ORIGINAL ARTICLE

Locally forced convection in sub-kilometre scale simulations with the Unified Model and WRF

M. Jucker[1,2] | T. P. Lane[3,2] | C. L. Vincent[3,2] | S. Webster[4] | S. A. Wales[3,2] | V. Louf[5]

[1]Climate Change Research Centre, University of New South Wales, Sydney, Australia
[2]Australian Research Council Centre of Excellence for Climate Extremes, Australia
[3]School of Earth Sciences, The University of Melbourne, Melbourne, Australia
[4]U.K. Met Office, Exeter, United Kingdom
[5]School of Earth, Atmosphere and Environment, Monash University, Melbourne, Australia

Correspondence
Martin Jucker, Climate Change Research Centre, University of New South Wales, Sydney, Australia
Email: publications@martinjucker.com

Present address
Climate Change Research Centre, University of New South Wales, Sydney, Australia

Funding information
Australian Research Council, Grant Number: CE170100023 and FL150100035

This study evaluates the performance and benefits of kilometre and sub-kilometre scale convection permitting simulations over tropical Australia. Focusing on an extended Monsoon break period we can directly compare Unified Model (UM) and Weather Research and Forecasting model (WRF) simulations to CPOL radar observations and soundings. We show that the two models have different behaviour, and both are different to observations. Whereas WRF produces daily squall lines whether or not they occurred in observations, the UM primarily generates small but intense storms. The UM and WRF produce qualitatively different surface density currents at different times in the diurnal cycle. Once the density currents are present, the models also show different behaviour in relation to convective initiation. While higher resolution helps in the distribution of total precipitation over the domain, most characteristics do not change with higher resolutions, and model difference are always larger than resolution differences. While CAPE/CIN does not seem to be important to explain model differences, our findings point to the evolution of density currents in the boundary layer as most important source of model errors and differences.

Keywords — Convection permitting simulations, Unified Model, Weather Research and Forecasting Model, Tropical convection, High resolution simulations, Numerical Weather Prediction models, Cold pools, Density currents

Abbreviations: NWP: Numerical Weather Prediction, UM: Unified Model, WRF: Weather Research and Forecasting model, LFC: Level of free convection, CAPE: Convective Available Potential Energy, CIN: Convective Inhibition

1 | INTRODUCTION

Convection permitting models are now in widespread use for numerical weather prediction. Furthermore, they are often used as a "truth" to assess the performance of climate or regional models in terms of convective behaviour (e.g. Song and Zhang, 2018; O'Gorman and Dwyer, 2018). The general idea is that the higher the resolution, the better the representation of otherwise parameterised sub-grid processes. However, these high resolution models also have their biases, and many processes still have to be parameterised unless the simulation uses resolution of the order of tens of meters. The general aim of this study is to understand some of the biases in two common convection-permitting modelling systems and their sensitivity to model resolution.

Due to the computational cost of convection permitting simulations, previous studies mainly concentrate on specific events in the historical record for model evaluation (e.g Skamarock et al., 1994; Weisman et al., 1997; Lean et al., 2008; Dauhut et al., 2015; Hassim et al., 2016). These events are usually squall lines or other extreme events of historical importance, and this approach has the advantage of showing how well a given model (or model configuration) performs in high-impact situations. Another strategy is to apply tracking methods and then use statistical approaches to study model behaviour, such as 3D reflectivity object tracking (e.g. Caine et al., 2013). Stein et al. (2015) developed a similar statistical diagnostics tool for the evaluation of Numerical Weather Prediction (NWP) models using rainfall and 3D radar reflectivity data, but also focused on one day of shallow convection and one day of deep convection over the U.K. (Keat et al., 2019) then used the same technique to analyse much longer time periods over southern Africa and found that their model produced a realistic diurnal cycle in rainfall at 1.5km horizontal resolution, but that there were still considerable biases in storm size and intensity.

Extreme precipitation or wind events are often forced by larger scale atmospheric dynamics, and it is difficult to determine whether the response to large scale forcing or local processes within the high-resolution domain are the cause of observed biases (e.g. Vincent and Lane, 2016; Peatman et al., 2015; Wapler et al., 2010). In addition, high-resolution modelling on the kilometre scale has become standard for NWP and therefore everyday operational forecasting, not just high-impact events. Therefore, we will concentrate on a period where large scale dynamical forcing is minimal to study the performance of kilometre and sub-kilometre scale convection permitting simulations. This is similar to an earlier study by Caine et al. (2013), although rather than focusing on storm statistics, we will concentrate on the causes of differences in observed rainfall and that produced by convection permitting models. Here, the role of relatively cold surface air in triggering convection has recently come into focus. In particular the creation of surface cold pools resulting from downdrafts and evaporation around existing convective activity has been linked to the triggering of new convection and its organisation into larger systems (e.g. Tompkins, 2001; Schlemmer and Hohenegger, 2014). A similar phenomenon is the land-sea breeze where relatively cold air moves from the ocean over warmer land during the day (sea breeze), and from the cooler land over the sea at night (land breeze) (Pielke, 1974). We summarise these mechanisms under the term 'density currents' in this paper.

As we want to evaluate the behaviour of high-resolution models on short time scales (one focus will be the diurnal cycle) and small spatial scales, it is important to have a high quality observational product to compare model output to. For this reason, we defined our domain according to the Darwin C-band polarimetric (CPOL) radar data domain. Besides the advantage of having negligible orographic forcing in this domain, CPOL data is conveniently available in 10 minute steps and 2.5km grid spacing (Louf et al., 2019). As to large scale dynamical forcing, we know that the weather over Darwin is mainly impacted by the Australian monsoon and its regimes (Pope et al., 2009; May and Ballinger, 2007). For our purposes, we will concentrate on the 'Monsoon Break/Moist Easterly (5)' regime, where large scale forcing is weakest and precipitation is largely determined by the diurnal cycle (Pope et al., 2009). As a consequence, local processes determine the evolution of tropical convection on a day-to-day basis in these simulations.

After introducing the model setups and data which forms the basis of our study, we will examine the main biases of the convection-permitting simulations compared to observations (Section 3), and then perform an in-depth analysis of the thermodynamic structure, diurnal cycle, and the impact of finer horizontal grid spacing in Section 4.

2 | MODEL SETUP AND DATA

As mentioned earlier, we are interested in a time period when local processes are more important to the convective life cycle than large scale atmospheric forcing. Over Darwin, whether or not this is the case is largely determined by the phase of the Australian Monsoon. The phase with the weakest large scale forcing is the Monsoon Break (phase 5 of Pope et al. (2009)), and one of the longest periods of consecutive days of Monsoon Break during the CPOL data period (1998-2017) is February 2006. We will concentrate on the six days between February 14-19 2006.

Observational data comes from the Darwin CPOL radar dataset (Louf et al., 2019) on a 2.5 km grid and a radius of 150km, which includes 3D radar reflectivity and derived rainfall rates every ten minutes. In addition, we will make use of the sounding data at Darwin airport (00 and 12UTC daily) as made available by the University of Wyoming website `http://weather.uwyo.edu/upperair/sounding.html`.

This work is a part of a community effort within the Unified Model Convection Working Group to address model biases and systematic errors for the development of the next generation numerical weather prediction models. Therefore, we concentrate on simulations run with the tropical configuration of the Regional Atmosphere version 1 (RA1T, Bush et al., 2019) of the nesting suite of the UK Met Office's Unified Model (UM, Lean et al., 2008) v10.6, with several nests spanning the kilometre and sub-kilometre scales (4km-145m). As the convective development cycle involves many variables for which no measurements are available, the same simulations are also performed with the Weather Research and Forecasting (WRF) model v3.9 (Skamarock et al., 2008) for comparison to a different model. Both models are setup such that the coarsest nest is driven by ERA-Interim reanalysis data (Dee et al., 2011), and nests without convective parameterisation use horizontal grid spacings of 4, 1.33 km and 444m, as shown in Fig. 1a. Both models are run with 80 vertical levels, with the model top for WRF at 26km, and at 38.5km for the UM.

The UM has an additional domain with horizontal grid spacing of 145m, and the WRF setup includes a 12km horizontal grid spacing domain with convective parameterisation. The physics setup of the UM is set by the 'RA1T' configuration described in Bush et al. (2019). The physics setup for both models is listed in Table 1. The WRF physical parameterisations are the same as in Vincent and Lane (2016) as these choices were found to behave best over the Maritime Continent. We did not change any of the microphysics or other schemes in this study as the UM's philosophy is to provide an 'as-is' configuration and the emphasis of this work is to evaluate the most up-to-date configuration rather than a study of which scheme behaves best under which conditions.

To allow for an appropriately long spinup of the nests while ensuring the models don't drift too far from observations, we ran every day independently. For each day, we initialised the models at 06UTC, and ran for 42 hours until 00UTC of the day after. The first 18 hours were discarded as spinup, leaving exactly 24 hours between midnight UTC and midnight UTC the next day. This accounts for slight discontinuities in the rainfall timeseries of Fig. 2 at 00UTC, but thanks to the choice of time period, rainfall was always close to zero during that time. In addition to these daily restarted simulations, we also ran free running simulations for the entire period with each model (dotted lines in Fig. 2). The biases and timing errors discussed later were also present in the free-running simulations, implying that the initialisation method was not the dominant contributor.

