Convective initiation and storm life-cycles in convection-permitting simulations of the Met Office Unified Model over South Africa

Received: 01 October 2018 | Revised: 18 January 2019 | Accepted: 22 January 2019

DOI: 10.1002/qj.3487

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

Convective initiation and storm life cycles in convection-permitting simulations of the Met Office Unified Model over South Africa

William J. Keat[1] | Thorwald H. M. Stein[1] | Elelwani Phaduli[2] | Stephanie Landman[2] | Erik Becker[2] | Mary-Jane M. Bopape[2] | Kirsty E. Hanley[3] | Humphrey W. Lean[3] | Stuart Webster[4]

[1]Department of Meteorology, University of Reading, Reading, UK
[2]South African Weather Service, Pretoria, South Africa
[3]Met Office at Reading, Department of Meteorology, University of Reading, Reading, UK
[4]Met Office, Exeter, UK

Correspondence
William Keat, Department of Meteorology, University of Reading, Earley Gate, Reading, Berkshire, RG6 6BB, UK.
Email: william.keat@reading.ac.uk

Funding information
Newton Fund

Convective initiation is a challenge for convection-permitting models due to its sensitivity to sub-km processes. We evaluate the representation of convective storms and their initiation over South Africa during four summer months in Met Office Unified Model simulations at a 1.5-km horizontal grid length. Storm size distributions from the model compare well with radar observations, but rainfall in the model is predominantly produced by large storms (50 km in diameter or larger) in the evening, whereas radar observations show that most rainfall occurs throughout the afternoon, from storms 10–50 km in diameter. In all months, the modelled maximum number of storm initiations occurs at least 2 hr prior to the radar-observed maximum. However, the diurnal cycle of rainfall between the model and observations compares well, suggesting that the numerous storm initiations in the simulations do not produce much rainfall. Modelled storms are generally less intense than those in radar observations, especially in early summer. In February, when tropical influences dominate, the simulated storms are of similar intensity to observed storms. Simulated storms tend to reach their peak intensity in the first 15 min after initiation, then gradually become less intense as they grow. In radar observations, storms reach their peak intensity 15 min into their life cycle, stay intense as they grow larger, then gradually weaken after they have reached their maximum diameter. Two November case studies of severe convection are analysed in detail. A higher resolution grid length initiates convection slightly earlier (300 m as opposed to 1.5 km) with the same scientific settings. Two 1.5 km simulations that apply more subgrid mixing have delayed convective initiation. Analysis of soundings indicates little difference in the convective indices, suggesting that differences in convection may be attributed to the choice of subgrid mixing parameters.

KEYWORDS
Africa, convection-permitting models, convective initiation, model evaluation, radar observations, storm life cycles

1 | INTRODUCTION

Skilful prediction of the timing and location of convective storms reduces our risk from hazards such as lightning, hail and wind gusts, and thus forms one of the principal goals of numerical weather prediction. A "step change" in predicting convection occurred with the adoption of kilometre-scale or convection-permitting models (CPMs) (e.g., Clark *et al.*, 2016) by operational forecasting centres. CPMs show better skill compared to models that are run with convection parameterised for predicting the location of intense rainfall events (e.g., Lean *et al.*, 2008). When the model resolution is increased further to sub-km horizontal grid lengths, the physical representation of convective storms compares better

with radar observations and retrievals, including the updraft strength and size and the resulting rainfall intensity (Stein *et al.*, 2015). Despite these improvements, however, CPMs still often lag observations in terms of convective initiation (e.g., Kain *et al.*, 2008).

Convection initiation is a multifaceted problem covering a range of scales and processes, for example:

1. Relevant surface characteristics such as soil moisture, vegetation and orography can vary significantly at scales of less than 1 km (e.g., Weckwerth *et al.*, 2011; Taylor *et al.*, 2011).
2. Boundary-layer features that affect convection initiation, such as moisture gradients and convergence lines, occur at scales up to a few km and may evolve in a matter of minutes (e.g., Weckwerth and Parsons, 2006).
3. Instability in the atmosphere and the effects of developing cumulus clouds also occur on scales of O(1 km) and can evolve quickly (e.g., Morcrette *et al.*, 2007).
4. The structure of layers of convective inhibition (CIN) varies between models and will affect the simulated convection (e.g., Kain *et al.*, 2017).

CPMs operate in the boundary-layer "grey zone", as the grid length used is approximately the same as the size of the large eddies that characterise the convective boundary layer. Due to the importance of km-scale processes, which are difficult to represent adequately in CPMs, we wish to investigate to what extent the timing and location of convection initiation improves at sub-km horizontal grid lengths compared to a CPM at 1.5 km grid length. Although improvements in predicting the onset of rainfall at sub-km resolutions have been shown (e.g., Hanley *et al.*, 2015), we will aim to relate this (potential) improvement to the physical representation of smaller-scale features including storm sizes and CIN layers.

This study will focus on convection initiation in the Highveld region of South Africa. During the summer (November–February), this region is, due to convection, prone to intense rainfall and flooding as well as related hazards including hail, lightning and tornadoes (Simpson and Dyson, 2018). These storms sometimes occur as a result of synoptic-scale systems such as tropical–temperate systems (Harrison, 1984), which are sometimes associated with cut-off lows or upper air troughs (Hart *et al.*, 2013). Air-mass thunderstorms caused by intense surface heating in association with mesoscale and local effects are also common (Tyson and Preston-Whyte, 2000).

To improve forecasting of severe convective events, the South African Weather Service (SAWS) has routinely been running CPMs using the Met Office Unified Model (MetUM) since 2016 at grid lengths of 4 and 1.5 km. A 333 m grid-length configuration of the MetUM is also run by SAWS over a small domain centred on O. R. Tambo International Airport (ORTIA) to support aviation forecasting. None of the CPMs run by SAWS operationally use data assimilation, so our evaluation will focus on the physical characteristics of convection and CIN and the distribution of timings and locations of convection initiation, rather than the predictive skill of forecasting individual events.

The initiation of convection can be identified from observations in a number of ways, typically using ground-based weather radar observations or geostationary satellite measurements. Wilson and Mueller (1993) combined the identification of convergence lines in radar data from enhanced reflectivity and converging Doppler winds with the identification of convective storm enhancement, dissipation and movement in order to nowcast the timing and location of convection out to 30 min. Weather radars do not normally detect convective clouds until they have produced sufficiently large precipitating hydrometeors, so convection initiation likely occurs prior to the first radar echo. Mecikalski and Bedka (2006) used GOES visible and infrared satellite imagery to detect potential convective clouds prior to their detection by radar. They defined convection initiation as the first echo above 35 dBZ and managed to observe convective development in the satellite imagery out to 1 hr prior to this radar-defined convection initiation time. A similar satellite-based nowcasting tool, Cb-TRAM (Zinner *et al.*, 2008), has been developed using the Spinning Enhanced Visible and Infrared Imager (SEVIRI), which identifies growing, mature and dissipating thunderstorms. Since we are mostly interested in precipitating convection, we will use a radar-based definition of convection initiation. Such a definition has the benefit of being applicable to model diagnostics (i.e., forward-simulated radar reflectivity) for a like-with-like comparison of convection initiation statistics.

The article is organised as follows. The SAWS 1.5 km operational runs are presented in Section 2. The observational data and methodology for identifying convection initiation from SAWS radar data are subsequently presented in Section 3, followed by a description of the tracking methodology (Stein *et al.*, 2014) and considerations. A climatology of storm size distributions and rainfall contributions in the SAWS 1.5 km operational runs and radar observations from November 2016 to February 2017 is presented in Section 4. Section 5 presents convective initiation statistics over the same period and Section 6 presents the storm life cycles. Biases highlighted in these sections are analysed for two case studies using different model configurations at 1.5-km and 333-m. These case studies will be described in Section 7, where we compare the representation of convective storm sizes and life cycles as well as their initiation in the 1.5 km and 333 m simulations against radar observations. Additional analysis of soundings will be used to infer what model improvements in convective initiation may be due to the representation of small-scale variations. We will draw our conclusions and discuss potential improvements in Sections 8 and 9.

