Mixing length controls on high resolution simulations of convective storms

Kirsty E. Hanley[a,b*], Robert S. Plant[a], Thorwald H. M. Stein[a], Robin J. Hogan[a], Humphrey W. Lean[b], Carol Halliwell[b], Peter A. Clark[a]

[a]*Department of Meteorology, University of Reading, Reading, UK*

[b]*MetOffice@Reading, University of Reading, Reading, UK*

[*]Correspondence to: K. E. Hanley, MetOffice@Reading, Meteorology Building, University of Reading, Reading, RG6 6BB, UK. E-mail: kirsty.hanley@metoffice.gov.uk

We perform simulations of several convective events over the southern UK with the Met Office Unified Model (UM) at horizontal grid lengths ranging from 1.5 km to 200 m. Comparing the simulated storms on these days with the Met Office rainfall radar network allows us to apply a statistical approach to evaluate the properties and evolution of the simulated storms over a range of conditions. Here we present results comparing the storm morphology in the model and reality which show that the simulated storms become smaller as grid length decreases and that the grid length that fits the observations best changes with the size of the observed cells. We investigate the sensitivity of storm morphology in the model to the mixing length used in the subgrid turbulence scheme. As the subgrid mixing length is decreased, the number of small storms with high area-averaged rain rates increases. We show that by changing the mixing length we can produce a lower resolution simulation that produces similar morphologies to a higher resolution simulation. Copyright © 0000 Royal Meteorological Society

1. Introduction

Convective storms are a crucially important forecasting problem in the UK, not least because of the flooding they can cause. In recent years, many operational forecast centres, including the Met Office, have begun running order 1 km gridlength models for short-range weather forecasting where convection is represented explicitly rather than by a convection parameterisation. A number of studies (Lean *et al.* 2008; Kain *et al.* 2008; Weisman *et al.* 2008; Schwartz *et al.* 2009; Kendon *et al.* 2012) have shown that such models yield qualitively more realistic precipitation fields and are quantitatively more skilful than lower resolution simulations with parameterised convection. However, these gridlengths are unable to fully resolve the individual convective elements (e.g. Bryan *et al.* 2003) leading to convection still being under resolved (hence they are referred to as "convection-permitting" models rather than "convection-resolving"). This leads to significant shortcomings in the nature of the convective clouds simulated at these resolutions. For example, cells in the Met Office's current 1.5 km gridlength UK model tend to be too large with too much heavy rain and not enough light rain (e.g. McBeath *et al.* 2013), and tend not to organise into mesoscale complexes as observed, illustrating our lack of understanding of the nature of small-scale mixing and microphysical processes.

Numerical weather prediction (NWP) models need to represent the subgrid-scale mixing that is unable to be resolved on the model grid. With gridlengths of the order 10 km it is reasonable to suggest that subgrid mixing is best represented by both a convection scheme and a 1D boundary layer scheme, although several studies (Pearson *et al.* 2013; Holloway *et al.* 2013) have shown that large-scale convective organisation is better represented at gridlengths of 12 km when convection is explicit rather than parameterised. At gridlengths of order 1 km, a convection scheme is no longer appropriate and a traditional 1D boundary layer scheme starts to break down

as the horizontal gridlength approaches the depth of the boundary layer. In this case the subgrid mixing may be better represented by a subgrid turbulence scheme. The two turbulence-closure models widely used in both large-eddy simulations (LES) and NWP models are either a Smagorinsky-type first-order-closure scheme based on Smagorinsky (1963), or a one-and-one-half-order scheme using a prognostic equation for turbulent kinetic energy. High resolution (i.e. convection-permitting) versions of the Met Office's Unified Model (UM) use a Smagorinsky-type scheme (Halliwell 2007) where the eddy viscosity coefficient, κ is defined as:

$$\kappa = (c_s \Delta)^2 \, S \, f_m(Ri) = \lambda_0^2 \, S \, f_m(Ri) \qquad (1)$$

where

$$S^2 = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2. \qquad (2)$$

Here c_s is an empirically determined constant, Δ is the maximum horizontal grid length, $f_m(Ri)$ is a Richardson number dependent stability function that reduces the mixing length close to the surface and λ_0 is the mixing length.

Extensive studies using the Smagorinsky model to study the effects of subgrid mixing in LES of turbulent boundary-layer flow have been conducted (e.g. Mason and Callen 1986; Mason and Brown 1999; others?). Mason and Callen (1986) viewed c_s as a ratio of a mixing length scale λ_0 to a grid scale, $c_s = \lambda_0/\Delta$. They found that simulations with a fixed gridlength showed a strong dependence on c_s: larger c_s produced flows with smooth features whereas low c_s simulations tended to suffer from grid-scale noise. By examining such sensitivities, Mason (1994) concluded that a value of $c_s \sim 0.2$ gives a solution where the simulated eddies are well resolved. Mason and Brown (1999) showed that at a fixed mixing length, λ_0, a simulation with higher resolution (i.e. larger c_s) gives a more realistic solution. Canuto and Cheng (1997) proposed a subgrid-scale model that takes into account stratification and shear, giving a value of $c_s \sim 0.11$. They concluded that even c_s is often

treated as a constant, its value is actually dependent on a combination of physical processes that differ from flow to flow, meaning c_s should ideally be a dynamical variable that varies depending on the flow.

Takemi and Rotunno (2003) investigated the effect of subgrid mixing in squall line simulations using a mesoscale cloud model (the Weather Research and Forecasting (WRF) model) with order 1 km grid spacing. They investigated the sensitivities of the simulated cloud systems to the constant c_s used in the Smagorinsky subgrid turbulence scheme. Following Mason (1994) they found that simulations with smaller c_s produce solutions with a lot of grid-scale noise while those with larger c_s give excessively smoothed cells. They concluded that the optimum value for c_s is $0.25 - 0.3$. Holloway *et al.* (2013) compared 4 km grid length UM simulations of the Madden-Julian oscillation using a 3D Smagorinsky mixing scheme to a simulation using 2D Smagorinsky mixing (with vertical mixing done by a 1D boundary layer scheme). They found that the simulation with 3D Smagorinsky mixing better retained the large-scale convective organisation.

In this paper we gather statistics on the representation of convection in high resolution versions of the UM over a range of conditions by comparing the surface rainfall features in the model storms with the UK Met Office 1 km radar composite. The main aim is to investigate the sensitivity of storm morphology in the model to the mixing length used in the subgrid turbulence scheme. The work is part of a larger project called DYMECS (DYnamical and Microphysical Evolution of Convective Storms) in which a large database of detailed convective storm properties has been obtained by automatically tracking cells with the Chilbolton Advanced Meteorological Radar, located in the south of the UK (see Fig. 1).

The paper is organised as follows. Section 2 describes the UM and the configurations used. Section 3 describes the local and large scale conditions of the cases chosen. A statistical approach is applied in Section 4 to evaluate the properties of the simulated storms compared with those in

the radar composite. In Section 5 results from sensitivity experiments varying the mixing length used in the subgrid turbulence scheme are presented. Finally, a summary and discussion of our findings is presented in Section 6.

2. Model description and method

The numerical experiments are performed using the Met Office's Unified Model (UM) Version 7.8. The UM is the Met Office's operational numerical weather prediction (NWP) model and is used to provide global and regional deterministic and ensemble forecasts. The model solves non-hydrostatic, deep-atmosphere dynamics using a semi-implicit, semi-Lagrangian numerical scheme (Davies *et al.* 2005). In the horizontal, the model uses a regular latitude-longitude grid with Arakawa C staggering and for limited area configurations, the pole of the grid is rotated such that the domain is approximately centred on the equator, in order to minimise changes in gridlength across the domain. Charney–Phillips staggering is used in the vertical along with a terrain-following hybrid-height vertical coordinate. The model uses the surface-layer scheme of Best *et al.* (2011), the radiation scheme of Edwards and Slingo (1996), the mixed-phase cloud microphysics scheme of Wilson and Ballard (1999) and the non-local boundary-layer scheme of Lock *et al.* (2000). The model also includes a convection scheme based on Gregory and Rowntree (1990), although this is switched off at gridlengths of 1.5 km and finer. High resolution ($\Delta x \approx 1$ km) versions of the model also use a subgrid turbulence scheme, as described in the Introduction. This scheme can be applied in just the horizontal or also in the vertical. When applied in the vertical, the non-local boundary-layer scheme is switched off and the local boundary layer scheme uses the diffusion coefficients calculated from the subgrid-turbulence scheme.