In addition, we found that the largest biases in both models are located over the mainland, and we therefore restrict all data to the domain shown in black in Fig. 1b. In particular, we not only concentrate on land points only, but we also remove the Tiwi Islands, as the local phenomenon of sea breeze convergence over the islands and the resulting nearly daily occurring (during Monsoon Break) 'Hector the Convector' storm has been studied in detail previously (e.g Dauhut et al., 2015, 2016) and it distracts from the model performance over the mainland.

3 | MOTIVATION AND INITIAL EXAMINATION

3.1 | Timing and intensity of rainfall

As a measure of general model performance, Fig. 2 shows domain (as defined by Fig. 1b) mean precipitation as a time series (left) and a diurnal composite (right) for CPOL (black), the UM (blue) and WRF (green) at two resolutions. The supplementary material also includes three-dimensional time evolving animations of the 444m simulations and the CPOL dataset. The reader is encouraged to consult these movies throughout the discussion in this paper. The CPOL time series confirms that precipitation during this period is determined by the diurnal cycle, as no continuous rain event connects to the next day. Another feature is that on February 18 a squall line crosses the domain, seen as large peak in CPOL rainfall a few hours after the diurnal peak that day. This is the reason for the second peak in the CPOL diurnal composite of Fig. 2b. In contrast, WRF (green) generates a squall line every day, with only little precipitation during the peak hours of the diurnal cycle in CPOL. The timing of the squall lines is such that a study of the one real squall line event described above would conclude that WRF captures observed rainfall very well. As a result of the predominance of squall lines in the WRF simulations, the diurnal cycle of rainfall peaks around 12:00UTC, about four hours after the diurnal cycle peak in observations. The UM shows a rather different behaviour: It generates heavy rainfall about two to three hours earlier than the CPOL peak, and then a second peak (which is just as large) during the night around 19:00UTC. We will see later that this behaviour is tightly linked to the generation of density currents. These nightly peaks have different amplitudes in the time series (highest on Feb 18), but this is mainly due to the applied land mask including more or less of the rainfall, rather than day-to-day variability (see supplementary animations and later discussion).

From Fig. 2 it is clear that the differences between models and observations are much larger than the differences between kilometre and sub-kilometre simulations of the same model. This is perhaps not surprising since the higher resolution nests are somewhat constrained by the coarser resolution simulations at the boundaries. Nevertheless, we will see later that even well away from the boundaries and after sufficient spinup, turbulence still acts similarly, which can be seen e.g. in the boundary layer structure. Another interesting point is that besides the many similarities, the early rainfall peak in the UM is even earlier in the 444m simulation than the 1.33km resolution runs, a result which agrees with the findings of Keat et al. (2019).

3.2 | Initiation characteristics

Fig. 3 (top) shows snapshots of precipitation averaged onto the CPOL grid for February 15, 06:00UTC (15:30LST) for (left) CPOL, (middle) UM 444m and (right) WRF 444m. Three points are evident from these snapshots (these are representative of any other day of the simulation period): 1) Both models capture the convection over the Tiwi Islands ('Hector') rather well (but note again that the Tiwi Islands were removed from all domain averages in this paper); 2) CPOL shows initiation over the mainland both near the coast and further inland; 3) whereas the UM misses coastal initiation, it has more localised rainfall inland and WRF has more coastal initiation and less inland rainfall at this time of the day.

3.3 | Evolution characteristics

Fig. 3 (bottom) shows similar snapshots as described above, but now six hours later, at 12:00UTC (21:30LST). CPOL observations show a region of decaying organised convection (this time of the day is past the diurnal cycle peak). While the UM also has a somewhat extended region of non-zero rainfall, there is a small centre of intense rainfall within that structure. WRF on the other hand has produced a squall line crossing the mainland.

Fig. 4 shows snapshots from the supplemental movies of both the UM (top) and WRF (bottom) at 06:00UTC on February 19. The coloured surface shows the potential temperature averaged over the lowest 1km, thus showing density currents (sea/land-breeze and cold pools) in blue shades. More details are given in the figure caption and movie descriptions. The important point here is that both snapshots are taken at the same time, but the sea breeze is much stronger and advances faster over the mainland in the UM than in WRF. This again shows that the two models show qualitatively different behaviour, and neither of the two is particularly close to CPOL observations.

4 | ANALYSIS OF MODEL BEHAVIOUR

We will now link the faster evolution of the density currents to them being stronger and deeper (Rotunno et al., 1988; Lafore and Moncrieff, 1989; Weisman and Rotunno, 2004), and more generally connect density currents and convective activity in the models.

4.1 | Thermodynamic structure

Unfortunately, the thermodynamic structure of the atmosphere cannot be inferred from radar measurements, and during this period soundings are only available for Darwin airport at midnight and midday UTC, which is 9:30 and 21:30 local time. This is due to the global synchronisation of soundings, and is somewhat unfortunate for a study like ours. Nevertheless, some insights can be gained from Fig. 5, which shows the potential temperature and specific humidity profiles for Darwin airport (black), and the closest land grid points for the UM (blue, left) and WRF (green, right). Each profile corresponds to one day.

Both models have a warmer surface than the observations at 00UTC, with a well mixed boundary layer up to about 600m, whereas the soundings suggest much more stable conditions. At this time, WRF shows a lot more day-to-day variability in the boundary layer than the UM, and the latter already produces a distinct superadiabatic surface layer. Above the boundary layer, the models match the soundings well, and the UM generally also matches the LFC with observations. Twelve hours later (12UTC, right), the boundary layer profiles are similar to the soundings, but only the UM has developed the surface inversion, whereas WRF still has a well-mixed boundary layer. Between the heights of 1-2km, the UM is consistently warmer than both WRF and the soundings. As a consequence, the LFC is also about twice as high as observed. In contrast, WRF still matches the observed profiles well.

The specific humidity profiles (second and fourth rows) suggest that both models have reasonable values near the surface, but the UM is more humid just above the boundary layer at 00UTC, and it is drier at 12UTC. Being warmer and drier matches the domain mean downward motion (Fig. 8) around that time (which we will discuss in more detail later). Having more moisture available in the morning (00UTC) might help in triggering convection earlier in the UM. There is no qualitative difference between 1.33 km and 444m resolution for each model.

The supplementary movies and snapshots of Figs. 6 and 4 clearly show that the two models produce very different potential temperature structures within the boundary layer, in particular the evolution of land/sea breezes and cold pools. Fig. 6 shows cross sections of potential temperatures along the transect depicted by the dashed line in Fig. 1b for 444m resolution runs at time 2006-02-19 08:00UTC (17:30 local time). The sea breeze is easily identified by the rapid change of colour from lighter greens to darker blues, and marked with a vertical dashed line gives the position of the sea breeze in the UM, and the dotted line for WRF. The sea breeze in the UM is stronger and deeper (darker blue colours close to the surface) than for WRF over the mainland (roughly right half of the plot), which agrees with theory laid out by Rotunno et al. (1988) and Weisman and Rotunno (2004), who link the strength and depth of density currents to their horizontal propagation speed (loosely related to the 'dam-break' problem). On the other hand, the cold pool behind the sea breeze front (darkest blue colours near the surface near the 50 km mark) is stronger in WRF than the UM.

Fig. 4 shows snapshots for the entire domain at time 2006-02-19 06:00UTC (again at 444m resolution). The updrafts (red volume) in the UM are well aligned with horizontal potential temperature gradients (colour shading). It is striking how much warmer the mainland is in the UM compared to WRF, and how the updrafts are much faster, more localised and distributed throughout the mainland. Also, behind the sea breeze front (within the blueish regions), one can see how the UM has produced multiple small precipitating storms with and after the passage of the sea breeze, whereas WRF only produces weak increases in reflectivity but no noteworthy storms nor precipitation. We will turn our attention to this phenomenon now.

Fig. 7 (top) presents a statistical analysis of these phenomena by plotting two-dimensional histograms of grid cells binned according to strength of potential temperature gradient and updrafts. The potential temperature gradient bins are constructed from hourly maxima of the average $|+\checkmark| = \sqrt{(@_x\checkmark)^2 + (@_y\checkmark)^2}$, computed at the grid scale, within the first kilometre above the surface. Vertical velocity is binned by taking the hourly maximum of $w > 0.01$m/s and averaged between 500m and 2km. The potential temperature gradient represents density currents as it maximises along the fronts of land-sea breezes and cold pools, whereas updraft velocity is used as a diagnostic for convective initiation. With the choice of temporal and spatial filtering this technique allows for both a certain time lag between the passage of density currents and the updrafts being forced above those currents.

