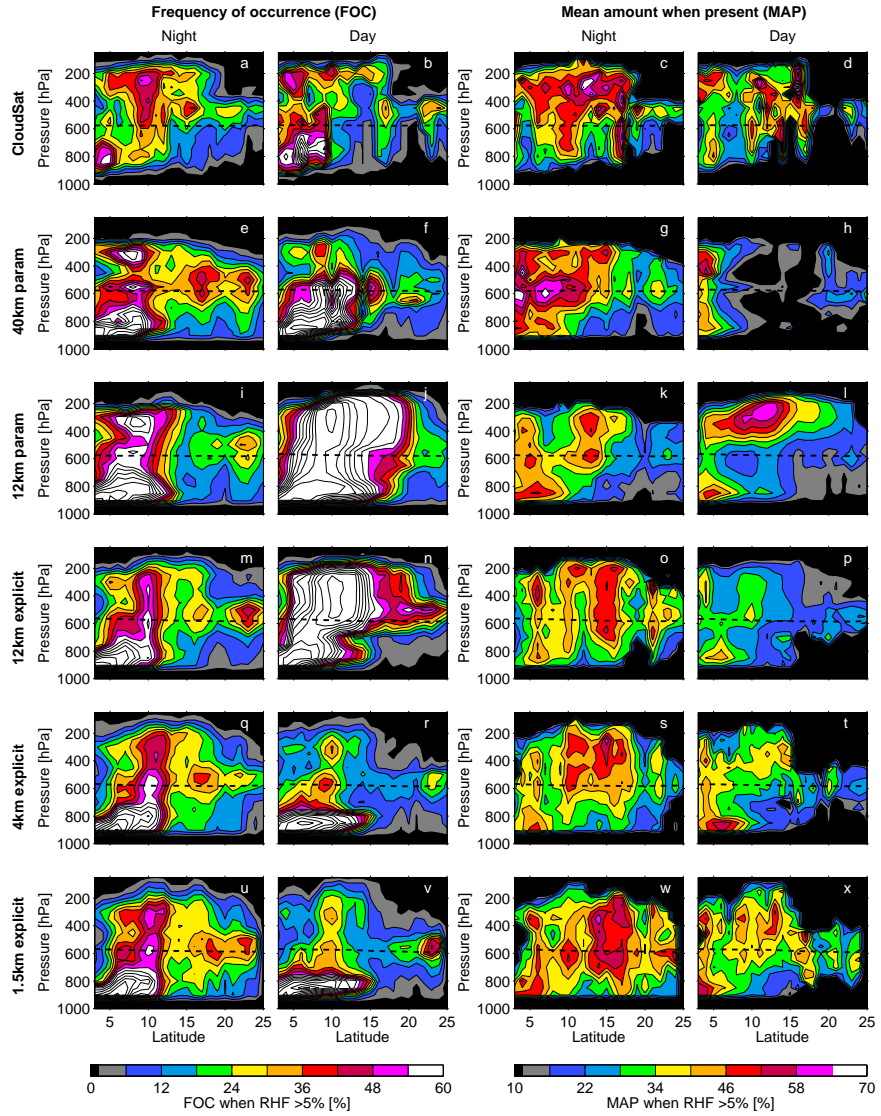
25

Figure 3: Radar hydrometeor frequency of occurrence (FOC) and mean amount when present (MAP) (RHF threshold $t = 5\%$) for nighttime overpasses and 0100 UTC model data (first and third column) and for daytime overpasses and 1300 UTC model data (second and fourth column). The dashed line indicates the mean pressure of the 0°C level. Frequency contours above 60% are spaced by 6%.
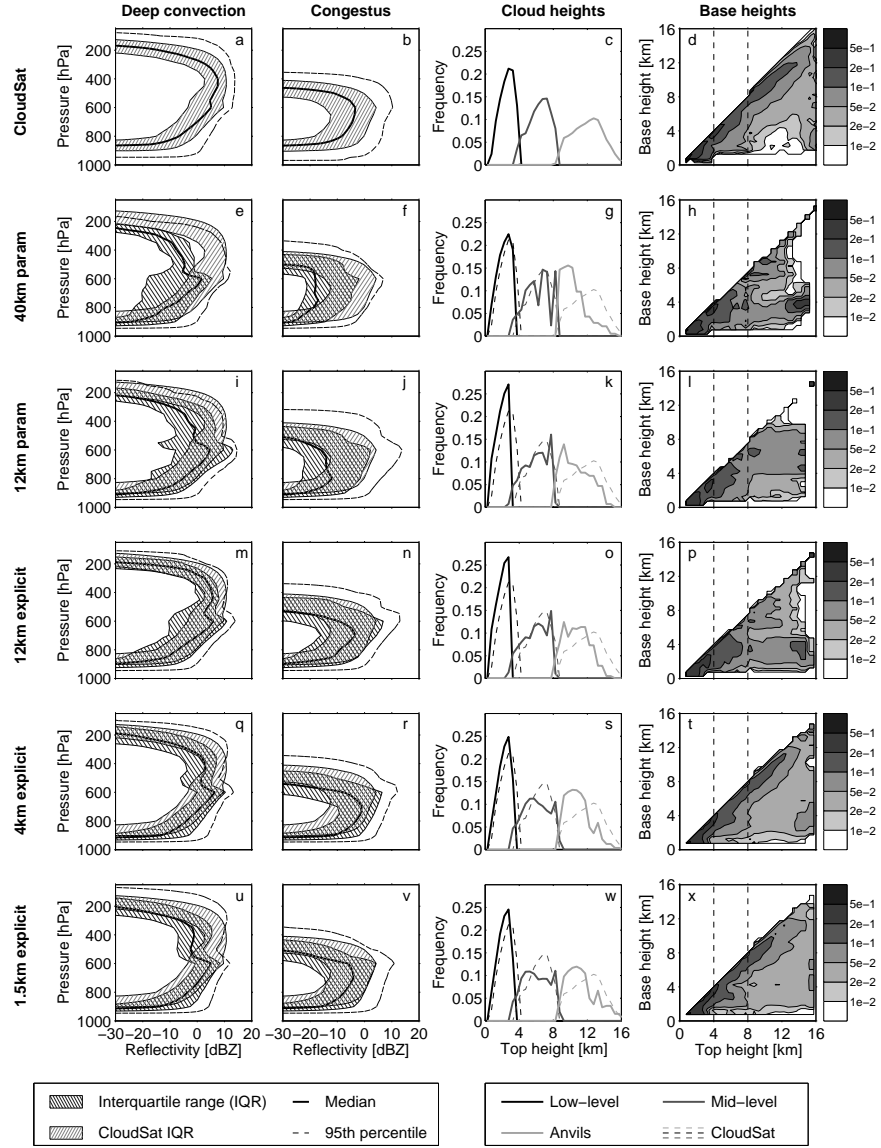
Figure 4: First and second column show the cumulative frequency of reflectivity with pressure for cumulonimbus profiles and congestus profiles, with the interquartile range (downward hatched area), the median (thick) and the 95th percentile; the CloudSat interquartile range is shown as an upward hatched area in the model panels. The third column shows the cloud-top height distribution for low-level cloud (black), mid-level cloud (dark gray), and anvil (light gray); the dashed lines in the model panels show the CloudSat values. The fourth column shows the probability of cloud-base height conditional on cloud-top height for low-level cloud, mid-level cloud, and anvil profiles, with contour intervals on a quasi-logarithmic scale.
