

# Statistical Assessment of Tropical Convection-Permitting Model Simulations Using a Cell-Tracking Algorithm

SIMON CAINE AND TODD P. LANE

*School of Earth Sciences and ARC Centre of Excellence for Climate System Science, The University of Melbourne, Melbourne, Australia*

PETER T. MAY

*Centre for Australian Weather and Climate Research, Melbourne, Australia*

CHRISTIAN JAKOB

*School of Mathematical Sciences and ARC Centre of Excellence for Climate System Science, Monash University, Melbourne, Australia*

STEVEN T. SIEMS AND MICHAEL J. MANTON

*School of Mathematical Sciences, Monash University, Melbourne, Australia*

JAMES PINTO

*National Center for Atmospheric Research,\* Boulder, Colorado*

(Manuscript received 15 September 2011, in final form 8 August 2012)

## ABSTRACT

This study presents a method for comparing convection-permitting model simulations to radar observations using an innovative object-based approach. The method uses the automated cell-tracking algorithm, Thunderstorm Identification Tracking Analysis and Nowcasting (TITAN), to identify individual convective cells and determine their properties. Cell properties are identified in the same way for model and radar data, facilitating comparison of their statistical distributions. The method is applied to simulations of tropical convection during the Tropical Warm Pool-International Cloud Experiment (TWP-ICE) using the Weather Research and Forecasting Model, and compared to data from a ground-based radar. Simulations with different microphysics and model resolution are also conducted. Among other things, the comparisons between the model and the radar elucidate model errors in the depth and size of convective cells. On average, simulated convective cells reached higher altitudes than the observations. Also, when using a low reflectivity (25 dBZ) threshold to define convective cells, the model underestimates the size of the largest cells in the observed population. Some of these differences are alleviated with a change of microphysics scheme and higher model resolution, demonstrating the utility of this method for assessing model changes.

## 1. Introduction

Recent improvements in computing capabilities have allowed many researchers and operational centers to

---

\* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

---

*Corresponding author address:* Todd Lane, School of Earth Sciences, The University of Melbourne, 3010 Melbourne, Australia.  
E-mail: tplane@unimelb.edu.au

run numerical models at sufficient resolution that convective processes can be treated explicitly. It has become common to configure models with horizontal grid spacings of  $O(1 \text{ km})$ , which may not truly resolve deep convection processes but permit them to evolve in a relatively realistic way. These so-called convection-permitting models are based on more realistic assumptions than coarser-resolution simulations with parameterized convection. However, they still require substantial validation to determine how realistic their representations of convective-scale processes actually are. The focus of this

study is on the application of an innovative object-based method of model evaluation at the convective scale.

Diagnosing the errors in convection-permitting models is the first step toward their improvement, but is also valuable in determining their limitations. Evaluation of models at the convective scale is especially important because these models have already formed the benchmark for modern high-resolution numerical weather prediction and are frequently used to inform the development of parameterizations. With the significant uptake of popular community models, like the Weather Research and Forecasting (WRF) Model, which are configured at convection-permitting resolution for a broad range of applications, identifying their weaknesses is critical for the user community.

Evaluating models at the convective scale is challenging. Moist convection has inherently low predictability and models have difficulty simulating the correct location and timing of individual convective clouds, which makes direct comparisons between observed and simulated processes problematic. [These difficulties have led to significant activity in the field of spatial verification methods, e.g., Gilleland et al. (2010).] Such problems may be exacerbated if comparisons rely on ground-based measurements with incomplete spatial coverage, such as rain gauges. Moreover, rainfall is a two-dimensional field that is the product of a sequence of processes and cannot easily be used to unambiguously determine differences between modeled and observed convective structures. Ground-based radar on the other hand provides three-dimensional observations of convective structures at frequent time intervals (typically about 10 min), and its continuous spatial coverage allows the statistics of a cloud population to be determined, which provides a useful basis for model evaluation.

Various methods have been used to take advantage of radar data for the purposes of model validation and evaluation. A standard approach is to compare contoured frequency with altitude diagrams (CFADs) derived from both model and observed data (Rogers et al. 2007; Van Weverberg et al. 2011; Caine 2009; Caine et al. 2009). Similar to the creation of CFADs, another approach is to calculate the coverage of an observed quantity and compare these measurements with model data. Illingworth et al. (2007) calculated cloud fraction from radar data and compared it with model results, while May and Lane (2009) calculated a statistical coverage product (SCP), which involves calculating the fractional coverage of reflectivity values greater than a chosen threshold, the fractional coverage of rain, snow, graupel, and hail, and the maximum reflectivity anywhere on the grid. One consequence of using fractional coverage or CFADs to validate a numerical model

is that the horizontal structure of individual convective elements is lost. To avoid the loss of horizontal structure an object-based methodology can be used. Davis et al. (2006) used an object-based approach to verify precipitation forecasts. In their analysis rain area and intensity were used as the basis for comparison between observed and simulated data.

The technique described in this paper differs from the previous authors' in that an automated cell-tracking algorithm is used to objectively identify, track, and provide information about the three-dimensional characteristics of individual storms or storm cells. Specifically, the method utilizes Thunderstorm Identification Tracking Analysis and Nowcasting (TITAN; Dixon and Wiener 1993), an established algorithm usually applied to radar data, to characterize and compare the statistics of convective cells observed by ground-based radar and simulated using the WRF Model. This paper builds upon the work of Pinto et al. (2007) in which TITAN was first used to analyze model results. By identifying individual convective cells as objects for the basis of comparisons, a large number of cell properties can be recorded, even over a relatively short time period. The statistical distributions of those properties in the model and observations are compared, avoiding some of the predictability issues or loss of spatial information inherent in other methods. Other authors have also used objective tracking and identification algorithms for comparing observations and model results (e.g., Wilcox 2003; Zhang et al. 1999). However, unlike TITAN, these other methods are two-dimensional and do not provide information about the vertical structure of individual convective cells, which is part of the focus of this study.

The aim of this study is to describe the application of TITAN for evaluating convection-permitting model simulations. To illustrate this method a series of model simulations of tropical convection observed during the Tropical Warm Pool-International Cloud Experiment (TWP-ICE) are compared to observations from the Bureau of Meteorology radar at Gunn Point near Darwin, Australia. The TWP-ICE experiment was conducted in the region surrounding Darwin during January and February 2006 (May et al. 2008). Over the course of the experiment synoptic conditions ranged from an active monsoon to a period with relatively suppressed convection, and finally to monsoon break conditions. The time period for this investigation (spanning 1200 UTC 8 February 2006 to 1200 UTC 13 February 2006) was chosen to coincide with the break conditions sampled during TWP-ICE, which occur when the low-level zonal winds revert to an easterly flow, advecting drier air to the region (Keenan and Carbone 1992). Convection during break periods tends to be smaller in scale (than convection

occurring during the monsoon), is locally forced, and often initiates along sea-breeze fronts. The localized nature of the convection within the break lends itself to the method because numerous convective systems evolve and decay within the radar range.

Other studies have analyzed convective-permitting simulations during TWP-ICE (e.g., Wapler et al. 2010; Caine 2009; Varble et al. 2011; Wang and Liu 2009; Fridlind et al. 2012; Wapler and Lane 2012), exploring aggregate properties of the convective processes and their sensitivity to physical parameterizations. In the spirit of providing a proof of concept of the TITAN analysis, the simulations presented herein are designed specifically with the comparison to radar in mind, and the simulations do not encompass the entire TWP-ICE period, only a 5-day period during the break regime. Even though the period of study is relatively short (constrained mostly by computational limitations and the length of the convective regime), it will be shown that these simulations expose important differences between the observed and modeled convective systems. Moreover, the method is also used to examine the effects of changes in model physics and resolution on the model outcomes.

The remainder of this paper is organized as follows: section 2 describes the data and methods used in the analysis with some further details provided in the appendixes. In section 3 results obtained using some traditional methods of model evaluation are presented along with an analysis of observed and simulated statistics of the convective cell properties defined by TITAN. The sensitivity of results to model microphysics and model resolution is also explored. Finally, the results and their implications are summarized in section 4.

## 2. Data and methods

### a. The Gunn Point radar

The model evaluation technique employed here revolves around comparison between the model simulations and observed reflectivity data from the Gunn Point radar in Darwin. The Gunn Point radar is a C-Band Polarimetric (CPOL) radar operating at a 5-cm wavelength, which has an uncertainty of 1 dBZ [see Keenan et al. (1998) for details]. Reflectivity data were obtained in Cartesian format consisting of 40 vertical levels spaced 500 m apart (0.5–20 km), with each vertical level containing  $121 \times 121$  grid points spaced 2.5 km apart; the temporal resolution of these data is 10 min, which is the time taken to complete a full 3D volume scan. The radar-derived rainfall rate is also determined using the algorithm described by Bringi et al. (2001), which was

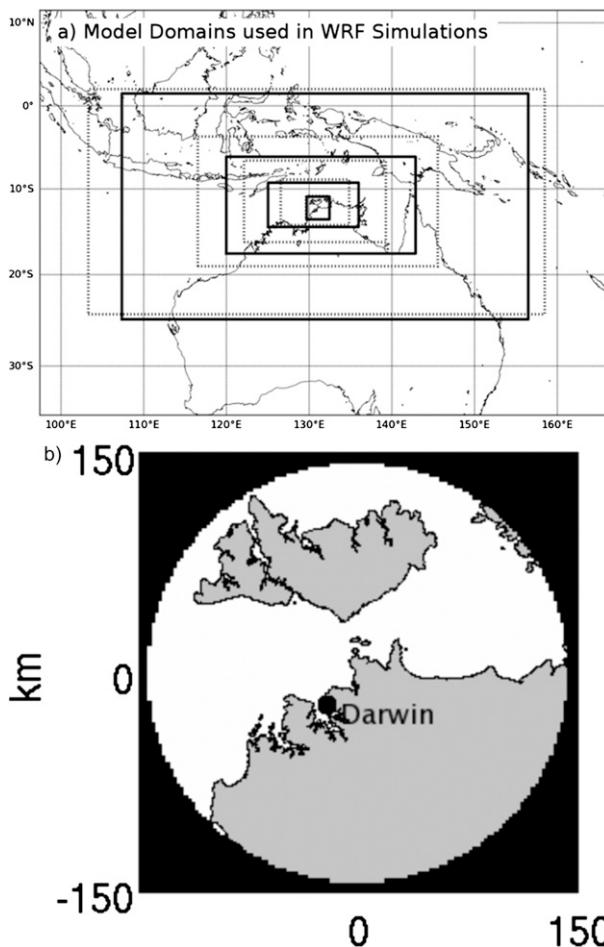


FIG. 1. (a) The spatial extent of the domains used in the ARW-WRF simulations. The configuration using four domains is represented by the solid lines and the configuration using five domains is represented by the dotted lines. (b) Example of the valid radar domain at 2.5-km altitude with circular radar mask applied.

developed and tested using data from the Gunn Point radar and evaluated against a rain gauge network. The algorithm takes advantage of the polarimetric variables available from the CPOL radar for a more accurate determination of rainfall.

### b. The Weather Research and Forecasting Model

In this study the Advanced Research Weather Research and Forecasting (ARW-WRF) Model, version 3.1.1 (Skamarock et al. 2008), is used to simulate tropical convection occurring over Darwin. Most of the analyses presented herein are derived from a nested simulation that uses four one-way nested domains. The locations of the outer domain and three (one-way) nested domains are shown in Fig. 1a (with solid lines) and have horizontal grid spacings equal to 33.75, 11.25, 3.75, and 1.25 km, respectively. These four domains have horizontal

dimensions of  $160 \times 90$ ,  $223 \times 115$ ,  $319 \times 157$ , and  $244 \times 244$  grid points, respectively. The dimensions of the innermost domain were chosen to encompass the range of the Darwin (Gunn Point) radar (Fig. 1b). Each domain uses 64 vertical levels with a 10-mb model top and a 5-km-deep upper-level absorbing layer.

The various physics packages common to all simulations include the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997), the Goddard shortwave radiation scheme (Chou and Suarez 1994), the Noah land surface model (Chen and Dudhia 2001), and the Mellor–Yamada–Janjić 2.5-level boundary layer scheme (Mellor and Yamada 1982; Janjić 2001). The Betts–Miller–Janjić cumulus parameterization scheme (Betts and Miller 1986) is used in the first two domains only. Domains 3 and 4, with their 3.75- and 1.25-km horizontal grids, are convection permitting and convective processes are thereby treated explicitly. Two different simulations are run with the high-resolution setup. The first uses the Purdue–Lin microphysics scheme (Lin et al. 1983; Rutledge and Hobbs 1984); this simulation is used for most of the paper and is referred to as simply the “Lin simulation.” The second configuration uses the Thompson microphysics scheme (hereafter referred to as the “Thompson simulation”) (Thompson et al. 2008). The outer domain of all simulations receives its initial and boundary conditions from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA-Interim) data (Dee et al. 2011); the 6-hourly ERA-Interim data is interpolated in time to provide the boundary conditions for the model. The simulations are initialized at 1200 UTC 8 February 2006 and end at 1200 UTC 13 February 2006: a period classified as belonging to the monsoon break regime (May et al. 2008). The first 12 h of these simulations are regarded as spinup time and excluded, leaving 4.5 days of simulation for the analysis. The performance of the simulation is, of course, affected by its 5-day length and biases could develop. An alternate approach might be to reinitialize the model periodically and/or include internal grid nudging. However, Wapler et al. (2010) showed for their simulations that the model performance (in terms of rainfall) was not necessarily better with a shorter simulation lead time. Furthermore, reinitializing the simulations would lead to discontinuities in the cell locations, disrupting the cell-tracking feature within TITAN. As discussed earlier, the sensitivity of model performance to such changes in simulation design are not the focus of this study and could be explored in further research.

### c. Microphysics schemes

The model simulations examined herein rely heavily on the details of the microphysics schemes used for both

the simulation of convection, and in determining the simulated radar reflectivity (derived later). Therefore, a brief description of the microphysics schemes used in the simulations is provided in this section.