During the period of interest, the deterministic operational nested suite consisted of four configurations: Global; North Atlantic and European (NAE); UK 4 km (UK4); and UK Variable-resolution (UKV). The UKV model is a
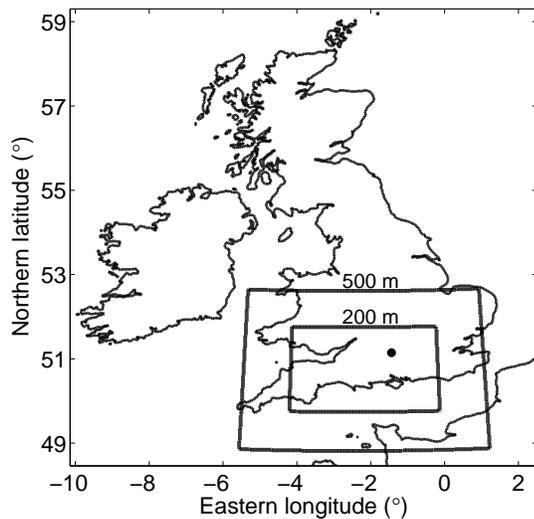
Figure 1. Model domains used. The outer domain is the fixed gridlength part of the UKV. The black dot indicates the location of the Chilbolton radar.

limited-area, variable resolution configuration of the UM nested within the 12 km gridlength NAE. The inner part of the domain covers the entire UK (shown in Fig. 1) and has a horizontal gridlength of 1.5 km. The smaller outer region has a horizontal gridlength of 4 km and in between there is a variable-resolution transition region. The variable resolution allows the boundaries of the UKV to be further from the UK at a cheaper computational cost than if the domain had a fixed resolution of 1.5 km. In the vertical, the UKV has 70 levels, the spacing of which increases quadratically with height up to the domain top at 40 km. The operational UKV was run at 0300, 0900, 1500 and 2100 UTC each day, with initial and boundary data provided by an NAE run initialised 3 hours earlier. A data assimilation cycle operated from $T - 2$ to $T + 1$ (where T is forecast run time in hours), and fields assimilated included surface and satellite-derived 3D cloud fractions and radar-derived surface rain rates.

To investigate how the representation of convection changes as the horizontal gridlength is decreased, a one-way suite of nested models have been run with gridlengths of 1.5 km, 500 m and 200 m, as shown in Fig. 1. All the models treat convection explicitly, i.e. without the use of a parameterisation scheme. For the UKV simulations, the

0400 UTC operational UKV analysis (the output of the 3-hour data assimilation cycle) was used as initial conditions, while lateral boundary conditions were provided by the 0000 UTC NAE forecast. The setup of the UKV is that which was operational at the start of DYMECS (summer 2011). The 500 m gridlength model, of domain size 500 × 425 km, also gets its initial conditions from the 0400 UTC operational UKV analysis and has boundary data provided by the UKV run just described. The western boundary of the domain is located 300 km west of Chilbolton and the southern boundary is 225 km south of Chilbolton (see Fig. 1). A 300 × 225 km 200 m gridlength model has been nested 50 km within the boundaries of the 500 m model (see Fig. 1). This model gets its inital and boundary data from the 500 m model run and is initialised 3 hours later than the other models (at 0700 UTC) to allow the 500 m model time to spin up. All models were integrated forward until 1900 UTC (15 hours for the UKV and 500m model, 12 hours for the 200m model), using timesteps of 50 s, 10 s and 6 s in the UKV, 500 m and 200 m models respectively.

The configuration of the 500 m and 200 m models is based on the high resolution UM simulations performed by Vosper *et al.* (2013) and is very similar to the UKV, but with a few differences. Unlike the UKV, which has 70 vertical levels, both of the nested models have 140 vertical levels (corresponding to a spacing of ≈ 75 m at 1 km above ground level compared to 150 m in the standard 70 level set). Motivation for increasing the number of vertical levels in these models was provided by early results running with 70 levels where there was a strong tendency for the precipitation field in the 200 m model to have excessive small-scale structure, in particular banded lines of precipitation (Halliwell *et al.* 2013). Increasing the vertical resolution was shown to reduce the amount of small-scale precipitation and eliminate the bands in the precipitation field, although they remain in the vertical velocity field to a lesser extent. The precipitation fields from the 500 m model showed very little change when increasing from 70 to 140

levels, and therefore it can be assumed that using 140 levels in the UKV would have very little impact on the results.

Another difference between the models is the critical relative humidity (RH_{crit}) profile used by the cloud scheme. The UM assumes that a grid box contains some cloud when the relative humidity within the grid box exceeds RH_{crit}. The UKV uses $RH_{crit} = 0.91$ in the lowest few layers with a gradual decrease to 0.8. On the assumption that there should be less subgrid variability in humidity in smaller grid boxes, the higher resolution models use $RH_{crit} = 0.97$ in the lowest few layers, decreasing smoothly to 0.9 at ≈ 3.5 km.

The final difference worth noting is that the UKV uses the Smagorinsky subgrid mixing scheme in the horizontal with vertical mixing done by the boundary layer scheme (2D subgrid turbulence), whereas the 500m and 200m models apply the subgrid mixing scheme in both the horizontal and the vertical (3D subgrid turbulence). In section 5 the UKV has also been run using the 3D subgrid mixing scheme.

2.1. Radar rainfall composite

The observations used for model verification come from the Met Office radar rainfall composite (Harrison *et al.* 2011). The radar reflectivity data consists of 5-minute scan sequences of four elevations from the 15 C-band radar across the UK, at a resolution of 600m in range and $1°$ in azimuth. Within the 200m model domain, no land point is further than 100km away from a radar. The current rainfall retrieval uses only single-polarisation radar data, thus rainfall rates are estimated from an empirical relationship between radar reflectivity Z ($mm^6 m^{-3}$) and rainfall rate R ($mm\ hr^{-1}$) (Harrison *et al.* 2011):

$$Z = 200R^{1.6}\,, \qquad (3)$$

which is the Marshall *et al.* (1955) relationship derived for mid-latitude stratiform rain.

Several steps are incorporated in the rainfall-estimation quality control procedure to correct for radar artefacts, including noise filtering, clutter identification, and beam blockage (Harrison *et al.* 2009). For each radar, the Gunn and East (1954) rainfall-rate-attenuation relationship is used:

$$A = 0.0044R^{1.17}\,, \qquad (4)$$

with attenuation A in dB and a maximum correction of a factor two increase in rain rate. A simple parameterisation of the vertical profile of reflectivity including bright band and orographic growth simulates the equivalent radar reflectivity close to the ground from which the surface rainfall rate is estimated (Harrison *et al.* 2009). For each radar at every hour, a single adjustment factor is then applied to all surface rainfall rates, based on comparisons with rain-gauge data over a time period that can range from the last hour to several days. The radar rainfall composite is then generated on a $1 \times 1km$ grid from these adjusted estimates.

A similar attenuation-corrected relationship was shown to have a mean absolute error of 31% for rain rates above $3mmhr^{-1}$ compared to hourly rain-gauge estimates (Bringi *et al.* 2011). In this paper, a rain-rate threshold of $4mmhr^{-1}$ is used and a minimum rainfall area of $10km^2$. The results were not sensitive to small relative changes in the rainfall rate threshold used. Where quantitative statements regarding the radar composite are made in this paper, these will be related to uncertainties in the rainfall estimates.

