The three-dimensional morphology of simulated and observed convective storms over southern England

THORWALD H. M. STEIN *

Department of Meteorology, University of Reading, United Kingdom

ROBIN J. HOGAN

Department of Meteorology, University of Reading, United Kingdom

KIRSTY E. HANLEY

MetOffice@Reading, Reading, United Kingdom

JOHN C. NICOL

National Centre for Atmospheric Science, Department of Meteorology, University of Reading, United Kingdom

HUMPHREY W. LEAN

MetOffice@Reading, Reading, United Kingdom

ROBERT S. PLANT

Department of Meteorology, University of Reading, United Kingdom

PETER A. CLARK

Department of Meteorology, University of Reading, United Kingdom

CAROL E. HALLIWELL

MetOffice@Reading, Reading, United Kingdom

* *Corresponding author address:* Thorwald Stein, Department of Meteorology, University of Reading, Lyle Building, PO Box 243, Reading, RG6 6BX, United Kingdom.

E-mail: t.h.m.stein@reading.ac.uk

ABSTRACT

A set of high-resolution radar observations of convective storms has been collected to evaluate such storms in the UK Met Office Unified Model during the DYMECS project (Dynamical and Microphysical Evolution of Convective Storms). The 3-GHz Chilbolton Advanced Meteorological Radar was set up with a scan-scheduling algorithm to automatically track convective storms identified in real-time from the operational rainfall radar network. More than 1,000 storm observations gathered over fifteen days in 2011 and 2012 are used to evaluate the model under various synoptic conditions supporting convection. In terms of the detailed three-dimensional morphology, storms in the 1500-m grid-length simulations are shown to produce horizontal structures a factor 1.5–2 wider compared to radar observations. A set of nested model runs at grid lengths down to 100m show that the models converge in terms of storm width, but the storm structures in the simulations with the smallest grid lengths are too narrow and too intense compared to the radar observations. The modelled storms were surrounded by a region of drizzle without ice reflectivities above 0 dBZ aloft, which was related to the dominance of ice crystals and was improved by allowing only aggregates as an ice particle habit. Simulations with graupel outperformed the standard configuration for heavy-rain profiles, but the storm structures were a factor 2 too wide and the convective cores 2 km too deep.

1. Introduction

The forecasting of convective storms is a fundamental issue in numerical weather prediction (NWP) models. A number of operational forecast centers now run NWP models at convection-permitting resolution of order 1 km (e.g. Lean et al. (2008); Baldauf et al. (2011)). Models at such resolutions perform better in terms of the diurnal cycle of convection over land and the distribution of rainfall rates compared to coarser NWP models, which are typically run with a convection parameterization scheme (e.g., Weusthoff et al. (2010)). However, even at these high resolutions, NWP models frequently have difficulty accurately representing convection. For instance, convection-permitting simulations may precede or lag observations in terms of convective initiation (Kain et al. 2008; Clark et al. 2013), fail to develop organized convection beyond the mesoscale (Holloway et al. 2012; Pearson et al. 2013), produce wide-spread light precipitation when it is not observed (Lean et al. 2008), or organize precipitation into fewer larger cells when widespread showers are observed (Baldauf et al. 2011). In order to improve model representation of convection, a better understanding of the morphological behaviour of convective storms is required from both models and observations.

A number of recent studies have analysed high-resolution model performance in convective situations using storm-tracking methods in radar rainfall data (May and Lane 2009; Herbort and Etling 2011; McBeath et al. 2014). In particular, radar data gathered in the Tropical Warm Pool-International Cloud Experiment (TWP-ICE, May et al. (2008)) has been used in several model-intercomparison studies that evaluate model ice and precipitation from convective storms (Varble et al. 2011; Fridlind et al. 2012; Caine et al. 2013). However, these studies were mostly restricted to macrophysical characteristics, such as rainfall areas or cloud-top heights, and were restricted by brief observation periods. As part of the DYMECS project (Dynamical and Microphysical Evolution of Convective Storms), this paper presents a combined statistical analysis of the morphology (height-varying width and intensity) of convective storms in models and observations in southern England.

During the DYMECS project, volume scans of convective storms were collected over forty days in 2011–2012 with the Chilbolton Advanced Meteorological Radar (CAMRa, Goddard et al. (1994a)). The use of radar to construct three-dimensional storm structures is well-established, and long-standing algorithms exist that generate storm statistics (e.g., Dixon and Wiener (1993); Steiner et al. (1995); Potts et al. (2000)). However, CAMRa's beamwidth of 0.28° allows for analysis of storm structures on finer scales than with conventional radars, which have beamwidths of the order of 1° or more. Furthermore, the minimum detectable signal is approximately −10 dBZ at 50 km and 0 dBZ at 150 km, so that the analysis can focus on the ice cloud and anvil structures of storms, in addition to the precipitating cores. These high-quality radar data are fundamental in providing a thorough evaluation of the morphology of storms in high-resolution models.

The paper is organized as follows: section 2 describes how storms were tracked in real-time using CAMRa and how three-dimensional volumes were reconstructed for model evaluation. The Met Office model configurations used in this paper are described in section 3, with a focus on the cloud and precipitation schemes. The first set of results concern the three-dimensional structure of storms and are presented in section 4, including an analysis of anvil occurrence in southern England. The three-dimensional structures reveal a discrepancy between the model ice cloud and precipitation, which is investigated further by conditioning vertical profiles of reflectivity on rainfall rates in section 5. Finally, a discussion of the results is given in section 6.

2. Observations

CAMRa is a 3-GHz (S-band) dual-polarization Doppler radar, calibrated with an uncertainty of less than 0.5 dB (Goddard et al. 1994b). Its large 25 m antenna results in a very high spatial resolution and high sensitivity, but also limits the scan rate to $2^{\circ}\mathrm{s}^{-1}$, making it unsuitable for 360°-volume scans for the purpose of studying convection. Instead, a real-time storm-tracking and scan-scheduling procedure was developed in the DYMECS project to automatically steer the radar to scan regions of interest as described below. This enabled the radar to be operated unmanned on forty separate convective days.

a. *Real-time tracking and storm selection*

The tracking algorithm developed specifically for DYMECS provides real-time information on the location of rainfall features relative to Chilbolton, as well as the speed and direction of propagation of these features. The UK Met Office radar composite provides rainfall estimates on a 1 km horizontal grid and is updated every 5 minutes; this will be referred to as the rainfall composite and was used as the rainfall input for the tracking algorithm. The rainfall is estimated from the Met Office network of C-band radars, which are calibrated regularly to rain-gauge data (Harrison et al. 2011).

For a rainfall composite image at time t_i (with dimensions 400 km $\times$ 400 km, centered on Chilbolton), the tracking algorithm goes through several steps outlined below.

i. Rainfall features are labelled using the *local table method* (Haralick and Shapiro 2002). In this method, a label matrix $\mathcal{L}$ is generated line-by-line and left-to-right, labelling individual pixels if their rain rate is above a given threshold. For each line an equivalence table registers whether a new region $\mathcal{S}$ is adjacent to existing regions in the previous line, and is then used to set the region label of $\mathcal{S}$ to the lowest identifier of all its adjacent regions. If adjacent to more than one region, further equivalences are resolved by repeating the routine right-to-left and bottom-up. When tracking with CAMRa, this method was typically applied using a minimum feature size of 4 km^2 and a typical rainfall-rate threshold of 1 mm hr^{-1}.

ii. To track features from one rainfall-composite image to the next, a velocity field is required to project the features identified at time t_i to t_{i+1}; the method described below is based on the "tracking of radar echo with correlations" (TREC, Rinehart and

Garvey (1978)). To construct this velocity field, $\mathcal{V}(t_i, t_{i-1})$, the cross-correlation of the rainfall images at times t_i and t_{i-1} is calculated using the two-dimensional fast Fourier transform, for 50×50 km boxes, each box separated by 25 km. The displacement associated with the maximum correlation is recorded for each box, resulting in x- and y-displacement fields at 25 km grid length. After the removal of outliers beyond two standard deviations from the mean, both x- and y-displacement fields are linearly interpolated to the 400 km $\times$ 400 km grid. The velocity field is then generated from these displacements, taking into account the time difference between the two images (typically five minutes).

iii. Each labelled storm in the label matrix $\mathcal{L}(t_i)$ is advected using its average displacement from $\mathcal{V}(t_i, t_{i-1})$, after which the advected label matrix is compared against the label matrix for the next time step, $\mathcal{L}(t_{i+1})$, for overlapping storms to keep track of pre-existing storms. For this purpose, an overlap fraction threshold of 0.6 is used, as is standard in the TITAN storm-tracking method (e.g., Dixon and Wiener (1993); Han et al. (2009)).

iv. For each storm at time t_i, a list of properties is constructed, including whether it is the result of a merger or break-up of storms from time t_{i-1}, leading to a database of storms with detailed information on storm history and characteristics, including mean and maximum rainfall rate, rainfall area, as well as the location in radar coordinates relative to Chilbolton and the direction of propagation.

A second algorithm uses this real-time storm information to issue automated radar-scanning commands to CAMRa. The two main components of this second algorithm are the storm-prioritization scheme and the scan-scheduling strategies. The storm prioritization scores each storm by its size (the area of surface rainfall rate above the threshold), maximum rain rate, and azimuthal width in polar coordinates, whilst scores are reduced for properties such as radial distance to the radar (too close or too far) and azimuthal separation from the storms currently prioritized (due to the slow scan rate of the radar). New storms are only added to the list of prioritized storms if a slot is available, as priority goes to storms currently being scanned in order to capture their full evolution. Eventually, a list of three to five of the highest-scored storms is constructed and scan commands are issued in the following order:

i. Group storms by proximity, e.g. if storms are close or even overlap in azimuth, they can be scanned simultaneously.

ii. For each group of storms, perform range-height indicator (RHI) scans through the locations of a number of maximum rainfall rates above 4 mm hr^{-1}, typically 1–2 maxima per prioritized storm.

iii. For each group of storms, perform a set of stacked plan-position indicator (PPI) sector scans, spaced at least $0.5°$ in elevation, to obtain storm volumes.

Such a cycle for a single group of storms typically lasts 5–15 minutes using CAMRa, during which time the storm positions are updated with tracking information based on the latest radar composite. For the fifteen days analysed in this study, 362 volume scans were completed, containing more than a thousand storm volumes.

b. Storm volume reconstruction

The minimum detectable signal of CAMRa is approximately 0 dBZ at a range of 150 km and will be used as the reflectivity threshold for volume reconstruction. Three-dimensional volumes are constructed from sets of PPI scans by transforming the CAMRa polar coordinates to Cartesian for each individual scan, then horizontally shifting the data to a communal base time (usually the time of the rainfall composite image preceding the first PPI scan) using the velocities calculated from the cross-correlation of the rainfall composite images and assuming that the entire storm moved at a common velocity for the duration of the scanning procedure. The individual PPI scans are then concatenated and re-gridded with radar reflectivities linearly interpolated in dBZ-space on to a regular Cartesian grid (333× 333× 100 m, comparable to the radar resolution of 300 m in range and 250 m resolution in azimuth at 50 km). The horizontal shift is not expected to introduce errors to the horizontal cross-sectional area as each coordinate is shifted using the same constant velocity; the shift mainly reduces the apparent tilt induced by scanning a storm volume while it moves. Errors introduced by the linear interpolation are expected be minimal for horizontal cross-sections as the grids are of comparable resolution, whilst linear interpolation in dBZ-space in the vertical will smooth out cloud edges, though the latter should not impact our results as the statistical analysis will be performed on a coarser vertical resolution.

A volume scan regularly contains multiple storms (see previous section), which need to be distinguished to identify their individual heights and widths. Therefore, a threshold of 4 mm hr^{-1} is used to identify individual storms in the rainfall composite contemporaneous to the volume scan and subsequently in the CAMRa data. Although a rainfall rate threshold alone is not sufficient to distinguish between convective and stratiform rainfall, the 4 mm hr^{-1} threshold is approximately equivalent to a 33 dBZ reflectivity threshold and should therefore encompass convective rainfall areas traditionally identified with thresholds between 35–40 dBZ (Biggerstaff and Listemaa 2000).