The Lin microphysics scheme is a single-moment bulk microphysical parameterization that simulates five forms of water (cloud water, cloud ice, rain, snow, and graupel). Cloud water and cloud ice are assumed to have negligible terminal velocities because of their small size, and are therefore nonprecipitating. Of the three precipitating fields (rain, snow, and graupel) particle-size distributions are assumed to follow a Marshall–Palmer or exponential equation of the following form:

$$N(D) = N_0 e^{(-\lambda D)}, \quad (1)$$

where  $N(D)$  is the number concentration of particles with diameter  $D$ ,  $N_0$  is the intercept parameter, and  $\lambda$  is the slope parameter that is defined as

$$\lambda_x = \left( \frac{\pi N_{0,x} \rho_x}{\rho_{\text{air}} q_x} \right)^{1/4}, \quad (2)$$

where  $\rho_{\text{air}}$  is the density of dry air; and  $\rho_x$ ,  $q_x$ , and  $N_{0,x}$  are the density, mixing ratio, and intercept parameter of the hydrometeor category (with subscript  $s$ ,  $g$ , and  $r$  corresponding to snow, graupel, and rain, respectively). In version 3.1.1 of the ARW-WRF Model these parameters are  $N_{0,s} = 3 \times 10^6$ ,  $N_{0,g} = 4 \times 10^6$ ,  $N_{0,r} = 8 \times 10^6$ ,  $\rho_s = 100 \text{ kg m}^{-3}$ ,  $\rho_g = 400 \text{ kg m}^{-3}$ , and  $\rho_r = 1000 \text{ kg m}^{-3}$ .

The Thompson microphysics scheme simulates the same five classes of water as the Lin scheme (cloud water, cloud ice, rain, snow, and graupel). However, as well as predicting the mixing ratio for these five species, this version of the Thompson scheme also predicts the cloud ice and rain number concentration (i.e., it is a double-moment scheme for both cloud ice and rain).

This configuration of the Thompson microphysics scheme uses an exponential distribution to represent the particle-size distributions for the hydrometeor species, with the exception of snow, which is the sum of an exponential and gamma function (Thompson et al. 2008). Furthermore, for rain the Thompson scheme has an intercept value  $N_{0,r}$  and a slope parameter  $\lambda_r$  dependent on both the rainwater mixing ratio  $q_r$  and the rain number concentration  $N_r$  [see Eqs. (B4) and (B5) of appendix B]. The Thompson scheme represents graupel in a similar fashion to the Lin scheme, with the exception that the intercept parameter  $N_{0,g}$  is a function of the graupel mixing ratio, rather than simply a constant as in the Lin scheme. In version 3.1.1 of the ARW-WRF Model the intercept value for graupel was updated [as seen in Eq. (B10)] and is now expressed in a similar fashion to the

rain intercept parameter described in Thompson et al. (2008). Finally, in the Thompson scheme the particle-size distribution for snow is represented by the sum of an exponential and gamma function shown in Eq. (B11), which is a function of both temperature and mixing ratio. For additional differences between the two microphysics schemes see Lin et al. (1983), Chen and Sun (2002), Thompson et al. (2004, 2008), Field et al. (2005), and Morrison and Pinto (2005).

#### *d. Simulated radar reflectivity*

Radar reflectivity (dBZ) is not a prognostic variable within the ARW-WRF Model. As radar reflectivity is proportional to mass squared, for the liquid phase (where the particles are assumed to be spherical) it is straightforward to calculate reflectivity as it is simply the integral of the sixth moment of the drop size distribution (see appendixes A and B). For the ice phase, reflectivity is still generally weighted to the sixth moment of the particle-size distribution, but for example, assumptions on the ice density are also important. Therefore, to compare data from the numerical model to those observed by a radar, a conversion between simulated variables and radar reflectivity is required. Many authors have developed such conversion algorithms (Anagnostou and Krajewski 1997; Chandrasekar and Bringi 1987; Koch et al. 2005) and the concept behind them is that particle-size distributions (specific for a given microphysics parameterization scheme) can be used to calculate the combined surface area that a distribution of particles would present to an unattenuated radar beam. The algorithms used herein are slightly different to account for the differences in the microphysical parameterizations, and details of these algorithms are provided in appendixes A and B.

Figure 2 shows example observed and (Lin) simulated reflectivities on two different days during the simulated period. These figures highlight the development of “Hector” thunderstorms (Figs. 2c,d) over the Tiwi Islands, the two islands to the northwest of the mainland, and localized convection forming over the mainland (Figs. 2a,b). Qualitatively, the model simulations appear to provide a realistic representation of the convection and its spatial characteristics.

#### *e. Cell-tracking and identification method*

The technique used in this study exploits the cell-tracking algorithm found in the TITAN software package (Dixon and Wiener 1993). TITAN was first developed in the 1980s at the National Center for Atmospheric Research (NCAR) to identify and track convective cells in radar data (NCAR 2012).

TITAN has been used for a variety of applications including nowcasting, case studies, climatology studies,

and weather modification studies (Krauss and Santos 2004; May and Ballinger 2007; Potts et al. 2000; Vallgren 2006; Horvath et al. 2008). TITAN defines a storm as a contiguous region of reflectivity at or above a user-defined threshold. Once identified, TITAN then approximates the defined region as either an ellipse or polygon, the vertical overlap of these regions defines the three-dimensional structure of the storm. For simplicity, in this study storms are approximated as ellipses. The number of storms identified by TITAN is sensitive to a minimum storm volume parameter. May and Ballinger (2007) used a minimum storm volume of  $30 \text{ km}^3$  to eliminate noisy data, in this study a larger minimum storm volume of  $100 \text{ km}^3$  was chosen to exclude storms that are poorly resolved by the model (although it is conceivable that there are some storm geometries that remain poorly resolved). Figure 3 shows an example of the storms identified by TITAN, in this example a threshold of 45 dBZ is used to define the storms. For a given storm identified by TITAN (defined by a specific dBZ threshold) a number of quantities are calculated and tracked through its lifetime, examples include the location of the storm, the length of the major and minor axes of the storm ellipse, the orientation of the storm, the maximum height of the storm, the maximum size of the storm region, and the speed and direction of the storm propagation.

The application of TITAN to radar data is well established. However, in order to use the cell-identification algorithm within TITAN as a tool for model assessment, the model output must be manipulated into a form compatible with the TITAN software and consistent with the radar data (Pinto et al. 2007). Three-dimensional model microphysical quantities are stored every 10 min for the length of each simulation, which matches the temporal resolution of the Gunn Point radar data. After the model microphysical variables are converted to simulated reflectivity, the data are interpolated and averaged to the same grid specifications as the data from the Gunn Point radar. Thus, for each model time period the data are interpolated onto 40 equally spaced vertical levels ranging from 0.5 to 20 km (500-m spacing), and averaged onto a Cartesian grid with a spacing of 2.5 km in each horizontal direction (which by design is an integer multiple of the model grid spacing). At each vertical level circular masks are applied to limit the simulated reflectivity data to the same area observable by the radar (e.g., Fig. 1b). Finally, the data are converted to meteorological data volume (MDV) format, the native input format for TITAN. The TITAN algorithm is then applied to this processed model data (i.e., the TITAN procedures are completed “offline”). At this stage the source of the data is not relevant to TITAN and the algorithm proceeds

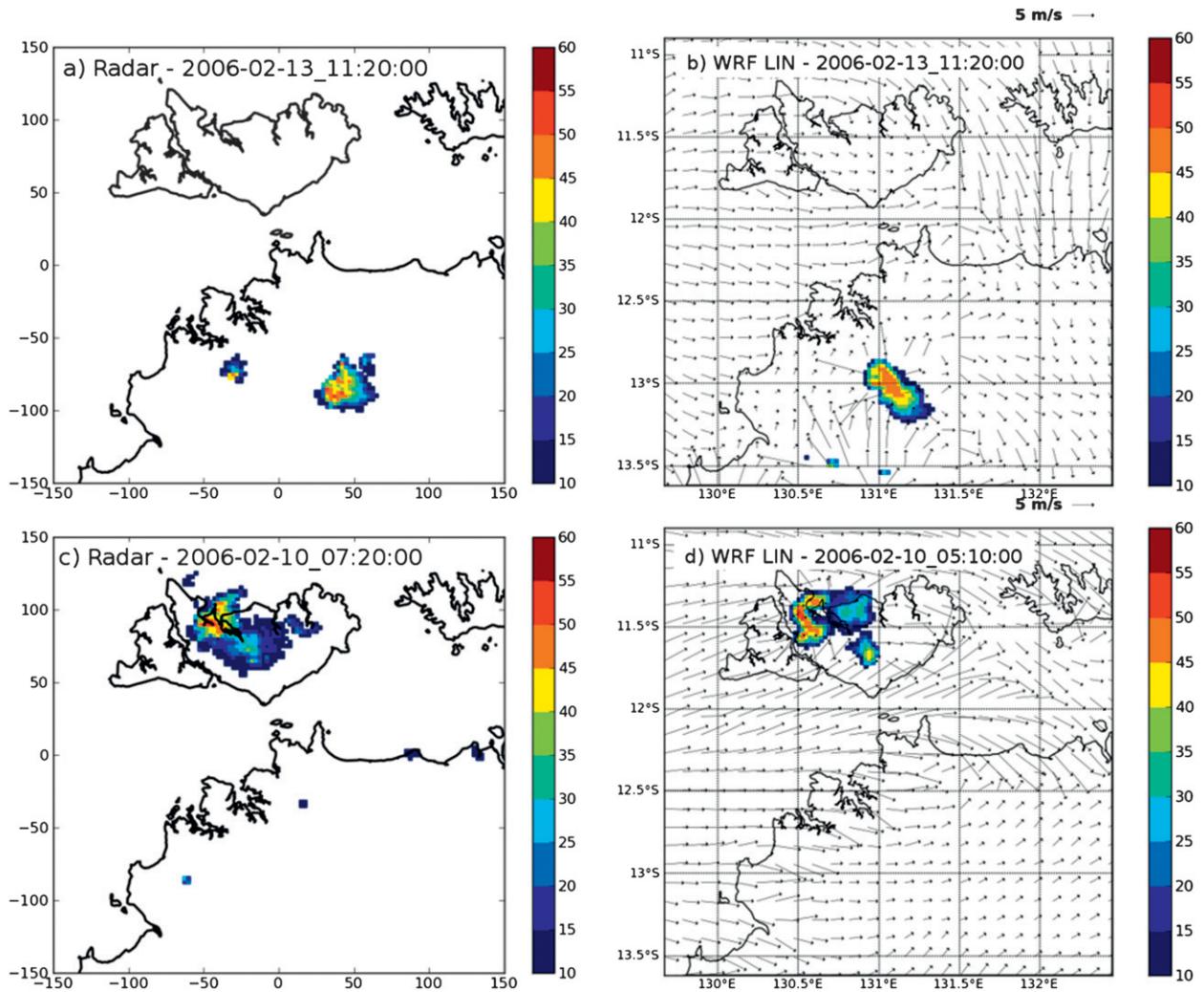


FIG. 2. Example (a),(c) observed and (b),(d) simulated reflectivity fields at 2.5 km. Note that (a) and (b) are at the same time, but (c) and (d) differ by 2 h, 10 min. Wind direction and magnitude are shown by the black arrows in the simulations (10-m altitude), with the reference vector shown at the top right.

in exactly the same way as with the radar observations to identify, track, and provide information on the characteristics of the simulated storm cells.

Applying TITAN to the analysis period (4.5 days, 10-min interval data) gives sample sizes (number of cells) on the order of 1000–2000. This allows simulated storm properties to be compared with the observed storm properties in a statistical fashion.

As TITAN allows storms to be identified by choosing a reflectivity threshold, different properties of the storms can also be investigated by changing the reflectivity threshold. For example, a relatively low reflectivity threshold (for example 25 dBZ), would likely contain information about the larger storm complex, including stratiform precipitation. If, on the other hand, a reflectivity threshold of 45 dBZ is used to define the storms, then the cell

properties identified by TITAN will provide information primarily about the intense convective cores (May and Ballinger 2007). Such statistics are examined in the next sections.

### 3. Results

A comparison of observed and simulated convective cell properties determined using TITAN will be examined shortly. However, before introducing these results some more traditional analysis is performed to contrast some of the results shown in later sections. Figures 4a,b show the domain-averaged rain rate and accumulated precipitation, derived from the radar and obtained from the Lin simulation (both are calculated over the range of the radar). These figures show that the model approximately

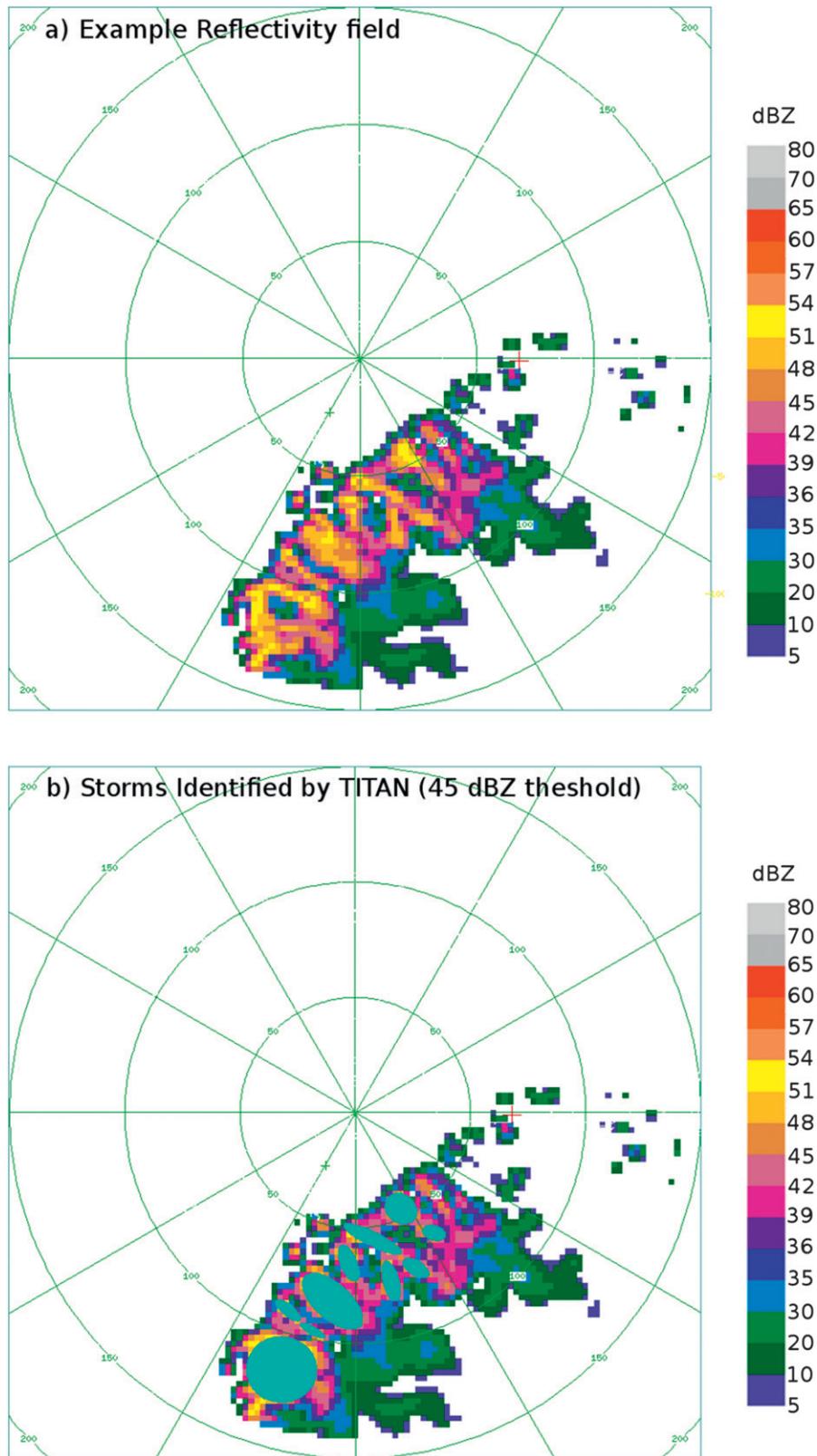


FIG. 3. (a) Sample reflectivity field derived from the Lin simulation at 1040 UTC 9 Feb 2006. (b) As in (a), but with green shading denoting the storms identified by TITAN using a 45-dBZ threshold.

