POLARIMETRIC RADAR CONVECTIVE CELL TRACKING REVEALS LARGE SENSITIVITY OF CLOUD PRECIPITATION AND ELECTRIFICATION PROPERTIES TO CCN

A Dissertation

by

JIAXI HU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Renyi Zhang
Co-Chair of Committee,	Daniel Rosenfeld
Committee Members,	Jianhua Huang
	Robert Korty
Head of Department,	Ping Yang

August 2018

Major Subject: Atmospheric Sciences

ABSTRACT

Hypotheses have been proposed for decades about cloud condensation nuclei (CCN) aerosol effect on delaying the warm rain process, invigorating deep convective cloud vertical development, and enhancing mixed-phase process. Observational support has been only qualitative with mixed results due to lack of regional measurements of CCN, while simulations have not produced a robust consensus. Quantitative assessment of these relationships became possible with the advent of CCN retrievals from satellites; when combined with measurements by polarimetric radar and Lightning Mapping Array (LMA), tracking convective cells observed on radar and examining precipitation properties throughout the cells' life cycle has permitted the study of the impact of CCNs on cloud and precipitation characteristics.

We composited more than 2000 well-tracked cells in the Houston region and stratified them by CCN, convective available potential energy (CAPE) and urban/rural locations. The analyzed cell properties include reflectivity (Z), differential reflectivity (Z_{DR}) and LMA data. The results show that added CCN to deep convective clouds delays the initiation of precipitation by up to 20 minutes. Added CCN invigorate the convection until saturation near CCN = 1000 cm^{-3}; increasing CCN from ~400 to an optimum of ~1000 cm^{-3} increases lightning activity by an order of magnitude. A further increase of CCN decreases lightning rates. Adding CAPE enhances lightning only under low CCN <500 cm^{-3}. Urban area enhances lightning for the same CCN only under low CCN conditions. Urban heat island cannot explain this observation. In summary, CAPE is

essential for the initiation of deep convection. It has been believed that CAPE and lightning are positively related. This is indeed the case when CAPE is low. But when CAPE is high, which means that deep convection is already in progress, aerosols dominate the lightning activity. These insights lead to refinement of the physical hypotheses which provide impetus for a field campaign in the Houston area.

DEDICATION

This dissertation is specially dedicated to my beloved grandpa, Xianyuan Guo who deceased in 2018.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my committee chair, Dr. Renyi Zhang, for his guidance throughout the course of this research. I would also express my deepest thanks to my co-chair, Dr. Daniel Rosenfeld, for all the knowledge and unselfish support throughout my Ph.D. years. My committee members, Dr. Jianhua Huang, Dr. Robert Korty are also highly appreciated for their valuable comments and constructive advises during my research study.

I want to thank my colleagues, the department faculty and staffs for making my time at Texas A&M University a great experience. Special thanks go to Dr. Alexzander Ryzhkov, Dr. Pengfei Zhang, Dr. Dusan Zrnic and Dr. Jeffery Snyder from the National Severe Storm Laboratory (NSSL), who offered me substantial helps during my visit to NSSL. I also benefit a lot from the collaboration with Early Williams from MIT to analyze the results from my cloud cell tracking result. I would also thanks Richard Weitz for the data support of the Houston lightning mapping array.

I also want to extend my gratitude to the National Science Foundation and Dr. Richard Orville to provided me a three-year funding support.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professor Renyi Zhang [advisor], Daniel Rosenfeld [co-advisor] and Robert Korty of the Department of Atmospheric Sciences and Professor Jianhua Huang of the Department of Statistics.

All other work conducted for the dissertation was completed by the student independently.

This work was made possible in part by National Science Foundation under Grant Number 446021. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Science Foundation.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

Page

Page

1. INTRODUCTION

Until now, only global statistics have been presented to qualitatively address the relationships between aerosols, convective available potential energy (CAPE), and lightning. Aerosols can have profound impacts on cloud microphysics, precipitation formation and cloud electrification. Cloud condensation nuclei (CCN) nucleate more numerous and smaller cloud droplets that are slower to coalesce into raindrops (Gunn and Phillips, 1957). On the other hand, cloud water that was not converted to rain is lifted to above the 0° C level and converted into ice hydrometeors. The phase change from liquid to solid releases latent heat of freezing that invigorates convection and intensifies the precipitation (Rosenfeld et al., 2008). Enhanced mixed-phase process promotes cloud electrification (Latham, 1981), because charging occurs when graupel and ice crystals collide in supercooled water clouds, and the strong updrafts separate vertically the oppositely charged graupel and ice crystals (Reynolds et al., 1957; Takahashi, 1978; Williams, 1988). Therefore, the role of CCN in enhancing mixed phase processes and updrafts was also proposed to enhance cloud electrification (Rosenfeld and Lensky, 1998; Williams et al., 2002). Further support for the role of aerosols in cloud invigoration and electrification has been provided by a number of other studies (Altaratz et al., 2014; Andreae et al., 2004; Khain et al., 2005; Khain et al., 2008; Rosenfeld et al., 2014; Rosenfeld et al., 2008; Tao et al., 2012; Tao et al., 2007; van den Heever et al., 2006; van den Heever et al., 2011), while other studies supported the role of CAPE (Bang and Zipser, 2016; Blanchard, 1998; Emanuel, 1997; Williams et al., 2002; Williams et al., 1992;