To include drizzling parts of the storm and possible anvil cloud, all (rain and no-rain) pixels in the rainfall composite within 25 km of a labelled storm are given the same identifier as the storm if it is their nearest storm. The storm-neighboring regions thus generated are then interpolated to the surface-only Cartesian grid associated with the volume scan using the nearest-neighbor method. For each volume scan, values outside a storm's neighboring region are excluded when reconstructing that storm volume. The storm volume is then constructed bottom-up, starting with the location of the rainfall feature identified in the rainfall composite. At each vertical level, areas with radar reflectivity above 0 dBZ are identified and all such areas overlaying any part of the storm identified in the level directly below are included in the storm volume. This way, any unconnected cloud or rainfall features in the storm-neighboring region are excluded from the three-dimensional reconstruction of the storm if they did not overlap the storm at any vertical level, whereas expansive cloud and anvil regions are included if within the storm-neighboring region and attached to the storm.

Not all volumes observed are considered for this study as many storm-neighboring regions are partly observed either due to the actual storm being close to the edge of the azimuthal swath observed, or being too close to the radar and cut off by the scan with highest elevation. For a storm to be considered, firstly, the lowest-elevation scan must observe the storm at an altitude below 2 km, whilst the highest-elevation scan must overshoot the storm: that is, no values above 0 dBZ should appear above the labelled region in the highest-elevation scan. Secondly, the scanned depth of the storm (the maximum height minus the minimum height of $Z \geq 0$ dBZ) divided by the number of individual scans with dBZ ≥ 0 must be less than 1 km, to ensure an adequate representation of the vertical storm structure. Combined with the minimum PPI spacing of 0.5° in elevation, this constraint implies that storms beyond 100 km in range are typically excluded. Finally, using the storm-neighboring region described above, storms are only included if at least two-thirds of the storm-neighboring region falls within the azimuth swath scanned by the radar.

3. Models

The model simulations in this study were performed with the Met Office Unified Model (UM) Version 7.8. For all DYMECS cases, the UM was run using the configuration of the Met Office 1500-m forecast model (UKV) that was operational between 20 July 2011 and 17 January 2012. The UKV is a limited-area model nested within the Met Office North Atlantic and European (NAE) model of 12 km grid length. It has a horizontal grid length of 1500 m in the inner domain covering the UK and Ireland and 4 km grid length in the outer domain with a variable grid length in the transition region. This variable grid allows the UKV to run over a larger domain without the need of an intermediate, separate model. The UKV runs without a convection parameterization scheme and has 70 vertical levels with a top at 40 km; the stretched vertical grid has spacing of approximately 100 m at 1 km height and 300 m at 8 km. The DYMECS simulations of the UM at 1500-m grid length — using the UKV grid configuration — were initialised from the 0400 UTC operational UKV analysis (the output of a three-hour data-assimilation cycle) with lateral boundary conditions provided by the 0000 UTC NAE forecast.

Additional simulations were analysed for 25 August 2012, namely one-way nested UM configurations at 500-m and 200-m horizontal grid length (Hanley et al. 2014), while a 100 m grid-length simulation was one-way nested within the 200-m model; these three simulations were run with 140 vertical levels. All simulations were analysed on a 200×200 km grid centered on Chilbolton, apart from the 100 m grid-length simulation, which was analysed on a 140×140 km grid centered on Chilbolton. Furthermore, at 1500-m grid length and at 200 m grid length, a configuration was run with prognostic graupel (used operationally in the UKV since 16 January 2013) as well as a configuration with all ice set to aggregates.

Sub-grid mixing in the 1500-m grid length simulations was treated using the Lock et al. (2000) first-order non-local boundary-layer scheme with local, moist Richardson-number-based vertical mixing above the diagnosed boundary layer and a Smagorinsky-Lilly-based horizontal mixing scheme with a mixing length of 300 m, which also takes moist Richardson number into account. The higher-resolution configurations used the Smagorinsky-Lilly-based scheme in 3D, with a ratio of mixing length to grid scale of 0.2. For further details regarding the model configuration and the high-resolution simulations in the DYMECS project, see Hanley et al. (2014).

The UM uses a single-moment microphysics scheme (Wilson and Ballard 1999) with mixing ratios of cloud ice and cloud liquid as prognostic variables, since then developed to include prognostic rain; graupel is available as an additional prognostic variable but is only included in this study where explicitly mentioned. The large-scale precipitation scheme contains a diagnostic split between ice crystals and aggregates, both of which are modelled with a gamma distribution to describe particle sizes (Cox (1988), see Table 1); precipitating ice is diagnosed from the ice mixing ratios but does not exist as a separate prognostic. The mass-diameter relationship for aggregates is based on Locatelli and Hobbs (1974) and for crystals on Mitchell (1996) (see Table 1). The rain particle-size distribution is based on Marshall and Palmer (1948), and the graupel parameterization follows from Ferrier (1994).

The UM has an option to treat crystals and aggregates as separate prognostic variables, but this has not been used in this study. The diagnostic split between ice crystals and aggregates assumes that the cloud-ice mixing ratio $q_{\rm cf}$ is separated between these two habits using (Wilkinson et al. 2011):

$$f_{\rm a} = 1 - \exp\left\{-0.0384\,[T - T_{\rm ct}]\,\frac{q_{\rm cf}}{q_{\rm cf0}}\right\}\ , \tag{1}$$

with $f_{\rm a}$ the fraction of aggregates, $T_{\rm ct}$ the temperature at the top of the cloud layer, and $q_{\rm cf0} = 10^{-4}$ kg kg^{-1}. For an ice mixing ratio of $q_{\rm cf} = q_{\rm cf0}$, this fraction is less than 50% at temperatures within 18°C of the cloud top temperature. For precipitating clouds with ice-cloud tops within 3 km of the 0°C level, this implies that at all heights, the majority of the ice mass is assumed to be crystals, in contrast with observations (e.g. Field (1999)). It will be shown in section 4 that this leads to lower reflectivities in parts of the ice cloud of convective storms than observed. Therefore, a separate run has been included in this study for which the diagnostic split between the two ice particle habits is turned off, and all ice is treated as aggregates; this will be referred to as the “no crystals” configuration. It should be noted that aggregates and crystals have different fall-speed-diameter relationships, which for a given ice water content leads to higher precipitation rates when all ice is assumed to be aggregates than when some of it is crystals. For the simulations considered in this study, this led to a 10% increase in domain-averaged precipitation in the “no crystals”-configuration during the peak of convective activity compared to the standard run.

For this study, the model hydrometeor fields have been converted to radar reflectivities, to enable like-with-like analysis against the radar data (McBeath et al. 2014). The reflectivity forward model assumes Rayleigh scattering for the radar wavelength considered ($\lambda = 10$ cm) and is described in appendix A. Due to the long wavelength, effects of attenuation were not considered in the forward model. Storm volumes from the model were reconstructed from the simulated reflectivities following the method outlined in section 2b.

a. Representativeness of DYMECS cases

The observational strategy during the DYMECS project was to scan for several hours on any day for which the UKV forecast showed rainfall from convective cells in the vicinity of Chilbolton. This non-discriminatory approach to scanning days has resulted in a varied selection of convective weather situations, including post-frontal storms and isolated convection, but lacking mesoscale convective systems, which are relatively rare in the UK (Lewis and Gray 2010). The observation period includes April 2012, which had 2.3 times the average rainfall for England and Wales (Kendon et al. 2013). The subset of fifteen cases discussed in this paper were selected as they had a 0°C level at least 1 km above the surface and a substantial number of storms extending above the 0°C level within 100 km of the radar; the cases in this subset are well-distributed among the DYMECS period.

We expect that the results of the model evaluation presented in this paper are representative of convective storms in the UK and could be extended to the mid-latitudes in general. Furthermore, the evaluation of the 1500 m UM simulation for 25 August 2012 leads to qualitatively similar conclusions to this model evaluation for the remaining fourteen cases. This suggests that although this single case is not representative of the DYMECS period, the performance of each model configuration for this case may be assumed as typical for its general representation of convective storms.

4. Three-dimensional structures

In this section, the models are evaluated against the observed storm structures in terms of quartiles of storm radius with height for different reflectivity thresholds. Only storms with rain rates of at least 4 mm hr^{-1} over a contiguous surface area of 4 km^2 are included in the statistics. Storm heights are considered relative to the height of the 0°C level, which for each case was determined from the height of the 0°C-isotherm at 1300UTC at the grid point nearest to Chilbolton in the 1500 m UM simulation (see Table 2). For a previous version of the UM at 12-km resolution, a similar derivation of freezing level height using wet-bulb temperature had a root-mean-squared error less than 200 m (Mittermaier and Illingworth 2003), which will be assumed an upper bound for this error in the UM version used in this study. The centering of height on the 0°C level allows for a clear distinction between ice cloud and rainfall when storms from different days are combined.

The reflectivity thresholds used for the analysis are 0 dBZ, 20 dBZ, and 40 dBZ, chosen to represent the structure of ice cloud and anvil, the stratiform part of the storm, and the convective core, respectively. In terms of rainfall rates, assuming $Z[mm^6m^{-3}] = 200R^{1.6}$ (Marshall and Palmer 1948), these thresholds relate to approximately 1 mm day^{-1} (drizzle), 1 mm hr^{-1} (light rain), and 12 mm hr^{-1} (heavy rain). In terms of the frozen part of the storms, using the relationships between ice water content, reflectivity, and temperature from Hogan et al. (2006) at $-20\,°C$, 0 dBZ relates to ice water contents of approximately 0.05 g m^{-3}; 20 dBZ to about 0.8 g m^{-3}; and 40 dBZ to 12 g m^{-3}. It should be noted that Waldvogel et al. (1979) used a 45-dBZ threshold at 1.4 km above the 0°C level for hail detection, so that observed ice cloud with reflectivities higher than 40 dBZ can be assumed to contain graupel.

Within the database of storm structures, storms are separated by the cloud-top height above 0°C level, so that rather than cloud-top height, ice-cloud depth (ICD) is reported. ICD categories are set at ≤ 4 km for "shallow" storms, 4–6 km for "intermediate" storms, and > 6 km for "deep" storms; the two thresholds relate to temperatures of approximately

$-25°C$ and $-40°C$, respectively. These thresholds were chosen to obtain a proportional split between categories in terms of number of observed storms for a clear distinction between the rarer deep storms and the more common shallow storms. As shown in Table 2, over these fifteen DYMECS cases, 63% of the observed storms with rain rates above 4 mm hr^{-1} fall in the shallow category, 31% are storms of intermediate height, and 6% are deep. The storms identified in the UM simulations were categorized using the same ICD thresholds as for the observations.

a. Storm statistics over all fifteen DYMECS cases

Figure 1 shows the interquartile range of equivalent radius with height at different reflectivity thresholds, calculated from the storm structures observed during the DYMECS cases listed in Table 2. The equivalent radius of a reflectivity threshold at a given height is defined as the radius of the circle with an area equivalent to that of the storm region above the reflectivity threshold at that height.

The storm structures in the UM at 1500-m grid length (bottom row) are broader than the observed structures for all ICD categories, as well as for each reflectivity threshold and at nearly all heights. The model quartiles show a broadening from the top down with a sharp increase in width at the 0°C level, especially for the 0-dBZ contours, indicating a large area of drizzle surrounding the storms, despite a lack of cloud with $Z \geq 0$ dBZ aloft; this feature will be analysed further in section 5. The model median and 75th percentiles of the 40-dBZ threshold (panel f) do not persist as far into the frozen part of the cloud as observed with the radar. However, the median equivalent radius of the 40-dBZ threshold in the observations (panel c) suggests that these cores are comparable in size to the model grid length of 1500 m and are therefore unlikely to be represented well by the model in this configuration, whereas higher resolution models should start to resolve features at 1-km scales.