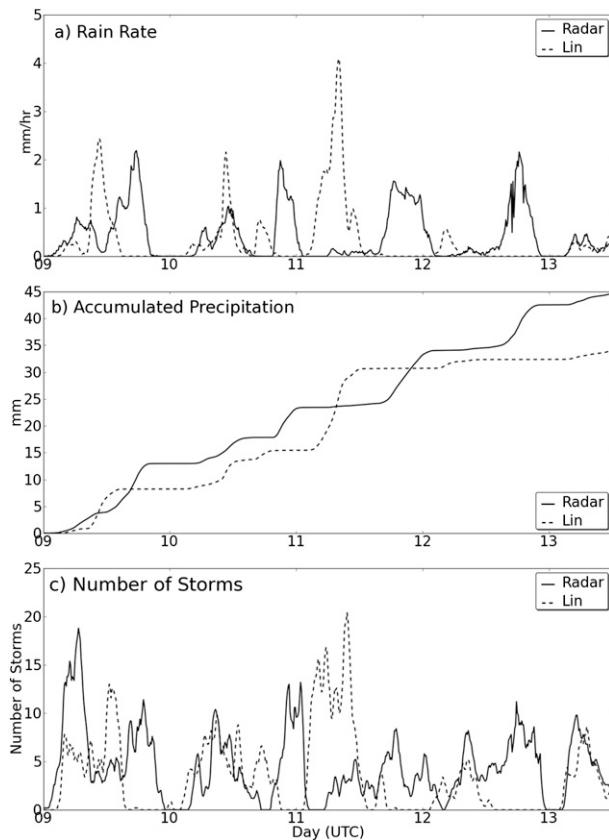


FIG. 4. (a) Domain-averaged rain rate derived from Lin simulation and radar data. (b) Domain-averaged accumulated precipitation. (c) Number of storms identified by TITAN using a 25-dBZ reflectivity threshold.

reproduces the number of large precipitation events, their duration, and their intensity. However, there are differences in the timing and intensity of individual events. During the first two days the Lin simulation domain-averaged precipitation and rain rate track the observations well. However, the model produces a large event on 11 February, whereas the observations show two smaller events before and after the modeled one. Despite this difference, by the end of the period the simulation and observations are in good agreement (in terms of rain rate) in the last 12 h of the simulation. The differences between observed and simulated precipitation are likely due to a combination of errors in model initial and boundary conditions as well as errors in physical parameterizations (e.g., Wapler et al. 2010; Wapler and Lane 2012). Further efforts with ensembles would expose the relative importance of these errors.

#### a. Pixel-based evaluation

One possible reason for differences between observed and simulated precipitation is inadequate representation of the properties of convection within the model, leading

to differences in the three-dimensional distribution of hydrometeors. Figures 5a,b show statistics of echo-top heights at various reflectivity values derived from the radar data and the Lin simulation. These statistics were created using a simple pixel-based approach, where for each horizontal grid point in the domain the maximum height of a given reflectivity value is calculated. For most reflectivity thresholds the spread of echo-top heights is overestimated in the Lin simulation, seen as an overestimation in both the interquartile range (IQR) and the height of the 95th percentile. Median-simulated echo-top heights are underestimated at all but the highest reflectivity thresholds. Nonetheless, the median in both the model and the observations show a similar decreasing trend with reflectivity threshold.

Figures 5c–f show histograms of echo-top height for the radar and the Lin simulation, for two different reflectivity thresholds (25 and 45 dBZ). [Here and throughout the manuscript these two thresholds are used to distinguish between the properties of intense convective cores (45 dBZ) and broader stratiform regions (25 dBZ).] At the 25-dBZ threshold the observations show a broad peak between 7 and 10 km, followed by a rapid drop of in frequency above 10 km. In contrast, the Lin simulation shows a sharp peak around 5 km and a shallow dropoff in frequency with height. At the 45-dBZ threshold the observed and simulated echo-top height distributions are more similar, with the simulated peak occurring around the correct height. However, once again the radar observations show a rapid decrease in frequency with increasing height (above 7 km) that is not found in the simulations. The number of samples that contribute to each distribution is also displayed in Fig. 5 (shown in the top-right corner of Figs. 5c–f). Note that the number of samples (total reflectivity area) is less in the simulated data for both reflectivity thresholds compared to the observations, and is substantially lower for the 25-dBZ threshold. Thus, the total storm area is underestimated by the model. The large difference in sample sizes seen here also highlights a potential problem with pixel-based evaluation of echo-top height. Namely, the size and number of discrete convective cells are inseparable and determining the underlying cause of differences between two datasets (e.g., model and observations) is difficult. In the following sections it will be shown that the TITAN approach of analyzing cells as individual objects can, among other things, delineate errors associated with cell size.

#### b. TITAN statistics

The TITAN algorithm is applied to the 4.5-day simulation period to identify the convective cells in the simulations and observations. The time series of the number of identified convective cells in the radar and

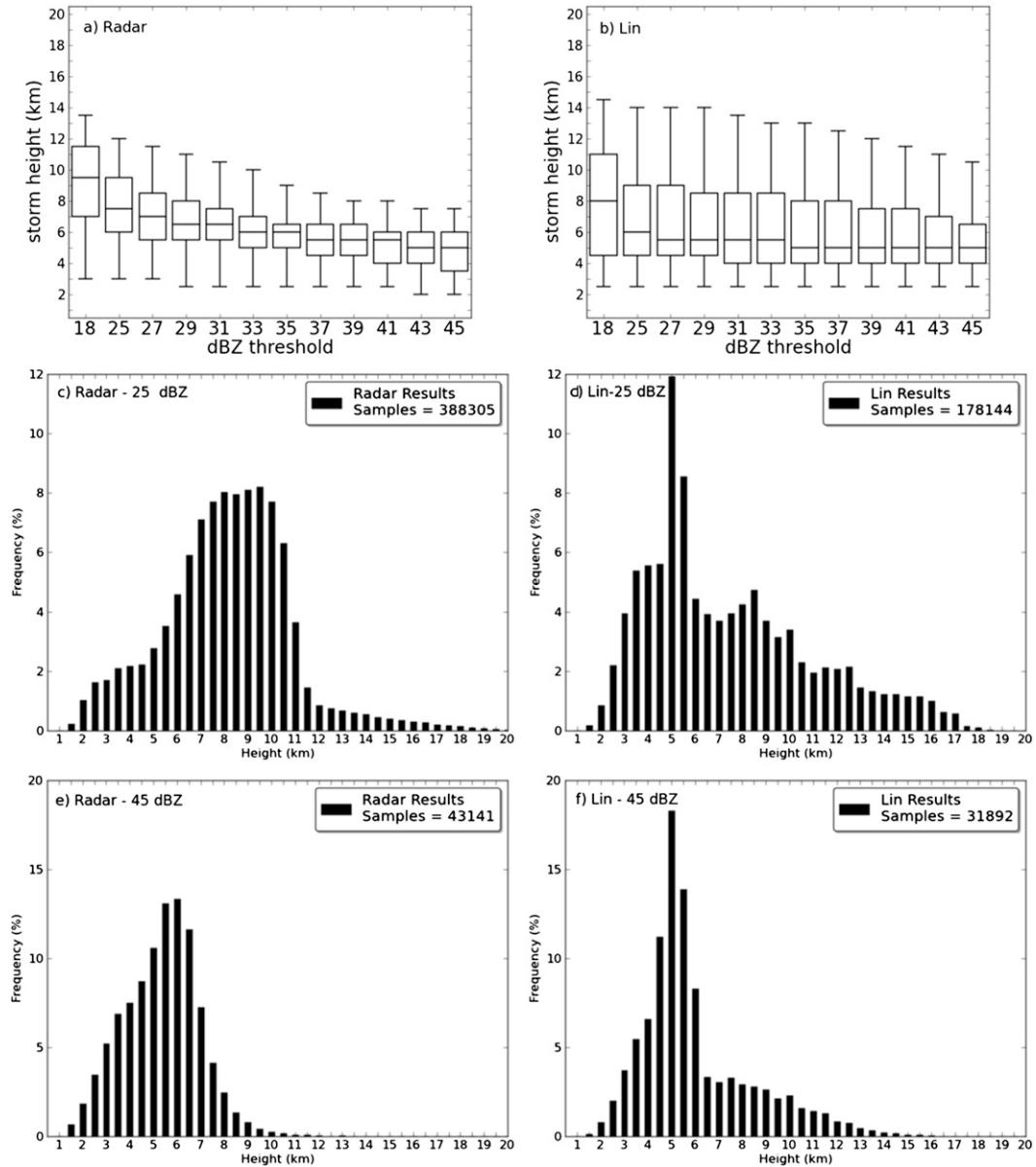


FIG. 5. Echo-top height statistics for (a) the radar and (b) the Lin simulation, where the statistics have been derived using a traditional pixel-based approach for determining echo-top height. In the box-and-whisker plots the interquartile range is displayed by the box, and the whiskers extend to the 5th and 95th percentiles. (c)–(f) Histograms of echo-top height for the same data when a threshold of 25 dBZ is used [(c) radar, (d) Lin], and when a 45-dBZ threshold is used [(e) radar, (f) Lin].

the Lin simulation (using a 25-dBZ threshold) are shown in Fig. 4c. Although there is some correspondence between rain rate (Fig. 4a) and the number of cells, there is not a direct relationship. For example, in some cases there is good agreement between the modeled and simulated rain rate but not the number of cells, and vice versa.

Figure 6 shows observed and simulated statistics of maximum storm height and cell size produced by TITAN, as a function of the reflectivity threshold used to define

the cells. Here, cell-top height (or simply storm height) is defined as the maximum height of the reflectivity threshold used to define the storm cell, and storm size is defined as the downward projection of the maximum storm dimensions. Note that this analysis circumvents the problem of biases in storm size affecting the echo-top height statistics seen in Fig. 5 because the TITAN analysis allows the storm height and storm size properties to be analyzed independently.

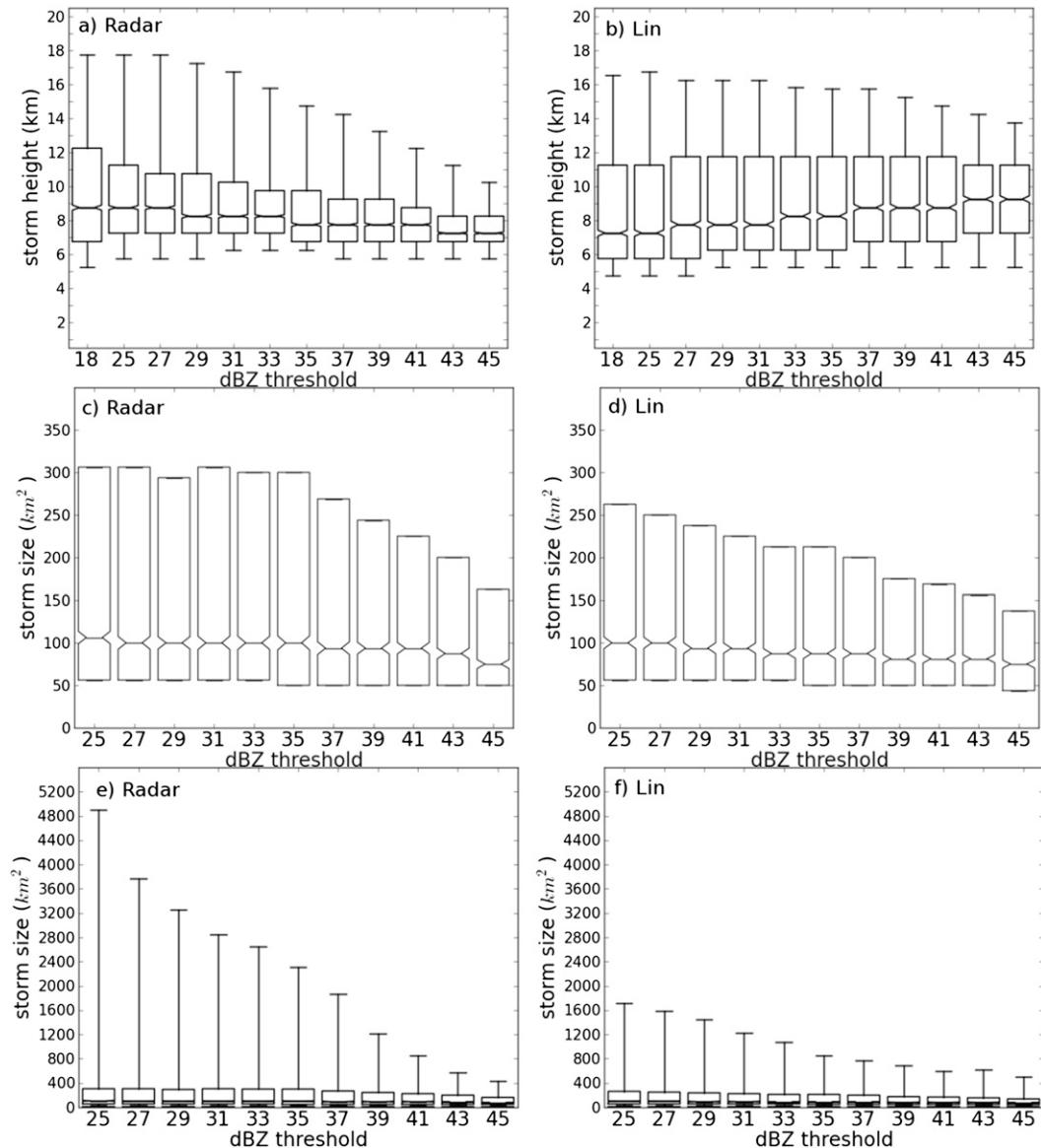


FIG. 6. Box-and-whisker plots derived from all storms identified by TITAN during the analysis period for different reflectivity thresholds. The interquartile range is displayed by the box, and the whiskers extend to the 5th and 95th percentiles. Cell-top height statistics for the (a) radar and (b) Lin data. (c),(d) The storm size statistics, with only the median and interquartile range displayed. (e),(f) As in (c),(d), but that the whiskers are also displayed.