Williams and Satori, 2004). A recent study showed strong evidence that increased aerosol in a remote marine environment resulting from ship emissions can more than double the amount of lightning over the shipping lanes compared to the background for the same CAPE (Thornton et al., 2017). An extensive study of convective features based on the Topical Rainfall Measuring Mission showed an equal weight for CAPE and aerosol optical depth (Stolz et al., 2015) on cloud electrification processes on a global data base.

The purpose of this study is to go further and look into individual convective cells and the way that CAPE and CCNs affect their time-height evolution of dynamics, hydrometeors and the resulting electrification. This is being done while controlling the differences in thermodynamic conditions, land/ocean contrast, and urban/rural area. The measurements included a composite of Next-Generation Radar (NEXRAD) radar, Suomi NPP satellite observation, lightning mapping array (LMA) and reanalysis sounding data around the Houston region, which allows accurate convective cloud sampling and selection to understand aerosol effect under various conditions.

Here we examine the impact of an aerosol perturbation incurred by Houston (Fig. 1.1) on the observed CCN, cloud microstructure, evolution of hydrometeors and cloud electrification by using a novel suite of tools developed for this objective. Houston is selected because it is a major source of anthropogenic aerosols on a relatively clean background in a climatic regime of deep tropical convective clouds during summer. The Houston area and the vicinity are covered well by a polarimetric NEXRAD radar and LMA (Fig.1.1). Houston is a major metropolis (population ~2.2 million) with an extensive petroleum industry; Houston hosts 14% of the refinery capacity of the USA (Corporation,

2018) Houston, located by the Texas Gulf coast, experiences various conditions that include oceanic clean, continental pristine and continental polluted air masses. These conditions are well presented within the coverage of a single NEXRAD radar (KHGX Houston/Galveston radar site) and the Houston LMA detection range.

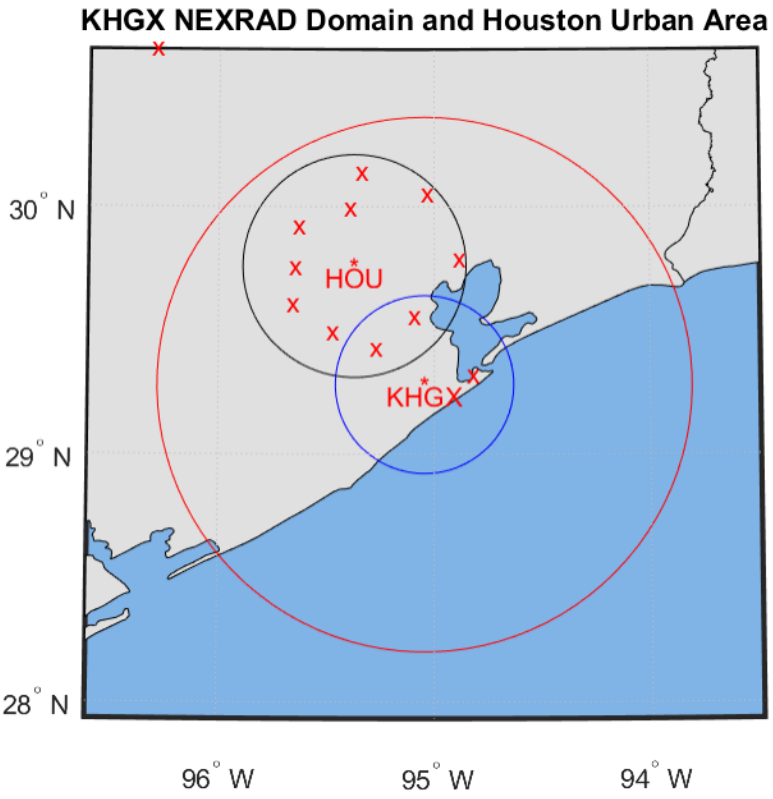

Figure 1.1 The domain of this study. KHGX represents the location of Houston/ Galveston radar. The red crosses represent the 12 Houston LMA sites used in this study. Also shown are the range circles of 40 (blue circle) and 120 km (red circle) from the radar, which were the lower and upper bounds for selecting the tracked cells. The Houston urban area is shown by the 50 km radius black circle.