For the UM (top left), there are two regimes: One is where the primary maximum is located, which is at low +✓ and low vertical velocity w, and corresponds to stochastic turbulence. The second regime is an almost linear relation between +✓ and w creating a long tail in the distributions of both w and +✓. The first, stochastic turbulent regime is also present in the histogram for WRF (top right), but there is an important qualitative difference to the UM: There is a secondary peak in w at potential temperature gradients of about 0.25–0.5K/km (between the black vertical lines). In this range, w is mostly independent of +✓, suggesting that some different process is at work compared to what is happening in the UM. For potential temperature gradients greater than about 0.5K/km, WRF shows again a similar behaviour to the UM in that there is an almost linear relation between w and +✓.

Looking at the spatio-temporal distribution of the three identified +✓ regimes above, we show in the bottom half of Fig. 7 snapshots of +✓, filtered by (left) +✓ < 0.25K/km, (centre) 0.25K/km < +✓ < 0.5K/km, and (right) +✓ > 0.5K/km, for the UM (left half) and WRF (right half). Three snapshots are shown: (top) 02UTC, (middle) 12UTC, and (bottom) 22UTC. As with Fig. 3, these are examples of a given day (February 14), but they are representative of all days.

It is clear that for the UM, the pre-convective state (02UTC) is characterised by weak +✓ over the ocean and very steep +✓ over the land, as the filter in +✓ appears almost equivalent to a land-sea mask. Given the linear relationship between w and +✓ discovered before, we conclude that these temperature gradients over the land initiate convection too early and spread all over the land masses (note that even though Fig. 7 only shows co-occurrence, the supplementary animations firmly suggest causality). The convection is vigorous and efficient in annihilating the temperature gradients, as already by 12UTC almost exclusively weak +✓ is present, effectively making the entire land mass one giant cold pool. During this time, the mean vertical motion over land is downward (at least in the lower troposphere, see Fig. 8). This cold pool expands over the ocean at night (22UTC) as a strong land breeze and creates intense convective activity when the cold pools from the different land masses collide. This explains the second late peak in diurnal rainfall in Fig. 2. Throughout the diurnal evolution, the mid-range potential temperature gradients do not play any major role for the UM.

In stark contrast, WRF only has steep +✓ close to the coast early in the day, which is closely related to the sea breeze (rightmost column, 02UTC). Over the rest of the land, intermediate +✓ dominates at this pre-convective point in time. Thus, convection does not initiate as vigorously as in the UM early in the day and is mainly associated with the incoming sea breeze (again we refer to the supplemental animations and the snapshots in Fig. 4). Around midday UTC (evening local time), there are important potential temperature gradients associated with the squall line, but there are also intermediate values of +✓ throughout the domain, which, according to the histograms discussed above, can generate a wide variety of updraft strengths. This might favour the creation of larger scale convective systems, but further analysis is required for such conclusions. The late land breeze front (22UTC) is similar to that of the UM but somewhat weaker and therefore results in less precipitation over the ocean during the night.

4.2 | Diurnal cycle

Fig. 7 discussed above suggests that there is a qualitative difference in how the two models are responding to density currents. Fig. 8 explores another possible relation, namely vertical velocity and both CIN and CAPE (which are routinely used as proxies for convective activity). For a broader analysis, Fig. S2 in the supplementary material shows the same plot but with reflectivity instead of vertical velocity and 'MCAPE/MCIN' . The difference is that MCAPE/MCIN are computed from the location of maximum √$_e$ within the lowest 3 km at each grid point, while CAPE/CIN are computed from the surface everywhere, but the conclusions remain unaltered. As with all other domain averages, only grid points over the mainland contribute to these plots.

As expected, the UM shows much larger values of w from the surface to about 15km between about 03–07UTC (about 12–16local) and a secondary peak late at night/early morning (18:30–21UTC/04–06:30local). In contrast, WRF has weaker (but still early) w which then persists until the passage of the squall line 10:00–15:00UTC. There is also a clear difference in depth of the early shallow w peak (starting around 02UTC), which reaches about 2.5km in the UM but is much flatter in WRF and of longer duration. On one hand, the longer time scales and shallower depths of convective development again suggest a broader spectrum of processes determining w in WRF than in the UM, but on the other hand, it also fits with the picture that low-level density currents in the UM determine its convective behaviour: First, they produce stronger low-level updrafts which facilitate initiation of deep convection, then cold pools form and expand, resulting in the secondary updraft peak at around 20:00UTC.

At the time of the day when WRF preferentially produces intense squall lines, the UM produces domain-averaged downdrafts, which are probably linked to the vigorous convection around 06 UTC and already start around 09UTC at lower levels and reach the entire atmosphere above the boundary layer by 15 UTC. This explains why it is drier and warmer than both WRF and the Darwin airport soundings as discussed in Fig. 5. It is interesting that this time period falls exactly into the time of the day when the CPOL diurnal rainfall peaks.

Both CAPE and CIN are quite similar between WRF and the UM up to about 09UTC (18:30 local time), which includes the time around 05–07UTC when the UM produces vigorous convective activity but not WRF. This makes it difficult to attribute the different behaviour to differences in CAPE and/or CIN – if anything, CAPE is slightly smaller and CIN slightly larger for the UM than for WRF, which would make the UM less likely to produce deep convection than WRF. Later in the day, the passage of the squall line in WRF reduces CAPE and at the same time increases CIN , which explains why there is no convective activity during the night when the land breezes from the mainland and the Tiwi Islands collide (which is causing the second peak in the UM). Note that there are slightly different choices in variables for this analysis, but we found that similar conclusions hold when compositing the diurnal cycle of reflectivity rather than vertical velocity, and using 'MCAPE/MCIN' rather than surface CAPE/CIN. This is shown in the supplementary material.

In an attempt to better understand the differences between the two models at around 06UTC, Fig. 9 shows the same profiles as Fig. 5, but at 06UTC (without the observed soundings due to their unavailability). The vertical profiles for (left) specific humidity and (right) potential temperature of both models are shown on the same plot for easier direct comparison.

For both resolutions, the humidity (left column) is very similar between the two models, including at the surface. Above the boundary layer, one could make a case that the UM is somewhat warmer (right column), but not on all days. At sub-kilometre scale resolution (bottom row), the UM has a deeper boundary layer than WRF, about 700m compared to about 300–400m. This combined with the discussion of the +√ vs. w relationship in Fig. 7 and the discussion of the soundings in Fig. 5 might hint to differences in boundary layer behaviour, suggesting that future work should concentrate on either further increasing resolution (both horizontal and vertical) to properly resolve boundary layer physics, and/or investigate the behaviour of WRF with different boundary layer schemes.

4.3 | Horizontal resolution sensitivity

After studying the diurnal cycle, we devote the last section to the examination of the full six-day simulations as a whole, with an emphasis on the effects of increased horizontal resolution.

So far, we have seen that the differences between the UM and WRF are much larger than the differences between simulations with different horizontal resolution of the same model. Focusing on the UM only, we show in Fig. 10 the 6-day rainfall accumulation (i.e. total precipitation) over the six-day period investigated in this work. We now also include a 145m resolution simulation which we only performed with the UM and within a smaller domain. The smaller domain was chosen to be around an area of complex coastlines as shown in Fig. 1 to check whether it would produce more coastal convection as seen in CPOL.

All resolutions have much higher peak total precipitation than CPOL, and extended regions where little to no rain falls over the entire simulation period. The radar observations (top left) show a much different picture, where some rain falls everywhere within the domain over the six days. Together with the findings of Section 3.3 and Fig. 11 described below, this suggests a dominance of single small scale storms in the UM, whereas in observations larger systems pass over the domain. This is similar to the conclusions of Stein et al. (e.g. 2015); Keat et al. (e.g. 2019) who applied storm size statistics to UM and radar data.

Fig. 11 analyses total precipitation in a different way, by creating histograms of number of grid points within the domain which contain a given total precipitation for the 24hours after spinup. For this analysis, all model data was conservatively re-gridded onto the CPOL grid. There is a qualitative difference between the 1.33km simulation (green) and all other simulations (and CPOL): Whereas the kilometre-scale simulation shows a peak at zero precipitation (i.e. most of the grid cells remain dry), all other curves do not have any completely dry cells, and show a high proportion of light rain. As resolution increases, the peak of the PDF gradually moves towards the peak location of CPOL. However, both the 444m and 145m simulations have a much larger tail with high amounts of rainfall than the CPOL dataset. Closer inspection reveals that the tails originate from the second rainfall peak in the diurnal cycle (Fig. 2), which are mostly due to colliding density currents as described above. Thus, increasing resolution does give some improvement on the low rainfall side of the distribution, but the precipitation bias linked to density currents is exacerbated. Hence, we should be cautious about the benefits of horizontal resolutions of the order of 100m. We conclude from our work that there are systematic model errors which do not vanish with higher resolution, and although there seems to be a step change between 1.33km and 444m here, resolutions of the order of 100m show little improvement over those of the order of 500m. Therefore, it is probably better to investigate those model errors with sub-kilometre but not O(100 m) simulations as they show the same errors but are less demanding in both CPU time and storage requirements.