The results in Fig. 6 include all storms identified by TITAN between the 0000 UTC 9 February 2006 and 1200 UTC 13 February 2006. Note that no temporal information has been included in Fig. 6, and each cell identified by TITAN is treated as an independent sample regardless of whether TITAN has identified the storm previously. In this way the bulk statistics of the storm cell properties are examined over the analysis period. There are 2859 (2220) cells identified by TITAN in the radar data (Lin simulation) when a reflectivity threshold of 25 dBZ is used to define the cells, and 1150

(1037) cells identified in the radar data (Lin simulation) when 45 dBZ is the threshold. Thus, despite the underrepresentation in total simulated storm area noted in Fig. 5, the Lin simulation and the radar data possess a similar number of convective cells.

Figures 6a,b show the cell-top height statistics for the radar and Lin simulation, respectively (note again that this statistic is different from the simple pixel-based echo-top statistic seen in Fig. 5 because it is denoting the distribution of the tops of discrete cells). Statistics obtained when a reflectivity threshold of 18 dBZ is used to

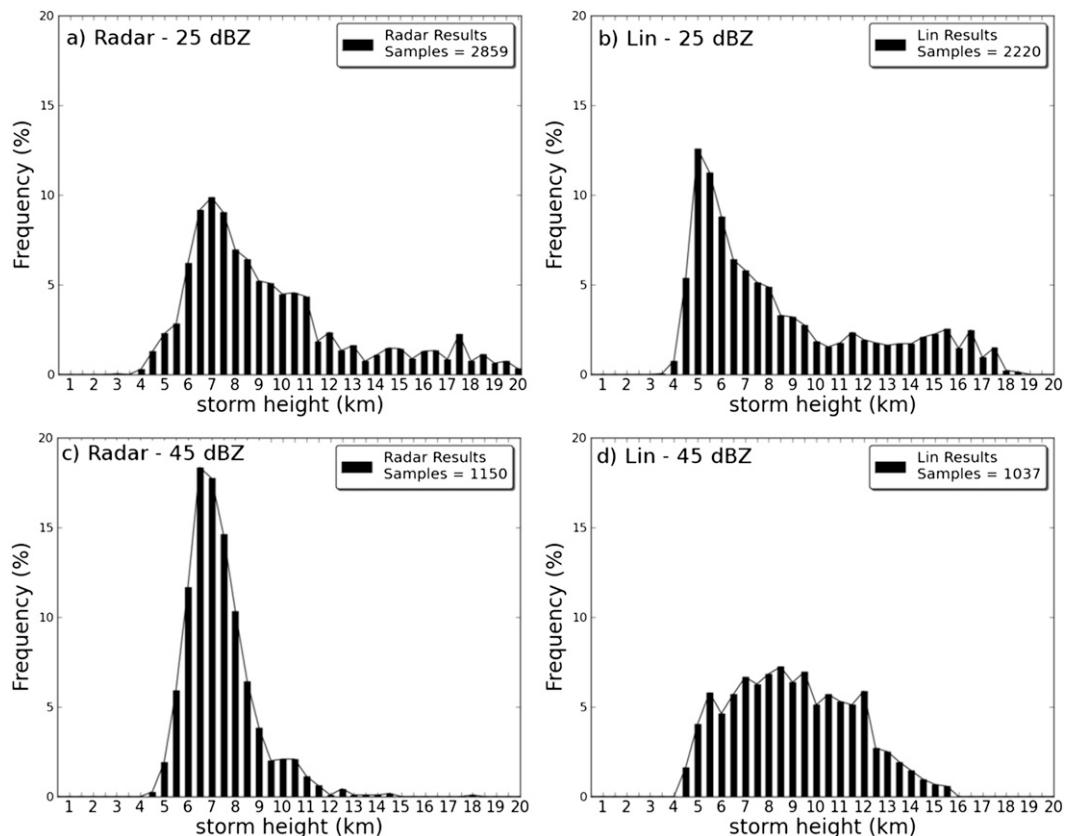


FIG. 7. Histograms of cell-top height when the 25-dBZ threshold is used to define the storms in the (a) radar and (b) Lin data. (c), (d) As in (a), (b), but a 45-dBZ threshold is used. The number of cells identified for these thresholds is also displayed in the legend.

define the cells have been included in these two figures as previous authors have used the 18-dBZ contour a proxy for cloud-top height (Magagi and Barros 2004; Heymsfield et al. 2000; Dupree et al. 2006).<sup>1</sup> There are substantial differences between the simulated and observed storms. The simulated interquartile range is large and approximately constant over the reflectivity thresholds in the model, while it is small and decreasing with the reflectivity threshold in the observations. Median cell-top heights show opposite trends for the observed and simulated storms; the median cell-top height of the observed (simulated) storms is a maximum (minimum) at the lowest reflectivity threshold and gradually decreases (increases) as the reflectivity threshold is increased. At the largest reflectivity thresholds the cell tops are too high. These differences are in contrast to the median trends

<sup>1</sup> Note that 18 dBZ corresponds to the detection limit of the precipitation radar on board the Tropical Rainfall Measuring Mission (TRMM) satellite and will generally be an underestimate of true cloud-top height.

versus reflectivity threshold exposed using the traditional approach (cf. Fig. 5), which appears to show better agreement with the observations. However, those traditional statistics combine cell sizes and heights together, and opposing errors in these two properties can compensate one another.

To further elucidate the differences between the observed and simulated distribution of storm-top heights, the histograms for the 25- and 45-dBZ thresholds are shown (Fig. 7). For the 25-dBZ threshold (Figs. 7a,b) both the simulated and observed height distributions appear qualitatively similar. However, although the peak in the modeled distribution and its median (Figs. 6a,b) are approximately 2 km lower than observed, the modeled distribution features a greater occurrence of cells between approximately 11 and 16 km high. At the 45-dBZ threshold the median of the simulated distribution is approximately 2 km too high (Figs. 6c,d) and the distribution is too broad and flat (Figs. 7c,d). Note that errors in the model reflectivity conversion algorithm might explain part of the relatively large differences near the melting level (~5 km); however, the differences at

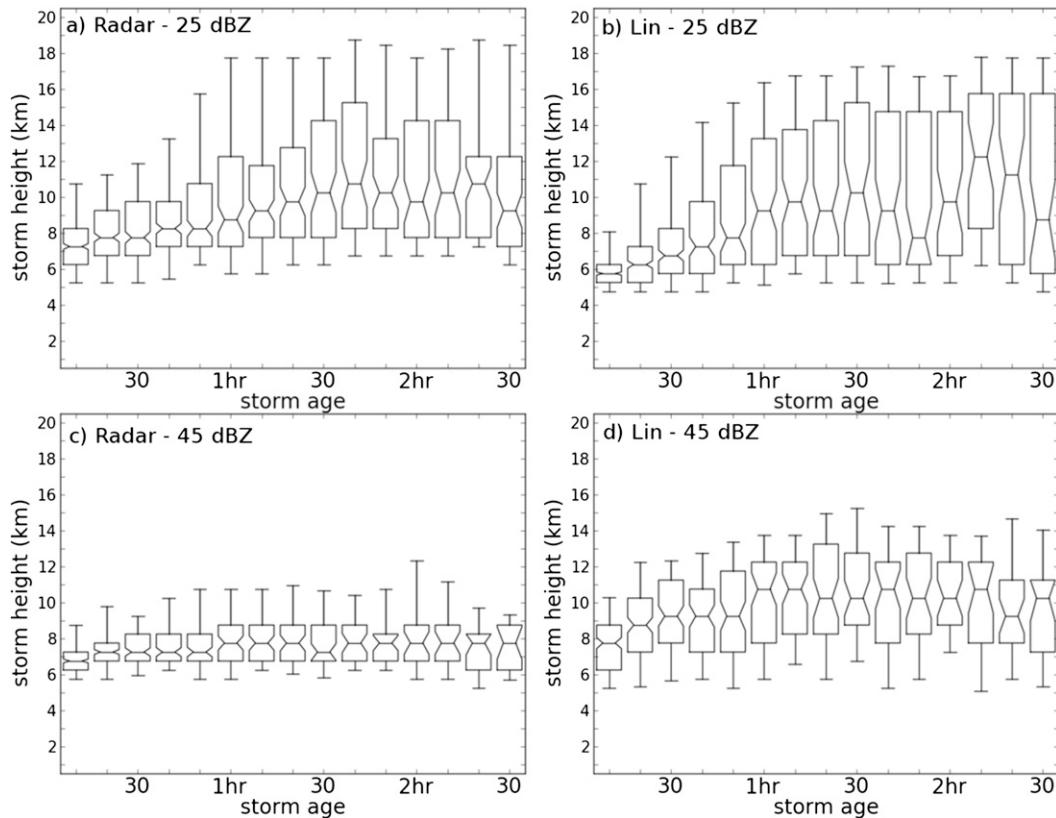


FIG. 8. Cell-top height statistics as a function of storm age. (a),(b) The results of the radar and Lin data when the 25-dBZ threshold is used to define the storms. (c),(d) As in (a),(b), but a 45-dBZ threshold is used.

higher altitudes are most likely attributable to actual differences between the modeled and observed cell characteristics. Thus, these comparisons show that the simulated convection produces an overrepresentation of deep convective cells, which is worse at higher reflectivity thresholds (more intense cells).

Figures 6c,d show the median and IQRs of storm sizes for both the radar and the Lin simulation. The median storm sizes and 25th percentile compare reasonably well in both datasets. However, the main difference occurs at the upper percentiles; the 75th percentile is approximately 50–100 km<sup>2</sup> smaller in the model than observed. The differences for the larger cells are further highlighted in Figs. 6e,f (where the whiskers extend to the 95th percentile) and exceed 1000 km<sup>2</sup>. These figures show that the Lin simulation is failing to represent the size of the larger storms, which is consistent with the simulation underestimating total area (Fig. 5) while producing a similar number of convective cells. The disagreement in size is worse at the lower reflectivity thresholds, suggesting increased difficulty representing the stratiform area of storms compared to the size of intense convective cores. Finally, TITAN also records

the horizontal geometry of the (elliptical) cells and the analysis found the modeled cells to have lower eccentricity than the observed cells (not shown). Thus, this comparison shows that the simulated cells are too small and too circular, in addition to being too deep.

In addition to the geometry of the detected cells, the TITAN algorithm also records temporal information about the cell age. The cell age corresponds to either the time since the cell first initiates or the time since a cell entered the analysis domain. Although not shown here, the shape of the age distributions of cells in the Lin simulation are similar to those observed by radar, with only minor differences for the youngest and oldest cells. These similarities suggest that the different cell geometries discussed above are not directly related to differences in the age of cells (e.g., there is not a simulated dominance of older cells that might be expected to be deeper).

Further details of the differences between the observed and simulated cell-top height distributions are exposed by considering how they vary with cell age (Fig. 8). For the 25-dBZ threshold (Figs. 8a,b) during the first 10 min of detection (the first bin) the simulated storms

have a lower median cell-top height and interquartile range than observed. The observed median cell-top height increases gradually to 10 km (at ages of about 1 h, 40 min), compared to the more rapid simulated increase to 10 km (at about 1 h). The interquartile range increases more rapidly with age in the simulation, and at almost all ages the interquartile ranges are larger in the simulation than the observations. These differences are exacerbated at the 45-dBZ threshold (cf. Figs. 8c,d). Although some of these differences might be explained by changes in simulated storm morphology as they enter the highest resolution domain, these results show that the simulated storms deepen more rapidly than observed and provide further evidence that the simulated cell tops are predominantly too high.

In summary, the results of this section elucidate specific differences between the observed and simulated storms. In general the Lin simulation produces convective cells that are predominantly too tall and horizontally too small. These two errors are combined together and can compensate one another using traditional pixel-based echo-top statistics, but are separable using the TITAN method. The model simulation had substantial difficulty representing the area of storms using the 25-dBZ threshold. This means the stratiform area is misrepresented, which is controlled, in part, by the cloud microphysics scheme and may be sensitive to the specific scheme used. The sizes of intense convective cores are better represented in the model, albeit still too small, presumably because their sizes are more constrained by the dynamics. Nonetheless, the characteristics of the convective cores (too tall and too small) are consistent with insufficient entrainment and detrainment, which could be a result of the model resolutions used. Underestimating mixing between the intense convective cores and their environment would lead to an overestimation of the updraft strength, and therefore the height of the convective cells. Moreover, underestimating detrainment could also lead to an underproduction of stratiform precipitation at midlevels and hence storm size at the lower reflectivity thresholds. These hypotheses will be investigated in more detail in the next two sections.

### *c. Sensitivity to microphysics*

The previous section identified errors in the representation of convective cell properties, and suggested that errors at lower reflectivity thresholds might be explained by certain aspects of the microphysics parameterization. Indeed, many previous studies have demonstrated how convective cell properties are sensitive to microphysics (e.g., Bryan and Morrison 2012; Morrison and Milbrandt 2011; Morrison et al. 2009; Varble et al. 2011, etc.). In this section, the sensitivity of

the convective cell properties to cloud microphysics is examined; the ARW-WRF Model is integrated a second time with the same setup described previously, except that the Thompson microphysics scheme is used instead of the Lin microphysics scheme.

The distributions of cell-top heights and storm sizes as a function of reflectivity threshold for the Thompson simulation are shown in Figs. 9b,d. For ease of comparison the results for the Lin simulation have been reproduced in Figs. 9a,c, and the median observed cell-top heights are also shown in Figs. 9a,b. Figure 9 shows that the greatest difference produced by changing to the Thompson microphysics scheme occurs in the distribution of cell-top heights when reflectivity values below 35 dBZ are used to define the storms. While the median and 25th percentile are similar between the Lin and Thompson simulations, the height of the 75th percentile is reduced considerably for the Thompson case. Above 35 dBZ there is less difference between the cell-top heights of the Lin and Thompson simulations, except the IQR has reduced in the Thompson simulation, which also improves its agreement with the observations (cf. Fig. 6a). However, as with the Lin simulation, the Thompson simulation shows an increase in the median cell-top height with increasing reflectivity threshold, which is opposite to what is observed. Nonetheless, the largest beneficial change in the cell-top height has occurred in the representation of the heights of stratiform cloud and/or weak convection (i.e., the lower reflectivities).

Figures 9c,d show that despite the improvements in cell-top height (at lower reflectivities), storm sizes are very similar in the Lin and Thompson simulations. The cumulative distribution function (CDF) of cell sizes at the 25- and 45-dBZ thresholds (not shown) also illustrates that the size distributions between the Lin and Thompson simulations are in close agreement to one another and changing the microphysics does not provide an improved representation of cell sizes. This result suggests that errors in cell sizes might also be explained by other model characteristics, such as model resolution (see the next section).