3. Case overview

The UKV has a number of deficiencies in its representation of convection. To highlight some of these, Fig. 2 shows typical convective rainfall fields from the UKV for a case of widespread small-scale showers and a case with deeper convective cells, and the corresponding radar composite images. Fig. 3a shows that in the shower case (20th April 2012) a low pressure centre was situated on the east coast of the UK. Scattered showers developed during the morning, propagating northeastwards. Scans from the Chilbolton
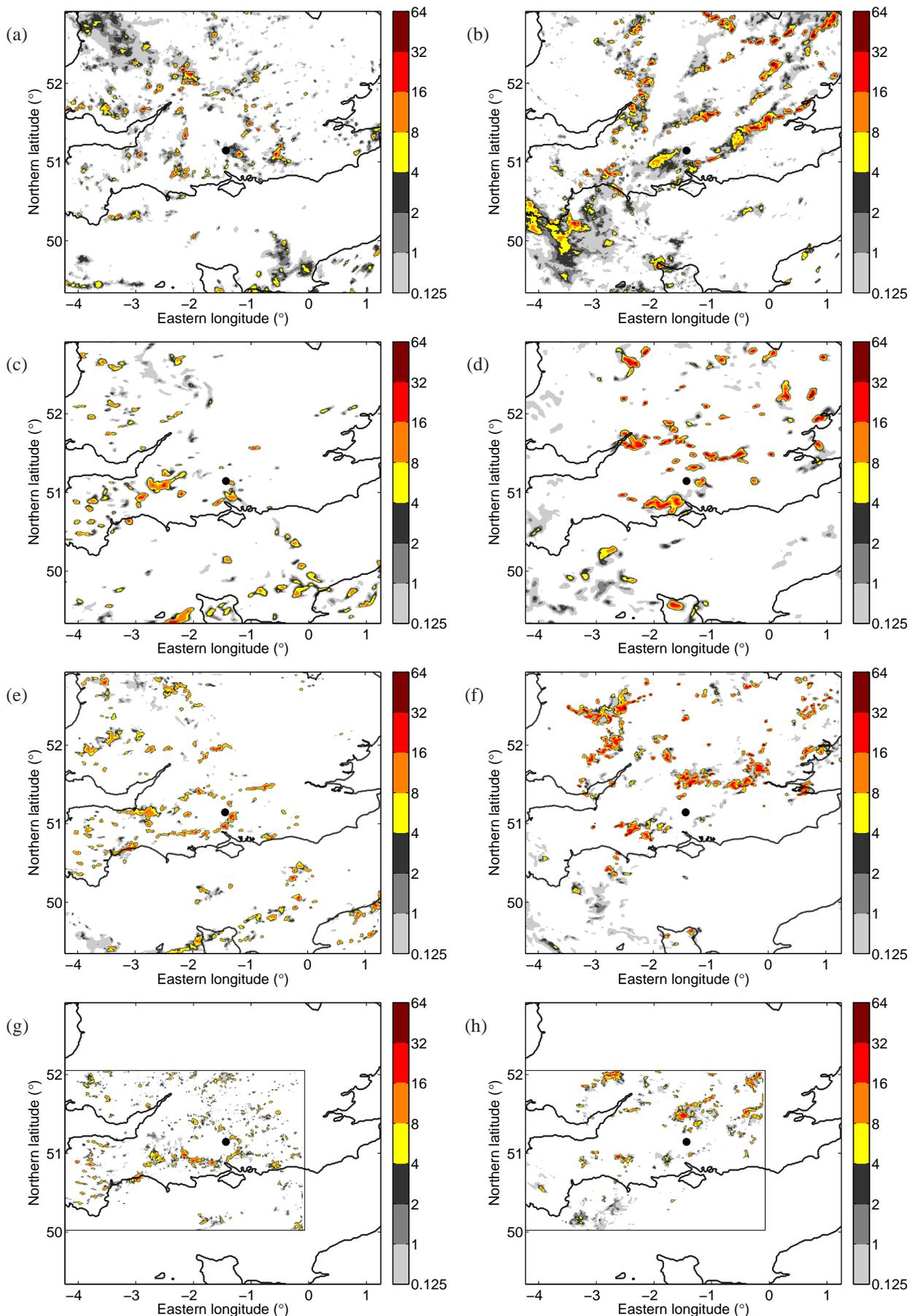
Figure 2. Rain-rate in $\mathrm{mm\ hr^{-1}}$ for a shower case (11UTC 20th April 2012, left) and a large storm case (15UTC 25th August 2012, right) from (a) and (b) the radar composite, (c) and (d) the UKV, (e) and (f) the 500 m model and (g) and (h) the 200 m model. All rainrates have been aggregated to the 1.5 km UKV grid.
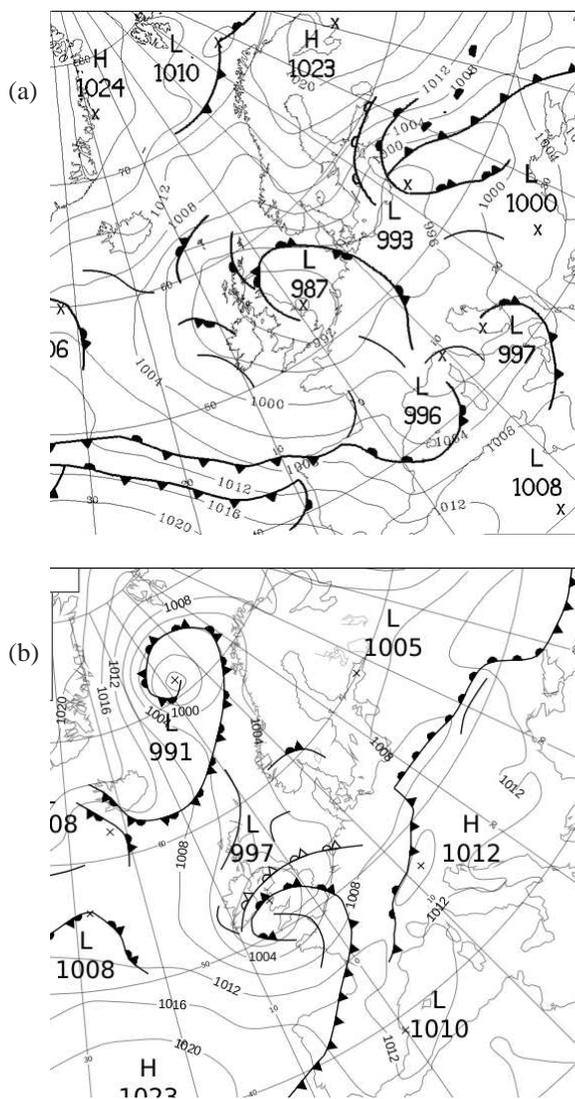
Figure 3. Met Office surface pressure analysis at 00 UTC for (a) 20th April 2012 and (b) 25th August 2012.

radar on this day showed that the storms reached up to 6 km. For the deeper storm case (25th August 2012) widespread showers and thunderstorms developed over the southern UK associated with a low over the Irish sea moving eastward (Fig. 3b). Scans from Chilbolton showed that the storms reached 10 km; heavy rain was observed throughout the day and thunder and lightning was widely reported from 14 UTC. In both cases, the UKV produces cells that are too intense and too far apart with too little light rain (Figs. 2c and d). Neither case produces enough small cells and the deeper storm case seems to be missing the large scale organisation that is evident in the radar (Fig. 2b). These problems likely result from the convection being under-resolved at this gridlength.

To investigate whether these problems are improved by increasing the horizontal resolution, these cases have been run using the 500m and 200m models described in Section 2. For the shower case (Figs. 2e and g), as the gridlength is decreased the number of small cells looks to be increasing, with possibly too many small cells in the 200m model. Compared with the radar (Fig. 2a) the cells are still too intense with not enough light rain. For the deeper storm case (Figs. 2f and h) the number of small cells also appears to increase as the gridlength is decreased, although not as much. Again, the 500m model cells are too intense compared with the radar, but this is less evident in the 200m model. There also seems to be more organisation at these resolutions than in the UKV.

Fig. 4 shows the domain-averaged precipitation over the 300×225 km domain of the 200m model obtained from all three models and the radar composite for these two cases. On 20th April (Fig. 4a), all three models do a reasonable job of simulating approximately the correct amount of domain-averaged precipitation as well as the precipitation evolution. The 200m model is best at capturing the correct initiation time and early precipitation, which suggests that the convection is better resolved at this gridlength, although between 13 and 16 UTC it over-predicts the precipitation. On 25th August (Fig. 4b), convection initiates earlier as the grid length is reduced, with the 200m model again capturing the correct initiation time and amount of precipitation. The increase in the observed precipitation at about 14 UTC is due to a large area of precipitation that moves into the south-west corner of the domain. This feature is missed by the UKV (see Fig. 2), and therefore the 500m and 200m models, which receive their lateral boundary conditions from the UKV, also miss it.

In the next section, two sets of DYMECS cases have been analysed: a set of three cases with smaller, shallower cells (tops below 6 km) from April 2012 (referred to as "shower" cases) and a set of three cases with larger, deeper cells reaching above 8 km (referred to as "larger storm" cases). Table 1 shows thermodynamic properties for each

	11th April 2012 12Z	20th April 2012 09Z	24th April 2012 09Z	7th August 2011 00Z	11th July 2012 09Z	25th August 2012 09Z
LCL (hPa)	862	898	893	948	931	931
LFC (hPa)	833	819	848	931	865	920
LNB (hPa)	514	682	788	396	616	341
CAPE (J kg^{-1})	95	18	15	115	89	142
CIN (J kg^{-1})	-2.82	-13.40	-6.91	-0.44	-15.0	-0.22
PW (mm)	11.6	10.9	10.1	20.8	16.8	22.3