For both the model simulations and the radar observations, the medians of the 0-dBZ and 20-dBZ thresholds suggest that the deepest storms are marginally larger than those in the intermediate ICD category (e.g., López (1976)), although both overlap in interquartile range at 1 km above the 0°C level. The shallow storms however are shown to be narrower, with median equivalent radius at 1 km at the 0-dBZ and 20-dBZ thresholds a factor 1.5 smaller than these radii for intermediate storms.

b. Sensitivity to model horizontal grid length

During the DYMECS case of 25 August 2012, a large number of storms with ICD > 6 km were observed; these were the tallest storms in absolute height over all cases considered, reaching up to 10 km above mean sea level. This case was chosen to study the possible improvement in storm structures with decreasing model grid length and with different settings in the ice microphysics scheme. Figure 2 shows the storm structures for this case as observed by the radar and simulated in the UM at 1500-m and 500-m grid length; Figure 3 shows the structures simulated in the UM at 200-m and 100-m grid length.

For shallow storms, the radii at 1 km in the 1500-m simulation are a factor 2–3 larger than observed. However, the observed shallow storms have a median radius smaller than 3 km, unlikely to be represented well by simulations at 1500-m grid length. At 500-m grid length, these storms are still a factor 2 larger than observed, but shallow storms in the 200-m and 100-m grid-length simulations are of similar size to those observed. Out of the four simulations with the standard ice microphysics, the 100-m model appears to best match the observations for shallow storms.

There is a tendency towards narrower storm structures as model grid length decreases from 1500 m (second row, Figure 2) to 500 m (third row, Figure 2) to 200 m (first row, Figure 3). The 500-m simulation best represents the 0-dBZ equivalent radius in intermediate and deep storms, which at 1 km above the 0°C level are a factor 1.5–2 wider than observed in the 1500-m grid-length simulation. The median radii of these storms in the 200-m simulation are narrower than those observed (first row, Figure 2), by factors of 1.8 and 1.4 at 1 km above the 0°C level, respectively. Hanley et al. (2014) show that in the 200-m simulation, storms with equivalent radius below 5 km typically have higher average rain rates than observed, which is reflected in the storm morphology for instance by the wider radius of the 40-dBZ and 20-dBZ contours relative to the 0-dBZ contour in the shallow and intermediate storms, compared to the observations. The convective cores in deep storms are represented well by the 200-m simulation and slightly too broad in the 500-m and 1500-m simulations, though again it should be noted that their observed scales are smaller than 3 km.

The 100-m model (Figure 3, bottom row) has storm structures that are similar to the 200-m simulation. This suggests that the representation of bulk properties as represented by these metrics has become independent of model resolution; the simulated morphology of convective storms in the Met Office models has "converged" at 200-m grid length. These two models also represent the width of the 40-dBZ contour in deep storms well, suggesting that at grid lengths of 200 m or smaller, convective cores can be resolved. These results are consistent with previous studies of convection in high-resolution simulations (e.g., Bryan et al. (2003); Bryan and Morrison (2011)). However, the convective cores in the shallow and intermediate storm structures in these models are larger and more frequent than observed, confirming that these storms are too intense (Hanley et al. 2014). Furthermore, Hanley et al. (2014) show that the rainfall-area size distribution for a given grid-length varies with the mixing length chosen for the subgrid turbulent mixing scheme, which suggests that the invariance of the model at high resolutions is sensitive to model formulation.

For each of the simulations in Figures 2 and 3, the 0-dBZ contour increases by a factor 1.2 (typically 2–3 km) across the 0°C level, as seen in Figure 1. This feature is therefore likely a result of the ice-microphysics parameterization and cannot obviously be resolved by increasing the model resolution, although it becomes less distinct in the 100-m simulation.

c. Sensitivity to model ice microphysics

Additional storm-structure statistics for the case of 25 August 2012 are shown in Figure 4 for UM configurations with changes to the ice-microphysics parameterization. The major difference between the no-crystals simulations in the top two rows of Figure 4 and the standard configurations in Figures 2 and 3 appears in the precipitating part of the storms. The no-crystals runs do not have a noticeable (sharp) increase in median equivalent radius of the 0-dBZ contour across the 0°C level and are therefore more similar to observed storms in this respect. This difference between the no-crystals and the standard configuration is noticeable at both 1500-m grid length and at 200-m grid length, providing further proof that this feature is due to ice-microphysics parameterization and not model resolution. Note that the no-crystals simulations do not show any improvement in median storm structure at other heights, although a recategorization can be noted, as substantially more storms are in the deep category in the no-crystals simulations than in the standard configuration at both 1500-m grid length and at 200-m grid length. This is due to higher reflectivities near cloud tops, as the ice now consists solely of aggregates instead of a mixture of crystals and aggregates, and thus the 0-dBZ contour may reach higher for the same ice water content in the no-crystals simulation than in the standard configuration.

The effect of the no-crystals configurations around the 0°C level could be expected, because for the same ice water content, the no-crystals configuration will have higher forward-modelled reflectivities than the standard configuration, as aggregates have replaced crystals; this also holds for the simulation with prognostic graupel for low ice water contents, where graupel will not be present. The relationship between the reflectivities of ice and rain will be investigated further in section 5.

The simulations with graupel have the expected effect of a deeper and broader core exemplified by the 40-dBZ contour compared to the standard configuration. At both 1500-m grid length and at 200-m grid length, the cores in the graupel simulations are too deep in all three storm categories and too broad. The deep storms in the 1500-m graupel simulation have also widened 0-dBZ and 20-dBZ contours compared to the 1500-m standard configuration in Figure 2f. This difference is not obvious when comparing deep storm structures in the 200-m simulations, though the 20-dBZ contour is broader near cloud top in the graupel simulation and reaches above 6 km, deeper than the standard configuration. Using these metrics, the graupel simulations perform worse than the standard UM configuration for the case of 25 August 2012.

d. Anvil occurrence

The statistical evaluation in Figures 1, 2, 3, and 4 masks the occurrence of anvil cloud. To study anvil occurrence, a storm is defined to have an anvil when the ratio between its maximum 0-dBZ equivalent radius above 2 km above the 0°C level and the equivalent radius at 1 km above the 0°C level is at least 1.05; this ratio will be referred to as the anvil factor. The masking of anvils in the figures mentioned above is due to several contributing factors, for example the varying heights of anvil over all the DYMECS cases, the varying anvil characteristics during a storm life cycle, and a generally low frequency of anvil occurrence or generally low anvil factors over southern England. In order to study the anvil characteristics for the DYMECS project, in this section the analysis is confined to only those three-dimensional structures which exhibit an anvil cloud. The analysis is performed for 25 August 2012. All storms with ICD $>$ 4 km are considered, so both intermediate and deep storms contribute to the statistics.

In Figure 5, the probability density of anvil factors is shown, as well as the anvil probability for given times of day, averaged over a three-hourly window. The anvil factor distribution appears exponential in all model configurations, with the 1500-m simulation failing to produce anvil factors above 1.7, although only a single larger anvil was observed on that day. None of the models reproduce the daily cycle of anvil probability of occurrence very well. This cycle appears lagged by 2–3 hours in all three simulations compared to the observed peak at 1300UTC. The 1500-m simulation shows a morning peak, possibly due to spin-up from the model initiation time at 0400UTC.

5. Vertical profiles of reflectivity factor

The strong increase of equivalent radius across the 0°C level, which is not seen in the observations nor in the no-crystals simulation, suggests that the microphysical relationship between ice and rain as exhibited by their reflectivities is different in the standard model configuration compared to observations. Probability distribution functions (PDFs) of reflectivity versus height were constructed from vertical profiles of reflectivity (similar to "contoured frequency by altitude diagrams", Yuter and Houze Jr (1995)) conditioned on the mean reflectivity value observed between 0.2–1 km below the 0°C level ("rain reflectivity"). For a single vertical profile, the first level at which Z $<$ 0 dBZ was considered the cloud top; unconnected layers above, for instance due to an overhanging anvil, were thus excluded. Shear or other dynamical features that may affect the reflectivity structure inside a storm were ignored. The PDFs were conditioned on rain reflectivities between 0–5 dBZ, 20–25 dBZ, and 40–45 dBZ to evaluate the UM against radar observations under different rainfall conditions.

a. Storm statistics over all fifteen DYMECS cases

The PDFs of reflectivity versus height using the data from all cases listed in Table 2 are shown in Figure 6 for the radar observations (top row) and the UM at 1500-m grid length (bottom row). The drizzling profiles (left) show more frequent ice cloud with Z $\geq$ 0 dBZ in the observations than in the UM, highlighted by the 75th percentile of reflectivity. The lack of ice cloud in drizzling profiles agrees with the large drizzle region surrounding the storms in the 1500-m simulations in Figure 1. The light-rain profiles (middle) show a similar model error of too few ice reflectivities above 0 dBZ, highlighted by the quartiles at lower values compared to observations, particularly above 1 km. The heavy-rain profiles (right) show the model 25th, 50th, and 75th percentiles within 5 dB of the observed values, though all drop below 0 dBZ too soon, suggesting that heavy rainfall results from relatively more shallow profiles in the model compared to observations. The underlying PDFs for heavy rain show another discrepancy between model and observations, with observed values above 40 dBZ up to 4 km above the 0°C level, whereas the model only rarely produces such reflectivities above the 0°C level and only up to 2 km, which agrees well with the structure of convective cores discussed in section 4.

The differences in these PDFs between the 1500-m model and the radar observations are most striking in the ice-cloud part of the drizzle and light-rain profiles. The low frequency of ice reflectivities above 0 dBZ at higher levels can be partly explained by the diagnostic split between ice crystals and aggregates. Using equation (1) and the derivation in appendix A, it can be shown that at $T = -10$°C and for an ice mixing ratio of 10^{-4} kg kg^{-1}, an increase in fraction of aggregates from 0.1 to 0.2 (0.9 to 1.0) will increase reflectivities by 1.76 dBZ (0.76 dBZ). This should mostly affect precipitating profiles with low cloud-ice tops, which in the simulations with standard ice microphysics will have more than 50% of their mass as ice crystals.

b. Sensitivity to model ice microphysics and horizontal grid length

Figures 7, 8, and 9 show the PDFs of reflectivity versus height for the 25 August 2012 case as observed by the radar and simulated in the UM at the same grid lengths and configurations analysed in Figures 2, 3, and 4. For the drizzling profiles (left columns), none of the model configurations produce high enough reflectivities in the ice part to generate a similar distribution to the observations. However, for the no-crystals simulation at 1500-m grid length in Figure 9, more than 25% of drizzling profiles have ice reflectivities above 0 dBZ up to nearly 2 km. Whilst this is still below the height observed for drizzling profiles, it suggests an improved relationship between ice reflectivities and rain reflectivities. The impact of the no-crystals configuration on drizzling profiles is reduced in the 200-m grid-length simulation.

The model PDFs for light rain (middle columns of Figures 7, 8, and 9) indicate a bi-modal distribution of (1) a shallow mode (ICD $\leq$ 4 km) with low ice reflectivities (Z $<$

20 dBZ) and (2) a deeper stratiform mode (ICD $>$ 4 km) with relatively high ice reflectivities (Z $\geq$ 20 dBZ), both roughly distinguished by the 75th percentile. The observed PDF instead exhibits a broad peak, associating light-rain profiles with higher ICD than in the models. As with the drizzling profiles, compared to the standard configuration, the no-crystals simulations in Figure 9 have a slight increase of the height where the 75th percentile reaches 0 dBZ. This is due to higher reflectivities from aggregates near cloud-top, rather than the model simulating actually deeper clouds. For the same reason, the graupel simulations in Figure 9 have the 75th percentile remain above 0 dBZ at higher altitudes, though for both microphysical changes, this improvement is minimal in the 200-m grid-length simulations.

The PDFs of heavy-rain profiles (right columns of Figures 7, 8, and 9) show a reasonable representation in the models of the broad distribution of reflectivities with height, as the 75th percentile remains within 5 dB of the observed quartile for the standard configuration and the no-crystals simulations, though the 25th percentile and the median still drop below 0 dBZ about 1 km before the observed quartiles. The shallow mode dominates the heavy-rain profiles in the 200-m and 100-m simulations, as the medians drop below 0 dBZ at approximately 2 km, compared to approximately 5 km in observations; no improvement in the PDFs was found when the high-resolution models were analysed on a 1500-m horizontal grid. The dominance of a shallow mode in heavy-rain profiles in these simulations agrees with the morphology of shallow and intermediate storms in these models in Figure 3, which feature a prominent convective core. Again, the no-crystals configuration shows a slight increase in height for the different quartiles.