To further elucidate the changes in the cell-top height behavior, histograms at the 25- and 45-dBZ threshold are determined for the Thompson simulation (Figs. 10b,d) and reproduced for the Lin simulation (Figs. 10a,c). At the 25-dBZ threshold the peak in the Thompson histogram has moved to a higher altitude and broadened compared to the Lin histogram. Comparison of these figures with the observed histograms (dashed line in Fig. 10) shows that the Thompson histogram more closely resembles the observed distribution. At the higher (45 dBZ) threshold, the Thompson histogram is slightly

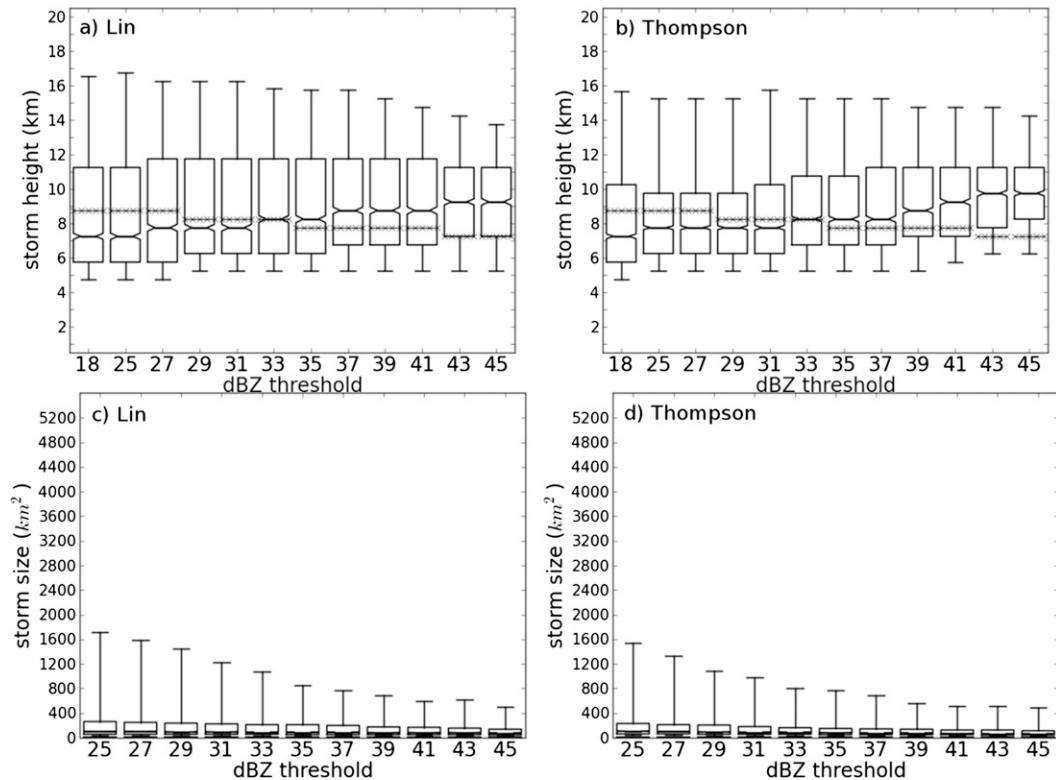


FIG. 9. Cell-top height statistics as a function of reflectivity threshold for (a) Lin and (b) Thompson simulations. For comparison the observed median cell-top height values are shown by the crosses. The storm size statistics are shown as a function of reflectivity threshold for (c) Lin and (d) Thompson.

narrower than Lin, but certainly not narrow enough to be classed as a good representation of the observations.

Statistical tests can also be applied to the TITAN-derived distributions to objectively determine whether changing the microphysics scheme (or any other model property) has improved the simulation in a statistically significant way. The two-sample Kolmogorov–Smirnov (K–S) test is applied to both the radar and Lin, and the radar and Thompson histograms for a variety of reflectivity thresholds (Table 1). The null hypothesis for the K–S two-sample test is that the two samples are drawn from the same continuous distribution. If the  $P$  value is high (close to 1) then the null hypothesis cannot be rejected, which can be interpreted to mean that the modeled distribution is drawn from the same distribution as the observations, implying statistically significant agreement.

From Table 1 it can be seen that the Thompson microphysics has significantly improved the cell-top height statistics for low reflectivity values. The  $P$  values for the K–S test are greater than  $\sim 0.9$  for reflectivity thresholds below 30 dBZ, indicating that the Thompson and radar-derived histograms are likely drawn from the same distribution at the 90% significance level. In contrast, the  $P$

values are small for the comparison between the radar and the Lin simulation cell-top height distributions, implying poor agreement between the two distributions. Table 1 also summarizes the findings regarding storm size; the ARW-WRF Model performs badly at representing the storm size distributions at low reflectivity values (in both the Lin and the Thompson simulations). However, at higher reflectivity thresholds ( $> \sim 40$  dBZ) both the Lin and the Thompson simulations provide a good representation of the storm size distribution, with the K–S test indicating the model and radar histograms are likely drawn from the same distribution at the 90% significant level. Thus, these tests show that at thresholds less than 30 dBZ the Thompson simulation provides a better representation than the Lin simulation of the observed cell-top height distribution; however, changing the microphysics scheme does not improve the storm size statistics.

To gain some insight into why the Thompson simulation better represents the observed echo-top height distributions at low reflectivities, Fig. 11 shows vertical profiles of hydrometeor mixing ratios for both simulations. Figures 11a,b show the domain-averaged mixing ratio for the Lin and Thompson simulation, respectively.

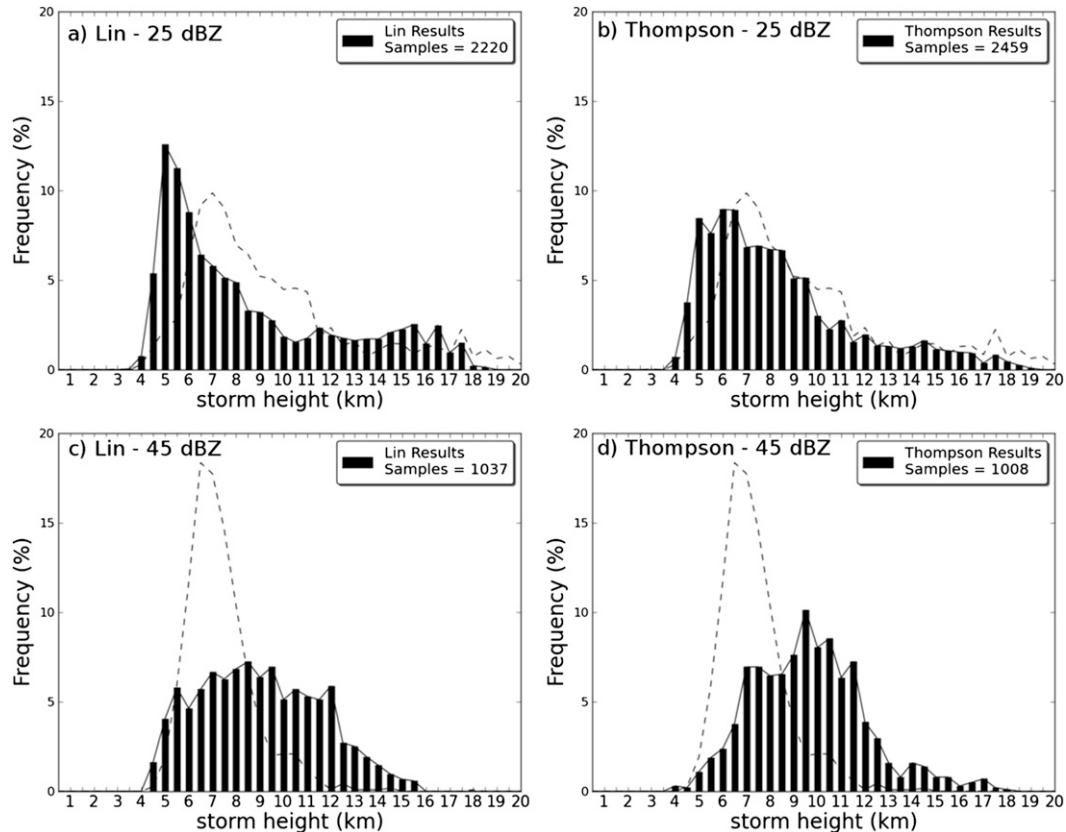


FIG. 10. Histograms of cell-top height for the (a) Lin and (b) Thompson simulations with a 25-dBZ threshold. (c),(d) As in (a),(b), but for a 45-dBZ threshold. For comparison the observed cell-top height distributions are shown by the dashed line.

The Thompson simulation has snow as the dominant hydrometeor category at higher altitudes, whereas the Lin simulation has graupel as the dominant category with very little snow. There is significantly more water at higher altitudes in the Thompson simulation than in the Lin simulation. Figures 11c,f also show the mixing ratios averaged over the 25- and 45-dBZ cells; this comparison demonstrates that the relative amplitudes of rain and graupel are similar in the Lin simulation, regardless of the reflectivity threshold. However, the Thompson simulation shows a change in the relative amplitudes of snow (the dominant ice category) and rain depending on the reflectivity threshold. This comparison suggests that Thompson simulation exhibits different microphysical behaviors for strong and weak cells, whereas the Lin simulation does not.

Recent studies using the ARW-WRF Model during TWP-ICE have compared the performance of these two microphysics schemes (e.g., Wu et al. 2009; Wang et al. 2009). Specifically, Wang et al. (2009) link the overrepresentation of cirrus in simulations with the Thompson scheme to an overrepresentation of snow. Moreover,

Wu et al. (2009) compare graupel from simulations with the Thompson and Lin schemes to polarimetric radar observations and find that the Lin (Thompson) scheme overrepresents (underrepresents) the amount of graupel. Given the large differences between the graupel and snow concentrations seen in the two simulations here (Fig. 11) and their similarity to the results of Wang et al. (2009, their Fig. 6), it is likely that similar biases are present (even though there may be differences in the implementations of the schemes). Moreover, given that graupel has a significantly higher terminal velocity than snow, it follows that the Lin simulation should precipitate more water mass than the Thompson simulation, with the Thompson simulation retaining more hydrometeor mass aloft. The domain-averaged hydrometeor sedimentation flux for both simulations is shown in Fig. 12, along with a comparison of the rain rates (during a period that the simulated timing of convection in both simulations agrees well with the observations). Indeed, the Lin simulation has a much larger flux of graupel than the Thompson simulation, the Lin simulation has a negligible flux of snow, and the flux of rain is

TABLE 1. Results of the Kolmogorov–Smirnov two-sample test comparing histograms of cell-top height and storm size derived from the ARW-WRF simulations to those derived from radar data. The test was applied between (left column) the radar and the Lin simulation, (middle column) radar and Thompson simulation, and (right column) radar and 0.417-km (Lin) simulation. Higher  $P$  values correspond to better agreement between modeled and observed distributions, and values greater than 0.7 are in bold.

dBZ	Type	$P$ value	$P$ value	$P$ value
		Lin	Thompson	Lin (0.417 km)
25	Height	0.361	<b>0.983</b>	<b>0.724</b>
29	Height	0.139	<b>0.983</b>	0.139
33	Height	0.080	<b>0.724</b>	0.139
37	Height	0.080	0.231	0.043
41	Height	0.080	0.231	0.139
45	Height	0.230	0.043	0.139
25	Size	0.000	0.000	0.000
29	Size	0.000	0.001	0.000
33	Size	0.042	0.022	0.007
37	Size	0.528	0.359	<b>0.949</b>
41	Size	<b>0.999</b>	<b>0.949</b>	<b>0.9999</b>
45	Size	<b>0.999</b>	<b>0.983</b>	<b>0.9997</b>

larger in the Lin simulation (Figs. 12a,b). Accordingly, the rain rates in the Thompson simulation are smaller than the Lin simulation (Fig. 12c). [Note that some studies (e.g., Morrison et al. 2009) note increased rain with two-moment schemes compared to similar one-moment schemes; however, in this case the differences in the representation of graupel and snow aloft probably makes the most important contribution to the differences in rain.]

The cell-top height distribution at the 25-dBZ threshold is significantly improved in the Thompson simulation compared to the Lin simulation. The distribution for the Lin simulation is too peaked and the altitude of the peak is too low, whereas the Thompson simulation has a broader distribution with a peak at higher altitudes (cf. Fig. 10). The peak in the 25-dBZ histogram in the Lin simulation corresponds to the intersection of the graupel and rain at about 5 km (Figs. 11c and 12a) and the mixing ratios and fluxes of rain (graupel) drop off rapidly with increasing (decreasing) altitude. In comparison, in the Thompson simulation (Figs. 11d and 12b) the mixing ratio and flux of rain extends higher above the melting level ( $\sim 5$  km); the mixing ratio and flux of graupel also extends well below the melting level. Thus, in the Thompson simulation both rain and graupel are less peaked in the vertical. The (supercooled) rain extends to approximately 9 km, causing the broader cell-top height distribution that agrees with the observations. In contrast, above the melting level the Lin simulation has no rain above 6 km; given that liquid water makes a stronger contribution to radar reflectivity than ice species, modifying the ability of the Lin scheme to retain supercooled

water might improve its representation of the cell-top height distribution at low reflectivity thresholds.

The 45-dBZ threshold cell-top height distributions for both simulations (Figs. 10c,d) show an overrepresentation of storm tops between about 10–15-km altitudes. This overrepresentation may be related to too much hydrometeor mass being transported aloft. Indeed, the Thompson simulation, with its greater mass of hydrometeors aloft, has a stronger peak in the histogram at these heights. As suggested earlier, the model may produce updrafts that are too strong and hence too tall due to insufficient entrainment, which would transport too much water aloft. This will be considered further in the next section.

#### d. Sensitivity to resolution

In the previous sections it was found that the simulated convective cells were predominantly too tall and too small, a problem that could only be partly remedied with changes to the model microphysics. One hypothesis is that at 1.25-km horizontal grid spacing, entrainment/detrainment is poorly represented (e.g., Bryan et al. 2003; Bryan and Morrison 2012) leading to biases in the updraft characteristics. This should result in unrealistically large updraft speeds and consequently higher cell-top heights than observed. For example, Blossey et al. (2007) found that changing horizontal resolution modified the strength of the convective overshoot, while in a study by Lang et al. (2007) CFADS of vertical velocity (their Fig. 8) indicated that when changing from 1000- to 250-m grid spacing the frequency of updrafts with speeds greater than  $5 \text{ m s}^{-1}$  decreased, as did the altitude these updrafts reached. Furthermore, underestimating detrainment and mixing might lead to an underestimation of the stratiform area as well. To test this hypothesis, another experiment was performed (using the Lin microphysics) with an additional nested domain included with horizontal grid spacing of 417 m. This fifth domain encompassed the same horizontal area as the 1.25-km domain in the previous simulations; to accommodate this, the dimensions of the transitional domains were increased. See Fig. 1a for the location of these new domains (shown with dotted lines). Specifically, the 33.75-, 11.25-, 3.75-, 1.25-, and 0.417-km domains have horizontal dimensions of  $180 \times 90$ ,  $283 \times 154$ ,  $502 \times 289$ ,  $718 \times 487$ , and  $745 \times 745$  grid points, respectively. All other parameters were kept the same as the Lin simulation described above.