A similar conclusion can be drawn from the kinetic energy spectra (following Skamarock (2004), Fig. 12) and the similarly computed spectra of +√ (Fig. 13). Again the differences between the models are larger than between resolutions. In particular, WRF shows a higher effective resolution (defined as the length scale below which the energy spectrum deviates from the power law of the inertial range) at equal grid spacing, and the UM has distinctly enhanced energy content at the smallest scales, 2?x, i.e. the right endpoints of the lines in Fig. 12 (?x denotes horizontal grid spacing). The latter is a known feature of high-resolution models (Errico, 1985; Skamarock, 2004), but it is much more pronounced for the UM (solid lines for the UM vs. dashed for WRF in Fig. 12). While Fig. 12 represents an average over the entire simulation, Fig. 13 shows the diurnal cycle of near-surface +√ spectrum for all model configurations. Note that here the spectra have been normalised to the time mean to show the diurnal cycle more clearly, i.e. we show $\log_{10}$ (T/hT i), where T is the spectrum of +√ within the lowest kilometre over the entire model domain as a function of time and h·i denotes time mean.

These spectra show how similar all six days are within each simulation and how increased resolution does not change the picture qualitatively. For the UM, each day the period of generally higher +√ spectrum (left column, red shading) is initiated with an increased peak at the smallest scales, seen by the downward curving (i.e. earlier appearance) of the right side of the red shading. In contrast, WRF does show the initial peak in the smallest scales (right column), but this does not directly connect with the evolution of the larger scales in the spectrum. This early peak is the convective activity related to the sea breeze (best seen in the supplemental movies), where the UM produces many independent small scale storms but WRF only produces a slight increase in reflectivity but not much rainfall. The primary peak in the UM appears during early afternoon local time , whereas it is much later in the day for WRF, where it is connected to the squall lines produced by that model. In this sense, the spectra of +√again show the tight relationship between low level density currents and rainfall over the mainland.

5 | SUMMARY AND CONCLUSIONS

We have run kilometre and sub-kilometre scale convection permitting simulations over several days and during a period of predominantly locally forced convective activity. The aim of this was to expose the triggering and growth of convection in such models, rather than studying how they behave when strongly forced by high impact/extreme events such as squall lines, as the latter can drown model-specific behaviour and biases within the response to large-scale forcing or other extreme events. Understanding such intrinsic behaviour should then better allow for bias corrections, which in turn makes models more reliable and better performing even in strongly forced situations. The domain centred around Darwin Australia has been setup in such a way that direct comparison to high quality radar observations and twice daily soundings was possible. The boundary conditions came from ERA-Interim reanalysis. Two different models, namely the Unified Model (UM) and the Weather Research and Forecasting (WRF) model, were run in nested configurations which were as similar as possible.

Even though configured with the same boundary conditions and during a time of locally forced convection, the two models behave very differently, producing different rainfall compared to each other and also the radar observations (Fig. 2). While some measures indicate a performance improvement with higher resolution, such as total precipitation over the domain (Fig. 11), the differences between resolutions with the same model are much smaller than the differences between models at same resolution, indicating that the main sources of the discrepancies to observations come from systematic errors in each model, rather than insufficient resolution of particular processes. One important exception to this might be the boundary layer, as even our sub-kilometre resolution simulations (at 444 m and 145 m) do not fully resolve the three-dimensional boundary layer physics. Indeed, we found indications that boundary layer physics is one of the most important sources of the different behaviour between the two models and observations (Figs. 5 and 9).

One obvious difference between the UM and WRF can be found in the simulated kinetic energy spectra (Fig. 12): The UM has a disproportionate amount of energy at the smallest (2?x) scales, and it also has a lower effective resolution than WRF. This might be related to the the abundance of small and individual objects of intense convective activity often observed in convection permitting simulations with the UM (Hanley et al., 2015) (and sometimes referred to as ‘blobbiness’).

A second difference is that the UM has a direct relationship between the depth and strength of density currents and vertical velocity, whereas WRF produces a range of updraft strengths above similar potential temperature gradients (Fig. 7). The UM produces strong density currents early during the day and over the land, resulting in early onset of convection. Interestingly, CAPE and CIN do not explain the different behaviour (Fig. 8).

This study cannot determine which of the two models is better performing for situations of locally forced tropical convection over the land, neither was it designed to do so. Rather, both models have intrinsic behaviour which they repeat day after day of simulation, and neither is particularly close to observations. Rather, it points to further necessary work to understand the effects of resolution, dynamical core and boundary layer parameterisations. For instance, there is an opportunity to test the different boundary layer schemes in WRF, and also the midlatitude (“M”) version of the Regional Atmosphere setups of the UM (Bush et al., 2019), and the impact of near-surface vertical resolution should be investigated. Either way, it seems clear that simply increasing model resolution does not resolve the biases in diurnal rainfall.

6 | ACKNOWLEDGEMENTS

This work has been supported by the ARC Centre of Excellence for Climate Extremes under grant CE170100023, and MJ has received additional support from ARC grant FL150100035. This work made extensive use of the *xarray* package (Hoyer and Hamman, 2017) and was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

REFERENCES

M. J. Best, M. Pryor, D. B. Clark, G. G. Rooney, R .L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding. The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. Geoscientific Model Development, 4(3):677–699, 2011. ISSN 1991-9603. doi: 10.5194/gmd-4-677-2011.

I. A. Boutle, J. E.J. Eyre, and A. P. Lock. Seamless stratocumulus simulation across the turbulent gray zone. Monthly Weather Review, 142(4):1655–1668, 2014. ISSN 15200493. doi: 10.1175/MWR-D-13-00229.1.

Mike Bush, Tom Allen, Caroline Bain, Ian Boutle, John Edwards, Anke Finnenkoetter, Charmaine Franklin, Kirsty Hanley, Humphrey Lean, Adrian Lock, James Manners, Marion Mittermaier, Cyril Morcrette, Rachel North, Jon Petch, Chris Short, Simon Vosper, David Walters, Stuart Webster, Mark Weeks, Jonathan Wilkinson, Nigel Wood, and Mohamed Zerroukat. The first Met Office Unified Model/JULES Regional Atmosphere and Land configuration, RAL1. Geoscientific Model Development Discussions, (June):1–47, 2019. ISSN 1991-959X. doi: 10.5194/gmd-2019-130.

Simon Caine, Todd P. Lane, Peter T. May, Christian Jakob, Steven T. Siems, Michael J. Manton, and James Pinto. Statistical Assessment of Tropical Convection-Permitting Model Simulations Using a Cell-Tracking Algorithm. Monthly Weather Review, 141(2):557–581, 2 2013. ISSN 0027-0644. doi: 10.1175/MWR-D-11-00274.1. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-11-00274.1.

Thibaut Dauhut, Jean-pierre Chaboureau, Juan Escobar, and Patrick Mascart. Large-eddy simulations of Hector the convector making the stratosphere wetter. Atmospheric Science Letters, 16(2):135–140, 4 2015. ISSN 1530261X. doi: 10.1002/asl2.534. URL http://doi.wiley.com/10.1002/asl2.534.

Thibaut Dauhut, Jean-Pierre Chaboureau, Juan Escobar, and Patrick Mascart. Giga-LES of Hector the Convector and Its Two Tallest Updrafts up to the Stratosphere. Journal of the Atmospheric Sciences, 73(12):5041–5060, 12 2016. ISSN 0022-4928. doi: 10.1175/JAS-D-16-0083.1. URL http://journals.ametsoc.org/doi/10.1175/JAS-D-16-0083.1.

D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656):553–597, 4 2011. ISSN 00359009. doi: 10.1002/qj.828. URL http://doi.wiley.com/10.1002/qj.828.

J. M. Edwards and A. Slingo. Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. Quarterly Journal of the Royal Meteorological Society, 122(531):689–719, 1996. ISSN 00359009. doi: 10.1256/smsqj.53106.

Ronald M. Errico. Spectra Computed from a Limited Area Grid. Monthly Weather Review, 113(9):1554–1562, 9 1985. ISSN 0027-0644. doi: 10.1175/1520-0493(1985)113<1554:SCFALA>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281985%29113%3C1554%3ASCFALA%3E2.0.CO%3B2.

Kirsty E. Hanley, Robert S. Plant, Thorwald H.M. Stein, Robin J. Hogan, John C. Nicol, Humphrey W. Lean, Carol Halliwell, and Peter A. Clark. Mixing-length controls on high-resolution simulations of convective storms. Quarterly Journal of the Royal Meteorological Society, 141(686):272–284, 2015. ISSN 1477870X. doi: 10.1002/qj.2356.

M. E. E. Hassim, T. P. Lane, and W. W. Grabowski. The diurnal cycle of rainfall over New Guinea in convection-permitting WRF simulations. Atmospheric Chemistry and Physics, 16(1):161–175, 1 2016. ISSN 1680-7324. doi: 10.5194/acp-16-161-2016. URL http://www.atmos-chem-phys.net/16/161/2016/.