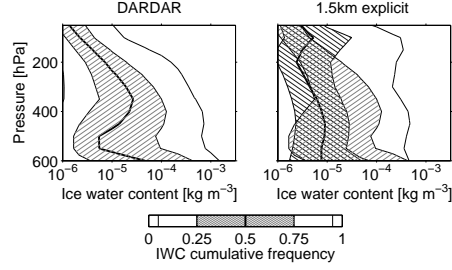
27

Figure 5: Contoured frequency with altitude (pressure) diagram for ice water content retrievals from CloudSat-CALIPSO observations (DARDAR) and grid-box mean ice water content in the 1.5-km simulation. Contours show the interquartile range (downward hatched area), the median (thick) and the 5th and 95th percentiles; the DARDAR interquartile range is shown as an upward hatched area in the model panel.
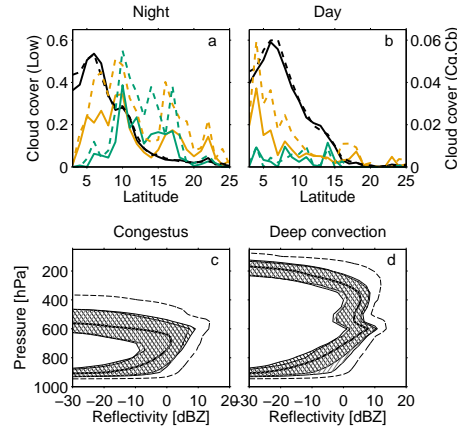

Figure 6: (a,b) Fractional cloud-type cover (FCC) as in Figure 2, but for the 4-km AB2012 simulation (solid) and the 4-km explicit simulation (dashed) restricted to 26th–29th July 2006, showing only low-level clouds (black), congestus (orange), and cumulonimbus (green); congestus and cumulonimbus FCC are scaled by a factor 10 (see right y-axis). (c,d) CFAD for cumulonimbus profiles and congestus profiles in the 4-km AB2012 simulation as in Figure 2; the upward hatched area shows the 4-km explicit simulation interquartile range, restricted to 26th–29th July 2006.

28