Table 1. Thermodynamic properties of each case obtained from observed soundings at Larkhill, 51.2 N, −1.8 E. (Note: the 7th Aug 2011 sounding is from Herstmonceux, 50.9 N, 0.32 E). Quantities shown are lifting condensation level (LCL), level of free convection (LFC), level of neutral bouyancy (LNB), convective available potential energy (CAPE), convective inhibition (CIN) and precipitable water (PW). CAPE and CIN are calculated from a mixed-layer parcel ascent.
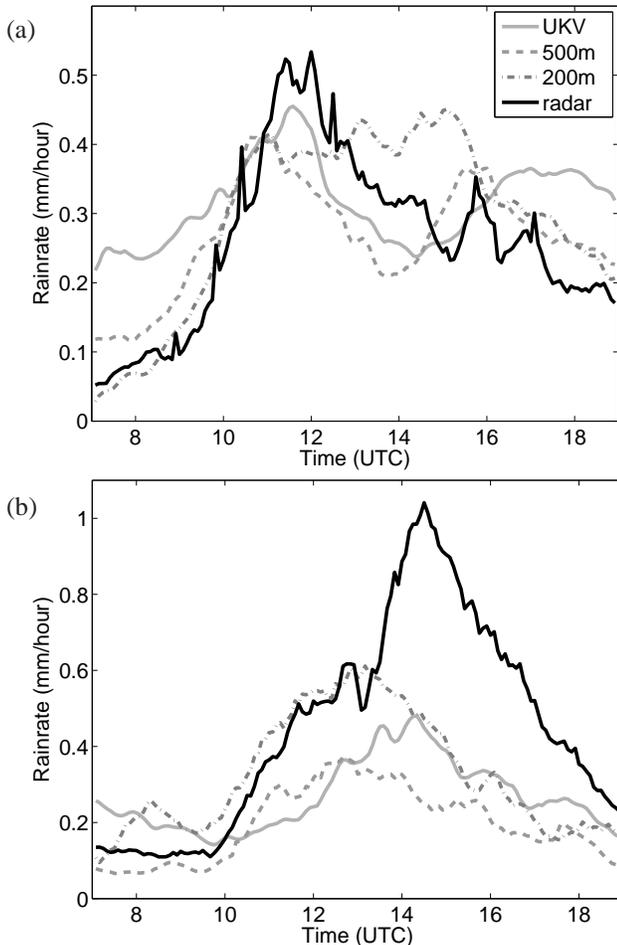

Figure 4. Domain averaged precipitation over the 300 x 225 km 200m model domain for the radar composite (black line), the UKV (grey solid line), 500m model (grey dashed line) and 200m model (grey dot-dashed line) for (a) 20th April 2012 and (b) 25th August 2012.

later time, 1200 UTC). This leads to larger values of mixed-layer convective available potential energy (CAPE) and the potential for deeper convective clouds. The larger storm cases also have higher values of precipitable water than the shower cases. The higher values of precipitable water in the larger storm cases can be understood by looking at the large-scale situation: synoptic charts (not shown) show that all six cases were associated with low pressure systems situated close to the UK; however, the three large cases all had high pressure situated over continental Europe bringing warm, moist air up to the UK from the south as in Fig. 3b. Therefore it appears that the depth of the convection in the shower cases was limited by the moisture at mid-levels as seen in other studies (e.g. Derbyshire *et al.* 2004).

4. Storm morphology

To quantify any systematic errors in the representation of convective precipitation, a number of statistics have been calculated. These have been calculated for both the shower cases and the larger storm cases. All cell statistics have been computed over the region covered by the 200 m model (see Fig. 1), with all model data first aggregated onto the 1 km horizontal gridlength radar composite grid. Cells have been identified in both the model and radar precipitation fields using a rainrate threshold of 4 mm hr^{-1} (representative of convective precipitation, see Fig. 2) and an area threshold of 10 km^2, to avoid including grid point storms from the 1.5 km gridlength UKV. We have checked that our conclusions are not qualitatively sensitive to the thresholds

of these cases. In general, the larger storm cases have a lower lifting condensation level (LCL) and level of free convection (LFC) and a higher level of neutral bouyancy (LNB) than the shower cases (the exception being the 11th April, where the sounding data were only available at a

chosen. To calculate statistics, precipitation data with a time resolution of five minutes have been used to capture cells throughout their entire life cycle. One caveat to this is that the same storm will count multiple times in the statistics.

One way to look at cell statistics is the distribution of cell sizes. Fig. 5 shows the distribution of storm equivalent diameter from the radar composite for the three shower cases (Fig. 5a) and the three larger storm cases (Fig. 5b). To easily compare the cases, which all have different numbers of storms, the number of storms in each bin has been normalised by the total number of storms throughout the day. Here, the equivalent diameter of a storm is defined as the diameter of a circle with the same area as the storm. The storm size distributions form two clusters: the
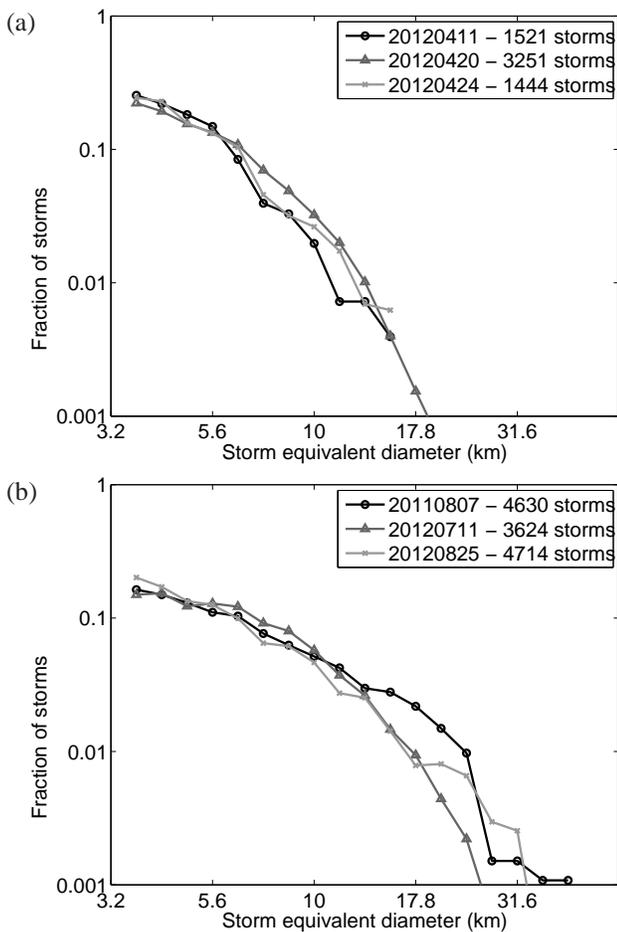

Figure 5. Distribution of storm equivalent diameter normalised by total number of storms from the radar composite for (a) the shower cases and (b) the large storm cases. A rain-rate threshold of $4 \ \mathrm{mm \ hr^{-1}}$ and an area threshold of $10 \ \mathrm{km^2}$ have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC.

shower cases have a larger fraction of small storms and a maximum diameter of approximately $18 \ \mathrm{km}$. In contrast, the

larger storm cases have fewer small storms than the shower cases and the maximum diameter is close to $30 \ \mathrm{km}$. This demonstrates that there are clear differences between the shower cases and the larger storm cases and therefore it is justifiable to look at the two subsets separately.

In order to understand the effect changing the resolution has on the distribution of storm size in the two different types of cases, Fig. 6 shows the distribution of storm equivalent diameter for the subset of three shower cases (Fig. 6a) and the subset of three larger storm cases (Fig. 6b). Comparing the UKV with the radar, it is clear that at this
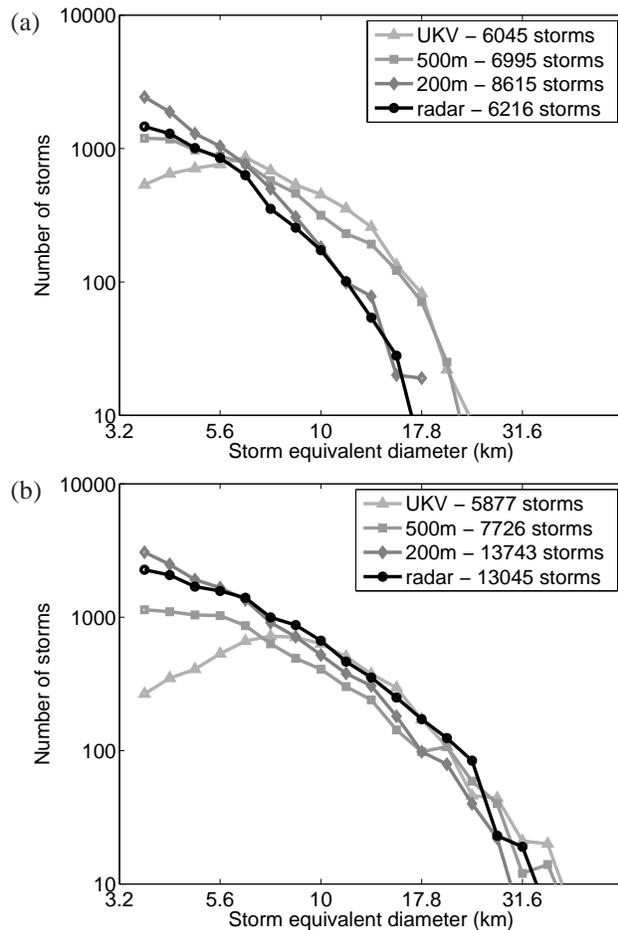

Figure 6. Distribution of storm equivalent diameter for (a) the shower cases and (b) the large storm cases. A rain-rate threshold of $4 \ \mathrm{mm \ hr^{-1}}$ and an area threshold of $10 \ \mathrm{km^2}$ have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC for the radar composite (black circles), the UKV (grey triangles), the 500m model (grey squares) and the 200m model (grey diamonds).

gridlength the model under-predicts the number of small cells for both the shower cases and the larger storm cases and that for the shower cases, with smaller convective cells, it produces too many larger cells. This agrees with the

"by-eye" comparison of the snapshot in Fig. 2. As the horizontal gridlength is decreased to 500 m, the number of small cells produced moves closer to that observed by the radar. Decreasing the gridlength to 200 m results in approximately the right number of small cells in the larger storm cases but now there are too many in the shower cases, as demonstrated in Fig. 2. The number of large storms at 200 m is decreased in both panels, leading to an improvement in the cases with smaller cells but a worsening in the cases with larger cells, where the UKV was already doing well. This effect is most noticeable in the largest case (25th August 2012) where the observations show storms with equivalent diameters exceeding 30 km whereas the largest equivalent diameters in the 200 m simulation are only 18 km (see Fig. 11f). This suggests that the 200 m model is missing some mechanism that produces more large cells when there are deeper storms. As suggested by Fig. 5, these results hold for each individual case as well as for the aggregated data.