Stephan Hoyer and Joseph J. Hamman. xarray: N-D labeled Arrays and Datasets in Python. Journal of Open Research Software, 5:1–6, 4 2017. ISSN 2049-9647. doi: 10.5334/jors.148. URL http://openresearchsoftware.metajnl.com/articles/10.5334/jors.148/.

William J. Keat, Thorwald H. M. Stein, Elelwani Phaduli, Stephanie Landman, Erik Becker, Mary-Jane M. Bopape, Kirsty E. Hanley, Humphrey W. Lean, and Stuart Webster. Convective initiation and storm life cycles in convection-permitting simulations of the Met Office Unified Model over South Africa. Quarterly Journal of the Royal Meteorological Society, 145(721):1323–1336, 4 2019. ISSN 0035-9009. doi: 10.1002/qj.3487. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3487.

J. P. Lafore and M. W. Moncrieff. A Numerical Investigation of the Organization and Interaction of the Convective and Stratiform Regions of Tropical Squall Lines. Journal of the Atmospheric Sciences, 46(4):521–544, 2 1989. ISSN 0022-4928. doi: 10.1175/1520-0469(1989)046<0521:ANIOTO>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1989)046%3C0521:ANIOTO%3E2.0.CO;2.

Humphrey W. Lean, Peter a. Clark, Mark Dixon, Nigel M. Roberts, Anna Fitch, Richard Forbes, and Carol Halliwell. 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, 2008. ISSN 0027-0644. doi: 10.1175/2008MWR2332.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2332.1.

A. P. Lock, A. R. Brown, M. R. Bush, G. M. Martin, and R. N.B. Smith. A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Monthly Weather Review, 128(9):3187–3199, 2000. ISSN 00270644. doi: 10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

Valentin Louf, Alain Protat, Robert A. Warren, Scott M. Collis, David B. Wolff, Surendra Raunyiar, Christian Jakob, and Walter A. Petersen. An integrated approach to weather radar calibration and monitoring using ground clutter and satellite comparisons. Journal of Atmospheric and Oceanic Technology, 36(1):17–39, 2019. ISSN 15200426. doi: 10.1175/JTECH-D-18-0007.1.

Peter T. May and Andrew Ballinger. The Statistical Characteristics of Convective Cells in a Monsoon Regime (Darwin, Northern Australia). Monthly Weather Review, 135(1):82–92, 2007. ISSN 0027-0644. doi: 10.1175/MWR3273.1. URL http://www.scopus.com/inward/record.url?eid=2-s2.0-33846658969&partnerID=tZOtx3y1.

Paul A. O'Gorman and John G. Dwyer. Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events. Journal of Advances in Modeling Earth Systems, 2018. ISSN 19422466. doi: 10.1029/2018MS001351.

Simon C. Peatman, Adrian J. Matthews, and David P. Stevens. Propagation of the Madden???Julian Oscillation and scale interaction with the diurnal cycle in a high-resolution GCM. Climate Dynamics, 45(9-10):2901–2918, 2015. ISSN 14320894. doi: 10.1007/s00382-015-2513-5.

Roger A. Pielke. A Three-Dimensional Numerical Model of the Sea Breezes Over South Florida. Monthly Weather Review, 102(2):115–139, 2 1974. ISSN 0027-0644. doi: 10.1175/1520-0493(1974)102<0115:ATDNMO>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281974%29102%3C0115%3AATDNMO%3E2.0.CO%3B2.

Mick Pope, Christian Jakob, and Michael J. Reeder. Regimes of the north Australian wet season. Journal of Climate, 22(24):6699–6715, 2009. ISSN 08948755. doi: 10.1175/2009JCLI3057.1.

Richard Rotunno, Joseph B. Klemp, and Morris L. Weisman. A Theory for Strong, Long-Lived Squall Lines. Journal of the Atmospheric Sciences, 45(3):463–485, 2 1988. ISSN 0022-4928. doi: 10.1175/1520-0469(1988)045<0463:ATFSLL>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%281988%29045%3C0463%3AATFSLL%3E2.0.CO%3B2.

Linda Schlemmer and Cathy Hohenegger. The formation of wider and deeper clouds as a result of cold-pool dynamics. Journal of the Atmospheric Sciences, 71(8):2842–2858, 2014. ISSN 15200469. doi: 10.1175/JAS-D-13-0170.1.

William C. Skamarock. Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. Monthly Weather Review, 132(12):3019–3032, 2004. ISSN 0027-0644. doi: 10.1175/MWR2830.1. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR2830.1.

William C. Skamarock, Morris L. Weisman, and Joseph B. Klemp. Three-Dimensional Evolution of Simulated Long-Lived Squall Lines, 1994. ISSN 0022-4928.

William C. Skamarock, Joseph B. Klemp, Jimy Dudhia, David O. Gill, Dale M. Barker, Michael G. Duda, Xiang-Yu Huang, Wei Wang, and Jordan G. Powers. A Description of the Advanced Research WRF Version 3. Technical report, 2008. URL https://opensky.ucar.edu/islandora/object/technotes:500.

Fengfei Song and Guang J. Zhang. Understanding and improving the scale dependence of trigger functions for convective parameterization using cloud-resolving model data. Journal of Climate, 31(18):7385–7399, 2018. ISSN 08948755. doi: 10.1175/JCLI-D-17-0660.1.

Thorwald H. M. Stein, Robin J. Hogan, Peter A. Clark, Carol E. Halliwell, Kirsty E. Hanley, Humphrey W. Lean, John C. Nicol, and Robert S. Plant. The DYMECS Project: A Statistical Approach for the Evaluation of Convective Storms in High-Resolution NWP Models. Bulletin of the American Meteorological Society, 96(6):939–951, 2015. ISSN 0003-0007. doi: 10.1175/BAMS-D-13-00279.1. URL http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-13-00279.1.

Adrian M. Tompkins. Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Cold Pools. Journal of the Atmospheric Sciences, 58(6):529–545, 3 2001. ISSN 0022-4928. doi: 10.1175/1520-0469(2001)058<0529:OOTCIL>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%282001%29058%3C0529%3AOOTCIL%3E2.0.CO%3B2.

Claire L. Vincent and Todd P. Lane. Evolution of the Diurnal Precipitation Cycle with the Passage of a Madden–Julian Oscillation Event through the Maritime Continent. Monthly Weather Review, 144(5):1983–2005, 2016. ISSN 0027-0644. doi: 10.1175/MWR-D-15-0326.1. URL http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0326.1.

Kathrin Wapler, Todd P. Lane, Peter T. May, Christian Jakob, Michael J. Manton, and Steven T. Siems. Cloud-System-Resolving Model Simulations of Tropical Cloud Systems Observed during the Tropical Warm Pool-International Cloud Experiment. Monthly Weather Review, 138(1):55–73, 2010. ISSN 0027-0644. doi: 10.1175/2009MWR2993.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2009MWR2993.1.

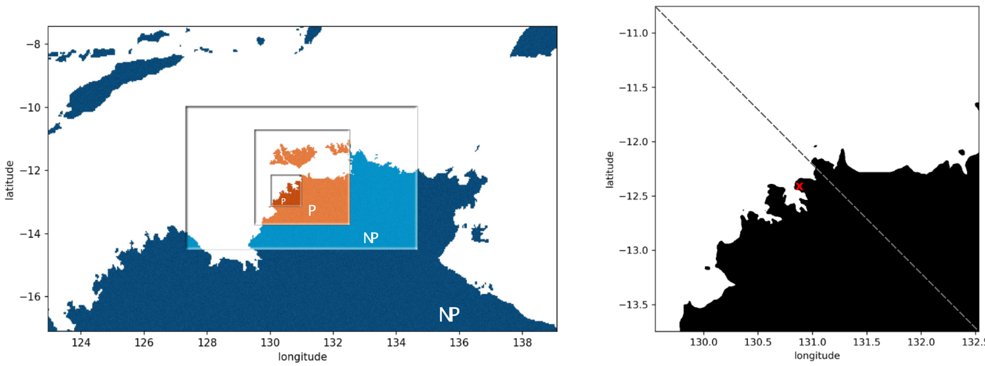
Morris L. Weisman and Richard Rotunno. "A theory for strong long-lived squall lines" revisited. Journal of the Atmospheric Sciences, 61(4):361–382, 2004. ISSN 00224928. doi: 10.1175/1520-0469(2004)061<0361:ATFSLS>2.0.CO;2.

Morris L. Weisman, William C. Skamarock, and Joseph B. Klemp. The resolution dependence of explicitly modeled convective systems. Monthly Weather Review, 125(4):527–548, 1997. ISSN 00270644. doi: 10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2.

Damian R. Wilson and Susan P. Ballard. A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Quarterly Journal of the Royal Meteorological Society, 125(557):1607–1636, 1999. ISSN 00359009. doi: 10.1256/smsqj.55706.

TABLE 1 WRF and UM physics setup.