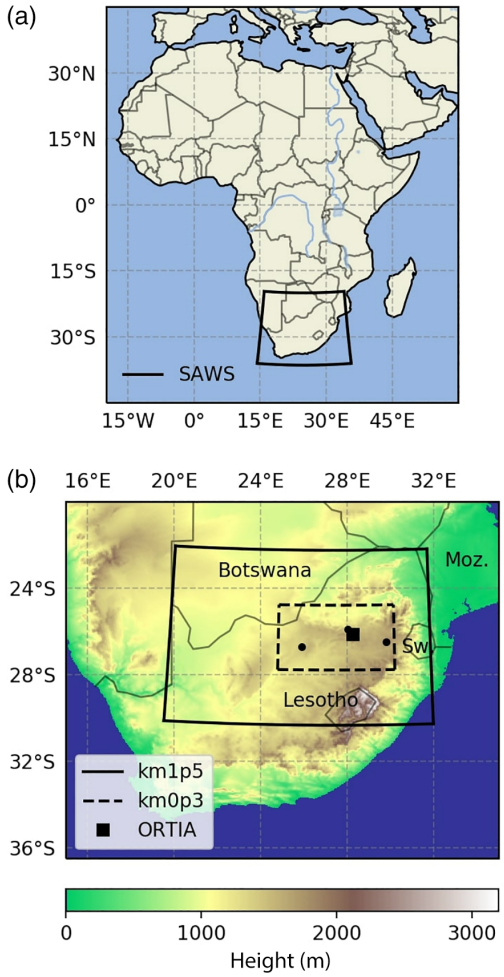

FIGURE 1 (a) Location in the context of Africa of the SINGV2p1 domain run by the South African Weather Service (SAWS). (b) Location of the domains for the 1.5 km and 333 m reruns, O. R. Tambo International Airport (ORTIA) and the SAWS radars used in this article (black circles); neighbouring countries are also identified (Sw. = Swaziland and Moz. = Mozambique). Colours indicate the orography in height above mean sea level [Colour figure can be viewed at wileyonlinelibrary.com]

2 | MODEL SIMULATIONS

All model simulations evaluated in this study were run with the MetUM. Global analyses are made available operationally by the Met Office four times daily (at 0000, 0600, 1200 and 1800 UTC) using the Global Atmosphere version 6.1 (GA6.1) science configuration (Walters *et al.*, 2017). The SAWS 1.5 km simulations used in this article are one-way nested in a GA6.1 simulation initiated from the 1800 UTC global analyses for a domain covering all of South Africa (see Figure 1). These simulations used a configuration of the MetUM developed specifically for the tropics, with 80 vertical levels with a model top of 38.5 km; these simulations will hereafter be referred to as SINGV2p1.

Additional simulations were performed with the MetUM for selected case studies over a smaller domain centred on ORTIA (see Figure 1). A 1.5 km grid-length simulation was one-way nested in a GA6.1 simulation, but initiated from 0000 UTC global analyses. A 333 m grid-length simulation was one-way nested inside the 1.5 km simulation and was also initiated at 0000 UTC. The 0000 UTC cycle point was chosen instead of the 1800 UTC point used by SAWS in order to reduce computational expense, while the next cycle point of 0600 UTC would not have allowed sufficient time for the nested simulation to spin up prior to the onset of convection. Note that the SINGV2p1 simulations were allowed 16 hr of spin-up (assuming onset of convection at 1000 UTC), whereas the additional simulations were allowed 10 hr for the purpose. Both the 1.5 km and 333 m simulations used the midlatitude science configuration of the regional atmosphere version 1.0, "RA1M", which will hereafter be referred to as RA1M-km1p5 and RA1M-km0p3. The recently developed tropical configuration, "RA1T", had not been tested at sub-km grid lengths at the time of writing, so only simulations at the 1.5 km grid length were performed with this configuration. Finally, reruns were also performed for the SINGV2p1 configuration, which were similar to the original runs from SAWS but on the smaller domain for a fair comparison against the other reruns. Sections 5 and 6 will evaluate the SINGV2p1 runs performed by SAWS, initiated at 1800 UTC, whereas Section 7 will evaluate the reruns, initiated at 0000 UTC.

Convection parameterisation is switched off in all the simulations. A summary of the differences in the science configurations between the model runs are listed in Table 1 and briefly described below. RA1T uses a prognostic large-scale

TABLE 1 Differences in configuration between the model simulations evaluated in this study. "BLP" stands for boundary-layer perturbations. All 1.5 km simulations were one-way nested in a GA6.1 simulation that was initiated from global analyses at the relevant cycle point. The 1800 UTC simulations were run by SAWS; domain differences are shown in Figure 1

Model	Cycle Point	Time step	Levels	Subgrid mixing length	Cloud scheme	BLP	Dates
SINGV2p1	1800 UTC	60 s	80	750 m	Smith		Nov. 1, 2016 to Feb. 28, 2017
SINGV2p1	0000 UTC	60 s	80	750 m	Smith	Yes	Nov. 9, 2016 & Nov. 12, 2016
RA1M-km1p5	0000 UTC	60 s	70	300 m	Smith	Yes	"
RA1M-km0p3	0000 UTC	12 s	70	60 m	Smith	Yes	"
RA1T	0000 UTC	60 s	80	300 m	PC2	No	"

cloud scheme (Wilson *et al.*, 2008), whereas the SINGV2p1 and RA1M-km1p5 configurations use a diagnostic large-scale cloud scheme (Smith, 1990). The RA1M configuration also uses stochastic boundary-layer perturbations that are applied to the potential temperature and moisture fields to counter the unrealistic smoothness of these files in kilometre-scale models, which is thought to be a contributing factor to the late initiation of explicit convection (Hanley *et al.*, 2015). In all configurations, the subgrid mixing is parameterised using a 3D Smagorinsky mixing scheme that is pragmatically blended with the boundary-layer scheme as described by Boutle *et al.* (2014). The unstable stability functions differ between the RA1M-km1p5 and RA1T configurations; the local free atmosphere mixing length is smaller in RA1T allowing more mixing in the boundary layer. The MetUM uses the Wilson and Ballard (1999) bulk microphysics scheme with prognostic liquid and ice cloud, as well as prognostic graupel and rain in the CPM simulations. Further details of the regional atmosphere configurations will be presented in Bush *et al.* (2019).

3 | RADAR DATA AND DIAGNOSTICS

In 2010, the SAWS radar network underwent a significant upgrade, during which ten METEOR 600 S-band radars were purchased. The majority of these replaced ageing C-band radars across the country; two were used to expand the network (Terblanche *et al.*, 2001). The radars have a maximum unambiguous range of 200 km. The radars are calibrated regularly during maintenance visits 1–2 times a year, but power failures or other issues may affect calibration between visits. Calibration of individual radars is therefore also monitored following Holleman *et al.* (2010), using the solar interference in the receiving channel and comparing it against the S-band signal monitored at the Dominion Radio Astrophysical Observatory in Canada and removing any mean bias.

The SAWS radar data are available every 6 min and are provided in polar co-ordinates. The data were processed using the University Corporation for Atmospheric Research (UCAR) Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN) software (Dixon and Wiener, 1993). Firstly, for each radar an 8-point bilinear interpolation of the radar fields onto constant altitude plan position indicators (CAPPIs) at 1 km horizontal resolution and 500 m in altitude was performed, referenced to height above mean sea level (AMSL). The CAPPIs were then merged to create a single 3D Cartesian field for the entire network. The TITAN merging algorithm was configured to select the maximum radar reflectivity for grid boxes covered by multiple CAPPIs. For every 6 min data file, an auxiliary mask was also generated by TITAN to indicate which radars were available at that time. Clutter was automatically removed using built-in TITAN functions. For each radar, a clutter map was generated on the CAPPI grid, based on the mean radar reflectivity over the course of a completely dry day. If the mean Z at a point exceeds 10 dBZ, that point is considered a candidate for clutter removal. For those points on a rainy day, if the observed Z is within 5 dBZ of the coinciding clutter Z, it is considered clutter and that data point is not considered in our analysis.

At the time the SINGV2p1 simulations were produced, only the surface rainfall rate was available from the model to identify convective storms and their initiation at high frequency (i.e., 5 min). For a like-with-like comparison, we calculate the rainfall rate from the observations using the Z–R relationship (Fulton *et al.*, 1998)

$$Z = 300R^{1.4}, \tag{1}$$

with R in mm/hr and Z in mm^6/m^3. We note that this Z–R relationship, while commonly used (e.g., historically for the continental U.S. NEXRAD system), is not validated for rainfall and radar observations in South Africa, although it has been used for quantitative rain-rate estimation over the Highveld region of South Africa (Becker, 2014). As was done for the U.S. NEXRAD system, we truncate any rain rates greater than 103.9 mm/hr to limit the effects of hail contamination that would lead to overestimates of the rain rate. In addition to this Z–R relationship, the Marshall–Palmer relationship of $Z = 200R^{1.6}$ (Marshall and Palmer, 1948) and $Z = 250R^{1.2}$, which was found to be most applicable to tropical rainfall (Rosenfeld *et al.*, 1993), is also used. The sensitivity of our results to the choice of the Z–R relationship is addressed in Section 5.

The input reflectivity for the Z–R relationship was from the 3.5-km CAPPI. For the radars considered, 3.5 km AMSL corresponds to approximately 2 km above ground level. The lowest elevation beam reaches 2 km above ground level at 125 km range, so any lower CAPPI would severely limit the range from the radars for rainfall estimates and thus overlap between the radars and the storm-tracking region.

Storm characteristics are derived by applying a flood-fill algorithm to identify individual rainfall objects, where we set a rainfall-rate threshold of 4 mm/hr. Of particular interest is the storm size, which we determine by calculating the storm-equivalent diameter (in km), $D_{eq} = 2\sqrt{A/\pi}$, where A is the area in km^2. We focus on diameter rather than area because it allows for a straightforward comparison with the model resolution.