Another way to look at cell statistics is to look at the area-averaged rainfall of each cell. Fig. 7 shows the distribution of storm area-averaged rainfall rate, again for a subset of three shower cases (Fig. 7a) and for a subset of three larger storm cases (Fig. 7b). Comparing the UKV with the radar shows that for both large storms and shower cases the UKV does not produce enough cells with high area-averaged rainrates and for the shower cases the UKV produces too many cells with small area-averaged rainrates. As the gridlength is decreased, the models tend to produce more cells with higher area-averaged rainrates. For the shower cases the 500 and 200 m models have too many cells with moderate rainrates but not enough at high rainrates. For the larger storm cases, both models have too few light rainrates but do a good job at representing the number of cells with high area-averaged rainrates.

These statistics can be combined to give a clearer picture of which cells are poorly represented by the models, for example to clarify whether the UKV is missing a lot of small, intense storms or a lot of small, weak storms.
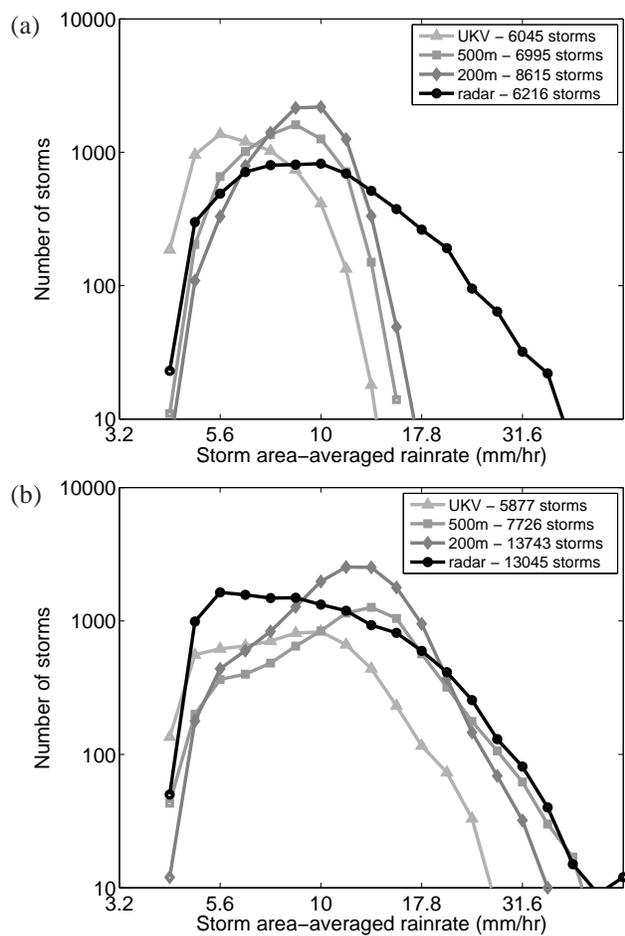

Figure 7. Distribution of storm area-averaged rain-rate for (a) the shower cases and (b) the large storm cases. A rain-rate threshold of 4 mm hr^{-1} and an area threshold of 10 km^2 have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC for the radar composite (black circles), the UKV (grey triangles), the 500m model (grey squares) and the 200m model (grey diamonds).

Figures 8 and 9 show 2-dimensional frequency distributions of storm equivalent diameter against storm area-averaged rainrate for the three shower cases and for the three larger storm cases respectively. For the shower cases (Fig. 8), none of the models produce enough small cells with high area-averaged rainrates. Compared to the radar, the UKV has a tighter spread with storm area-averaged rainrate increasing as storm equivalent diameter increases and a peak in the distribution which is shifted to too large cells and too low rainrates. In contrast, the radar has a larger spread in the distribution with many more cells of higher area-averaged rainrate at all cell sizes. As the grid length is decreased, the peak in the distribution is shifted towards smaller cells with higher rainrates. None of the models have as large a spread in the distribution as seen in the radar composite.
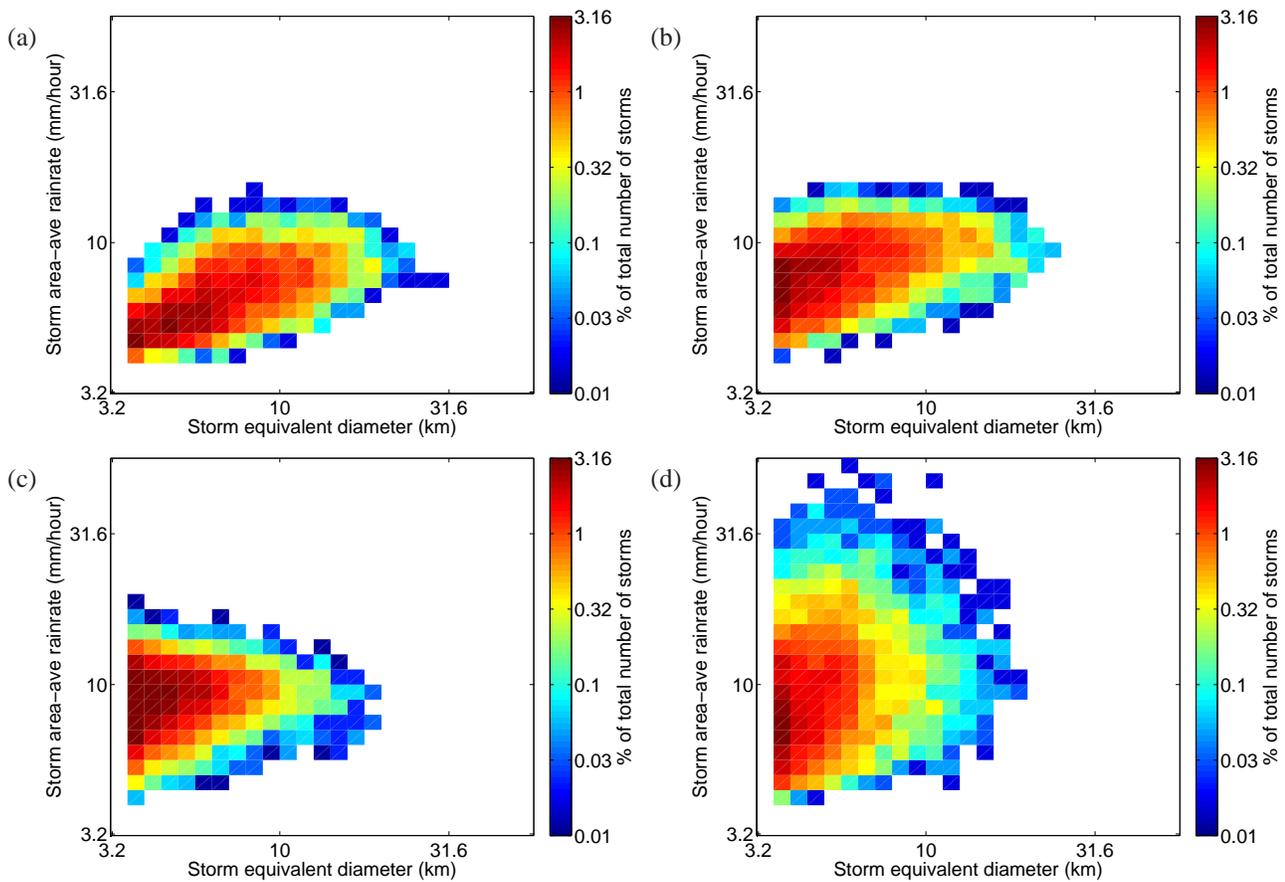
Figure 8. 2D distributions of storm equivalent diameter against storm area-averaged rain-rate for the shower cases. (a) UKV, (b) 500m, (c) 200m and (d) radar composite. Colours show the percentage of the total number of cells in each bin. A rain-rate threshold of $4\ \mathrm{mm\ hr^{-1}}$ and an area threshold of $10\ \mathrm{km^2}$ have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC.