Physics	WRF	UM
Version	3.9	10.6 / RA1T
Planetary boundary layer	MYJ TKE	Lock et al. (2000) and Boutle et al. (2014)
Surface layer	MO (Janjić Eta)	JULES (Best et al., 2011)
Microphysics	WSM 6–class graupel	based on Wilson and Ballard (1999)
Longwave radiation	RRTM	based on Edwards and Slingo (1996)
Shortwave radiation	Goddard	based on Edwards and Slingo (1996)
Cumulus (12km only)	Kain–Fritsch	N/A
Time step [s] (4 km, 1.33km, 444m, 145m)	10, $3\frac{1}{3}$, $1\frac{1}{9}$, N/A	120, 40, 12, 4
Vertical levels (top)	80 (25 km)	80 (38.5 km)
Boundary and initial conditions	ERA–Interim, updated every 6 hours.	
Run length, spinup	6 independent 42 hour simulations initiated at 06UTC.	
	First 18 hours discarded for spinup, last 24 hours used for analysis.	



(a) Domain setup. For both models, the outermost domain is nested inside ERA-Interim reanalysis, and the setup includes convection-permitting nests of horizontal grid spacings of 4km, 1.33km, 444m and 145m (UM only). WRF is set up with an additional outer nest of 12km with parameterised convection.

(b) Land mask used for "domain average" in this study (the islands where removed intentionally, see main text). The red cross marks Darwin Airport, where the soundings are taken. The gray dashed line denotes the cross section in Fig. 6.

FIGURE 1 Domain setup and land mask.

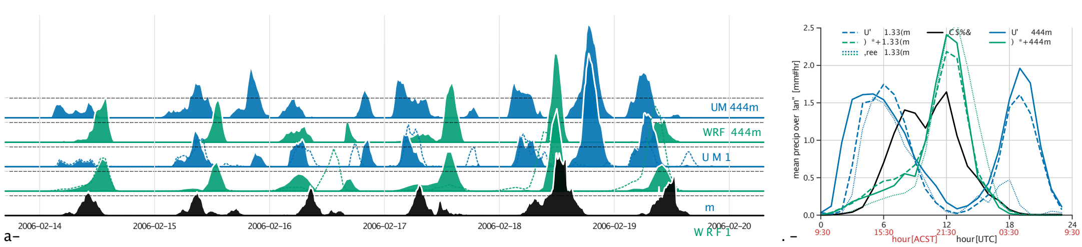


FIGURE 2 a) Domain mean (using landmask of Fig. 1b) rainfall rate timeseries and b) domain mean diurnal mean rainfall rates [mm/hr] for WRF (green) and the UM (blue), compared to CPOL data (black). The thin horizontal lines in a) mark the 90th percentile value of the CPOL timeseries (about 1.5mm/hour). The dotted lines show rainfall rates for the free running experiments.

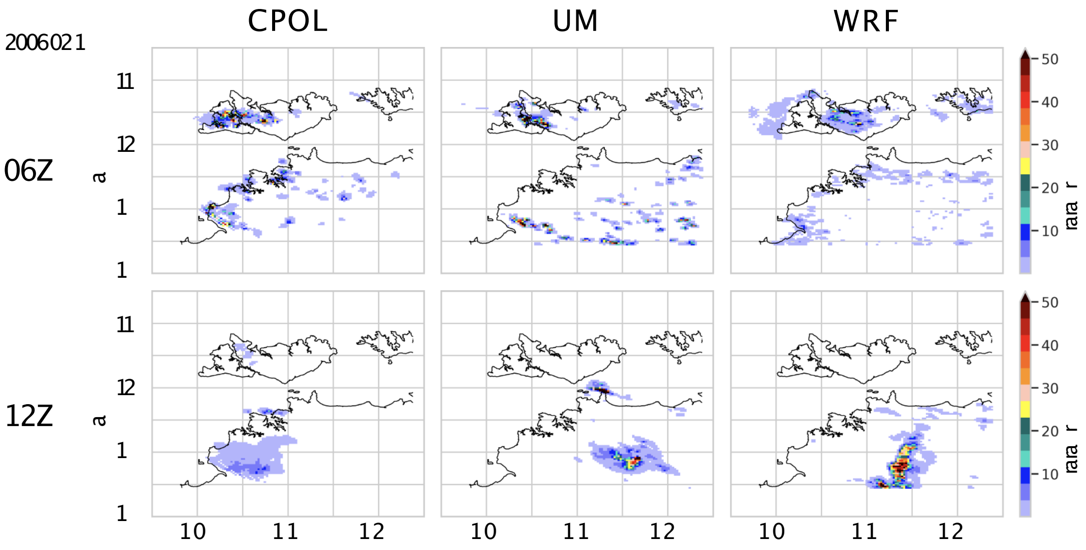
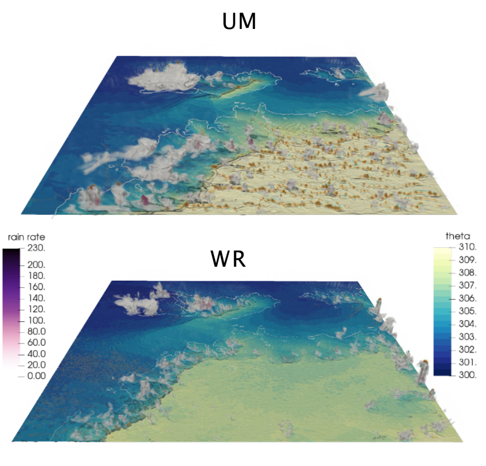
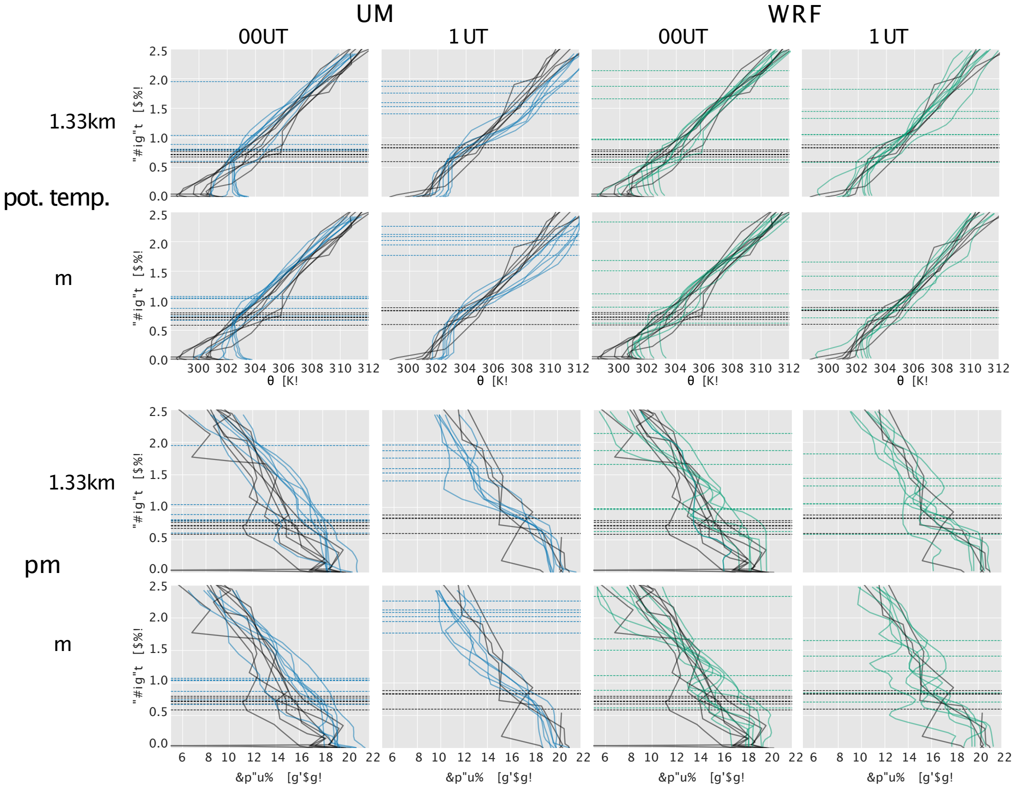


FIGURE 3 InstantaneousrainratesfromCPOL (left), Unified Model (centre)andWRF (right) in theafternoon(top) andat
night (bottom) at ?x=444 m andfor one given day. These snapshotsare similar for all days.



FIGU RE4 3D snapshotsofsurfacepotentialtemperature(colourshading),reflectivity (white volume; >10dBZ), CAPE (black contoursat 3,4,5x10^3J/kg), updrafts(red volume) and precipitation (purple “peaks”) for the UM (top) andWRF (bottom). The snapshotscorrespondto 06:00UTC on 2006-02-19.Animationsof the full six days of simulations can be found at https://youtu.be/xZmnmxOlsPEforthe UM, https://youtu.be/GqLip-bbLioforWRF and https://youtu.be/Dt3LEaRNfRE for CPOL.



FIGU RE 5 (First and second row) Vertical potential temperature profiles at Darwin airport (or the closest land grid point) for the UM (left, blue), WRF (right, green) and sounding data (black). (third and fourth row) Vertical specific humidity profiles at the same location. Dashed horizontal lines represent the level of free convection (LFC). Plotted is one line per day, at 00UTC (left) and 12UTC (right) to match the balloon soundings. Sounding data from the University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html).