3.1 | Storm tracking and convection initiation

After processing the radar data using the TITAN software, storms were then tracked using an algorithm developed by Stein *et al.* (2014). In brief, object motion vectors are calculated from the maximum cross-correlation between consecutive radar images, first on 50 × 50 km regions with a 25 km overlap, after which the 25 km gridded motion vectors are interpolated to the 1 km radar grid. Rainfall objects are individually labelled as described above and are

projected to the next time step using these motion vectors. At each time step, rainfall objects identified in the new radar image are compared against the projected rainfall objects from the previous image. If multiple new objects overlap a former object, then only the largest new object keeps the label and all smaller objects obtain a new label and are tagged as a "child". If two former objects merge to form a new object, the new object will inherit the label of the largest former object. The storm-tracking algorithm was applied to both the rainfall derived from the radar observations and the 5-min surface rain-rate output from SINGV2p1 for the entire period from November 2016 to February 2017.

Convection initiation is defined as the first time a contiguous area of 4 mm/hr > 10 km^2 is labelled. Any storm that is flagged as a "child" is not considered a new initiation event. A storm must have its central position removed from the domain boundary by at least half its equivalent diameter D_{eq} or by 15 km, whichever is largest. This is to ensure that it is a newly initiated storm, rather than a pre-existing storm entering the domain. For the life cycle statistics, this restriction is also applied to distinguish decaying storms from those leaving the domain.

Due to the 100% availability of the model data, every storm that existed in the domain between November 2016 and February 2017 was tracked and included in the following statistics. However, due to the intermittent radar coverage, some additional quality control is required to ensure a meaningful comparison. The newly identified storms must not occur immediately after the radar covering its location comes back online after an outage, as the storm may have existed previously during the time the radar was unavailable. Similarly, it must not have been last seen in a part of the domain that is masked in the next time step due to a drop-out of the radar.

Since we track storms using a rainfall-rate threshold of 4 mm/hr, this threshold also determines our definition of convection initiation. While we acknowledge that convective cloud will exist before the presence of the 4 mm/hr rainfall rate, our primary interest is in forecasting the timing of the intense rainfall as covered by our definition. Furthermore, this rainfall threshold is comparable to the 35-dBZ radar reflectivity threshold used by Mecikalski and Bedka (2006) to identify the timing of the initiation.

In addition to the aforementioned criteria, the intermittent radar coverage must also be accounted for in order to make a meaningful comparison of the initiation occurrence across the entire domain. For the spatial distribution of initiation events, the frequency of initiation occurrence for each month was multiplied by a factor

$$f = \frac{\text{Domain max. sampled frequency}}{\text{Sampled frequency}}. \quad (2)$$

The number of initiation events was then binned into a 0.2 × 0.2° latitude–longitude grid.

4 | STORM SIZE DISTRIBUTIONS

We first evaluate the SINGV2p1 performance in simulating the diurnal cycle of storm size distributions for the November, December, January and February (NDJF) 2016–2017 period. This climatological comparison helps to identify consistent biases in the timing of convection, storm sizes and storm duration. Figure 2 shows the number density distributions of the observed and SINGV2p1 storms as a function of the time of day (left two columns). There is a clear diurnal cycle of convection in both the observations and SINGV2p1, with larger numbers of storms in the early afternoon and an increase in large storms in the late afternoon. In the observations, this typically occurs between 1300 and 1600 UTC for all storm sizes. In contrast to observations, the maximum number of small storms (D_{eq} < 8 km) in SINGV2p1 occurs significantly and consistently earlier, typically between 1100 and 1200 UTC. Additionally, unlike in the observations, the timing of the peak number per D_{eq} bin in SINGV2p1 (illustrated by the black squares) tends to monotonically increase as D_{eq} increases, rather than occurring at approximately the same time (1300–1500 UTC). There does not seem to be a clear seasonality of the diurnal cycle. D_{50} is quite different across the months; the highest values in the evening in November are >64 km, compared to 32 km in February.

We also investigate the contribution of different storm sizes to the total amount of rainfall by computing the rainfall fraction per D_{eq} bin in Figure 2 (right two columns), only considering rainfall from identified storms. This is useful for revealing differences in the way SINGV2p1 represents storm intensity as a function of storm size. The black circles in these plots illustrate the storm diameters that equally divide the rainfall contributions from the smaller and larger storms, D_{50}, that is,

$$\int_0^{D_{50}} R(D)dD = \int_{D_{50}}^{\infty} R(D)dD\,, \quad (3)$$

where $R(D)$ is the total amount of rainfall produced by storms of diameter D. In this variable, SINGV2p1 behaves rather differently from the observed storms. Firstly, in the model, the largest contributions to the total rainfall are from the largest storms (D_{eq} > 32 km), where in the observations the greatest contribution comes from storms with diameters between 8 and 32 km. Rainfall is particularly dominated by larger storms in November. Secondly, the observed storms have a broad peak, both in the diameter space and in the time that encompasses noticeable contributions to the total rainfall after 1800 UTC, especially from storms smaller than 16 km. Instead, the model shows a narrower peak along a single line in the diameter–time space, from small storms at 1200 UTC to large storms at 1800 UTC. Overall, D_{50} occurs at larger D_{eq} at nearly all times in SINGV2p1. In November and December in particular, D_{50} increases far more rapidly with time once convective initiation has occurred in SINGV2p1 than in the observations. For example, in December, it takes 9 hr for D_{50} to increase from 16 to 32 km, while the same increase in

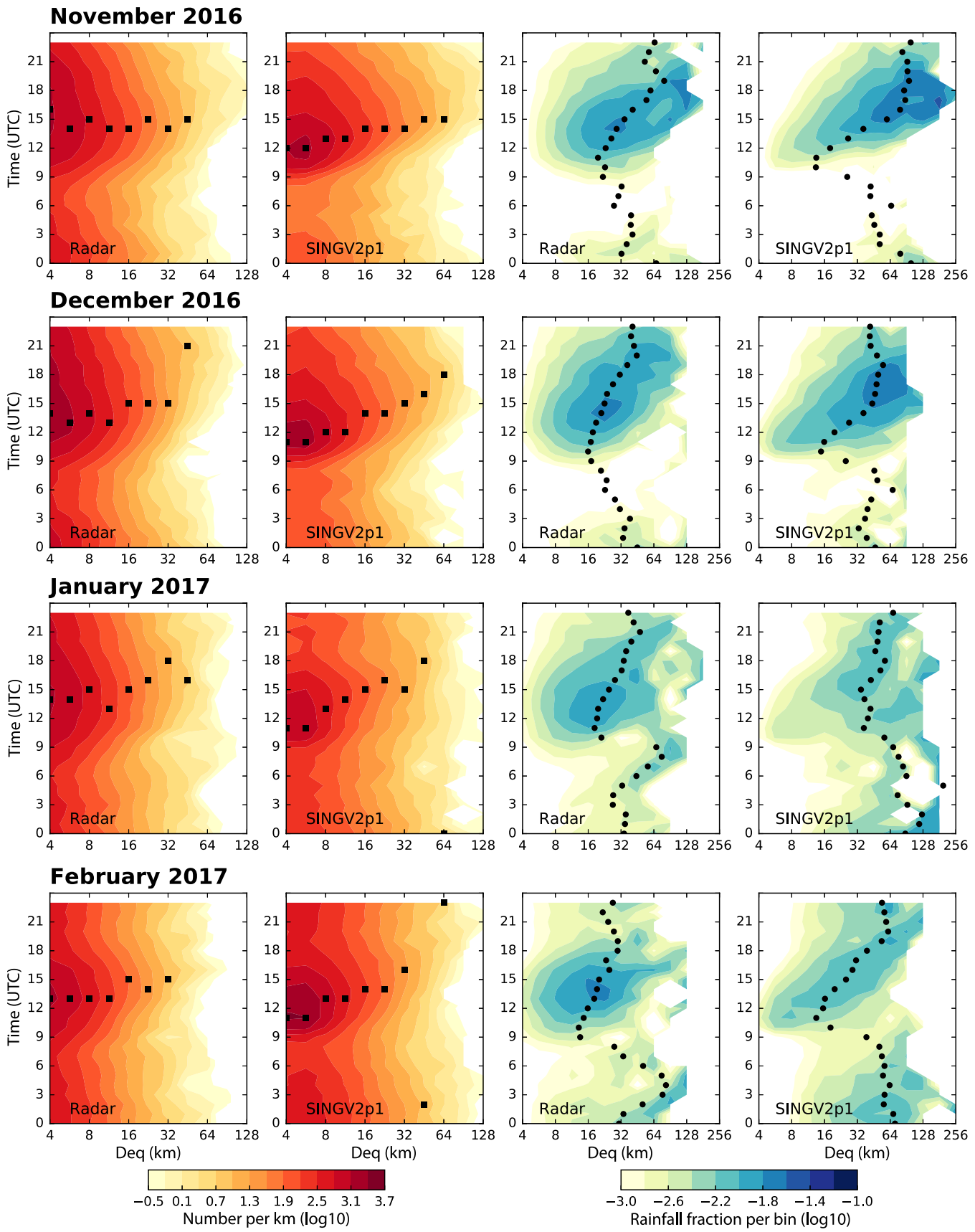

FIGURE 2 Diurnal cycle of storm size distributions observed by the radar and simulated by SINGV2p1 for the individual months (rows). Distributions are shown for every hour with D_{eq} in km using logarithmic bins of factor $\sqrt{2}$. First two columns: Number density distribution (colours, per km) and timing of the maximum number per D_{eq} bin (squares). Last two columns: Fraction of total rainfall per bin (colours) and D_{50} (circles; see the text and Equation 3 for a definition) [Colour figure can be viewed at wileyonlinelibrary.com]

SINGV2p1 takes just 3 hrs. This suggests that, in the model, large storms either become too intense or too numerous too quickly. There is also evidence in Figure 2 that the opposite may also hold, namely that the number of small storms decreases too rapidly in the model.