For all rain categories, the PDFs for the simulations with prognostic graupel resemble the standard configuration for low reflectivities, whilst for reflectivities greater than 20 dBZ, the graupel PDFs tail towards higher values. For heavy-rain profiles, the graupel simulation at 1500-m grid length best resembles the observed PDF out of all the model simulations. At 200-m grid length however, graupel is produced too frequently, leading to too high reflectivities at all quartiles. Interestingly, heavy-rain profiles in this simulation are associated with cloud-top heights similar to observations, as the median and 75th percentile remain above 0 dBZ up to approximately 4.5 km and 6 km. It can be concluded that the inclusion of graupel as a prognostic variable improves the reflectivity profiles for the heaviest precipitation, although with too high reflectivities in the 200-m graupel simulation, which agrees with the structure of convective cores discussed in section 4.

c. Relationship between ice and rain reflectivities

The discrepancy between ice and rainfall is investigated further by conditioning vertical profiles on the mean reflectivity in the ice part of the cloud, here defined as the mean reflectivity value between 1.2–2 km above the 0°C level. Assuming this "ice reflectivity" is a proxy for ice water content, the distribution of the rain reflectivity conditional on the ice reflectivity should indicate whether the models produce too high or too low reflectivities for given cloud-ice conditions. In Figure 10, the interquartile range for the conditional distribution is shown for observations (gray in all panels), the 1500-m simulation with standard ice microphysics, the 1500-m simulation with prognostic graupel, the 1500-m no-crystals simulation, and the 200-m simulation with the standard ice-microphysics set-up. Results for the 500-m and 100-m simulations are similar to those for the 200-m simulation and are therefore not shown. The single-moment microphysics scheme in the UM allows for a derivation of a relationship between ice aggregates and rain reflectivities using a constant flux assumption (see appendix B); this relationship is also indicated in Figure 10.

The 1500-m standard configuration (panel a), the simulation with prognostic graupel (panel b), and the 200-m simulation (panel d) frequently produce too high rain reflectivities for conditions of low ice reflectivities ($Z_{\mathrm{ice}} < 20$ dBZ). For these ice reflectivities, the median rain reflectivity for these three models is only 2 dB above the observations, equivalent to an increase in rainfall rate with a factor less than 1.5, but the 75th percentile is typically 5 dB higher than observed, equivalent to a rainfall rate increase by a factor of more than 2. The no-crystals simulation has all three quartiles approximately 5 dB lower than observed at ice reflectivities below 5 dBZ, following the slope of the constant-flux relationship, equivalent to a rainfall rate decrease by a factor 2. This suggests that the no-crystals simulation is not an obvious improvement over the standard configuration in terms of the relationship between cloud-ice and rain, although the 200-m no-crystals simulation has all three quartiles within 2 dB of the observations for this range (not shown).

For ice reflectivities between 20–30 dBZ, the no-crystals configuration shows a similar interquartile range to the standard configuration simulations, as all three follow the slope derived using the constant-flux assumption. This is expected as aggregates will dominate the ice mass at these reflectivities with the standard ice-microphysics parameterization. For the graupel simulation, at the highest ice reflectivities, the cloud is likely a mixture of (mostly) aggregates and graupel, so that a given ice reflectivity in the graupel simulation relates to a smaller ice water content than if all ice were aggregates. Thus, for a given ice reflectivity, a lower rain reflectivity is generated than if no graupel were included in the model. The same result is obtained for the 200-m graupel simulation (not shown).

6. Discussion and Conclusions

This study has presented a unique evaluation of convective storms over southern England simulated by the Met Office models. Radar volume scans targeted at individual storms have been used to simultaneously evaluate the three-dimensional storm morphology as well as the vertical distribution of hydrometeor concentrations inside such storms. The Met Office forecast model at 1500-m grid length (UKV) was evaluated against radar observations made with the 3-GHz Chilbolton radar, which included more than 1,000 storms observed over fifteen days in 2011–2012 during the DYMECS project. For 25 August 2012, a day where many storms reached heights of 10 km, the model was run at convection-permitting resolutions ranging from 1500 m horizontal grid length down to 100 m, and with simulations studying sensitivity to ice-microphysics parameterization. Radar reflectivities were forward-modelled from the model hydrometeor fields for a like-with-like comparison.

Individual storm structures were identified using a 4 mm hr^{-1} rainfall-rate threshold and, using the cloud-top height ($Z > 0$ dBZ), these were categorized into shallow, intermediate, and deep structures. Models and observations alike showed a tendency for storm width to increase by a factor of 1.5 from shallow to intermediate structures, but the increase from intermediate to deep storms was negligible. The models at 1500-m grid length produced storm structures that, at 1 km above the 0°C level, were a factor 1.5–2 broader than observed; this factor did not depend on whether graupel was used as a prognostic variable, or whether all ice was modelled as aggregates. For all three storm categories, the models produced narrower median storm structures with decreasing grid length, although the 200-m and the 100-m simulations were hardly distinguishable.

The 1500-m simulations did not represent the width and depth of convective cores ($Z >$ 40 dBZ) in the deepest storms very well, though observations showed that these cores have typical widths comparable to the 1500-m grid length. The 1500-m simulation with prognostic graupel produced convective cores that were a factor 3 wider than observed and 2–3 km taller; at 200 m grid length with graupel, the cores were still a factor 1.5 too wide and 2–3 km taller than observed. The 200-m and 100-m simulations adequately represented the median structure of convective cores, which suggests that model representation of convective storms has "converged" at 200-m grid length, confirming expections for the simulation of moist convection (Bryan et al. 2003). However, the cloud structures ($Z \geq 0$ dBZ) in the 200-m and 100-m simulations are slightly narrower than those observed for all three storm categories, and particularly the shallow and intermediate storms are too intense, in agreement with Hanley et al. (2014), who showed that for storms with radius less than 5 km, the 200-m simulation produced storm-averaged rainfall rates a factor 3 higher than observed.

At all resolutions, the modelled storms showed an increase in radius across the 0°C level of up to 5 km due to a drizzle region without cloud-ice aloft, which did not appear in the observed structures. The "no-crystals" simulations at 1500-m and 200-m grid length, which had all ice set to aggregates instead of a mixture of ice crystals and aggregates, produced median storm structures more similar in shape to those observed, without a drizzle region. This suggests that the drizzle region without cloud-ice aloft was due to crystals dominating shallow cloud tops, which led to ice reflectivities below 0 dBZ, but could still generate rain reflectivities above 0 dBZ. When ice reflectivities were conditioned on the rain reflectivity, the 1500-m no-crystals simulation had cloud-tops above 2 km above the 0°C level in approximately 25% of all drizzling profiles, closer to the observed frequency than all other models; this improvement was not apparent in the 200-m no-crystals run. For light-rain and heavy-rain profiles, the no-crystals simulations showed little difference with the standard configuration simulations. The 1500-m simulation with graupel compared well with observations for heavy-rain profiles, but the 200-m graupel simulation generated reflectivities around 40 dBZ too frequently.

When decreasing the horizontal grid length in the simulations with standard microphysics, the PDFs remained broadly similar, which agrees with results from Lang et al. (2007), who compared reflectivity PDFs from 1-km and 250-m grid length simulations for a case of tropical convection. Lang et al. (2007) and Lang et al. (2011) reported reflectivity distributions that were disjointed across the melting layer in model simulations, similar to our findings, and showed how changing the representation of graupel processes in their model provided a better comparison with their observed PDFs. Similar changes could improve the graupel PDFs for the DYMECS case studied, particularly if it would reduce the frequency of high reflectivities in drizzle and light-rain profiles.

In the PDFs of reflectivity versus height, all model configurations showed a prominent shallow mode (a 0-dBZ cloud top within 2 km above the 0°C level) contributing to the PDF for light-rain profiles ($20 \leq Z_{\mathrm{rain}} < 25$), which was not observed; in the 200-m and 100-m simulations, this mode also became prominent in the heavy-rain profiles. The existence of the shallow mode across all microphysics configurations and all resolutions suggests that this model error might be due to cloud-dynamics, such as turbulent mixing and entrainment processes. The shallow mode may also explain the lack of larger deep storms in the 200-m and 100-m simulations (see also Hanley et al. (2014)), as the intense rainfall from shallow storms acts as a moisture sink and could prevent these storms from deepening and broadening.

The presented analysis has focussed on ice processes in convective storms though there are hints that warm-rain processes can dominate convective rainfall in the UK. For instance, Figure 10 suggests that for ice reflectivities up to 20 dBZ, the interquartile range of rain reflectivities is of the order 10 dB or more, which corresponds to a range of rainfall rates varying by factors up to 4; the rain reflectivities are also higher than expected from a linear relationship with ice reflectivities, suggesting that warm-rain processes may enhance precipitation. Another open question regards the impact of resolution and microphysics parameterization on storm dynamics, for instance updraft strength and size, which may in turn affect the storm morphologies presented in this paper. The DYMECS data will allow us to evaluate storm dynamics with sets of RHI scans through the locations of convective cores, which in combination with the PPI volumes analysed in this paper will lead to joint analysis of storm morphology and dynamics. Finally, future work should focus on the temporal evolution of individual storm volumes as well as the storm population. The success rate of mid-level storms growing into deep and the time of day of peak storm growth are of particular interest and can be compared to similar results from tropical convection (Kumar et al. 2013, 2014). The radiative impact of the model delay in anvil occurrence requires further investigation too, though this will be of greater importance in for instance the tropics, as very few DYMECS cases involved frequent occurrence of anvils.

More research using the DYMECS cases will be conducted to evaluate the Met Office models under different synoptic conditions, as well as studies of model sensitivity to dynamics settings (e.g. Hanley et al. (2014)). Combined with other emerging data sets of convective storms (e.g. Tao et al. (2013)), the DYMECS data and the analysis presented in this paper will provide a modern test bed for the evaluation of convection-permitting models.

Acknowledgments.

We thank Paul Field, Ian Boutle, and Jonathan Wilkinson for useful discussion regarding the UM microphysics. CAMRa is operated and maintained by the Rutherford Appleton Laboratory. We are especially grateful to Darcy Ladd, Mal Clarke, and Alan Doo at the Chilbolton Observatory for their invaluable assistance with gathering the radar data. We acknowledge use of the MONSooN system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, which is a strategic partnership between the Met Office and the Natural Environment Research Council. The DYMECS project is funded by NERC (grant NE/I009965/1).

APPENDIX A

Forward model for radar reflectivities

The forward model for radar reflectivities from the UM microphysics (McBeath et al. 2014) assumes the Rayleigh scattering limit, because of the long CAMRa wavelength, so that reflectivity is considered proportional to mass squared (e.g. Hogan et al. (2006)):

$$Z_j = \mathcal{R}_j \int_0^\infty [M_j(D)]^2 \, n_j(D) dD \;, \tag{A1}$$

with j denoting the hydrometeor type and

$$\mathcal{R}_j = 10^{18} \frac{|K_j|^2}{0.93} \left(\frac{6}{\pi \rho_j} \right)^2 \;, \tag{A2}$$

with parameter values in Table 1. The mass-diameter relationship and particle size distribution are given by:

$$M_j(D) = a_j D^{b_j} \;, \tag{A3}$$

$$n_j(D) = N_{0j} \lambda_j^{\beta_j} D^{\alpha_j} \mathrm{e}^{-\lambda_j D} \;, \tag{A4}$$

with parameter values in Table 1.