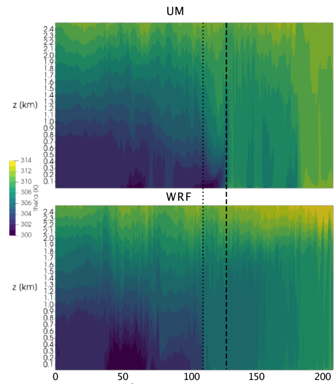
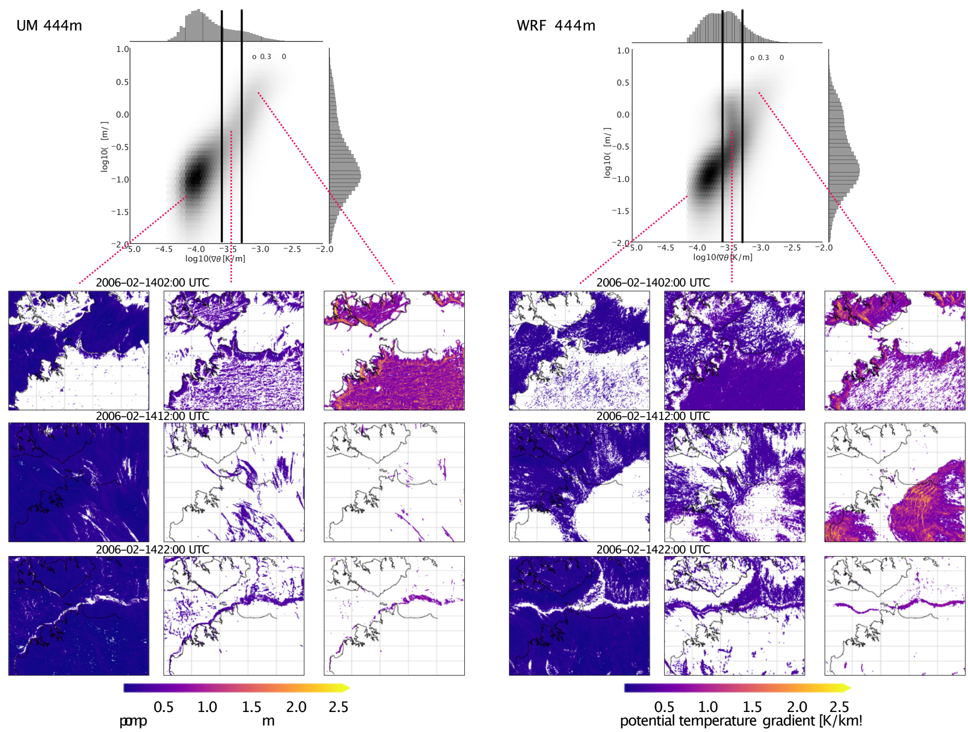


FIGURE 6 2D snapshots of 444m resolution potential temperature along a diagonal cross section going from the top left to the bottom right corner of the domain, as shown by the dashed line in Fig. 1b). UM is shown on top and WRF on the bottom, and both snapshots correspond to 08:00 UTC on 2006-02-19. The sea breeze has advanced faster in the UM (dashed vertical line) than WRF (dotted vertical line). It is also stronger and deeper, which is agreement with the theory of Weisman and Rotunno (2004). The intense cold pool behind the sea breeze front (darkest colours near the surface around 50km mark) is related to Hector over the Tiwi Islands.



(a) UM 444m (b) WRF 444m

FIGU RE 7 (top) Two-dimensionalhistogramsof updrafts(vertical velocity w > 0.01m/s) versus horizontal potential temperature gradient |+√|. Both axes are logarithmic. See text for details. (bottom) Domain snapshots for 2006-02-14andUTC 02:00 (prior to onsetof convection), 12:00 (after the passageof the sea breeze) and 22:00 (land breeze) of +√for (left) log10(+√) < @3.6(0.25K/km), (middle) @3.6 < log10(+√) < @3.3(0.5K/km) and (right) log10(+√) > @3.3. All days show very similar evolution. For the UM at 22UTC the middle and right panels look similar as there is essentially only one narrow region of significant |+√| along the edge of the land breeze.

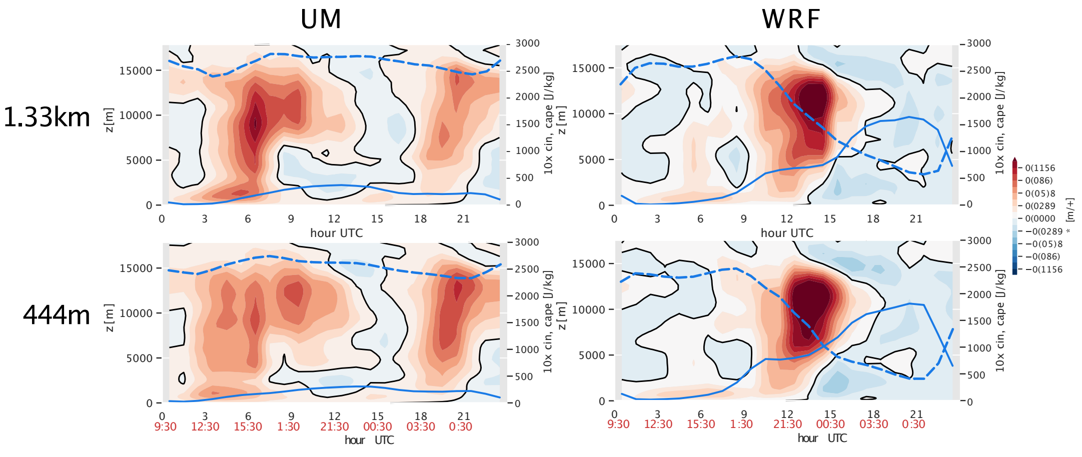


FIGURE 8 Diurnal composite of domain averaged vertical velocity (shading), surface CAPE (dashed blue line) and CIN (solid blue line) averaged over all mainland points for (left) the UM and (right) WRF and (top) 1.33km and (bottom) 444m resolution. Note that the values for CIN where multiplied by 10 to use the same y-scale.

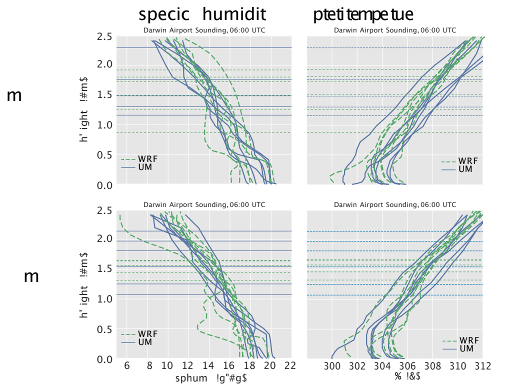
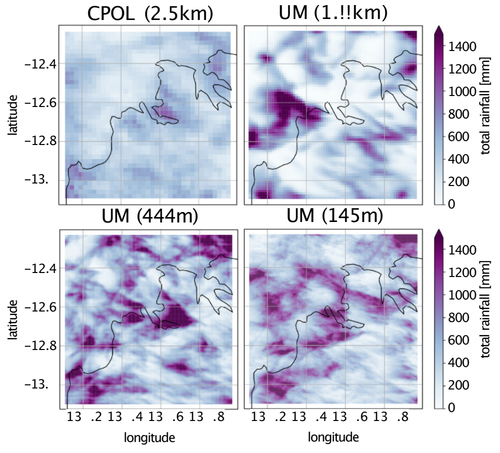


FIGURE 9 Soundings as in Fig. 5, but at 6:00UTC, where no balloon soundings are available. Specific humidity is shown on the left, and potential temperature on the right. The top row is for 1.33km and the bottom row for 444m resolution.



F I GU R E 10 Total precipitation over the six-day period investigated in this work within the smallest (and highest resolution) domain for (top left) CPOL radar observations and the UM at (top right) 1.33 km, (bottom left) 444 m, and (bottom right) 145 m resolutions. Note that these plots are shown on their native grids within the smallest domain which corresponds to the 145 m domain (minus 20 grid points on each side to remove boundary effects).

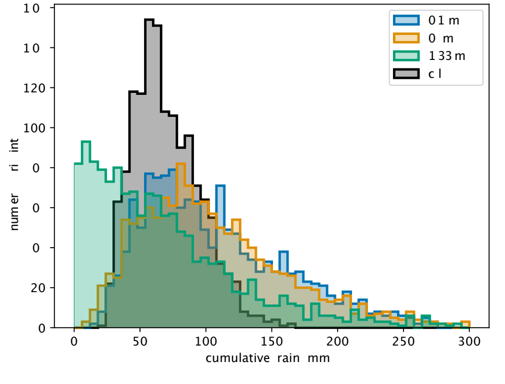
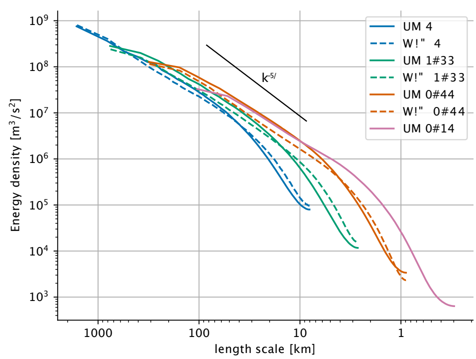


FIGURE 11 PDFs of UM (colours) and CPOL (black) total precipitation from Fig. 10, but now conservatively interpolated onto the CPOL 2.5 km grid. Only the 1.33 km simulation (green) shows a peak at zero precipitation, but all simulations have a much larger high-precipitation tail than the CPOL dataset (black).