5 | CONVECTIVE INITIATION STATISTICS OVER SOUTH AFRICA

Next, we focus on the timings and locations of convective initiation during the 4-month evaluation period. We restrict our analysis to initiation primarily associated with daytime heating, considering only storms that initiate between 0600 and 1800 UTC. Figure 3 shows the observed (black) and SINGV2p1 (red) timings of convective initiation as a percentage of total events (top row) and the domain average rain rates using the Z–R relationships of $300R^{1.4}$ (black), $200R^{1.6}$ (dashed green), $250R^{1.2}$ (dashed blue) and SINGV2p1 (red) for each month (bottom row) between 0600 and 1800 UTC. The characteristics of convective initiation between the observations and SINGV2p1 are quite different. In terms of the distribution of initiation events, the model peaks at around 1200 UTC, 1–2 hr before the observations. It also rapidly increases from around 1000 UTC, whereas the observed increase is much more gradual.

The timing and distribution of the observed domain average rain rates are well correlated with the number of observed initiation events for each of the months. However, in SINGV2p1, the domain average rainfall distributions are less correlated with convective initiation across the months. In November, despite the rapid increase in initiation events in SINGV2p1 between 1000 UTC and 1200 UTC, the domain average rainfall does not appreciably increase until 1200 UTC. Figure 2 shows that small (initiating) storms in SINGV2p1 tend to contribute little to the total rainfall. In December, Figure 2 shows that small (initiating) storms between 1000 and 1200 UTC contribute relatively more rainfall more quickly, leading to a better correlation of the number of initiation events and observed domain average rain rates during this time period. This could be because storms in November have a different intensity than those in December, or that storms in December are longer-lived in the model. It suggests that the characteristics of initiating storms and their subsequent growth are incorrectly represented in the model. In Section 7, we analyse simulations at 300 m and study vertical temperature profiles to understand these differences.

Using $Z = 300R^{1.4}$ and $Z = 200R^{1.6}$ produces approximately the same R; the use of $Z = 250R^{1.2}$ results in an increase of R. It is not the aim of this article to discuss the merits of these individual relationships in detail; their inclusion is to illustrate that our interpretation does not change significantly as a result of this choice.

Figure 4 shows the observed and SINGV2p1 frequency distributions as a percentage of the total number of initiation events for each month. In total, 15,515 initiation events fulfilled the criteria in the radar data and 15,092 storms did so in the model data. The largest number of observed convective initiation events was in December, with a total of 5,008 storms, whereas the largest number of model convective initiation events was 4,965, occurring in February. In general, initiation events in the model are more evenly distributed across the domain than in observations. The least amount of convective initiation appears to occur over high orography in the southeast (see Figure 1). The model agrees with the lack of initiation observed in the very far east of the domain.

6 | STORM LIFE CYCLES

In this section, we evaluate the life cycles of storms for each month, in terms of their size and intensity. We consider only the "dominant" storms in the life cycle analysis, meaning that storms must be initiated as original storms and not as "children" and they must terminate by dissipation, not by merging with another storm. Storms were binned by their lifetime (total duration for which they were tracked).

Focusing firstly on the shorter duration storms (1–2 hr) in Figure 5, we can see that the life cycles in SINGV2p1 are quite different from those observed. The peak median D_{eq} is ~12 km in November, decreasing slightly to 8 km in February; the timing of this peak in the life cycle is later in SINGV2p1. The peak D_{eq} in the model occurs later at 60, 50, 60 and 55 min compared to 42, 30, 42 and 36 min for November, December, January and February (NDJF), respectively, in the observations.

There is a significant difference in the time it takes to reach the peak storm average R. In SINGV2p1, it is achieved much earlier, after 15, 10, 10 and 10 min compared to 30, 30, 18 and 10 min for NDJF, respectively, in the observations. It is also noteworthy that the intensity of the model storms at initiation is close to their peak intensity; growth is preferentially in the form of becoming larger rather than more intense. This is in contrast to the observed storms, which tend to grow in intensity *and* size after initiation.

Similar conclusions can be drawn for the longer-lived (2–3 hr) storms. Overall, these storms grow to be more intense and larger than the short-lived storms. The magnitude of peak D_{eq} is similar in all months, at ~15 km. Peak storm average R in the model is again reached far sooner at 15, 20, 30 and 15 min compared to 42, 36, 56 and 30 min for NDJF, respectively, in the observations.

Although SINGV2p1 broadly captures the diurnal cycle of convection (shown in Figure 2), it clearly fails to accurately reproduce the characteristics of individual storms. Understanding the reason for this disagreement is key for the future development and improvement of CPMs. In the next section, we run 333 m simulations and investigate vertical temperature profiles to assess the impact on the representation of convective storms and convective initiation timing.

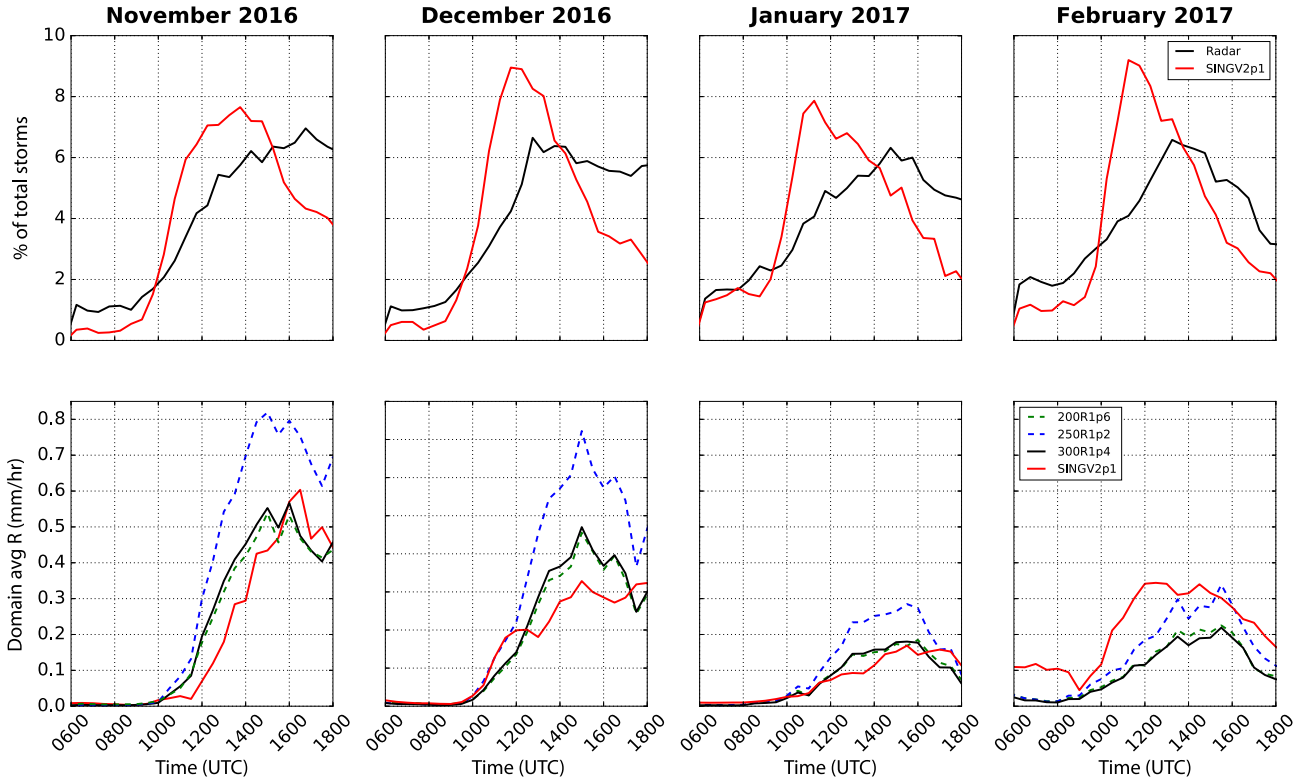

FIGURE 3 Top row: A comparison of the observed (black) and SINGV2p1 (red) timings of the convective initiation for each month. Bottom row: Domain average rain rates using the Z–R relationships of $300R^{1.4}$ (black), $200R^{1.6}$ (dashed green), $250R^{1.2}$ (dashed blue) and SINGV2p1 (red) for each month [Colour figure can be viewed at wileyonlinelibrary.com]

7 | SENSITIVITY TO MODEL CONFIGURATION

In this section, we present additional simulations of the MetUM in two case studies to investigate the impact of higher resolution on the representation of convective storms. For both cases, simulations were one-way nested in a GA6.1 simulation that was initiated from global analyses valid at 0000 UTC.