The λ_j can be derived through the in-cloud water content W_j from the model specific humidities q_j, that is, $W_j = q_j \rho_{\mathrm{air}} / \mathcal{C}_j$, with $\mathcal{C}_j$ the cloud fraction of hydrometeor type j. Since the water content is the integral of mass over the particle size spectrum,

$$W_j = \int_0^\infty M_j(D) n_j(D) dD \;, \tag{A5}$$

the following relationship between λ_j and W_j is obtained:

$$\lambda_j = \left[\frac{N_{0j} a_j \Gamma(b_j + 1 + \alpha_j)}{W_j} \right]^{\dfrac{1}{b_j + 1 + \alpha_j - \beta_j}} \;. \tag{A6}$$

Then, using this λ_j and combining equations (A1), (A3), and (A4), Z_j is obtained:

$$Z_j = \mathcal{C}_j \mathcal{R}_j N_{0j} a_j^2 \quad \text{(A7)}$$
$$\times \Gamma(1 + 2b_j + \alpha_j) \lambda_j^{-(1+2b_j+\alpha_j-\beta_j)} . \quad \text{(A8)}$$

This approach was followed for ice aggregates and crystals, graupel, and rain, using the parameter values in Table 1.

For liquid cloud, a constant number concentration over land was used of $N = 3 \times 10^8$ m^{-3} (Wilkinson et al. 2011), with the following particle size distribution:

$$n_{\rm liq}(D) = \theta D^2 \mathrm{e}^{-\lambda_{\rm liq} D} , \quad \text{(A9)}$$

so that

$$\theta = \frac{N}{2\lambda_{\rm liq}^3} . \quad \text{(A10)}$$

The liquid water content can be related to $\lambda_{\rm liq}$ using equations (A5), (A3), (A9), and (A10) to find

$$\lambda_{\rm liq}^3 = \frac{W_{\rm liq}}{60 N a_{\rm liq}} . \quad \text{(A11)}$$

Combining these with equation (A1), $Z_{\rm liq}$ can be derived:

$$Z_{\rm liq} = \mathcal{R}_{\rm liq} \frac{N a_{\rm liq}^2}{2} \Gamma(9) \lambda_{\rm liq}^6 = \mathcal{R}_{\rm liq} \frac{201.60}{N} W_{\rm liq}^2 . \quad \text{(A12)}$$

For liquid, the same a_j is used as for rain, namely $\pi \rho_{\rm liq}/6$ (see Table 1).

APPENDIX B

Derivation of the relationship between ice and rain reflectivities

Let us assume a constant mass flux between ice aloft and rainfall, that is $F_{\text{ice}} = F_{\text{rain}}$. To derive a relationship between ice and rain reflectivities (approximately 1 km above and below the 0°C level), we ignore dynamical and microphysical processes which may violate the constant-flux assumption (e.g. shear, riming) and we do not consider the radar bright band.

These fluxes can then be related to their particle size distributions as follows:

$$F_j = \int_0^\infty M_j(D)V_j(D)n_j(D)dD \ , \tag{B1}$$

where j denotes the hydrometeor type (ice or rain). Similarly, reflectivity can be related to the particle size distribution as shown by equation (A1). Thus, a relationship between F_j and Z_j can be established by solving for λ_j.

The velocity-diameter relationship for ice follows from Mitchell (1996), using the area-diameter relationship and Reynolds-Best relationships:

$$\begin{aligned} A_{\text{ice}}(D) &= r_{\text{ice}} D^{s_{\text{ice}}} \ , && \text{(B2)} \\ \text{Re} &= h_{\text{ice}} \text{Be}^{f_{\text{ice}}} \ , && \text{(B3)} \\ V_{\text{ice}}(D) &= h_{\text{ice}} \nu \left(\frac{2 a_{\text{ice}} g}{\rho_{\text{air}} \nu^2 r_{\text{ice}}} \right)^{f_{\text{ice}}} && \text{(B4)} \\ &\times \ D^{f_{\text{ice}}(b_{\text{ice}}+2-s_{\text{ice}})-1} \left(\frac{\rho_0}{\rho} \right)^{\mathcal{G}} \ , \end{aligned}$$

with $\mathcal{G} = 0.4$ and $\rho_0 = 1.0$ kg m^{-3}. For both aggregates and crystals, the UM parameters are $r_{\text{ice}} = 0.131$, $s_{\text{ice}} = 1.88$, $h_{\text{ice}} = 0.2072$, and $f_{\text{ice}} = 0.638$ (Wilkinson et al. (2011), following Mitchell (1996), all in SI units). For ice at 1–2 km above the 0°C level, an air temperature of -10°C is assumed, so that $\nu = 1.25 \times 10^{-5}$ m^2 s^{-1} and $\rho_{\text{air}} = 1.34$ kg m^{-3}.

For rain, the Abel and Shipway (2007) relation is used:

$$V_{\text{rain}}(D) = \left(\gamma D^{\delta} \text{e}^{-\mu D} + \eta D^{\epsilon} \text{e}^{-\sigma D} \right) \left(\frac{\rho_0}{\rho} \right)^{\mathcal{G}} \ , \tag{B5}$$

with $\gamma = 4854.1$, $\delta = 1.00$, $\mu = 195.0$, $\eta = -446.009$, $\epsilon = 0.782127$, and $\sigma = 4085.35$ (all in SI units).

The rain flux can be directly related to the reflectivity through λ_{rain} as follows:

$$\lambda_{\mathrm{rain}} = \left[\frac{C_{\mathrm{rain}} N_{0\mathrm{rain}} a_{\mathrm{rain}}^2 \Gamma(1+2b_{\mathrm{rain}})}{Z_{\mathrm{rain}}}\right]^{\frac{1}{1+2b_{\mathrm{rain}}}}, \tag{B6}$$

$$\begin{aligned} F_{\mathrm{rain}} = {} & a_{\mathrm{rain}}\gamma \left(\frac{\rho_0}{\rho}\right)^{\mathcal{G}} N_{0\mathrm{rain}}\Gamma(b_{\mathrm{rain}}+\delta+1) \\ & \times(\lambda_{\mathrm{rain}}+\mu)^{-(b_{\mathrm{rain}}+\delta+1)} \\ & + a_{\mathrm{rain}}\eta \left(\frac{\rho_0}{\rho}\right)^{\mathcal{G}} N_{0\mathrm{rain}}\Gamma(b_{\mathrm{rain}}+\epsilon+1) \\ & \times(\lambda_{\mathrm{rain}}+\sigma)^{-(b_{\mathrm{rain}}+\epsilon+1)} . \end{aligned} \tag{B7}$$

For ice, a similar relationship between the flux and reflectivity follows:

$$\begin{aligned} F_{\mathrm{ice}} = {} & a_{\mathrm{ice}} \left(\frac{\rho_0}{\rho}\right)^{\mathcal{G}} h_{\mathrm{ice}}\nu \left(\frac{2a_{\mathrm{ice}}g}{\rho_{\mathrm{air}}\nu^2 r_{\mathrm{ice}}}\right)^{f_{\mathrm{ice}}} N_{0\mathrm{ice}} \\ \times {} & \Gamma\left[b_{\mathrm{ice}} + f_{\mathrm{ice}}(b_{\mathrm{ice}}+2-s_{\mathrm{ice}})\right] \\ \times {} & \left[\frac{Z_{\mathrm{ice}}}{\mathcal{R}_{\mathrm{ice}} N_{0\mathrm{ice}} a_{\mathrm{ice}}^2 \Gamma(1+2b_{\mathrm{ice}})}\right]^{\frac{b_{\mathrm{ice}}+f_{\mathrm{ice}}(b_{\mathrm{ice}}+2-s_{\mathrm{ice}})}{1+2b_{\mathrm{ice}}}} , \end{aligned} \tag{B8}$$

with different values of a_{ice}, b_{ice}, and $N_{0\mathrm{ice}}$ for crystals and aggregates given in Table 1. Now, using the constant-flux assumption, a relationship between Z_{ice} and Z_{rain} can be obtained. This relationship, assuming that only aggregates contribute to Z_{ice}, is shown as a dotted line in Figure 10.

REFERENCES

Abel, S. J. and B. J. Shipway, 2007: A comparison of cloud-resolving model simulations of trade wind cumulus with aircraft observations taken during RICO. *Quarterly Journal of the Royal Meteorological Society*, **133 (624)**, 781–794.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, **139 (12)**, 3887–3905.

Biggerstaff, M. I. and S. A. Listemaa, 2000: An improved scheme for convective/stratiform echo classification using radar reflectivity. *Journal of Applied Meteorology*, **39 (12)**, 2129–2150.

Bryan, G. H., and H. Morrison, 2011: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Monthly Weather Review*, **140 (1)**, 202–225.

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, **131 (10)**, 2394–2416.

Caine, S., T. P. Lane, P. T. May, C. Jakob, S. T. Siems, M. J. Manton, and J. Pinto, 2013: Statistical assessment of tropical convection-permitting model simulations using a cell-tracking algorithm. *Monthly Weather Review*, **141 (2)**, 557–581.

Clark, P. A., K. A. Browning, R. M. Forbes, C. J. Morcrette, A. M. Blyth, and H. W. Lean, 2013: The evolution of an MCS over southern England. Part 2: Model simulations and sensitivity to microphysics. *Quarterly Journal of the Royal Meteorological Society*, n/a–n/a.

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

Dixon, M. and G. Wiener, 1993: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, **10 (6)**, 785–797.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *Journal of the Atmospheric Sciences*, **51 (2)**, 249–280.

Field, P. R., 1999: Aircraft observations of ice crystal evolution in an altostratus cloud. *Journal of the Atmospheric Sciences*, **56 (12)**, 1925–1941.

Fridlind, A. M., and coauthors, 2012: A comparison of TWP-ICE observational data with cloud-resolving model results. *Journal of Geophysical Research: Atmospheres*, **117 (D5)**, D05 204.

Goddard, J. W. F., J. Eastment, and M. Thurai, 1994a: The Chilbolton Advanced Meteorological Radar: A tool for multidisciplinary atmospheric research. *Electronics and Communication Engineering Journal*, **6**, 77–86.

Goddard, J. W. F., J. Tan, and M. Thurai, 1994b: Technique for calibration of meteorological radars using differential phase. *Electronics Letters*, **30**, 166–167.

Han, L., S. Fu, L. Zhao, Y. Zheng, H. Wang, and Y. Lin, 2009: 3D convective storm identification, tracking, and forecasting—An enhanced TITAN algorithm. *Journal of Atmospheric and Oceanic Technology*, **26 (4)**, 719–732.

Hanley, K. E., R. S. Plant, T. H. M. Stein, R. J. Hogan, J. C. Nicol, H. W. Lean, C. Halliwell, and P. A. Clark, 2014: Mixing length controls on high resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, (accepted).

Haralick, R. M. and L. G. Shapiro, 2002: *Computer and robot vision (volume I)*. Addison Wesley Longman.

Harrison, D. L., K. Norman, C. Pierce, and N. Gaussiat, 2011: Radar products for hydrological applications in the UK. *Proceedings of the Institution of Civil Engineers*, **165**, 89–103.

Herbort, F. and D. Etling, 2011: Post-frontal shower cells in the COSMO-DE model. a comparison with radar measurements. *Meteorologische Zeitschrift*, **20 (2)**, 217–226.

Hogan, R. J., M. P. Mittermaier, and A. J. Illingworth, 2006: The retrieval of ice water content from radar reflectivity factor and temperature and its use in evaluating a mesoscale model. *J. Appl. Meteorol.*, **45**, 301–317.

Holloway, C. E., S. J. Woolnough, and G. M. S. Lister, 2012: The effects of explicit versus parameterized convection on the MJO in a large-domain high-resolution tropical case study. Part I: Characterization of large-scale organization and propagation*. *Journal of the Atmospheric Sciences*, **70 (5)**, 1342–1369.

Kain, J. S., and coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing nwp. *Weather and Forecasting*, **23 (5)**, 931–952.

Kendon, M., T. Marsh, and S. Parry, 2013: The 2010–2012 drought in England and Wales. *Weather*, **68 (4)**, 88–95.

Kumar, V. V., C. Jakob, A. Protat, P. T. May, and L. Davies, 2013: The four cumulus cloud modes and their progression during rainfall events: A C-band polarimetric radar perspective. *Journal of Geophysical Research: Atmospheres*, **118 (15)**, 8375–8389.

Kumar, V. V., A. Protat, C. Jakob, and P. T. May, 2014: On the atmospheric regulation of the growth of moderate to deep cumulonimbus in a tropical environment. *Journal of the Atmospheric Sciences*, **71**, 1105–1120.