F I GU R E 12 Kinetic energy spectra for the UM (solid) and WRF (dashed) 4km (blue), 1.33km (green), 444m (red) and 145m (magenta) simulations. The short solid black line shows the $k^{@5/3}$ slope. The effective resolution (where the wave spectra begin to fall off) is higher for WRF, and the UM has a more pronounced energy surplus at 2?x. Spectra are computed following Skamarock (2004), and are averaged in space (full domain between 3 and 9 km height) and time.

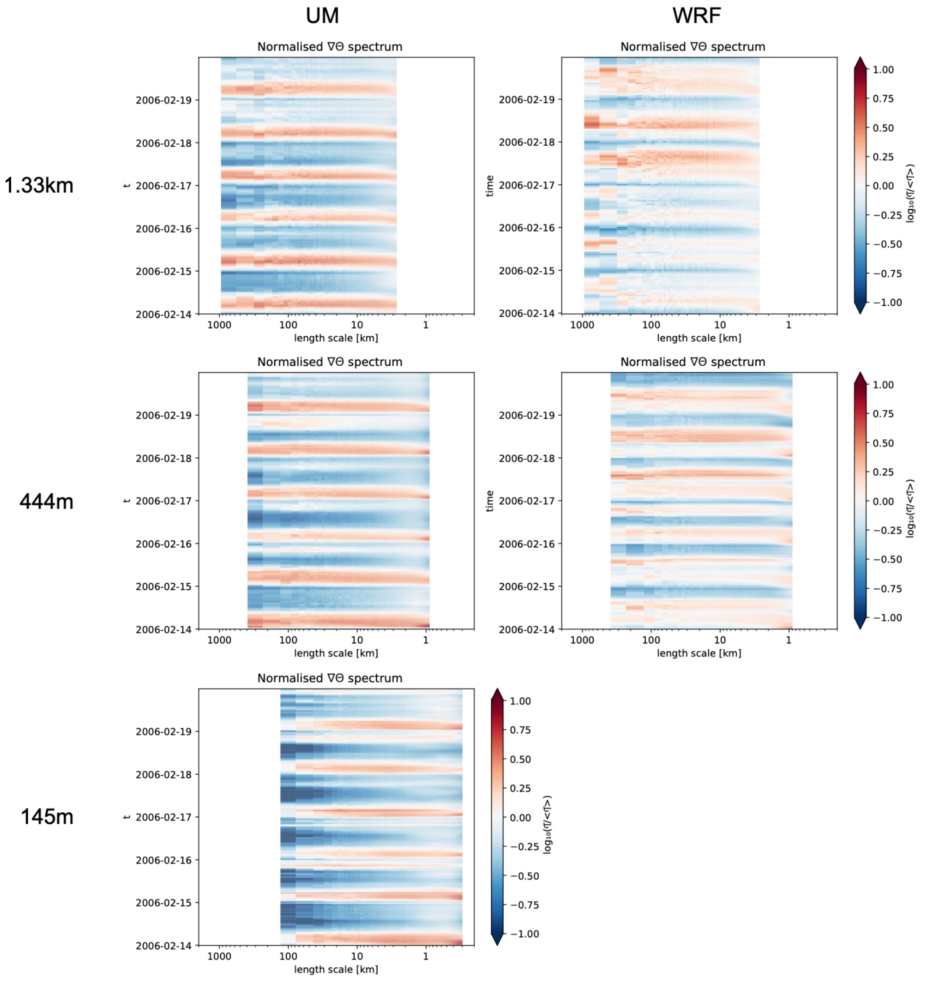
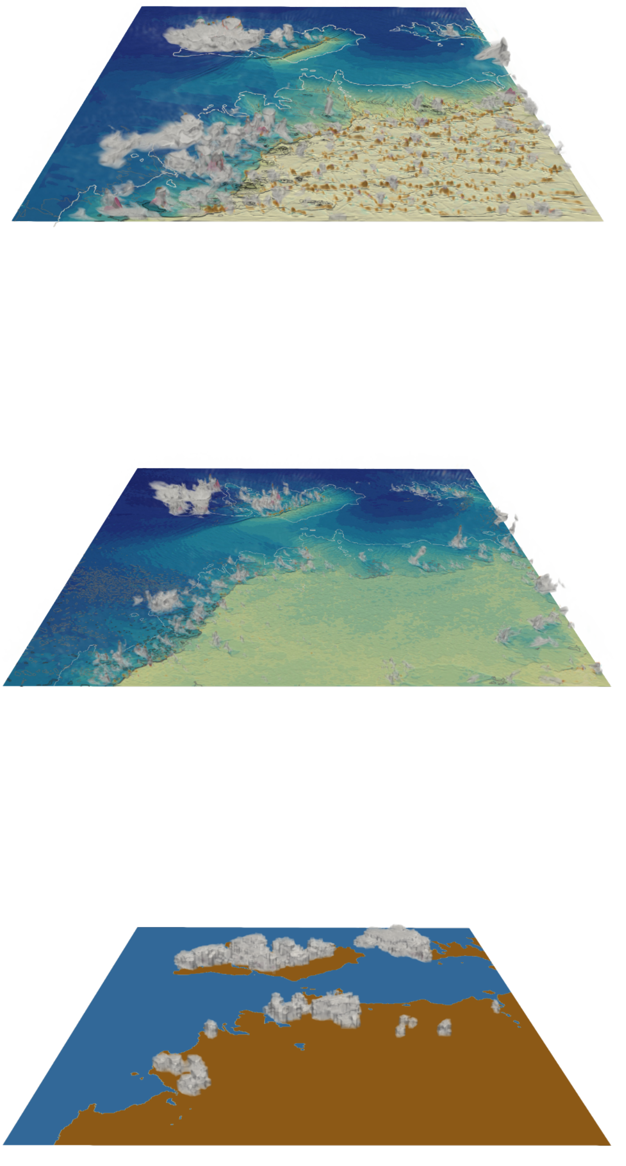


FIGURE 13 Hovmøller-like diagram of the evolution of the normalised spectrum of $\nabla\theta$ within the lowest 1 km for (left column) the UM and (right column) WRF, and (top) 1.33 km and (middle) 444 m and (bottom) 145 m.

SUPPLEMENTARY MATERIAL



FIGURES 1 Click on images to see movies on YouTube. The movies are full 6-day animations of the (top) UM and (middle) WRF 444m simulations showing surface potential temperature, reflectivity, vertical velocity, precipitation and CAPE. (bottom) The same but for the CPOL dataset, which means there is only reflectivity and precipitation. From these movies it is very clear how the two models behave very differently, and how both are different again from observations.

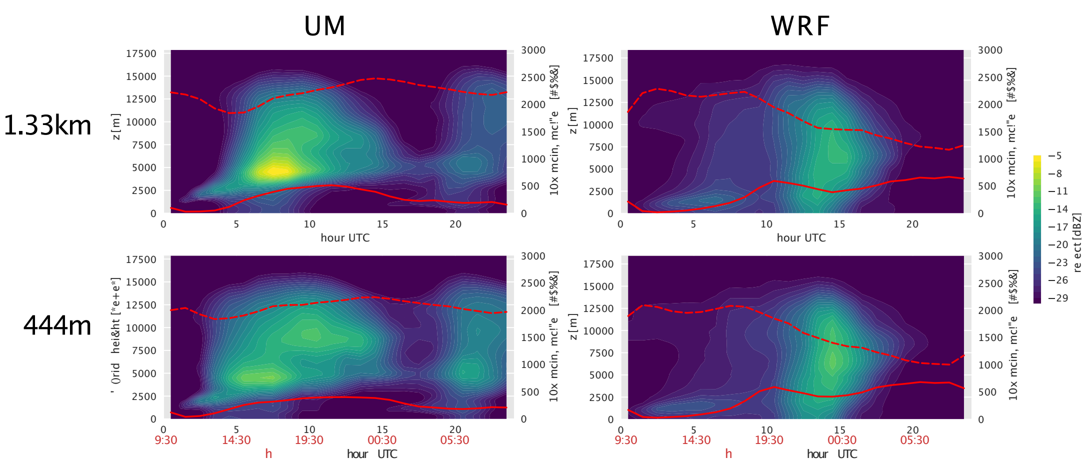


FIGURE S2 Same as Fig. 8, but showing the diurnal composite of reflectivity [dBz] rather than vertical velocity, and MCAPE/MCIN instead of CAPE/CIN.