On November 9, 2016, a surface trough triggered widespread convection that caused severe flooding across Gauteng province; almost 90 mm of rainfall accumulation was recorded at ORTIA in less than 90 min (Simpson and Dyson, 2018). A similar convective outbreak occurred 3 days later, on November 12, 2016. As described in Section 2, these simulations consist of the RA1T, RA1M-km1p5 and RA1M-km0p3 configurations of the MetUM, as well as reruns of the SINGV2p1 configuration for the reduced domain size. Each model was subsampled to the size of the RA1M-km0p3 domain so that the same areas could be compared. The differences between the model configurations are described in Table 1.

Figure 6 shows the number density distributions and the rainfall fractions per bin for each configuration for November 9 and November 12. Firstly, on November 9, there are quite large differences between the storm size distributions. All of the models, with the exception of RA1T, produce too many small storms. The timings of the peak number of storms with $D_{eq} < 16$ km in RA1M-km1p5 all occur between 0800 and 0900 UTC; this compares to a monotonic increase with D_{eq} with all other configurations. None of the models accurately reproduce the observed convective storm characteristics. For example, between 0900 and 1400 UTC in the RA1M-km1p5 and RA1M-km0p3 models, D_{eq} is 10 and 6 km, respectively, compared to 32 km for the observed storms. All configurations produce too-large storms in the evening that contribute far too much to rainfall, consistent with the November storms in Figure 2.

Similar behaviour is seen on November 12, although there is better consistency between the configurations in terms of the storm size distributions; the models are more similar to each other than they are to the observations. Storms appear far too early in all configurations. The RA1M-km0p3 configuration in particular produces too many small storms ($D_{eq} < 8$ km). The rainfall contributions of smaller storms in the models are much greater than observed. The timing of the peak number of storms of all sizes is too early, but for all configurations the timing of the peak occurrence is increasingly later for increasing storm size. This is not the case in the observations, where the timing of the peak number of storms is more variable with D_{eq}, but typically between 1500 and 1700 UTC.

Figure 7 shows the distribution of initiation events occurring on each day and the domain average R for each day. For November 9, the same statistics are also shown but for the 1800 UTC initialised run. For November 9, the RA1M-km1p5

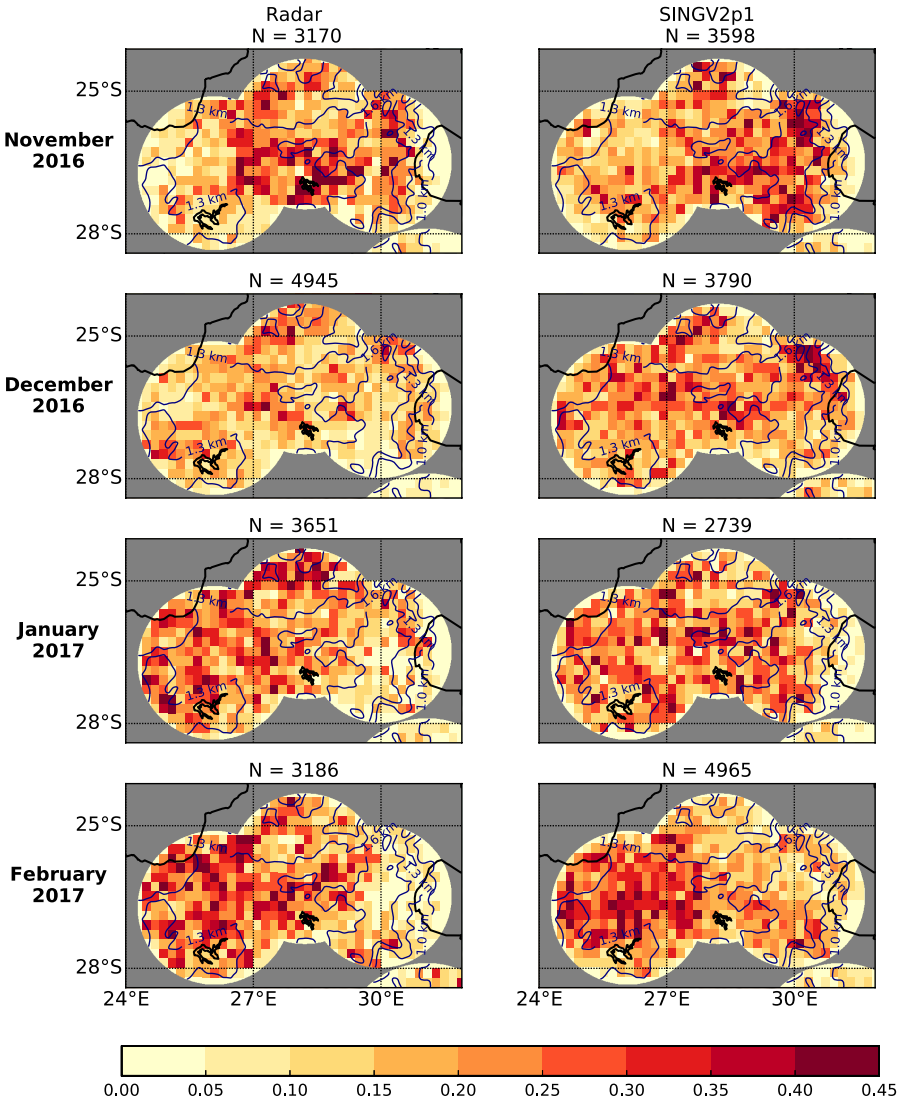

FIGURE 4 Locations of convective initiation observed (left) and in SINGV2p1 (right) expressed as a percentage of the total number of observed storms for November, December, January and February (NDJF) 2016–2017. Dark blue contours illustrate the orography at 1, 1.3 and 1.6 km [Colour figure can be viewed at wileyonlinelibrary.com]

and RA1M-km0p3 models initiate too soon, as indicated by both the number of initiation events and the increase in the domain average *R*. SINGV2p1 initiates slightly too late, but exhibits the same delay in the domain average *R* relative to the number of initiation events seen in Figure 3. RA1T in this case initiates about 2 hr too late.

For the November 12 case, all models initiate too soon. RA1M-km0p3 initiates about 1 hr earlier than RA1M-km1p5 and RA1T, which are about 1 hr earlier than observations. Interestingly, the distributions of the domain average *R* for the RA1T and SINGV2p1 models agree very well with the radar; the RA1M-km1p5 and RA1M-km0p3 domain average *R* is almost at a minimum, while the other models and observations are peaking at 1600–1700 UTC.

7.1 | Vertical temperature profiles

In order to understand the behaviour of the models with regard to convective initiation, we now look at vertical temperature profiles. Given a correct model state, one would expect CPMs to initiate convection *late* due to being unable to resolve the very small initial plumes, and this fits with the fact that anything that makes the fields smoother (increasing grid length, reducing perturbations, increasing diffusion) makes the model initiate later, as is generally seen in Figure 7. In this context it is harder to explain why the models are sometimes seen to initiate *too early* compared to the observations. One possible explanation is that there is a lack of convective inhibition (CIN) in the model profiles (as noted for a case on the U.S. Great Plains by Hanley and Lean (2019). A second, perhaps more speculative explanation is too-strong stochastic perturbations applied to the model fields. Here, we use model soundings to investigate this. No observations of the pre-convective environment are available; however, it is still interesting to study the differences between the models.

Figure 8 shows the model temperature profiles from the RA1M-km1p5 at Irene, Gauteng at 0600 UTC for the two cases. These profiles were chosen to sample the