Lang, S., W.-K. Tao, J. Simpson, R. Cifelli, S. Rutledge, W. Olson, and J. Halverson, 2007: Improving simulations of convective systems from TRMM LBA: Easterly and westerly regimes. *Journal of the Atmospheric Sciences*, **64 (4)**, 1141–1164.

Lang, S. E., W.-K. Tao, X. Zeng, and Y. Li, 2011: Reducing the biases in simulated radar reflectivities from a bulk microphysics scheme: Tropical convective systems. *Journal of the Atmospheric Sciences*, **68 (10)**, 2306–2320.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Monthly Weather Review*, **136 (9)**, 3408–3424.

Lewis, M. W. and S. L. Gray, 2010: Categorisation of synoptic environments associated with mesoscale convective systems over the uk. *Atmospheric Research*, **97 (1–2)**, 194–213.

Locatelli, J. D. and P. V. Hobbs, 1974: Fall speeds and masses of solid precipitation particles. *Journal of Geophysical Research*, **79 (15)**, 2185–2197.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Monthly Weather Review*, **128 (9)**, 3187–3199.

López, R., 1976: Radar characteristics of the cloud populations of tropical disturbances in the Northwest Atlantic. *Monthly Weather Review*, **104 (3)**, 268–283.

Marshall, J. S. and W. M. K. Palmer, 1948: The distribution of raindrops with size. *Journal of Meteorology*, **5 (4)**, 165–166.

May, P. T. and T. P. Lane, 2009: A method for using weather radar data to test cloud resolving models. *Meteorological Applications*, **16 (3)**, 425–432.

May, P. T., J. H. Mather, G. Vaughan, K. N. Bower, C. Jakob, G. M. McFarquhar, and G. G. Mace, 2008: The Tropical Warm Pool International Cloud Experiment. *Bulletin of the American Meteorological Society*, **89 (5)**, 629–645.

McBeath, K., P. R. Field, and R. J. Cotton, 2014: Using operational weather radar to assess high-resolution numerical weather prediction over the British Isles for a cold air outbreak case-study. *Quarterly Journal of the Royal Meteorological Society*, **140 (678)**, 225–239.

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

Mittermaier, M. P. and A. J. Illingworth, 2003: Comparison of model-derived and radar-observed freezing-level heights: Implications for vertical reflectivity profile-correction schemes. *Quarterly Journal of the Royal Meteorological Society*, **129 (587)**, 83–95.

Pearson, K. J., G. M. S. Lister, C. E. Birch, R. P. Allan, R. J. Hogan, and S. J. Woolnough, 2013: Modelling the diurnal cycle of tropical convection across the 'grey zone'. *Quarterly Journal of the Royal Meteorological Society*, doi:10.1002/qj.2145.

Potts, R., T. Keenan, and P. T. May, 2000: Radar characteristics of storms in the Sydney area. *Monthly Weather Review*, **128 (9)**, 3308–3319.

Rinehart, R. E. and E. T. Garvey, 1978: Three-dimensional storm motion detection by conventional weather radar. *Nature*, **273**, 287–289.

Steiner, M., R. A. Houze Jr, and S. E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *Journal of Applied Meteorology*, **34 (9)**, 1978–2007.

Tao, W.-K., and coauthors, 2013: Precipitation intensity and variation during MC3E: A numerical modeling study. *Journal of Geophysical Research: Atmospheres*, **118 (13)**, 7199–7218.

Varble, A., and coauthors, 2011: Evaluation of cloud-resolving model intercomparison simulations using TWP-ICE observations: Precipitation and cloud structure. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **116 (D12)**.

Waldvogel, A., B. Federer, and P. Grimm, 1979: Criteria for the detection of hail cells. *Journal of Applied Meteorology*, **18 (12)**, 1521–1525.

Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Monthly Weather Review*, **138 (9)**, 3418–3433.

Wilkinson, J., D. Wilson, and R. Forbes, 2011: The large-scale precipitation parametrization scheme. *Unified Model Documentation Paper*, **26**.

Wilson, D. R. and S. P. Ballard, 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, S. E. and R. A. Houze Jr, 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.

List of Tables

Table 1. List of UM microphysical parameters (Wilkinson et al. 2011). Here, $T_{\mathrm{max}} = \max[T(^\circ\mathrm{C}), -45^\circ\mathrm{C}]$. In the derivation of the Z_{ice}–Z_{rain} relationship, $T_{\mathrm{max}} = -10^\circ$C is used.

Parameter	Description	Units	Rain	Aggregates	Crystals	Graupel
$\lvert K\rvert^2$	dielectric factor	kg^2 m^{-6}	0.93	0.174	0.174	0.174
ρ	particle density	kg m^{-3}	1000	917	917	500
a	-	kg m^{-b}	523.599	0.0444	0.587	261.8
b	-	-	3	2.1	2.45	3.0
N_0	intercept parameter	m^{-4}	8×10^6	$2 \times 10^6 \mathrm{e}^{-.1222 T_{\mathrm{max}}}$	$40 \times 10^6 \mathrm{e}^{-.1222 T_{\mathrm{max}}}$	5×10^{25}
α	-	-	-	-	-	2.5
n_b	-	-	-	-	-	3.0

TABLE 2. List of fifteen DYMECS cases used in this study in year-month-day format. 0°C-level height in km is derived from the 1500 m model. The ice-cloud depth (ICD) columns refer to numbers of storms with a given ice cloud depth, that is the difference in kilometers between the maximum height of dBZ ≥ 0 and the 0°C-level height.

Date	0°C height	ICD ≤ 4 km	ICD 4–6 km	ICD > 6 km
20110807	2.19	45	17	1
20110818	3.18	13	-	-
20110823	3.59	56	2	-
20110826	2.30	53	39	2
20110827	1.98	51	1	-
20111103	2.45	67	15	7
20111104	1.96	27	8	4
20120411	1.10	14	42	6
20120418	1.17	11	23	22
20120420	1.02	46	85	-
20120424	1.22	31	43	-
20120711	2.10	115	59	9
20120718	2.78	85	6	-
20120806	2.34	98	3	-
20120825	2.67	27	20	24
Total	-	739	363	75

List of Figures

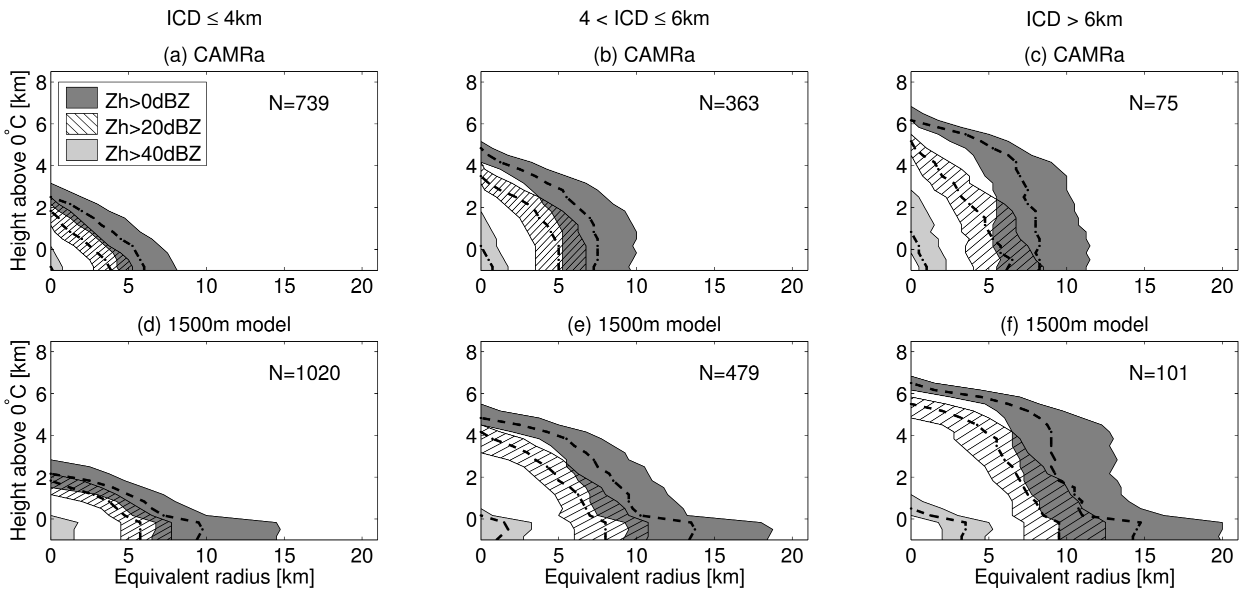

FIG. 1. Storm structures in observations (top row) and the Met Office model at 1500 m grid length (bottom row) for the DYMECS cases listed in Table 2, with height relative to the 0°C level and widths in equivalent radius as defined in section 4. A rain-rate threshold of 4 mm hr^{-1} and an area threshold of 4 km^2 were used to identify individual storms. Median equivalent radii are shown in thick dashed lines for different reflectivity thresholds, with thin lines either side indicating the 25th and 75th percentile radius; the interquartile range for the 0-dBZ, 20-dBZ, and 40-dBZ threshold are shaded dark gray, hatched, and shaded light gray, respectively. Storms are grouped by ice-cloud depth (ICD), namely below 4 km (left column), 4–6 km (middle), and above 6 km (right). The number of individual storms in each category is indicated in the top-right corner of each panel.

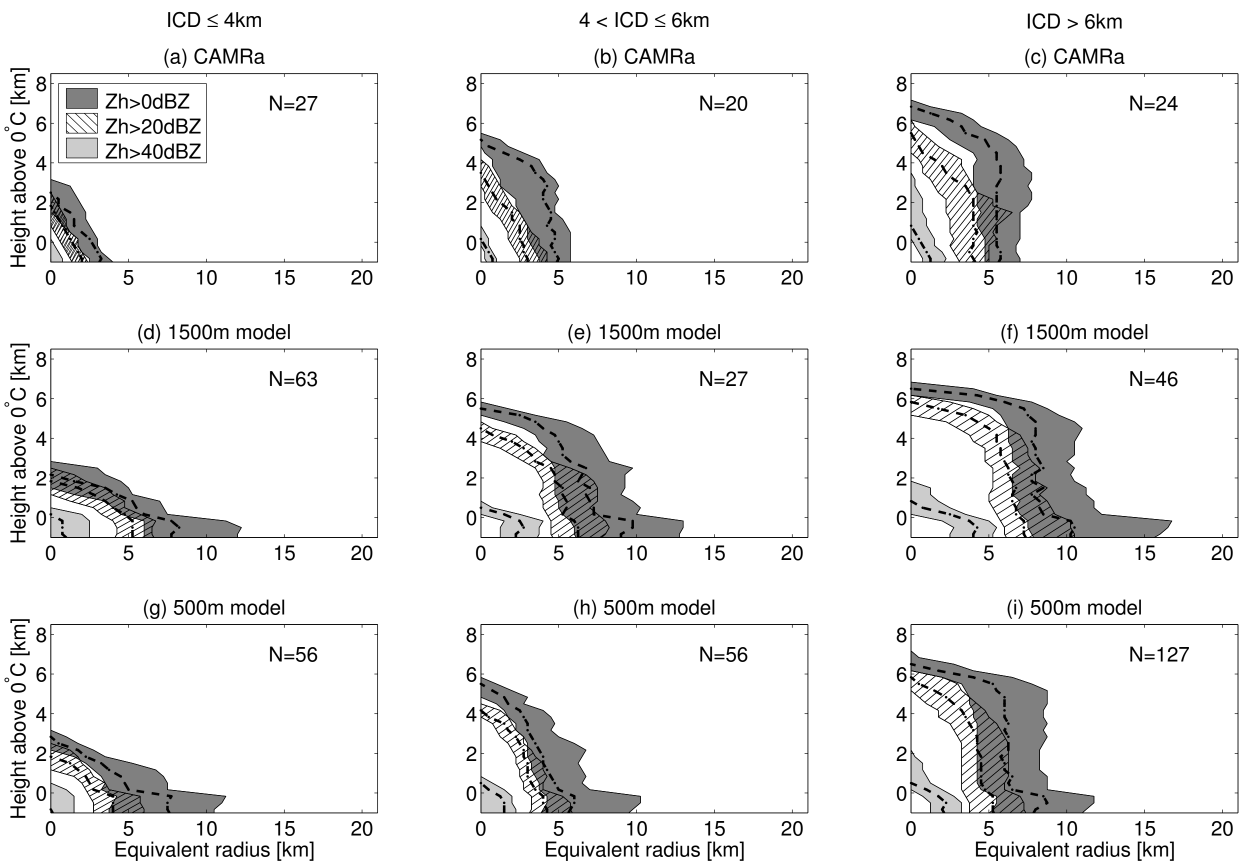

FIG. 2. As in Figure 1, but for the case of 25 August 2012. Rows show: observations (first), UM at 1500-m grid length (second), UM at 500-m grid length (third). The 500-m simulation was run with 140 vertical levels.