Figures 13a,b show the cell-top height and size distributions as a function of reflectivity threshold for the 0.417-km simulation. This new simulation shows notable improvements. Although the previously seen unrealistic trend of increasing median cell-top heights with

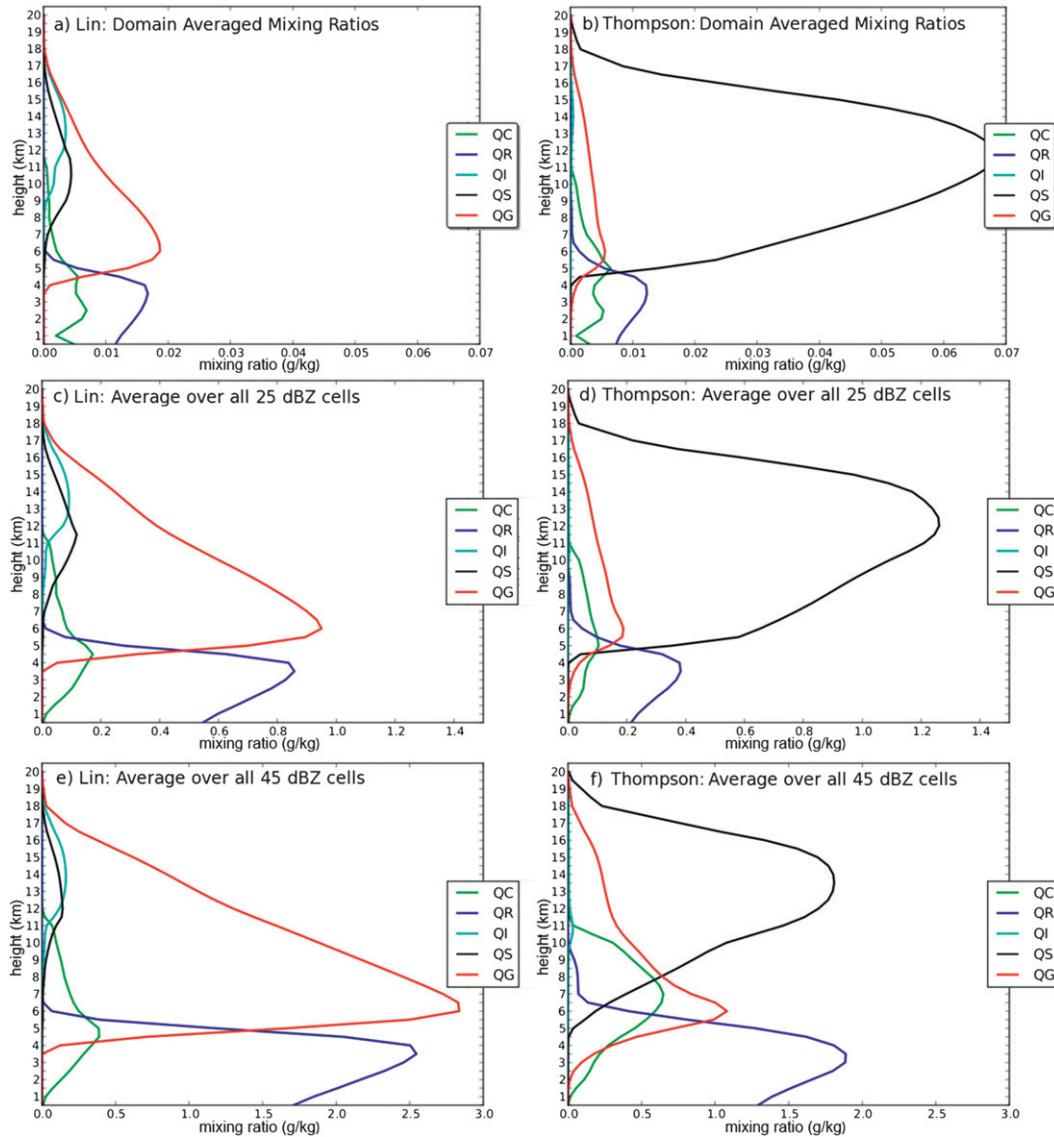


FIG. 11. The domain-averaged mixing ratios calculated over the 4.5-day simulation period for the (a) Lin and (b) Thompson simulations. (c),(d) As in (a),(b), but mixing ratios averaged over all 25-dBZ cells. (e),(f) As in (c),(d), but for 45-dBZ cells. Each panel shows the mixing ratios of cloud water (QC), rain (QR), cloud ice (QI), snow (QS), and graupel (QG).

reflectivity threshold is still present between 25 and 35 dBZ, above 39 dBZ median cell-top heights begin to decrease as the reflectivity threshold is increased, consistent with the observations. The storm size distribution has also improved; recall the major deficiency in the size distribution for the 1.25-km simulation was the underestimation of the size of the largest storms (Figs. 6e,f). In the 0.417-km simulation, although the median sizes do not change considerably compared to the 1.25-km simulation (not shown), the 95th percentile of the size has increased, e.g., at the 25-dBZ (45 dBZ) threshold it has increased from 1700 to 2300 km<sup>2</sup> (400 to 700 km<sup>2</sup>).

However, at the lower thresholds these values are still substantially smaller than observed. The K-S test demonstrates improvement in the size distribution (Table 1), with observed and simulated sizes at thresholds of 37 dBZ and higher being drawn from the same distribution at the 90% confidence level. Significant disagreement remains at lower thresholds.

The histograms of cell-top height show some improvement at 0.417-km grid spacing (Figs. 13c,d) compared to 1.25 km (Figs. 7b,d) with slightly lower cell-top heights. For example, at the 25-dBZ threshold the distribution is less peaked at ~5-km altitudes, which is in

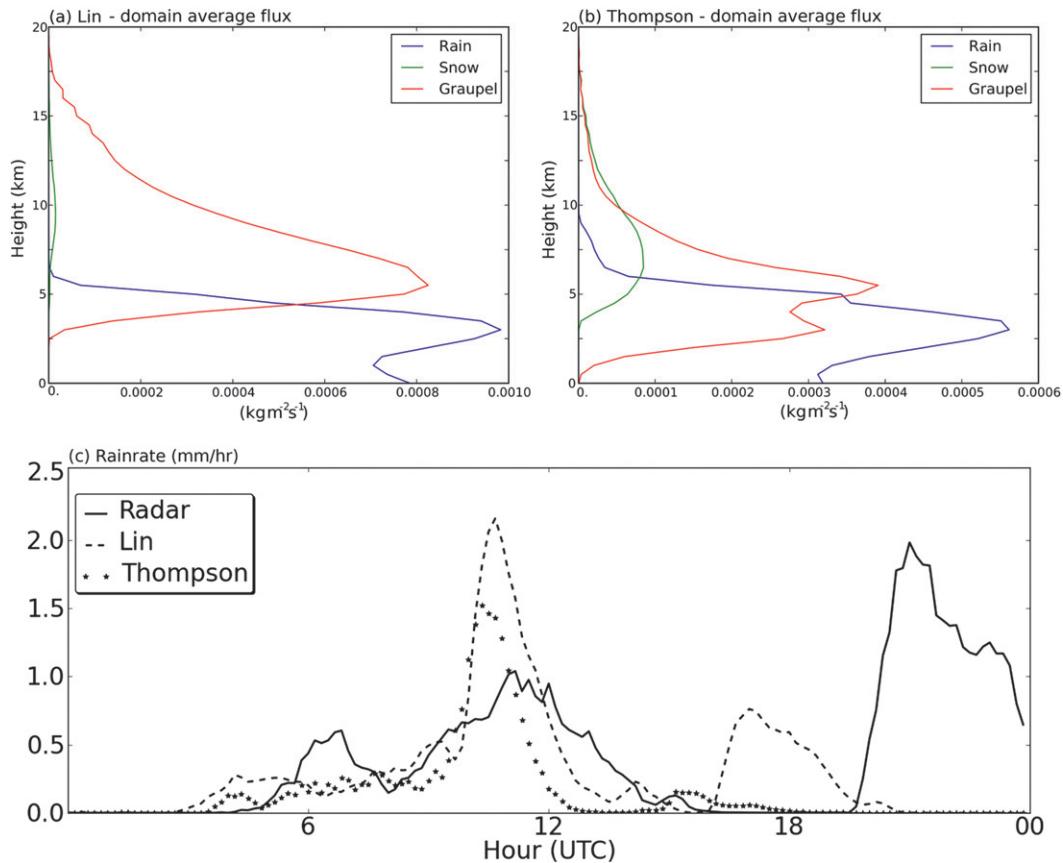


FIG. 12. The domain-averaged hydrometeor sedimentation flux calculated over the 4.5-day simulation period for (a) the Lin simulation and (b) the Thompson simulation; fluxes of rain, snow, and graupel are shown. (c) The rain rate from the observations, Lin, and Thompson simulations on 10 Feb 2006. Note the abscissae of (a) and (b) are different.

better agreement with the observations. Also, at the 45-dBZ threshold, the distribution is slightly more peaked around the 7–9-km altitudes at the expense of some of the higher-altitude contributions to the distribution. Specifically, there is a notable reduction in frequency between 10- and 13-km altitudes for the 45-dBZ threshold; this improves the comparison with the observations at these altitudes and supports the hypothesis described in the previous section that too much water mass was transported to these altitudes with the 1.25-km model. The K–S test (Table 1) demonstrates some improvement (although not nominally statistically significant) at the 25-dBZ threshold, with the simulated and observed cell-top heights drawn from the same distribution at the 70% confidence level. Significant disagreement remains at higher thresholds.

In an attempt to investigate whether increasing the horizontal resolution has led to an increase in entrainment, Figs. 14a,b show histograms of updraft velocity for the 1.25- and 0.417-km simulations at both 5- and 10-km altitudes. These figures show that at 5-km altitude, the

1.25-km simulation has a lower frequency of velocities above 12 m s<sup>-1</sup> than the 0.417-km simulation, whereas at 10-km altitude, the 1.25-km simulation has a higher proportion of velocities above 12 m s<sup>-1</sup>. The opposite relationship occurs below 12 m s<sup>-1</sup>. Figure 14c shows the ratio of the frequency of velocities at 10 and 5 km (Fig. 14b divided by Fig. 14a); while there are more strong updrafts at 10 km compared with 5 km in both simulations, the 1.25-km simulation has much greater relative frequency of strong updrafts at 10 km. Thus, it seems reasonable that increased entrainment and mixing in the 0.417-km simulation has led to a relative decrease in updraft velocities at higher altitudes, which ultimately reduces cell-top height. Note that vertical profiles of hydrometeor mixing ratios (from the 0.417-km simulation) averaged across the convective cells defined by the 25- and 45-dBZ thresholds (not shown) are qualitatively similar to those shown in Figs. 11c,e, indicating that differences found between the 1.25- and 0.417-km simulations are likely result of changed dynamics rather than microphysics.

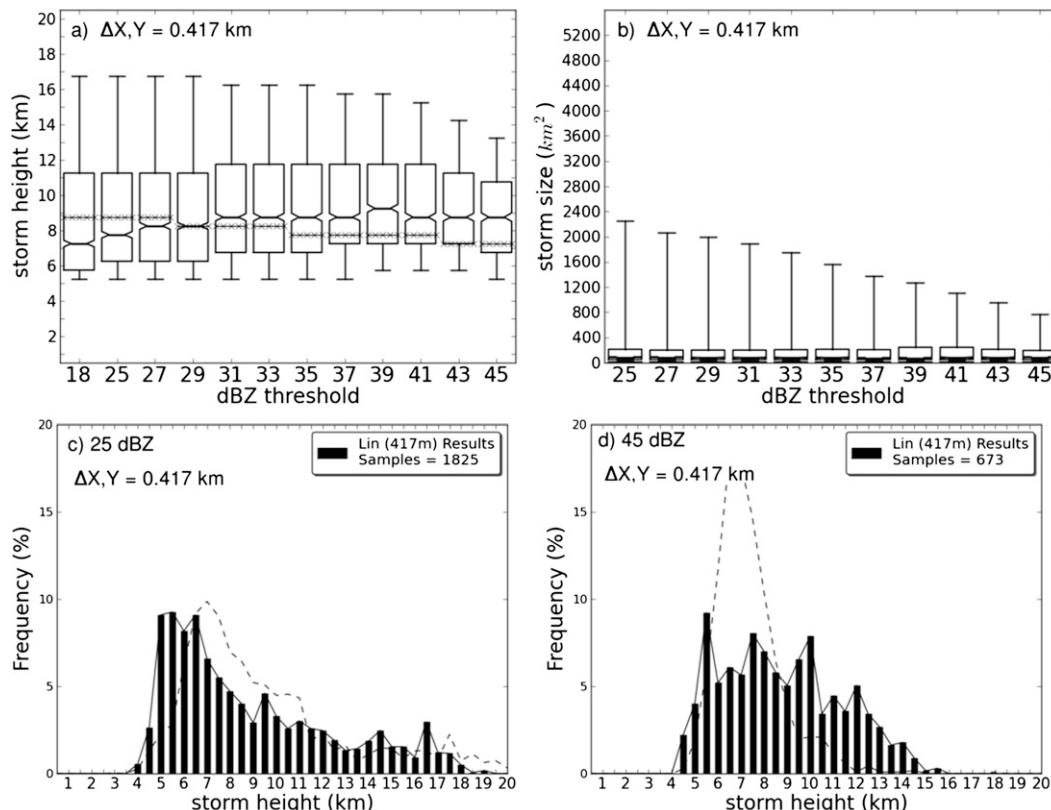


FIG. 13. (a) The cell-top height and (b) storm size statistics as a function of reflectivity threshold for the five-domain (0.417-km grid spacing) Lin simulation. For comparison the observed median cell-top height values are shown by the crosses in (a). Histograms of cell-top height when the thresholds of (c) 25 and (d) 45 dBZ are used to define the storms in the five domain (0.417-km grid spacing) simulation. For comparison the observed cell-top height distributions are shown by the dashed line.