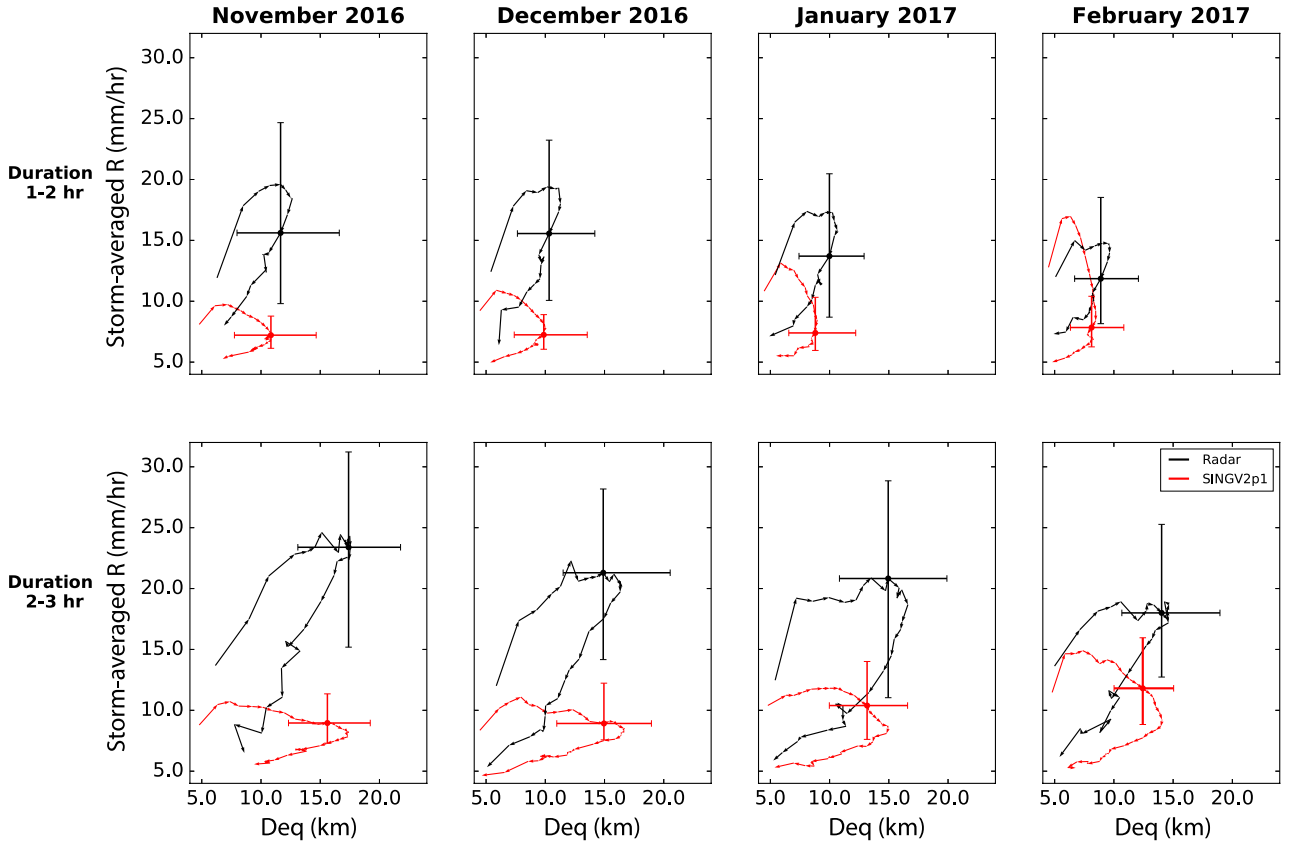

FIGURE 5 A comparison of median observed (black) and SINGV2p1 (red) storm life cycles, characterised by D_{eq} and the storm-averaged rainfall rate. The top row shows storms that lived for 1–2 hr; the bottom row shows storms that lived for 2–3 hr. Each arrow represents the change in average R and D_{eq} every 5 and 6 min for the SINGV2p1 and radar-observed storms, respectively; the direction of the arrow indicates the evolution of the storm properties in time. The error bars represent the interquartile ranges of the storm properties 1 hr into their life cycle [Colour figure can be viewed at wileyonlinelibrary.com]

pre-convective environment. Only one model profile is shown for each case because the different models all have very similar profiles. It is probably the case that these profiles are all very similar to that in the driving global model, although the latter profiles were not available. It is possible that the moisture field, in particular, might be modified in the regional models due to the surface scheme, but this does not appear to be the case here.

The November 12, 2016 case – where the models all initiate too early compared to the radar observations – has significant CIN in the profile which would have been a significant factor in the initiation time. It is possible that the early initiation is a result of the CIN being greater in the real atmosphere. We do not have much evidence for this from this study, but it is worth recording that the closest observational data, the 0000 UTC sounding (not shown) had a larger amount of CIN. It is noticeable that the profile for the November 9, 2016 case has very little CIN, and in this case the model without stochastic perturbations (RA1T) did initiate late, so the evidence is consistent with the two explanations for early initiation.

8 | DISCUSSION

The results shown here indicate that there are fundamental differences in the way in which convection is represented in SINGV2p1 compared to observations. This is true both within each day and throughout the 4-month evaluation spanning from November 2016 to February 2017.

While Figure 3 shows that the SINGV2p1 model is broadly able to recreate the distributions of the domain average R quite well, the way in which storms produce and distribute rainfall differs vastly from observations. This is particularly significant for operational forecasts of severe rainfall events. SINGV2p1 produces too many large storms and these contribute far more to the overall rain rate than the observed storms, indicated by D_{50} (black circles) in Figure 2. Furthermore, in Figure 3 (November), observing the domain average R in isolation would lead one to conclude that the initiation in SINGV2p1 is too late, while the number of initiation events indicates that this is not the case. Figure 2 suggests that this difference is due to a number of small storms initiating that do not contribute significantly to rainfall. Although these weakly precipitating storms may not be important for rainfall accumulation predictions, they are undesirable for weather forecasters trying to interpret the model output. Understanding this behaviour is fundamental to continuing model development.

Figure 3 indicates a notable lack of newly initiating storms beyond the time of peak initiation occurrence for all months. This is also indicated in the storm size distributions in Figure 2 by a lack of small storms after 1500 UTC. One

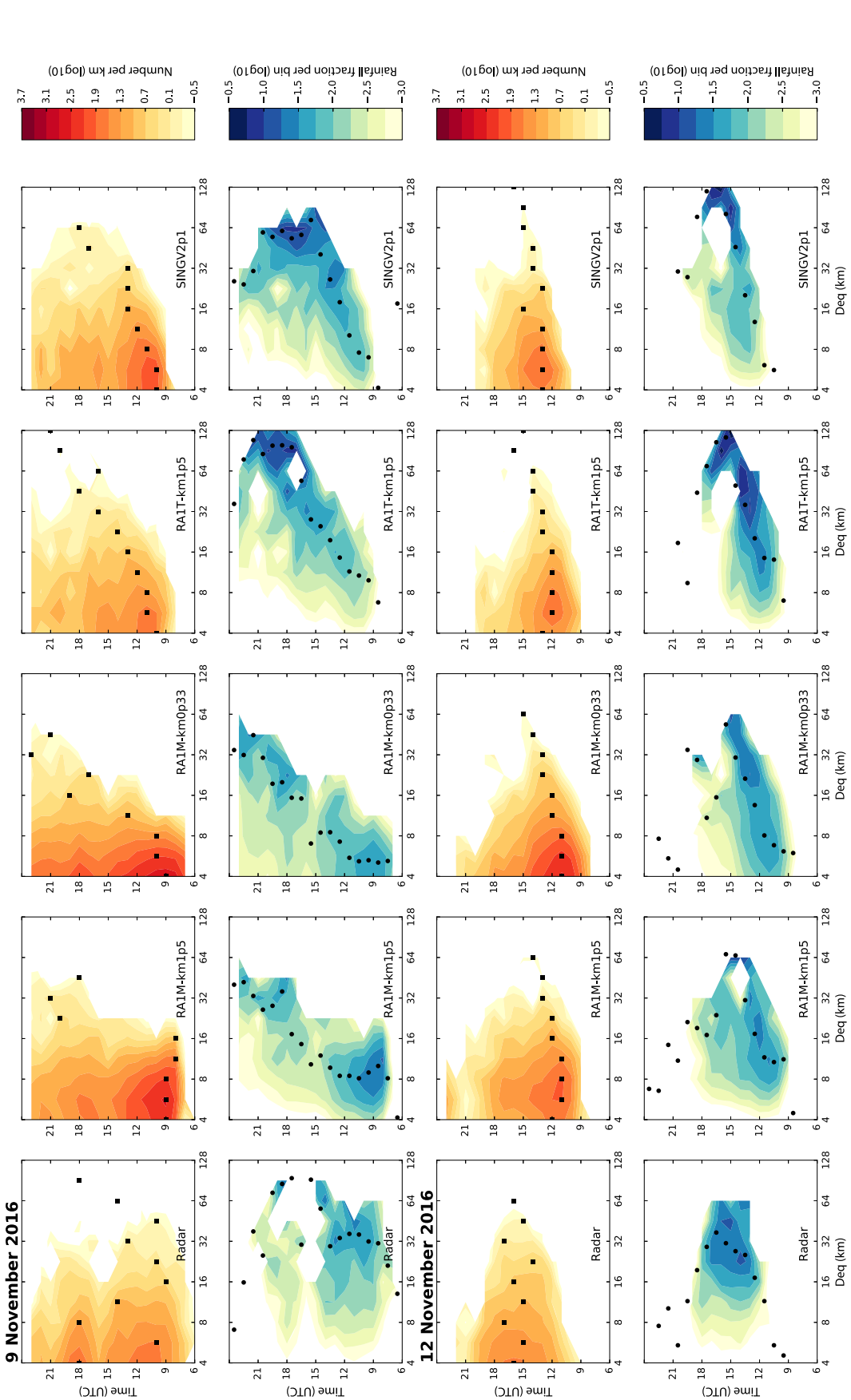

FIGURE 6 The number density distributions and the rainfall fractions per bin for the observed, RA1M-km1p5, RA1M-km0p3, RA1T and SINGV2p1 simulations for November 9 (top two rows) and November 12 (bottom two rows). First and third rows: Number density distributions (colours, per km) and timings of the maximum number per D_{eq} (squares). Second and fourth rows: Fraction of total rainfall per bin (colours) and D_{50} (circles; see Equation 2 for the definition) [Colour figure can be viewed at wileyonlinelibrary.com]