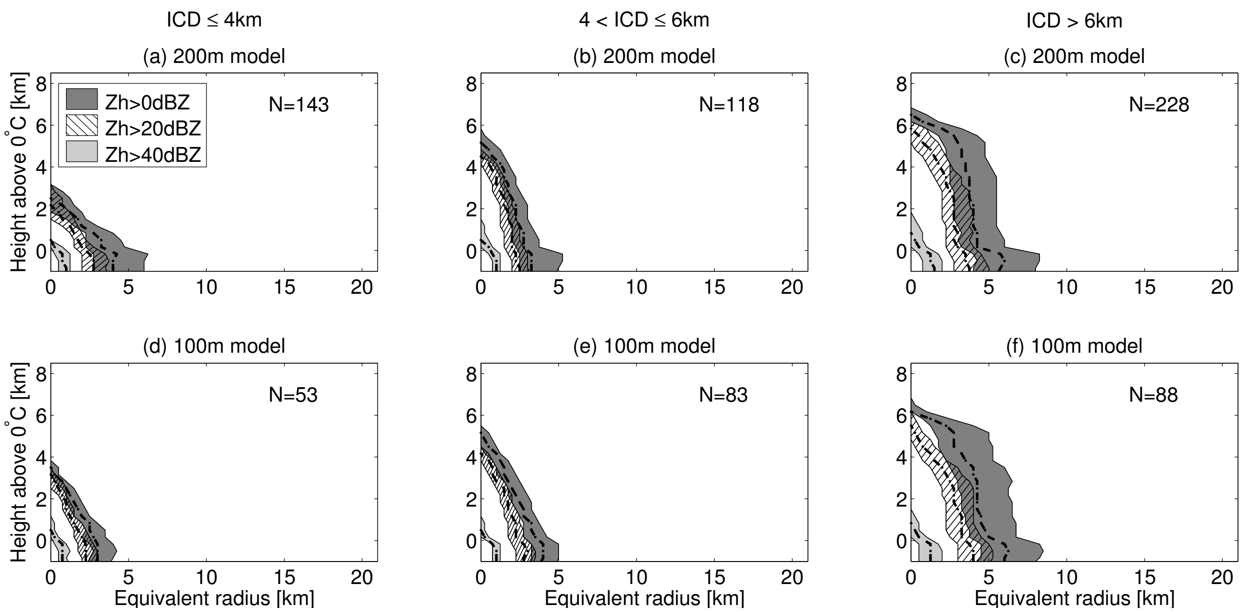

FIG. 3. As in Figure 2, but for the UM at 200-m grid length (first row) and the UM at 100-m grid length (second). Both simulations were run with 140 vertical levels. The 100-m simulation was analysed on a smaller domain of 140×140 km.

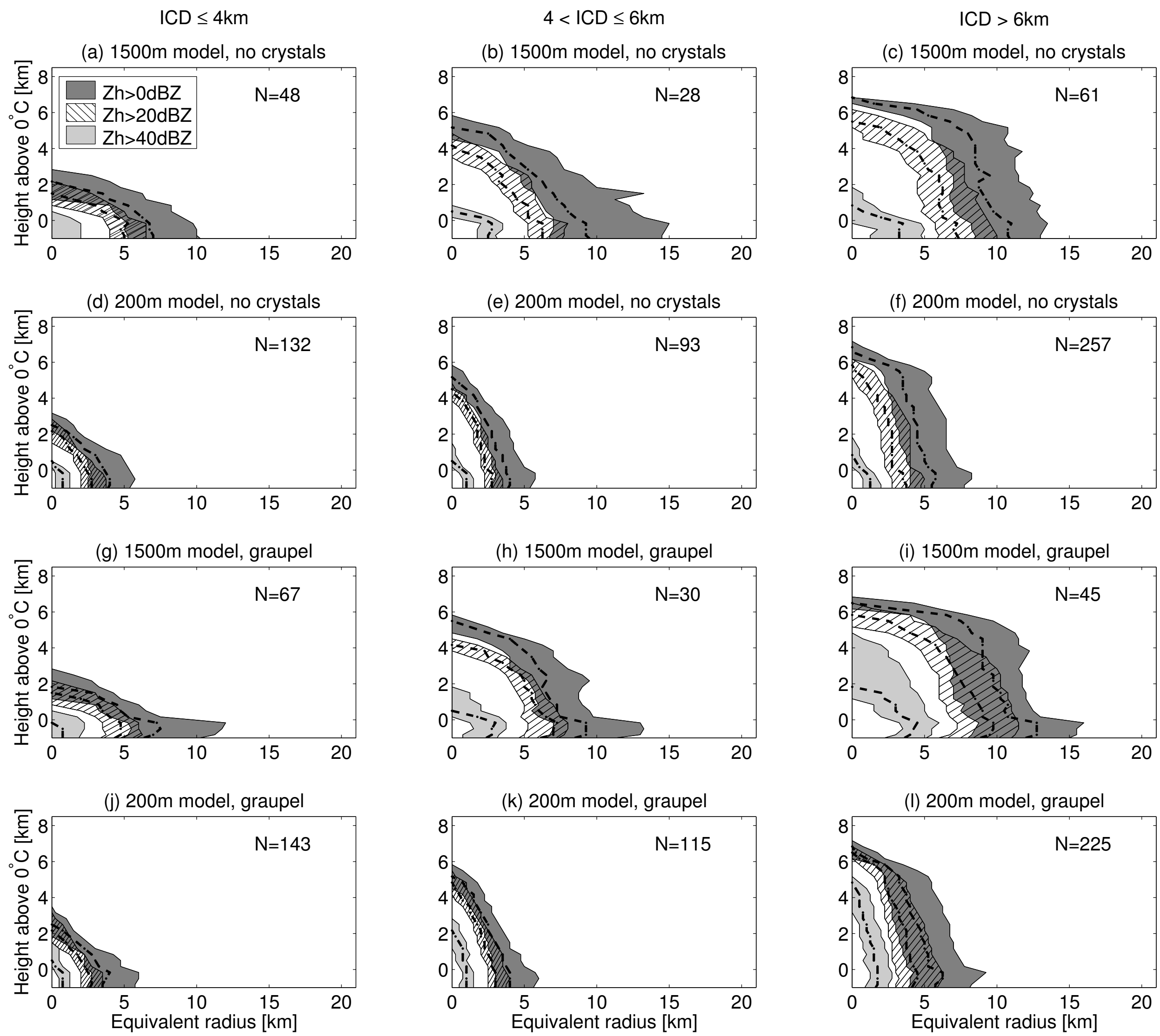

FIG. 4. As in Figure 2, but for the UM with all diagnostic ice set to be aggregates at 1500-m grid length (first row) and 200-m grid length (second); and the UM with graupel at 1500-m grid length (third) and 200-m grid length (fourth). Both 200-m simulations were run with 140 vertical levels.

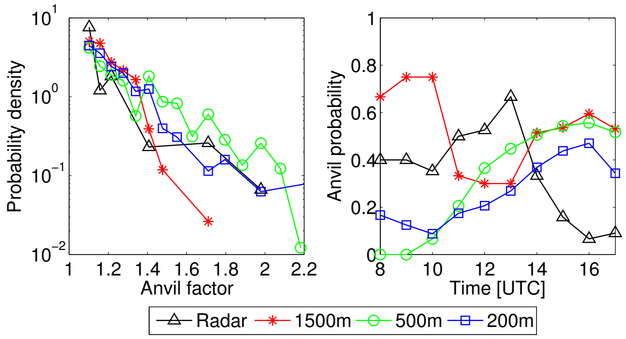

FIG. 5. Summary statistics of anvil structures for 25 August 2012 in the radar observations (black triangles), the UM at 1500-m grid length (red stars), the UM at 500-m grid length (green circles), and the UM at 200-m grid length (blue squares). The left panel shows the probability density of anvil factors above 1.05 with bin size of 0.05 and the right panel shows, for given times of the day, the probability that a storm has anvil factor greater than or equal to 1.05, using a three-hour running mean. Only storms with ICD $>$ 4 km are considered.

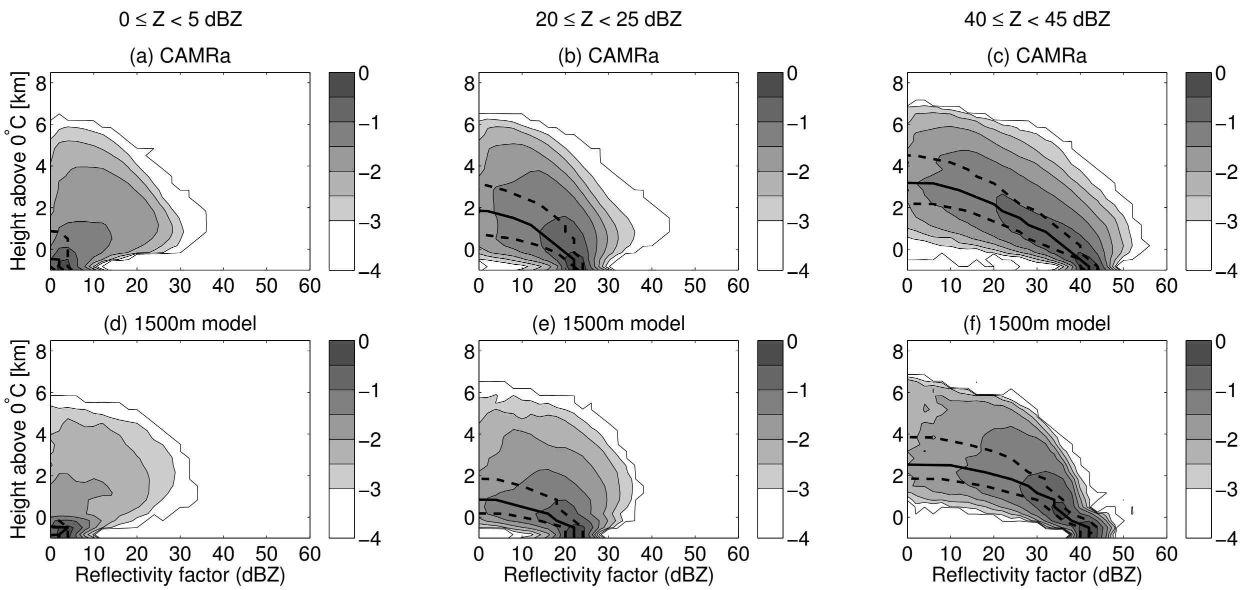

FIG. 6. PDFs of radar reflectivity factor versus height for observations (top row) and the UM at 1500-m grid length (bottom row) for all cases, with storm heights relative to the 0°C level and normalized probability density on a $\log_{10}$ scale in units $\mathrm{dB}^{-1}\mathrm{km}^{-1}$. Storms are grouped by mean reflectivity between 0.2–1.0 km below the 0°C level, namely 0–5 dBZ (left), 20–25 dBZ (middle), and 40–45 dBZ (right). Lines indicate the 25th, 50th (solid), and 75th percentile of reflectivity versus height.