In summary, these higher-resolution simulation results suggest that some of the errors in the height and size of the convective cells can be remedied with increased model resolution. However, even at grid spacings as small as 0.417 km, entrainment processes are not properly represented by the numerical model (e.g., Bryan et al. 2003; Bryan and Morrison 2012) and even higher model resolution is required. Such increases in resolution are beyond our current capabilities but will be the subject of continuing research.

#### 4. Summary and conclusions

This study has presented a method for comparing convection-permitting model simulations to radar observations using an innovative object-based approach. The method uses the automated cell-tracking algorithm TITAN to identify individual convective cells and determine their properties like depth, size, longevity, and shape. Cell properties are identified in the same way for model and radar data, facilitating comparison of their

statistical distributions. Examining the statistics of the cloud populations alleviates the difficulties caused by errors in timing and location of the simulated storms, and is an improvement over some other statistical methods because it retains information about the spatial characteristics of discrete convective elements. Moreover, the method allows properties of high-reflectivity regions (strong convective cores) and low-reflectivity regions (including stratiform clouds) to be evaluated separately.

To demonstrate the capabilities of the method, it was applied to a simulation of tropical convection during TWP-ICE, an experiment based around Darwin, Australia. The 4.5-day period used for the analysis occurred during a monsoon break regime, which is characterized by diurnally forced localized convective systems. A nested version of the ARW-WRF Model was used and comparisons of clouds simulated on its 1.25-km horizontal grid spacing domain were made with observations from Darwin's Gunn Point radar. Additional simulations with different microphysics and higher model resolution were also conducted.

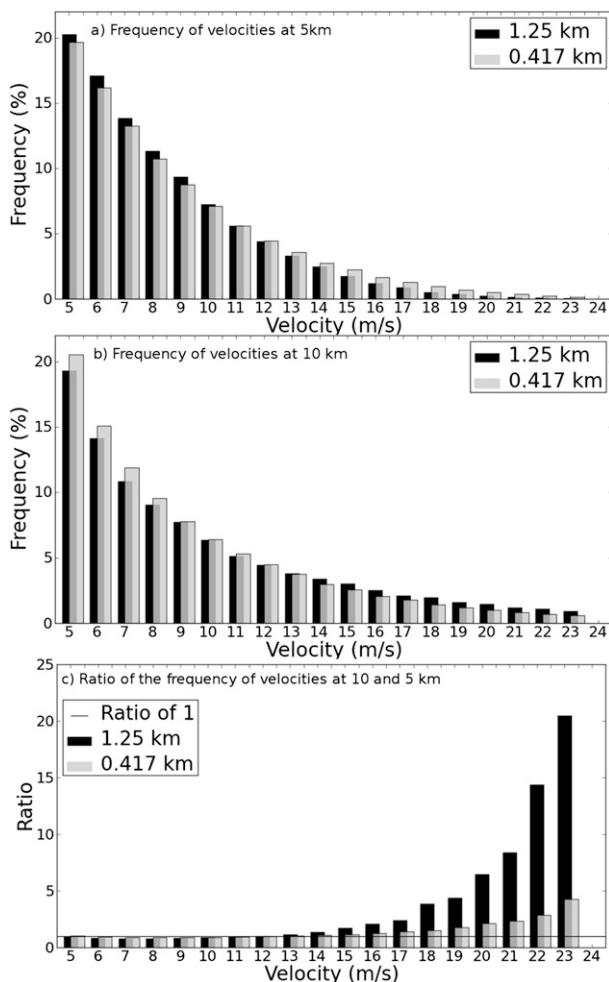


FIG. 14. Frequency of velocities for the 1.25- (black) and 0.417-km (gray) Lin simulations at (a) 5 and (b) 10 km. (c) The ratio of the frequency of velocities at 10 and 5 km are shown.

The comparisons between the model and the radar elucidated errors in the depth and size of convective cells. On average, simulated convective cells reached higher altitudes than the observations. Also, when using a low-reflectivity (25 dBZ) threshold to define convective cells the model underestimated the size of the largest convective cells. The most intense convective cores (reflectivity >45 dBZ) occurred over a narrow distribution of altitudes in the observations and the model produced a much broader distribution. Changing the model's microphysics parameterization from the Lin scheme to the Thompson scheme alleviated some of the problems exposed using a 25-dBZ threshold, yet did not improve the representation of the 45-dBZ cores. Moreover, increasing the model's horizontal resolution made a slight improvement to the height and size of the convective cores. These results are consistent with

a potential underestimation of entrainment in the 1.25-km resolution simulations.

This study has, however, compared modeling results to observations from only one radar, which restricts the geographic size of the comparison region. Computational resources also limited the size of our highest-resolution domains and the length of the period of interest. These limitations place some uncertainty on the comparisons of storms near the margins of the radar range/model domains, or complexes that extend over the entire radar range. While not shown here, similar analyses were attempted during the active monsoon region; TITAN analysis during the active monsoon using low reflectivity thresholds ( $\sim 25$  dBZ) were unusable because the reflectivity was often greater than this threshold everywhere. Nevertheless, this study has been able to demonstrate notable (and statistically significant) differences between the model and the observations during our period of interest. More importantly, statistically significant improvements were identified by a change in model configuration. The method could be readily extended to longer time periods and/or comparisons between simulations using a larger domain and observations from a network of radars with overlapping ranges.

In summary, we have presented a powerful method for evaluating model simulations, which is able to expose errors in the representation of convective structures in models with explicit representation of convective processes. The method is flexible and can be used to compare models with observations or utilized to evaluate changes in model configurations like resolution or physical parameterizations. Used in isolation, or in tandem with other established evaluation approaches, this convective cell-based approach provides an additional tool for assessing and improving deficiencies in cloud and modern numerical weather prediction models. Further application of this method to other models, scenarios, and more extensive datasets are planned for the future.

*Acknowledgments.* Thanks to Greg Thompson (NCAR) for providing his radar reflectivity conversion algorithm. The reflectivity conversion algorithm used to convert the Lin microphysics data to radar reflectivity was obtained from Read/Interpolate/Plot (RIP) graphics package. This work was funded in part by the Australian Research Council's Discovery (DP0770381) and Future Fellowships (FT0990892) programs, and by the Australian Research Council Centre for Excellence for Climate System Science (CE110001028). Computing facilities were provided by the Victorian Partnership for Advanced Computing (VPAC) and the National Computational Infrastructure facility (NCI). ECMWF ERA-Interim data used in this study were obtained from the ECMWF data

server. We thank the Editor (P. Harr) and the anonymous reviewers for their comments that helped improve the manuscript.

APPENDIX A

**Simulated Radar Reflectivity for the Lin Microphysics Scheme**

The algorithm used to convert the microphysical data for the Lin simulation was obtained from the data visualization package Read/Interpolate/Plot (RIP). The description of the reflectivity formulation for the Lin scheme below is based heavily on the work by Stoelinga (2005).

*a. Equivalent reflectivity factor for rain*

The equivalent reflectivity factor for rain  $Z_e$  is taken as the sixth moment of the size distribution. Thus, for a Marshall–Palmer distribution we have

$$Z_e = \int D^6 N_0 e^{-\lambda D} dD. \tag{A1}$$

Assuming the rain droplets are perfectly spherical and  $D \ll 1$  cm, after some calculation the equivalent reflectivity for rain becomes

$$Z_{e,r} = \frac{\Gamma(7)N_{0,r}}{\lambda^7}, \tag{A2}$$

where  $\Gamma(n)$  is the Euler gamma function, substituting Eq. (2) for  $\lambda$  and noting  $\Gamma(7) = 720$  we get

$$Z_{e,r} = \frac{720(\rho_{\text{air}}q_r)^{7/4}}{N_{0,r}^{3/4}(\pi\rho_r)^{7/4}}. \tag{A3}$$

Equation (A3) describes the equivalent reflectivity factor for rain in terms of known constants and the predicted mixing ratio of rain.

*b. Equivalent reflectivity factor for snow and graupel*

To obtain the equivalent reflectivity factors both snow and graupel, Eq. (A3) is modified with the assumption that the snow and graupel particles are spheres of solid ice. Equation (A4) relates the diameter that a solid ice sphere would have to obtain the same reflectivity as a particle of snow (or graupel) that does not have constant density:

$$D_{\text{solid}} = D_s \left( \frac{\rho_s}{\rho_i} \right)^{1/3}, \tag{A4}$$

where  $\rho_i$  is the density of ice and  $\rho_s$  is the density of snow.

Substituting Eq. (A4) into Eq. (A1) and including a factor that accounts for ice having a different reflective capacity than water we get

$$Z_{e,x} = \frac{720N_{0,x}}{\lambda^7} \left( \frac{\rho_x}{\rho_i} \right)^2 \left( \frac{|K|_i^2}{|K|_w^2} \right), \tag{A5}$$

where  $|K|_i^2$  is the dielectric constant for ice (0.176) and  $|K|_w^2$  is the dielectric constant for liquid water (0.93) and the subscript  $x$  represents either snow or graupel. Finally, we can use the following equation to rearrange Eq. (A5)

$$\left( \frac{\rho_x}{\rho_i} \right)^2 \left( \frac{|K|_i^2}{|K|_w^2} \right) = \left( \frac{\rho_x}{\rho_w} \right)^2 \left( \frac{\rho_w}{\rho_i} \right)^2 \left( \frac{|K|_i^2}{|K|_w^2} \right) = 0.224 \left( \frac{\rho_x}{\rho_w} \right)^2, \tag{A6}$$

where  $\rho_w$  is the density of liquid water (1000 kg m<sup>-3</sup>) and  $\rho_i$  is the density of ice (917 kg m<sup>-3</sup>).

Thus, the reflectivity factor for snow and graupel are represented by Eqs. (A7) and (A8), respectively:

$$Z_{e,s} = \frac{720(\rho_{\text{air}}q_s)^{7/4}}{N_s^{3/4}(\pi\rho_s)^{7/4}} \times 0.224 \left( \frac{\rho_s}{\rho_w} \right)^2; \tag{A7}$$

$$Z_{e,g} = \frac{720(\rho_{\text{air}}q_g)^{7/4}}{N_g^{3/4}(\pi\rho_g)^{7/4}} \times 0.224 \left( \frac{\rho_g}{\rho_w} \right)^2. \tag{A8}$$

Last, to take into account the brightband effect due to melting ice particles, when snow or graupel is found above 0°C, these frozen particles are assumed to have a liquid outer skin and scatter the radar beam like liquid water. Consequently the factor of 0.224 in Eqs. (A7) and (A8) is removed for snow and graupel particles at temperatures above 0°C.

The reflectivity factors for rain, snow, and graupel [Eqs. (A3), (A7), (A8)] can then be added to obtain the total reflectivity factor for each model grid cell. As reflectivity factors are typically measured in mm<sup>6</sup> m<sup>-3</sup> the total reflectivity factor is multiplied by 10<sup>18</sup> to convert it from m<sup>6</sup> m<sup>-3</sup> to the standard units used for radar data. Finally, to convert the reflectivity factor to the standard unit (dBZ), the logarithm of the total equivalent reflectivity factor is multiplied by 10:

$$Z_e(\text{in dBZ}) = 10 \log_{10}[Z_e(\text{in mm}^6 \text{ m}^{-3})]. \tag{A9}$$

## APPENDIX B

and

**Simulated Radar Reflectivity for the Thompson Microphysics Scheme**

The algorithm used to convert the microphysical data to simulated radar reflectivity was obtained from G. Thompson (2011, personal communication).

The description of how to convert output from the Thompson scheme to simulated radar reflectivity described below is based on the formulation of the Thompson microphysics scheme found in version 3.1.1 of the ARW-WRF Model. Note that the implementation of the Thompson scheme differs depending on which version of the ARW-WRF Model is used.

*a. Equivalent reflectivity factor for rain*

In the Thompson scheme particle-size distributions (except for snow) are represented as a generalized gamma function of the following form:

$$N(D) = N_0 D^\mu e^{(-\lambda D)}. \quad (\text{B1})$$

In the case of rain,  $\mu = 0$  and Eq. (B1) reduces to the Marshall–Palmer distribution shown in Eq. (1), and consequently Eq. (A2) can be used to determine the reflectivity factor for rain. However, unlike the Lin scheme, in version 3.1.1 of the ARW-WRF Model the Thompson scheme predicts the rain number concentration as well as the rain mixing ratio (i.e., it is double moment in rain). A general description and the formulation of a double-moment scheme like the one found in the ARW-WRF version 3.1.1 of the Thompson scheme can be found in Morrison and Pinto (2005), Morrison et al. (2009), and Morrison and Milbrandt (2011). What is relevant to the discussion here is that both  $N_{0,r}$  and  $\lambda_r$  are a function of both the mixing ratio and number concentration values ( $q_r$  and  $N_r$ , respectively), where

$$\lambda_r = \left[ \frac{c_r N_r \Gamma(\mu_r + d_r + 1)}{q_r \Gamma(\mu_r + 1)} \right]^{1/d_r}, \quad (\text{B2})$$

where  $c_r = (\pi \rho_r)/6$  and  $d_r = 3$  and

$$N_{0,r} = \frac{N_r \lambda_r^{\mu_r + 1}}{\Gamma(\mu_r + 1)}. \quad (\text{B3})$$

Recalling  $\mu = 0$  and substituting the above values for  $c_r$  and  $d_r$  we get

$$\lambda_r = \left[ \frac{\pi \rho_r N_r \Gamma(4)}{6 q_r \Gamma(1)} \right]^{1/3} \quad (\text{B4})$$

$$N_{0,r} = \frac{N_r \lambda_r}{\Gamma(1)}. \quad (\text{B5})$$

Finally, substituting Eqs. (B4) and (B5) into Eq. (A2) yields

$$Z_{e,r} = \left( \frac{6}{\pi} \right)^2 \left( \frac{q_r^2}{N_r} \right) \frac{\Gamma(7) \rho_{\text{air}}}{[\rho_r \Gamma(4)]^2}. \quad (\text{B6})$$

*b. Equivalent reflectivity factor for graupel*

In the particle-size distribution function for graupel  $\mu = 0$  and so Eq. (B1) reduces to the classic Marshall–Palmer distribution once again. It is again assumed that the graupel particles are spheres of solid ice with constant density, and after including the factor that accounts for ice having a different reflective capacity than water we have

$$Z_{e,g} = \frac{\Gamma(7) N_{0,g} \left( \frac{\rho_g}{\rho_i} \right)^2 \left( \frac{|K_i|^2}{|K_w|^2} \right)}. \quad (\text{B7})$$

Here again  $|K_i|^2$  and  $|K_w|^2$  are the dielectric constants for ice and liquid water and  $|K_i|^2/|K_w|^2 = 0.189$ .