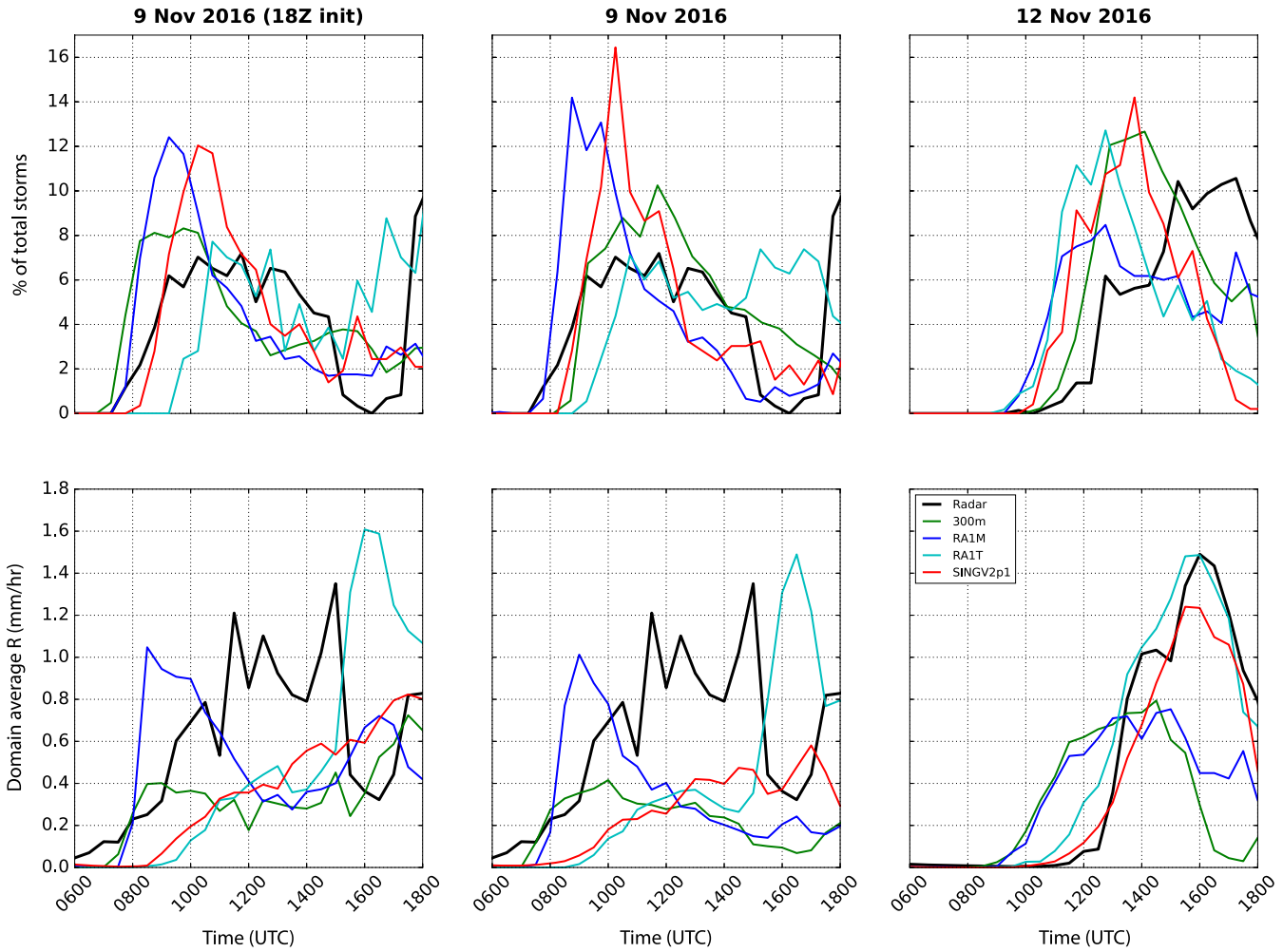

FIGURE 7 A comparison of the number of initiation events (top row) in the radar observations (black), RA1M-km0p3 (green), RA1M-km1p5 (blue), RA1T (cyan) and SINGV2p1 (red), and domain average rain rates (bottom row) for the November 9 and November 12 case studies [Colour figure can be viewed at wileyonlinelibrary.com]

possible reason for this could be the inability of the 1.5 km SINGV2p1 to resolve smaller-scale triggering mechanisms such as cold pools. There is some evidence of an increase in the number of small, evening storms in the RA1M-km0p3 case study runs.

Overall, SINGV2p1 exhibits less variability in terms of the timing of initiation and the characteristics of convective storms after initiation. In terms of the number of initiation events, SINGV2p1 shows a rapid increase at approximately the same time of day and peaks earlier than observed for all 4 months. There is also evidence that the storms in SINGV2p1 tend to have a preferred "mode" of growth once they have initiated: the timing of peak D_{eq} (black squares, Figure 2) tends to monotonically increase with D_{eq} rather than occurring at approximately the same time as observed. This could be indicative of a propensity for SINGV2p1 to merge storms and upscale them into larger storms, which then contributes too much rainfall. Consistent with this interpretation, we find that storms in the model do tend to be too long-lived (not shown here).

It is interesting that while convective initiation events take place at more similar times than observed (shown by the more "peaked" distributions in Figure 3), the locations of these initiation events tend to be more widespread than in the observations, where convection initiation tends to be more focused on the south and west of the domain.

Another substantial difference in the representation of storms is in their life cycles, shown in Figure 5. Although ultimately peaking at a similar D_{eq}, storms in SINGV2p1 tend to peak in intensity much sooner than observed storms and begin their lives with a storm average R much closer to their peak intensity. In order to distribute rain, they tend to grow larger rather than more intense. Observed storms, on the other hand, initially grow more intense and then become larger. The area encompassed by these curves represents the average total rainfall from storms; the fact that these areas are not too dissimilar despite these different intensity–area characteristics again highlights why it is important to look beyond the domain average R when evaluating CPMs.

The sensitivity of the convective initiation time in the models to physics set-up – namely the fact that that anything that tends to make the model's fields smoother also makes it later – is generally confirmed by the results here. We have less evidence for the timing of model initiation relative to

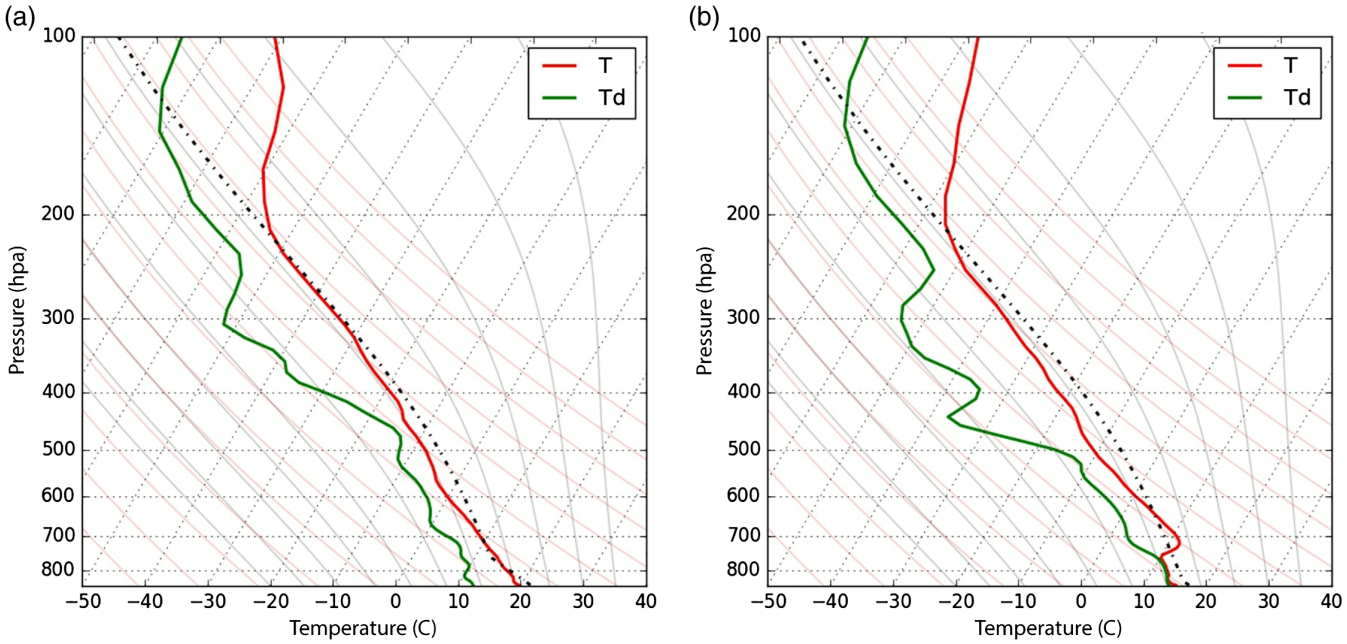

FIGURE 8 Vertical temperature profiles at Irene, Gauteng at 0600 UTC from the RA1M-km1p5 configuration for November 9 (left) and November 12 (right) [Colour figure can be viewed at wileyonlinelibrary.com]

observations, in particular why the models sometimes initiate early. It seems possible from the evidence presented here, albeit weakly because observed soundings were not available at the correct time, that a factor in early initiation could be the lack of CIN in the model profiles. This illustrates the importance of the driving model in understanding the behaviour of nested regional ones – the implication is that improving the physics in a regional model is never going to be sufficient on its own if it is inheriting errors from the larger-scale one driving it. Another factor could be too-strong perturbations applied to the model fields in the midlatitude configurations (RA1M-km1p5). It needs to be borne in mind that in the midlatitudes convection is often very severely under-resolved and tends to be completely missed if the fields are too smooth. This is why the midlatitude configurations tend to have stochastic perturbations applied, as well as lower diffusion in the form of a smaller subgrid mixing length. In this context it is interesting to consider the behaviour of the RA1M-km0p3 model in this study. It tends to initiate too early, but this is likely due to the stochastic perturbations or to the lack of CIN as opposed to an intrinsic problem with the model and, in fact, in many situations sub-km models do behave better (e.g., Stein *et al.*, 2015), as would be hoped.