For the larger storm cases, shown in Fig. 9, the UKV still has too narrow a distribution, missing the small cells with high rainrates and having too many medium-sized cells with moderate rainrates. As with the shower cases, as the gridlength is reduced the peak in the distribution shifts to smaller cells with higher rainrates, with the 200 m model producing too many small cells with high area-averaged rainrates. Again, for these cases all the models have less spread in the distribution than the radar composite; however the 500m and 200m models have more spread than the UKV.

The cell statistics presented here show that increasing the horizontal resolution does not necessarily improve the representation of the convective cells. While the UKV tends to predict cells that are too large and too intense, the higher resolution models produce smaller cells which are correct in cases where the cells are small in reality but are too small in cases of deeper convection when the cells should be bigger.

5. Mixing length sensitivities

The previous section highlighted some of the errors in representing convection in high resolution versions of the UM, namely the incorrect cell size and number. We would expect the subgrid turbulence scheme to have an effect on both of these properties of the cells. Subgrid-scale mixing in the UM is represented in the form of a Smagorinsky-type turbulence scheme, described in the Introduction. The scheme can be applied just in the horizontal, allowing the boundary layer scheme to mix in the vertical (this setup was used in the UKV simulations in Section 4). In this section the scheme has also been applied in the vertical, in all models, and the non-local part of the boundary layer scheme has been switched off. In this configuration, the same diffusion coefficient is used in both the horizontal and vertical diffusion schemes.
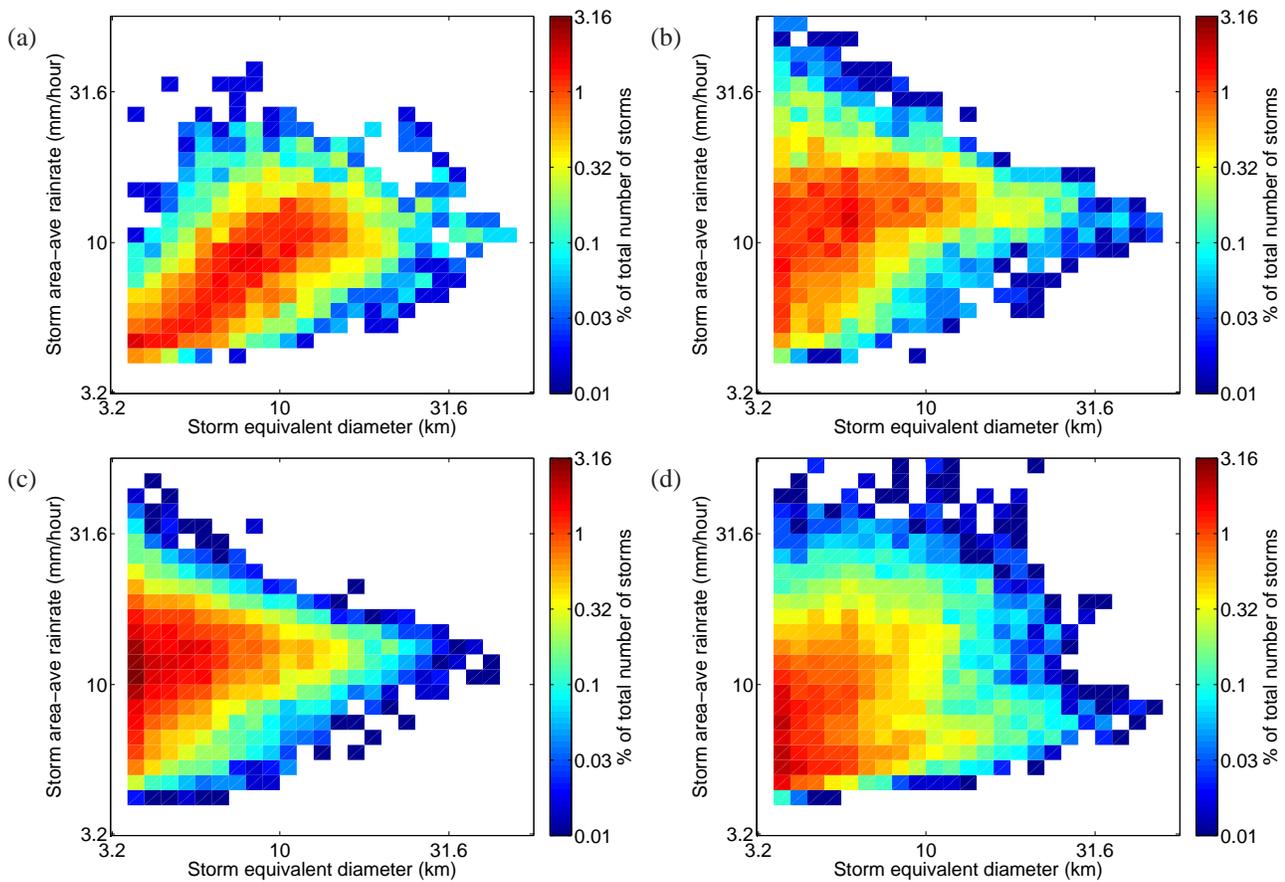
Figure 9. As Fig. 8 but for the larger storm cases.

The parameter c_s in Eqn. 1 controls the amount of subgrid mixing. Increasing c_s (and therefore the mixing length λ_0) increases the subgrid mixing which tends to smooth fields and reduce the number of small cells. The value of c_s used in the operational UKV is 0.2 (i.e. $\lambda_0 = 300$ m), and for this reason it was the value used in the default model configurations presented in Section 4. In this section we explore the sensitivity of the convective cells to the value of c_s.

Each model has been run with a mixing length of 300 m, 100 m and 40 m for one of the shower cases (20th April 2012) and one of the larger storm cases (25th August 2012). These cases were shown in Fig. 2. The mixing length values were chosen as they are the standard mixing length (corresponding to a value of 0.2 for c_s) in the UKV, 500m and 200m models respectively. As can be seen in Fig. 10, which shows the domain-averaged precipitation from the 500m model runs for each case, changing λ_0 does not have a very big effect on the overall amount of precipitation but it

does change the time of convective initiation. In both cases, convection initiates earlier as the mixing length is reduced. This is because reducing the mixing length makes it easier to trigger cells and reduces the amount of dry environmental air entrained into the moist convective updrafts.

Figure 11 shows distributions of storm-equivalent diameter for each model with varying mixing lengths. For the UKV, in both cases decreasing the mixing length decreases the amount of subgrid smoothing and increases the number of small cells, although not sufficiently to match the observed number. For the 20th April (Fig. 11a: shower case) decreasing the mixing length only affects the number of small cells whereas for the 25th August (Fig. 11b: large case), using a mixing length of 40 m does decrease the number of larger cells. For the 500m model, increasing the mixing length from 100 m to 300 m makes the distributions in both cases look more like the UKV - fewer small cells and more larger cells - whereas decreasing the mixing length to 40 m shifts both distributions to be more like
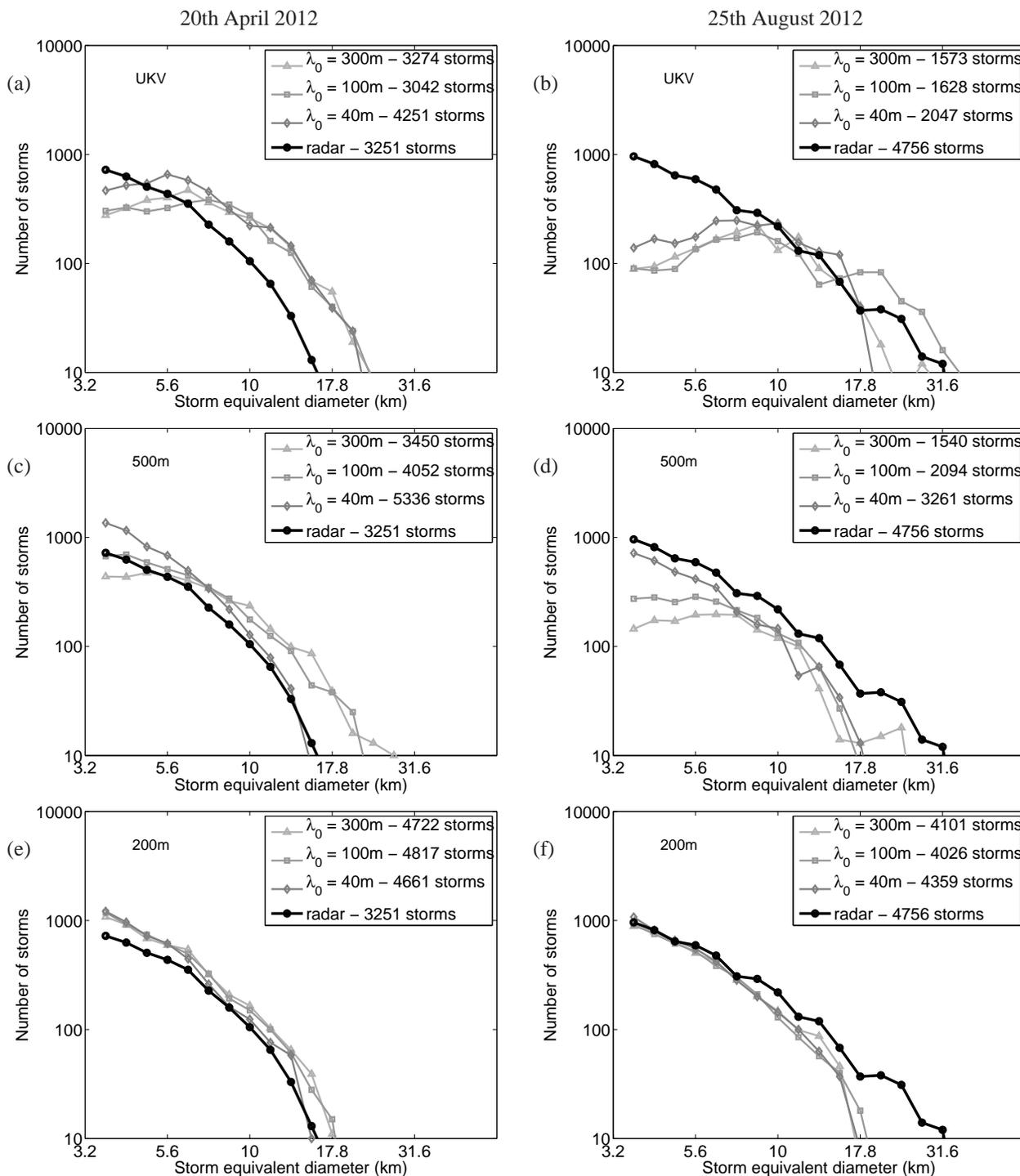
Figure 11. Distribution of storm equivalent diameter for the 20th April 2012 (left) and the 25th August 2012 (right) with mixing lengths of 300m (triangles), 100m (squares) and 40m (diamonds) for (a) and (b) the UKV, (c) and (d) 500m model and (e) and (f) 200m model. A rain-rate threshold of $4 \, \mathrm{mm} \, \mathrm{hr}^{-1}$ and an area threshold of $10 \, \mathrm{km}^2$ have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC.