In version 3.1.1 of the ARW-WRF Model in the Thompson scheme  $\lambda_g$  is described by the following equation:

$$\lambda_g = \left( \frac{\pi N_{0,g} \rho_g}{6 \rho_{\text{air}} q_g} \right)^{1/4}, \quad (\text{B8})$$

which when substituted into Eq. (B7) yields

$$Z_{e,g} = \frac{720 (\rho_{\text{air}} q_g)^{7/4}}{N_{0,g}^{3/4} (\pi \rho_g)^{7/4}} \times 0.189 \left( \frac{\rho_g}{\rho_i} \right)^2. \quad (\text{B9})$$

The intercept parameter for graupel takes on a form similar to the intercept parameter for rain described by Thompson et al. (2008):

$$N_{0,g} = \left( \frac{N_1 - N_2}{2} \right) \tanh \left[ \frac{(q_{g0} - q_g)}{q_{g0}} \right] + \left( \frac{N_1 + N_2}{2} \right), \quad (\text{B10})$$

where  $N_1 = 1 \times 10^4$ ,  $N_2 = 5 \times 10^6$  and  $q_{g0} = 0.15 \times 10^{-3}$ .

*c. Equivalent reflectivity factor for snow*

In the Thompson scheme the particle-size distribution for snow is represented by the sum of a gamma and exponential distribution, more specifically

$$N(D) = \frac{M_4^2}{M_3^3} \left[ \kappa_0 e^{-\frac{M_2}{M_3} \Delta_0 D} + \kappa_1 \left( \frac{M_2 D}{M_3} \right)^{\mu_s} e^{-\frac{M_2}{M_3} \Delta_1 D} \right], \tag{B11}$$

where  $\kappa_0 = 490.6$ ,  $\kappa_1 = 17.46$ ,  $\Delta_0 = 20.78$ ,  $\Delta_1 = 3.29$ ,  $\mu_s = 0.6357$ , and  $M_n$  is the  $n$ th moment of the distribution given by

$$M_n = \int D^n N(D) dD. \tag{B12}$$

However, for efficiency and simplicity the Thompson scheme relies on the power-law relationship between moments of the particle-size distribution as a function of temperature, developed by Field et al. (2005):

$$M_n = a(n, T_c) M_2^{b(n, T_c)}, \tag{B13}$$

where

$$\log_{10} a(n, T_c) = a_1 + a_2 T_c + a_3 n + a_4 T_c n + a_5 T_c^2 + a_6 n^2 + a_7 T_c^2 n + a_8 T_c n^2 + a_9 T_c^3 + a_{10} n^3$$

and

$$b(n, T_c) = b_1 + b_2 T_c + b_3 n + b_4 T_c n + b_5 T_c^2 + b_6 n^2 + b_7 T_c^2 n + b_8 T_c n^2 + b_9 T_c^3 + b_{10} n^3.$$

The constants  $a_x$  and  $b_x$  are not relevant to the discussion here, and can be found in Field et al. (2005) and Thompson et al. (2008).

However, to use Eq. (B13), the second moment of the distribution must be precalculated and used as a reference. In this case  $M_2 = q_s \rho_{\text{air}} 0.069^{-1}$ .

In the Thompson scheme the assumption that snow particles are spherical is dropped, instead the Thompson scheme assumes that the mass of the particles is proportional to the square of the particle diameter. Consequently radar reflectivity (for snow) is taken as the fourth moment of the particle-size distribution, where

$$M_4 = \int D^4 N(D) dD \approx a(4, T_c) \left( \frac{q_s \rho_{\text{air}}}{0.069} \right)^{b(4, T_c)}. \tag{B14}$$

Once again and including the factor to account for the difference in reflectance between water and ice we have the following:

$$Z_{e,s} \approx 0.189 \left( \frac{\rho_s}{\rho_i} \right)^2 a(4, T_c) \left( \frac{q_s \rho_{\text{air}}}{0.069} \right)^{b(4, T_c)}. \tag{B15}$$

A description of the algorithm that takes into account brightband effect due to the melting ice particles in the algorithm can be found in Blahak (2007).

Finally, the reflectivity factors for rain, snow, and graupel [Eqs. (B6), (B9), (B15)] can then be added to obtain the total reflectivity factor for each model grid cell and Eq. (A9) is used to convert the reflectivity factor to the standard unit (dBZ).

## REFERENCES

- Anagnostou, E. N., and W. F. Krajewski, 1997: Simulation of radar reflectivity fields: Algorithm formulation and evaluation. *Water Resour. Res.*, **33**, 1419–1428, doi:10.1029/97WR00233.
- Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single Column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709, doi:10.1002/qj.49711247308.
- Blahak, U., 2007: RADAR\_MIE\_LM and RADAR\_MIELIB—Calculation of radar reflectivity from model output. Internal Rep., Institute for Meteorology and Climate Research, University/Research Center Karlsruhe, 150 pp.
- Blossey, P. N., C. S. Bretherton, J. Cetrone, and M. Kharoutdinov, 2007: Cloud-resolving model simulations of KWJEX: Model sensitivities and comparisons with satellite and radar observations. *J. Atmos. Sci.*, **64**, 1488–1508.
- Bringi, V. N., G.-J. Huang, V. Chandrasekar, and T. D. Keenan, 2001: An areal rainfall estimator using different propagation phase: Evaluation using a C-band radar and a dense gauge network in the tropics. *J. Atmos. Oceanic Technol.*, **18**, 1810–1818.
- Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225.
- , J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, **131**, 2394–2416.
- Caine, S., 2009: The classification and simulation of precipitating convective regimes over Darwin, Australia. Ph.D. thesis, School of Mathematical Sciences, Monash University, 165 pp.
- , C. Jakob, S. Siems, and P. May, 2009: Objective classification of precipitating convective regimes using a weather radar in Darwin, Australia. *Mon. Wea. Rev.*, **137**, 1585–1600.
- Chandrasekar, V., and V. N. Bringi, 1987: Simulation of radar reflectivity and surface measurements of rainfall. *J. Atmos. Oceanic Technol.*, **4**, 464–478.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.
- Chen, S., and W. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan*, **80**, 99–118.
- Chou, M., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo 104606, 85 pp.
- Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597.

- Dixon, M., and G. Wiener, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A radar-based methodology. *J. Atmos. Oceanic Technol.*, **10**, 785–797.
- Dupree, W. J., M. Robinson, R. DeLaura, and P. Bieringer, 2006: Echo tops forecast generation and evaluation of air traffic flow management needs in the national airspace system. Preprints, *12th Conf. on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Amer. Meteor. Soc., P1.13. [Available online at <https://ams.confex.com/ams/pdfpapers/103707.pdf>.]
- Field, P. R., R. J. Hogan, P. R. A. Brown, A. J. Illingworth, T. W. Chouarton, and R. J. Cotton, 2005: Parametrization of ice-particle size distributions for mid-latitude stratiform cloud. *Quart. J. Roy. Meteor. Soc.*, **131**, 1997–2017, doi:10.1256/qj.04.134.
- Fridlind, A. M., and Coauthors, 2012: A comparison of TWP-ICE observational data with cloud-resolving model results. *J. Geophys. Res.*, **117**, D05204, doi:10.1029/2011JD016595.
- Gilleland, E., D. A. Ahijevych, B. G. Brown, and E. E. Ebert, 2010: Verifying forecasts spatially. *Bull. Amer. Meteor. Soc.*, **91**, 1365–1373.
- Heymsfield, G. M., B. Geerts, and L. Tian, 2000: TRMM precipitation radar reflectivity profiles as compared with high-resolution airborne and ground-based radar measurements. *J. Appl. Meteor.*, **39**, 2080–2102.
- Horvath, A., F. Acs, and A. T. Seres, 2008: Thunderstorm climatology analyses in Hungary using radar observations. *Quart. J. Hung. Meteor. Serv.*, **112**, 1–13.
- Illingworth, A. J., and Coauthors, 2007: CLOUDNET: Continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bull. Amer. Meteor. Soc.*, **88**, 883–898.
- Janjić, Z. I., 2001: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.
- Keenan, T. D., and R. E. Carbone, 1992: A preliminary morphology of precipitation systems in tropical northern Australia. *Quart. J. Roy. Meteor. Soc.*, **118**, 283–326.
- , K. Glasson, F. Cummings, T. S. Bird, J. Keeler, and J. Lutz, 1998: The BMRC/NCAR C-band polarimetric (C-POL) radar system. *J. Atmos. Oceanic Technol.*, **15**, 871–886.
- Koch, S. E., B. Ferrier, M. T. Stoelinga, E. Szoke, S. J. Weiss, and J. S. Kain, 2005: The use of simulated radar reflectivity fields in the diagnosis of mesoscale phenomena from high-resolution WRF model forecasts. Preprints, *11th Conf. on Mesoscale Processes*, Albuquerque, NM, Amer. Meteor. Soc., J4J.7. [Available online at <https://ams.confex.com/ams/pdfpapers/97032.pdf>.]
- Krauss, T. W., and J. R. Santos, 2004: Exploratory analysis of the effect of hail suppression operations on precipitation in Alberta. *Atmos. Res.*, **71**, 35–50.
- Lang, S., W. Tao, R. Cifelli, W. Olson, J. Halverson, S. Rutledge, and J. Simpson, 2007: Improving simulations of convective systems from TRMM LBA: Easterly and westerly regimes. *J. Atmos. Sci.*, **64**, 1141–1164.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.
- Magagi, R., and A. P. Barros, 2004: Estimation of latent heating of rainfall during the onset of the Indian monsoon using TRMM PR and radiosonde data. *J. Appl. Meteor.*, **43**, 328–349.
- May, P. T., and A. Ballinger, 2007: The statistical characteristics of convective cells in a monsoon regime (Darwin, northern Australia). *Mon. Wea. Rev.*, **135**, 82–92.
- , and T. P. Lane, 2009: A method for using weather radar data to test cloud resolving models. *Meteor. Appl.*, **16**, 425–432, doi:10.1002/met.150.
- , J. H. Mather, G. Vaughan, C. Jakob, G. M. McFarquhar, K. Bower, and G. G. Mace, 2008: The Tropical Warm Pool International Cloud Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 629–645.
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model from geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682.
- Morrison, H., and J. O. Pinto, 2005: Mesoscale modeling of springtime arctic mixed-phase stratiform clouds using a new two-moment bulk microphysics scheme. *J. Atmos. Sci.*, **62**, 3683–3704.
- , and J. Milbrandt, 2011: Comparison of two-moment bulk microphysics schemes in idealized supercell thunderstorm simulations. *Mon. Wea. Rev.*, **139**, 1103–1130.
- , G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007.
- NCAR, cited 2012: What is TITAN? [Available online at [http://www.ral.ucar.edu/projects/titan/home/whatis\\_titan.php](http://www.ral.ucar.edu/projects/titan/home/whatis_titan.php).]
- Pinto, J., C. Phillips, M. Steiner, R. Rasmussen, N. Oien, M. Dixon, W. Wang, and M. Weisman, 2007: Assessment of the statistical characteristics of thunderstorms simulated with the WRF model using convection-permitting resolution. Preprints, *33rd Conf. on Radar Meteorology*, Cairns, Australia, Amer. Meteor. Soc., 5.5. [Available online at <https://ams.confex.com/ams/pdfpapers/123712.pdf>.]
- Potts, R., T. Keenan, and P. May, 2000: Radar characteristics of storms in the Sydney area. *Mon. Wea. Rev.*, **128**, 3308–3319.
- Rogers, R. F., M. L. Black, S. S. Chen, and R. A. Black, 2007: An evaluation of microphysics fields from mesoscale model simulations of tropical cycles. Part I: Comparison with observations. *J. Atmos. Sci.*, **64**, 1811–1834.
- Rutledge, S. A., and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cloud-frontal rainbands. *J. Atmos. Sci.*, **41**, 2949–2972.
- Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp. [Available online at [http://www.mmm.ucar.edu/wrf/users/docs/arw\\_v3\\_bw.pdf](http://www.mmm.ucar.edu/wrf/users/docs/arw_v3_bw.pdf).]
- Stoelinga, M. T., cited 2005: Simulated equivalent reflectivity factor as currently formulated in RIP: Description and possible improvements. [Available online at [http://www.atmos.washington.edu/~stoeling/RIP\\_sim\\_ref.pdf](http://www.atmos.washington.edu/~stoeling/RIP_sim_ref.pdf).]
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542.
- , P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115.
- Vallgren, A., 2006: Statistical characteristics of convective storms in Darwin, northern Australia. M.S. thesis, Dept. of Applied

- Physics and Mechanical Engineering, Luleå University of Technology, 70 pp. [Available online at <http://epubl.ltu.se/1402-1617/2006/176/LTU-EX-06176-SE.pdf>.]
- Van Weverberg, K., N. P. van Lipzig, and L. Delobbe, 2011: Evaluation of moist processes during intense precipitation in km-scale NWP models using remote sensing and in-situ data: Impact of microphysics size distribution assumptions. *Atmos. Res.*, **99**, 15–38.
- Varble, A., and Coauthors, 2011: Evaluation of cloud-resolving model intercomparison simulations using TWP-ICE observations: Precipitation and cloud structure. *J. Geophys. Res.*, **116**, D12206, doi:10.1029/2010JD015180.
- Wang, W. and X. Liu, 2009: Evaluating deep updraft formulation in NCAR CAM3 with high-resolution WRF simulations during ARM TWP-ICE. *Geophys. Res. Lett.*, **36**, L04701, doi:10.1029/2008GL036692.
- Wang, Y., C. N. Long, L. R. Leung, J. Dudhia, S. A. McFarlane, J. H. Mather, S. J. Ghan, and X. Liu, 2009: Evaluating regional cloud-permitting simulations of the WRF model for the Tropical Warm Pool International Cloud Experiment (TWP-ICE), Darwin, 2006. *J. Geophys. Res.*, **114**, D21203, doi:10.1029/2009JD012729.
- Wapler, K., and T. P. Lane, 2012: A case of offshore convective initiation by interacting land breezes near Darwin, Australia. *Meteor. Atmos. Phys.*, **115**, 123–137.
- , —, P. T. May, C. Jakob, M. J. Manton, and S. T. Siems, 2010: Cloud-system-resolving model simulations of tropical cloud systems observed during the Tropical Warm Pool–International Cloud Experiment. *Mon. Wea. Rev.*, **138**, 55–73.
- Wilcox, E. M., 2003: Spatial and temporal scales of precipitating tropical cloud systems in satellite imagery and the NCAR CCM3. *J. Climate*, **16**, 3545–3559.
- Wu, J., A. D. Del Genio, M.-S. Yao, and A. B. Wolf, 2009: WRF and GISS SCM simulations of convective updraft properties during TWP-ICE. *J. Geophys. Res.*, **114**, D04206, doi:10.1029/2008JD010851.
- Zhang, G. J., D. Zurovac-Jevtic, and E. R. Boer, 1999: Spatial characteristics of the tropical cloud systems: Comparisons between model simulation and satellite observations. *Tellus*, **51A**, 922–936.