9 | CONCLUSION

The accurate prediction of convective initiation in convection-permitting models is challenging due to its sensitivity to sub-km processes. In this article we present a statistical evaluation of the performance of 1.5 km simulations of the Met Office Unified Model in representing convective storms and their initiation for a 4-month period over South Africa. We find that storm size distributions in the model compare well with observations; however, in the model, most of the rainfall is produced by larger storms (50 km in diameter or larger) during the evening, whereas radar observations show most rainfall occurring throughout the afternoon, from storms 10–50 km in diameter. On average, convective initiation in the model occurs about 2 hr sooner than the radar-observed maximum. In terms of their life cycles, model storms tend to reach their peak intensity within 15 min after initiation, compared to between 10 and 30 min for observed storms.

We also investigate the sensitivity of the convective initiation to model configuration for two cases of severe convection in November. A 300 m grid-length simulation initiates slightly earlier than a 1.5 km simulation with the same science settings. Two 1.5 km simulations that apply more subgrid mixing have delayed convective initiation. There is very little difference in the vertical profiles and convective indices in each of the model configurations. This suggests that the differences in convective initiation may be attributed to the choice of subgrid mixing parameters.

ACKNOWLEDGEMENTS

This work and two of its contributors (WK and TS) were supported by the UK–South Africa Newton Fund through the Met Office Weather and Climate Science for Service Partnership (WCSSP) South Africa. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk).

ORCID

William J. Keat https://orcid.org/0000-0002-8689-4847

REFERENCES

Becker, E.H. (2014) *Application of a quantitative precipitation estimation algorithm for the S-Band radar at Irene, South Africa*. PhD Thesis, University of KwaZulu-Natal, Durban..

Boutle, I., Eyre, J. and Lock, A. (2014) Seamless stratocumulus simulation across the turbulent gray zone. *Monthly Weather Review*, 142, 1655–1668.

Bush, M., Allen, T., Bain, C., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., Lean, H., Lock, A., Manners, J., Mittermaier, M., Morcrette, C., North, R., Petch, J., Short, C., Vosper, S., Walters, D., Webster, S., Weeks, M., Wilkinson, J., Wood, N. and Zerroukat, M. (2019) The Met Office Unified Model/JULES Regional Atmosphere and Land configurations (RAL) —Version 1. *Geoscientific Model Development*.

Clark, P., Roberts, N., Lean, H., Ballard, S.P. and Charlton-Perez, C. (2016) Convection-permitting models: a step-change in rainfall forecasting. *Meteorological Applications*, 23, 165–181.

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

Fulton, R.A., Breidenbach, J.P., Seo, D.-J., Miller, D.A. and O'Bannon, T. (1998) The WSR-88D rainfall algorithm. *Weather and Forecasting*, 13, 377–395.

Hanley, K.E. and Lean, H.W. (2019) 2.2-km Unified Model ensemble simulations of a case from the 2017 Hazardous Weather Testbed. in preparation. *Monthly Weather Review*.

Hanley, K.E., Plant, R.S., Stein, T.H.M., Hogan, R.J., Nicol, J.C., Lean, H.W., Halliwell, C. and Clark, P.A. (2015) Mixing-length controls on high-resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, 141, 272–284.

Harrison, M. (1984) A generalized classification of South African summer rain-bearing synoptic systems. *Journal of Climatology*, 4, 547–560.

Hart, N.C., Reason, C.J. and Fauchereau, N. (2013) Cloud bands over southern Africa: seasonality, contribution to rainfall variability and modulation by the MJO. *Climate Dynamics*, 41, 1199–1212.

Holleman, I., Huuskonen, A., Kurri, M. and Beekhuis, H. (2010) Operational monitoring of weather radar receiving chain using the sun. *Journal of Atmospheric and Oceanic Technology*, 27, 159–166.

Kain, J.S., Weiss, S.J., Bright, D.R., Baldwin, M.E., Levit, J.J., Carbin, G.W., Schwartz, C.S., Weisman, M.L., Droegemeier, K.K., Weber, D.B. and Thomas, K.W. (2008) Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Weather and Forecasting*, 23, 931–952.

Kain, J.S., Willington, S., Clark, A.J., Weiss, S.J., Weeks, M., Jirak, I.L., Coniglio, M.C., Roberts, N.M., Karstens, C.D., Wilkinson, J.M., Knopfmeier, K.H., Lean, H.W., Ellam, L., Hanley, K., North, R. and Suri, D. (2017) Collaborative efforts between the United States and United Kingdom to advance prediction of high-impact weather. *Bulletin of the American Meteorological Society*, 98, 937–948.

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

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

Mecikalski, J.R. and Bedka, K.M. (2006) Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Monthly Weather Review*, 134, 49–78.

Morcrette, C., Lean, H., Browning, K., Nicol, J., Roberts, N., Clark, P., Russell, A. and Blyth, A. (2007) Combination of mesoscale and synoptic mechanisms for triggering an isolated thunderstorm: observational case study of CSIP IOP 1. *Monthly Weather Review*, 135, 3728–3749.

Rosenfeld, D., Wolff, D.B. and Atlas, D. (1993) General probability-matched relations between radar reflectivity and rain rate. *Journal of Applied Meteorology*, 32, 50–72.

Simpson, L.-A. and Dyson, L.L. (2018) Severe weather over the Highveld of South Africa during November 2016. *Water SA*, 44, 75–85.

Smith, R. (1990) A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, 116, 435–460.

Stein, T.H.M., Hogan, R.J., Hanley, K.E., Nicol, J.C., Lean, H.W., Plant, R.S., Clark, P.A. and Halliwell, C.E. (2014) The three-dimensional morphology of simulated and observed convective storms over southern England. *Monthly Weather Review*, 142, 3264-3283.

Stein, T.H.M., Hogan, R.J., Clark, P.A., Halliwell, C.E., Hanley, K.E., Lean, H.W., Nicol, J.C. and Plant, R.S. (2015) The DYMECS project: a statistical approach for the evaluation of convective storms in high-resolution NWP models. *Bulletin of the American Meteorological Society*, 96, 939–951.

Taylor, C.M., Gounou, A., Guichard, F., Harris, P.P., Ellis, R.J., Couvreux, F. and De Kauwe, M. (2011) Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, 4, 430–433.

Terblanche, D., Pegram, G. and Mittermaier, M. (2001) The development of weather radar as a research and operational tool for hydrology in South Africa. *Journal of Hydrology*, 241, 3–25.

Tyson, P.D. and Preston-Whyte, R.A. (2000) *The Weather and Climate of Southern Africa*. Cape Town: Oxford University Press.

Walters, D., Boutle, I., Brooks, M., Thomas, M., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Gill, M., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J., Willett, M., Woolnough, S. and Prince, X. (2017) The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geoscientific Model Development*, 10, 1487–1520.

Weckwerth, T.M. and Parsons, D.B. (2006) A review of convection initiation and motivation for IHOP 2002. *Monthly Weather Review*, 134, 5–22.

Weckwerth, T.M., Wilson, J.W., Hagen, M., Emerson, T.J., Pinto, J.O., Rife, D.L. and Grebe, L. (2011) Radar climatology of the COPS region. *Quarterly Journal of the Royal Meteorological Society*, 137, 31–41.

Wilson, D.R. and Ballard, S.P. (1999) A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, 125, 1607–1636.

Wilson, J.W. and Mueller, C.K. (1993) Nowcasts of thunderstorm initiation and evolution. *Weather and Forecasting*, 8, 113–131.

Wilson, D.R., Bushell, A.C., Kerr-Munslow, A.M., Price, J.D. and Morcrette, C.J. (2008) PC2: a prognostic cloud fraction and condensation scheme. I: scheme description. *Quarterly Journal of the Royal Meteorological Society*, 134, 2093–2107.

Zinner, T., Mannstein, H. and Tafferner, A. (2008) Cb-TRAM: tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorology and Atmospheric Physics*, 101, 191–210.

How to cite this article: J. Keat W, H. M. Stein T, Phaduli E, *et al*. Convective initiation and storm life cycles in convection-permitting simulations of the Met Office Unified Model over South Africa. *Q J R Meteorol Soc.* 2019;1–14. https://doi.org/10.1002/qj.3487