the 200m model, with more small cells and fewer large cells. From these results, it can be expected that increasing the mixing length in the 200m model will smooth the precipitation fields and decrease the number of small cells while increasing the number of large cells. However, from

Fig. 11e and f we see that this is not the case. The reason for this is investigated in the remainder of this section.

To ensure numerical stability, the UM has a maximum value that it allows for the diffusion coefficients. In general, the model is run with a maximum value of the diffusion coefficient that is a quarter of the value required
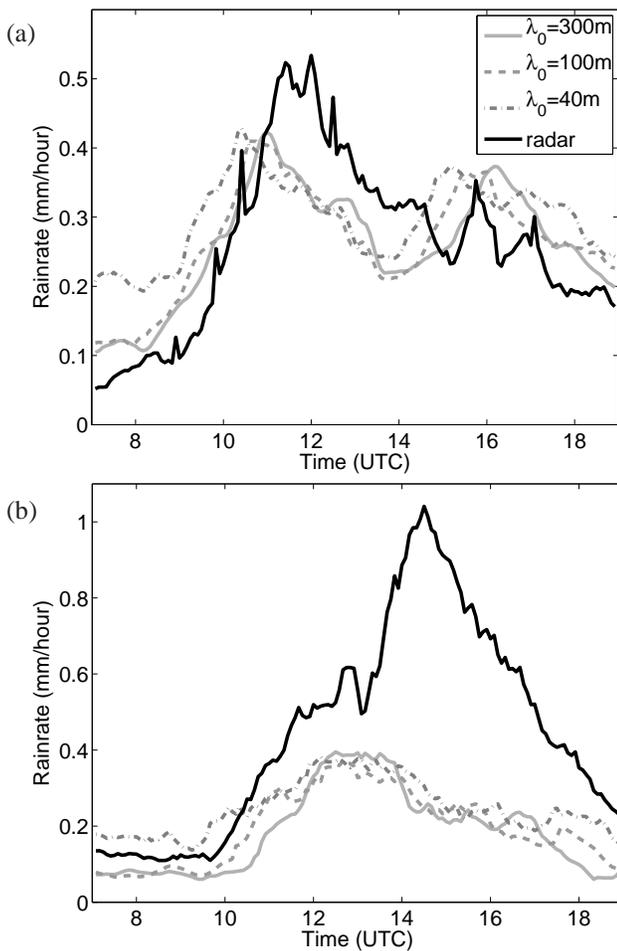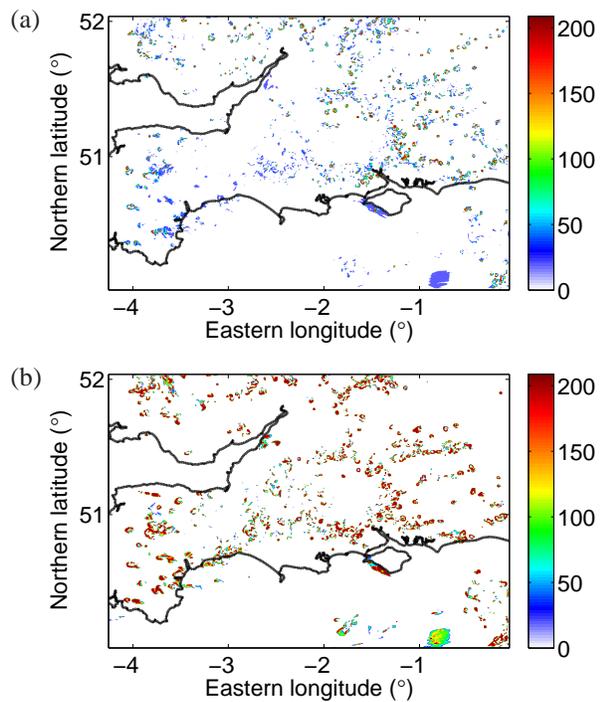
Figure 12. Smagorinsky scheme viscosity at 3 km agl at 11 UTC on 20th April 2012 from (a) the 200m model with a mixing length of 40 m and (b) the 200m model with a mixing length of 100 m.

Figure 10. Domain averaged precipitation over the 300 x 225 km 200m model domain for the 500m model with mixing lengths of 300m (grey solid line), 100m (grey dashed line) and 40m (grey dot-dashed line) and the radar composite (black line) for (a) 20th April 2012 and (b) 25th August 2012.

for the scheme to remain numerically stable. A possible explanation for why increasing the mixing length in the 200m model does not change the storm size distribution is that the model may be hitting this maximum value. To test this hypothesis we have looked at the viscosity, κ, computed by the Smagorinsky scheme in the standard 200m run with a mixing length of 40 m compared to the run with a 100 m mixing length for 20th April (Fig. 12). In both cases, the largest values of κ occur within cloud. For the run with the standard 40 m mixing length (Fig. 12a), κ is generally less than the maximum allowed value of $208 \text{ m}^2 \text{ s}^{-1}$. However, when the mixing length is increased to 100 m (Fig. 12b), the maximum value of κ is being reached in many locations. Increasing the mixing length further to 300 m produces a similar result for κ (not shown). This means that by

increasing the mixing length, we are not increasing the amount of subgrid mixing as much as we would expect. This explains why the storm size distributions shown in Figs. 11e and f have not really changed when the mixing length has been increased.

To determine whether the 200m model can be made to look more like the UKV just by increasing the amount of subgrid mixing one option would be to use a larger fraction of the maximum allowed value of the diffusion coefficient. The 200m model has been rerun for the larger storm case (25th Aug) using a mixing length of 100 m but with the fraction of maximum diffusion increased from 0.25 to 0.5. This has the effect of smoothing the precipitation field and decreasing the number of small storms, as well as slightly increasing the number of larger storms (Fig. 13). This change has made the 200m model look similar to the 500m model; but to try to make it more similar to the UKV we need to increase the subgrid mixing further. Simply increasing the mixing length to 300 m in this setup again hits the maximum value of the diffusion coefficient (not shown). Rather than increase the fraction of the maximum
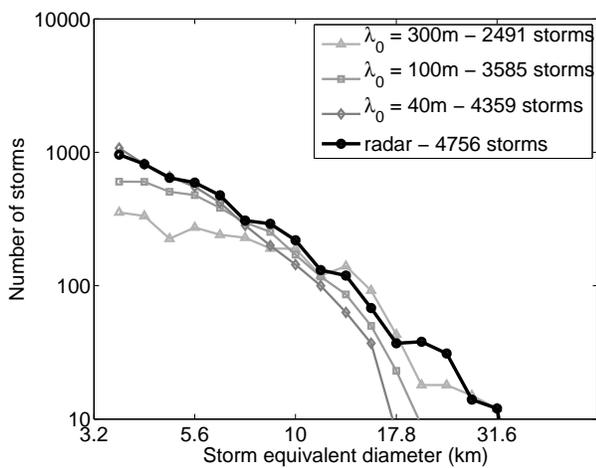
Figure 13. Distribution of storm equivalent diameter for the 200m model on 25th August 2012. The triangles show a mixing length of 300m, a timestep of 3 s and a maximum diffusion fraction of 0.5, the squares show a mixing length of 100m, a timestep of 6 s and a maximum diffusion fraction of 0.5 and the diamonds show the standard run (40m mixing length, 6 s timestep and maximum diffusion fraction of 0.25). The black line shows the radar composite distribution. A rain-rate threshold of 4 mm hr^{-1} and an area threshold of 10 km^2 have been used to identify storms. Data is every 5 minutes between 09 and 19 UTC.