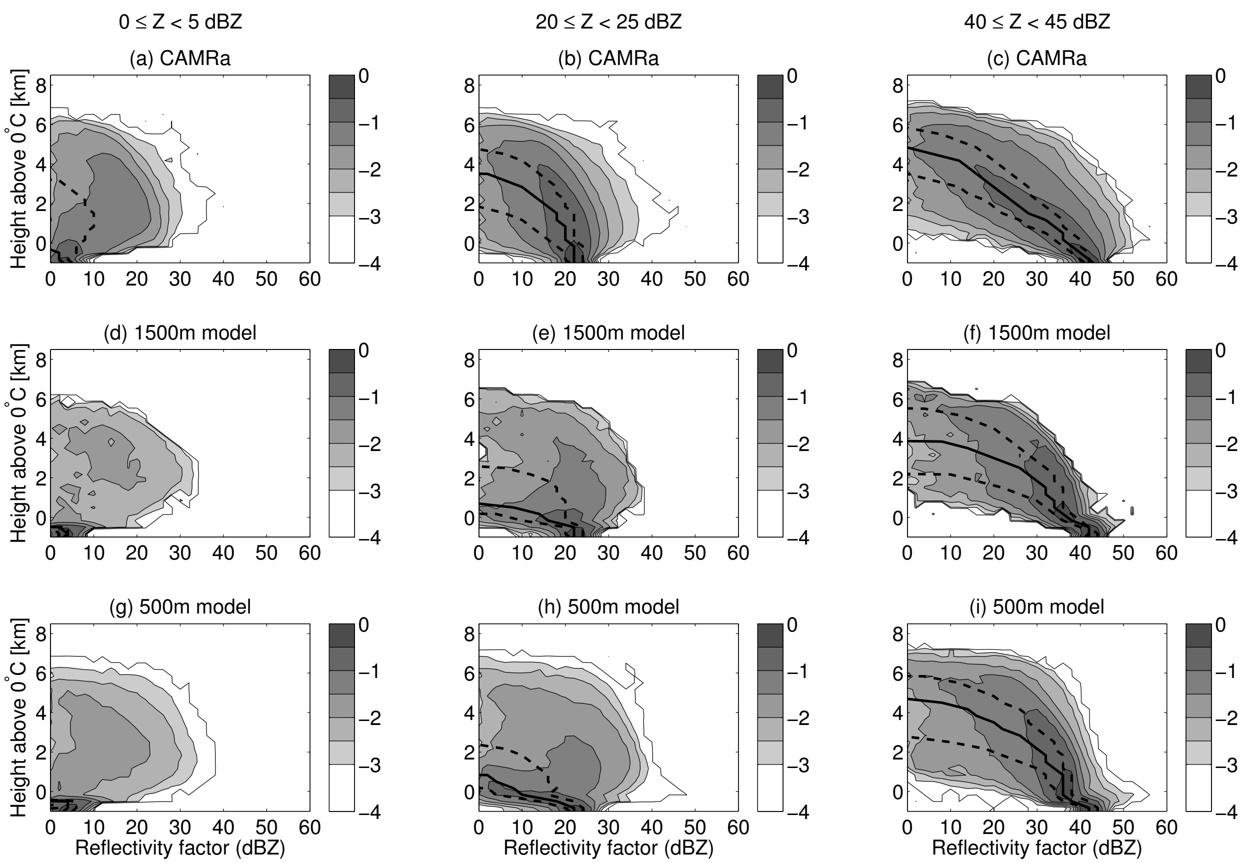

FIG. 7. As in Figure 6, but for the case of 25 August 2012. Rows are now in order: observations (first), UM at 1500-m grid length (second), UM at 500-m grid length (third). The 500-m simulation was run with 140 vertical levels.

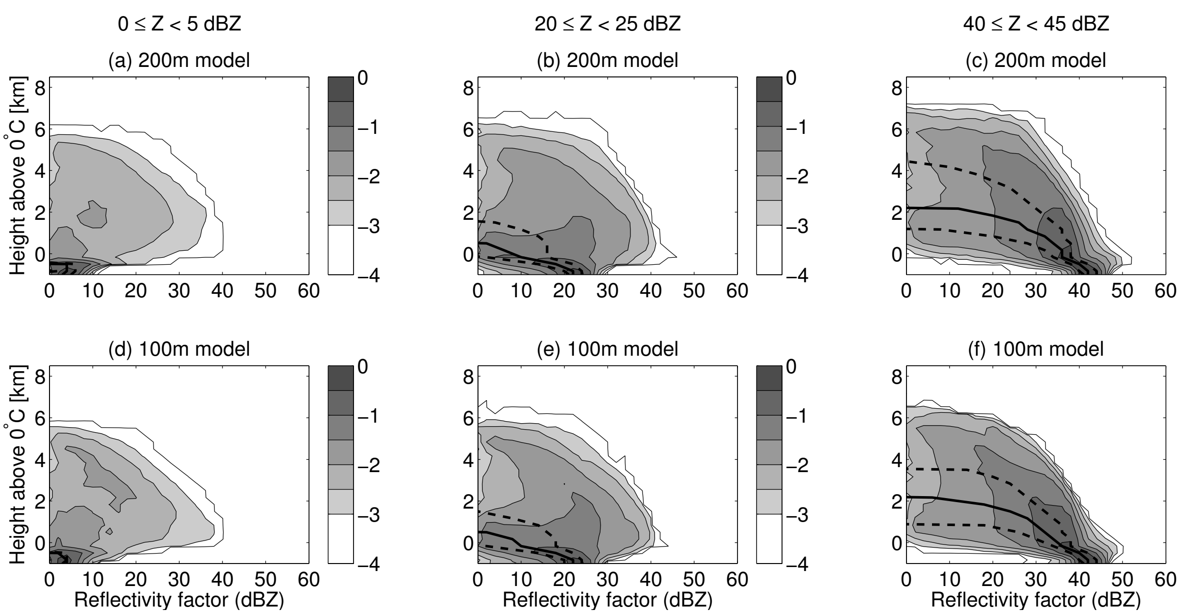

FIG. 8. As in Figure 7, but for UM at 200-m grid length (first row), UM at 100-m grid length (second). Both simulations were run with 140 vertical levels. The 100-m simulation was analysed on a smaller domain of 140 × 140 km.

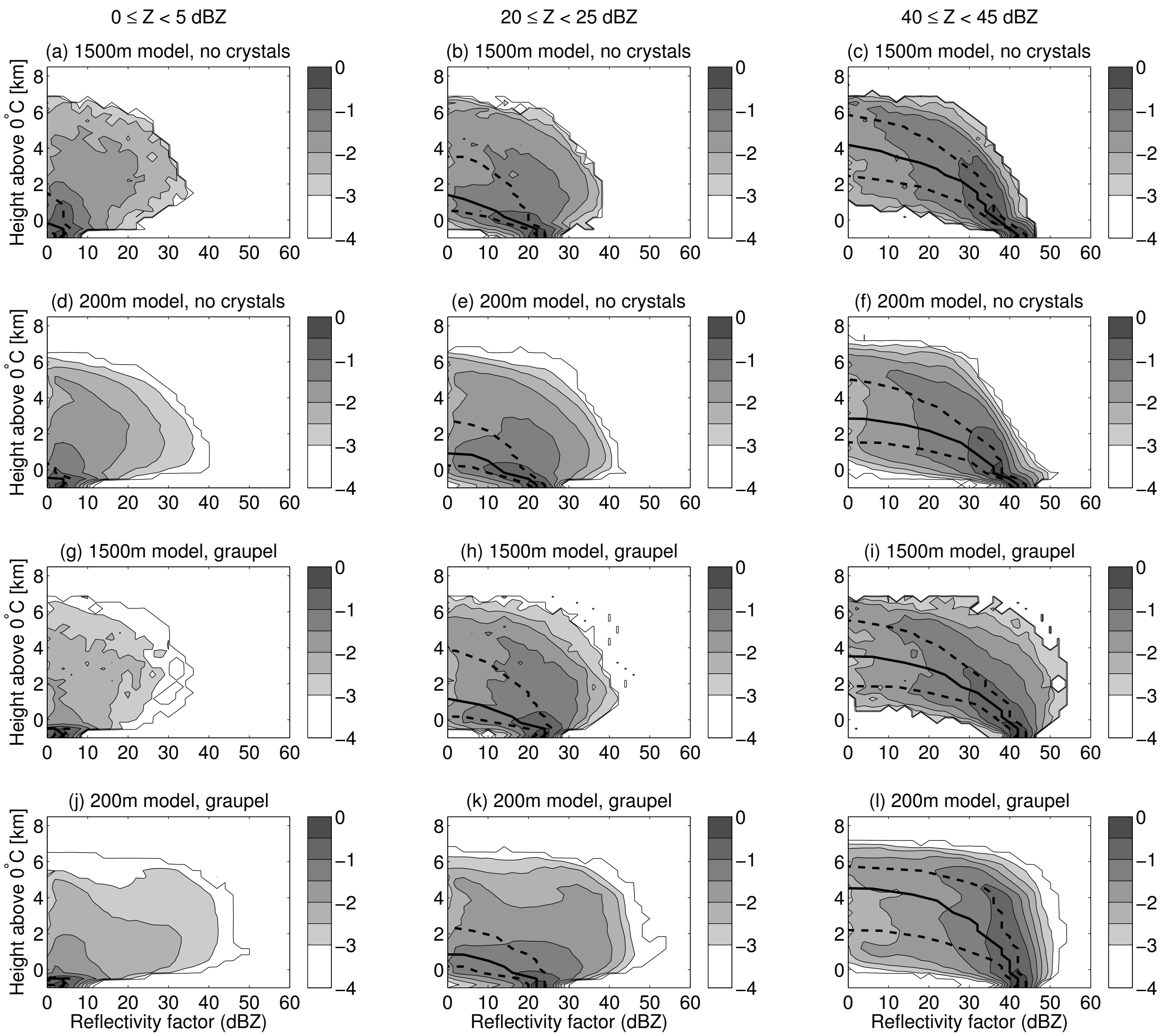

FIG. 9. As in Figure 7, but for the UM with all diagnostic ice set to be aggregates at 1500-m grid length (first row) and 200-m grid length (second); and the UM with graupel at 1500-m grid length (third) and 200-m grid length (fourth). Both 200-m simulations were run with 140 vertical levels.

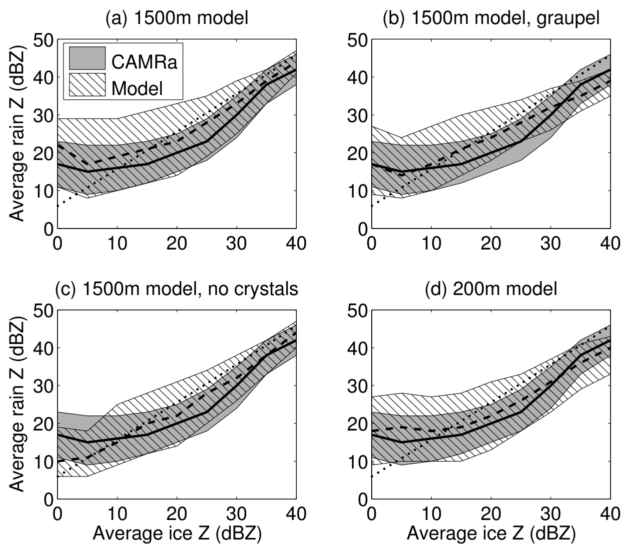

FIG. 10. The distribution of the reflectivity at 1 km below the 0°C level ("rain Z") preconditioned on the reflectivity at 1 km above the 0°C level ("ice Z"), for the case of 25 August 2012. Ice reflectivities are binned per 5 dB. The observed interquartile range is shown in each panel in dark gray, with the median in a thick solid line. Model interquartile range (hatched area) and median (thick dashed line) are shown for the UM at 1500-m grid length (panel a), the UM at 1500-m grid length including graupel (panel b), the UM at 1500-m grid length with all diagnostic ice set to be aggregates (panel c), and the UM at 200-m grid length (panel d). In all panels, the dotted line indicates the relationship derived from the model microphysics using a constant flux assumption for ice aggregates (see appendix B).