allowed value of the diffusion coefficient again, making it closer to the maximum value required for numerical stability, we can increase the maximum allowed value of the diffusion coefficient by decreasing the timestep. The line with triangles in Fig. 13 shows the 200m model run with a mixing length of 300 m, the fraction of maximum diffusion set to 0.5 and a timestep of 3 s. The storm size distribution is now closer to the UKV, the increased subgrid mixing having further smoothed the precipitation field leading to a reduction in the number of small storms and an increase in the number of larger storms.

6. Conclusions

A nested suite of high-resolution convection-permitting versions of the Met Office Unified Model with gridlengths of 1.5 km, 500 m and 200 m have been used to document the shortcomings of the representation of convection over a range of conditions. By comparing the model storms with those in the radar composite it was found that the 1.5 km gridlength UKV tended to produce cells that were too intense, too far apart and with not enough light rain. The UKV also failed to produce enough small storms in both shower cases and large storm cases. In shower cases it also

produced too many large cells. These problems suggest that convection is under-resolved at this gridlength. Reducing the gridlength to 500 m resulted in an increase in small cells and decreasing it further to 200 m tended to produce too many small cells. In contrast to the UKV, the 200 m model produced the right number of large cells in the shower cases but in the large storm cases it failed to produce large enough storms.

We tested the sensitivity of the model to the value of the mixing length used in the subgrid turbulence scheme, as this parameter is expected to play a role in determining the size of the convective cells. It was found that decreasing the mixing length in the UKV increased the number of small cells, but had little impact on the number of large cells. Increasing the mixing length in the 500 m model produced a cell size distribution similar to the UKV, whereas decreasing the mixing length produced a size distribution similar to the 200 m model. Due to stability restrictions, it was necessary to decrease the timestep in the 200 m model at the same time as increasing the mixing length. In doing so we were able to shift the tail of the storm size distribution towards storms with larger diameters to closer match the observations.

It has been shown here that the model that performs best is case dependent, with the 1.5 km gridlength UKV performing well in cases with large convective cells and the 200 m model performing well in cases with small scattered showers. Since we would generally expect increasing the horizontal resolution to improve the forecast, this implies that either we are missing some processes in the high resolution models that prevent them from producing large storms, and/or we have compensating errors in the UKV. We conclude from this work that rethinking the subgrid mixing formulation is the key to getting a good description of convective cloud properties at gridlengths on the order of 100 m to 1 km. Finally, while this study focuses only on analysing the horizontal structure in the surface rain rate field, further work to supplement it by examining the vertical structure of storms can be seen in Stein *et al.* (2013).

Acknowledgements

References

Best MJ, Pryor M, Clark DB, Rooney GG, Essery RLH, Mènard CB, Edwards JM, Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ. 2011. The Joint UK Land Environment Simulator (JULES), model description – part 1:energy and water fluxes. *Geoscientific Model Development* **4**: 677–699.

Bringi VN, Rico-Ramirez MA, Thurai M. 2011. Rainfall estimation with an operational polarimetric c-band radar in the united kingdom: Comparison with a gauge network and error analysis. *J. Hydrometeor* **12**(5): 935–954.

Bryan GH, Wyngaard JC, Fritsch JM. 2003. Resolution requirements for the simulation of deep convection. *Mon. Wea. Rev.* **131**: 2394–2416.

Canuto VM, Cheng Y. 1997. Determination of the smagorinsky-lilly constant c_s. *Phys. Fluids.* **9**: 1368–1378.

Davies T, Cullen MJP, Malcolm AJ, Mawson MH, Staniforth A, White AA, Wood N. 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.* **131**: 1759–1782.

Derbyshire SH, Beau I, Bechtold P, Grandpeix JY, Piriou JM, Redelsperger JL, Soares PMM. 2004. Sensitivity of moist convection to environmental humidity. *Q. J. R. Meteorol. Soc.* **130**: 3055–3079, doi:10.1256/qj.03.130.

Edwards J, Slingo A. 1996. Studies with a flexible new radiation code. Part I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.* **122**: 689–719.

Gregory D, Rowntree P. 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability dependent closure. *Mon. Wea. Rev.* **118**: 1483–1506.

Gunn KLS, East TWR. 1954. The microwave properties of precipitation particles. *Q. J. R. Meteorol. Soc.* **80**(346): 522–545, doi:10.1002/qj.49708034603.

Halliwell C. 2007. Unified Model Documentation Paper 28: Subgrid turbulence scheme. Technical report, Met Office.

Halliwell C, Lean H, Hanley K, Carter E, Stein T, McBeath K, Nicol J, Hogan R, Plant R. 2013. Progress report on comparisons of convection permitting models with observational data from DYMECS and other sources. Technical report, Met Office.

Harrison D, Scovell R, Kitchen M. 2009. High-resolution precipitation estimates for hydrological uses. *Water Management* **162**(2): 125.

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

Holloway CE, Woolnough SJ, Lister GMS. 2013. 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. *J. Atmos. Sci.* **70**: 1342–1369, doi:10.1175/JAS–D–12–0227.1.

Kain JS, Weiss SJ, Bright DR, Baldwin ME, Levit JJ, Carbin GW, Schwartz CS, Weisman ML, Droegemeier K, Weber DB, Thomas KW. 2008. Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting* **23**: 931–952.

Kendon EJ, Roberts NM, Senior CA, Roberts MJ. 2012. Realism of rainfall in a very high resolution regional climate model. *Journal of Climate* **25**: 5791–5806, DOI:10.1175/JCLI–D–11–00 562.1.

Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, Halliwell C. 2008. Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.* **136**: 3408–3424, doi:10.1175/2008MWR2332.1.

Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB. 2000. A new boundary layer micing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.* **128**: 3187–3199.

Marshall J, Hitschfeld W, Gunn K. 1955. Advances in radar weather. Elsevier, pp. 1 – 56.

Mason PJ. 1994. Large-eddy simulation: A critical review of the technique. *Q. J. R. Meteorol. Soc.* **120**: 1–26.

Mason PJ, Brown AR. 1999. On subgrid models and filter operations in large eddy simulations. *J. Atmos. Sci.* **56**: 2101–2114.

Mason PJ, Callen NS. 1986. On the magnitude of the subgrid-scale eddy coefficient in large-eddy simulations of turbulent channel flow. *J. Fluid Mech.* **162**: 439–462.

McBeath K, Field PR, Cotton RJ. 2013. Using operational weather radar to assess high-resolution numerical weather prediction over the British Isles for a cold air outbreak case-study. *Q. J. R. Meteorol. Soc.* : DOI:10.1002/qj.2123.

Pearson KJ, Lister GMS, Birch CE, Allan RP, Hogan RJ, Woolnough SJ. 2013. Modelling the diurnal cycle of tropical covection across the grey zone. *Q. J. R. Meteorol. Soc.* : DOI:10.1002/qj.2145.

Schwartz CS, Kain JS, Weiss SJ, Xue M, Bright DR, Kong F, Thomas KW, Levit JJ, Coniglio MC. 2009. Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.* **137**: 3351–3372.

Smagorinsky J. 1963. General circulation experiments with the primative equations. I: The basic experiment. *Mon. Wea. Rev.* **91**: 99–164.

Stein THM, Hogan RJ, Hanley KE, Clark PA, Halliwell C, Lean HW, Nicol J, Plant RS. 2013. The three-dimensional microphysical structure of convective storms over the southern United Kingdom. *Mon. Wea. Rev.* : submitted.

Takemi T, Rotunno R. 2003. The effects of subgrid model mixing and numerical filetering in simulations of mesoscle cloud systems. *Mon. Wea. Rev.* **131**: 2085–2101.

Vosper S, Carter E, Lean H, Lock A, Clark P, Webster S. 2013. High resolution modelling of valley cold pools. *Atmos. Sci. Let.* : doi:10.1002/asl2.439.

Weisman ML, Davis C, Wang W, Manning KW, Klemp JB. 2008. Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting* **23**: 407–437.

Wilson DR, Ballard SP. 1999. A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.* **125**: 1607–1636.