Algorithms that identify and track storms have been used to study the evolution of storms in various scenarios (Dixon and Wiener, 1993) and weather forecasting,

especially severe weather detection and short-term forecasting (Joe et al., 2004; Wilson et al., 2004). There are two well-known types of tracking algorithm, the cross-correlation method (Li et al., 1995; Rinehart and Garvey, 1978; Tuttle and Foote, 1990) and the centroid method (Bellon et al., 1980; Bjerkaas and Forsyth, 1979; Crane, 1979; Dixon and Wiener, 1993; Handwerker, 2002; Johnson et al., 1998). The strength of cross-correlation tracking algorithms is better identification of the shifting reflectivity echoes' speed and direction at the expense of inability to identify and track isolated cloud cells. The centroid tracking can identify isolated cells and provide temporal properties of storms (Johnson et al., 1998). But because centroid tracking algorithms depend on threshold values to identify storm cells, they can't capture the initial and ending stage of cloud cells and are inappropriate to identify and track stratiform clouds. The thresholding techniques can also make incorrect identification of cloud cells if there are multiple cells with similar strength within one cloud cluster.

An algorithm first developed by (Rosenfeld, 1987) for the study of cells that are embedded in convective rain systems is neither a centroid nor a cross-correlation tracking algorithm. The algorithm uses watershed principles to identify cloud cells by the outline of radar echoes and uses the largest common rain volume rate (measured in cubic meters of rain water which falls through the lowest elevation scan or through cloud base level [m^3/hr]) to track cells. This method overcomes the disadvantages of the centroid tracking thresholding and is able to identify cloud cells for both isolated and clustered cloud systems.

The goal of the MCIT algorithm is to create an automatic tool to identify and track cells from their first echo whether these are isolated or clustered, and to provide cloud cell lifecycle properties including polarimetric variables, lightning information, thermodynamic data. Of specific interest for future studies is to determine the influence of aerosol CCNs concentration and cloud microstructure retrieved from satellite data on the morphology, microphysical evolution and electrification of ordinary storms.

Detailed output datasets are described in stage 4 of the Methodology section, which addresses the assignment of cell lifetime and instantaneous properties.

The Multi-Cell Identification and Tracking (MCIT) algorithm (Hu et al., 2018) package is used in this study as an automatic tool to identify and track convective precipitation cells from their first radar echo until their loss of identity during their dissipation. The MCIT provides the cells' lifecycle properties including polarimetric quantities (NEXRAD data), lightning information (LMA source count) and thermodynamic data (sounding data). Here we quantify the influence of satellite-retrieved CCN concentration (Rosenfeld et al., 2016) on the morphology, microphysical evolution and electrification of ordinary convective storms. Each cloud cell is ascribed with relevant so-called "explaining" and "dependent" parameters from aforementioned data sources. The explaining parameters include cloud-base CCN concentrations (will use CCN for short in the paper), super saturation (SS), CAPE, cloud base height (CBH), and two location descriptions (one of Land/Coastal/Ocean and one of Urban/Rural). The dependent parameters include reflectivity (Z), differential reflectivity (Z_{DR}), LMA source count and additional parameters, as explained in the methodology part.

2. MULTI-CELL IDENTIFICATION AND TRACKING ALGORITHM

This section explains the MCIT algorithm in detail. It includes data pre-processing, cloud cell identification, cloud cell tracking and post processing. Specifically, the parameters and the differences between the MCIT algorithm and the algorithm in Rosenfeld (1987) are described.

2.1 Methodology

2.1.1 Definition of a cell

A convective cell is constituted by an updraft in which most of the condensation occurs in a vertical column. The cloud water is subsequently converted into precipitation; it falls through the updraft or replaces it with a downdraft leading to the demise of the convective cell. Therefore, a vertically-integrated column of radar-retrieved precipitation represented by VIL is a good marker for the core of a convective cell. VIL can be used as a quick and effective indicator of strong precipitation cells (Greene and Clark, 1972). VIL is calculated with a fixed Z-R relationship from the WDSS II software package developed by the National Severe Storm Laboratory (Lakshmanan et al., 2007). A cell peak is defined as a local maximum of VIL within a radar echo with at least 5 pixels (pixel resolution is 0.5 km * 0.5 km * 0.3 km). Coordination and resolution detail are discussed in STAGE 1 detailed in part b (*The MCIT algorithm*). The VIL map is smoothed by a mean filter over a rectangle of size 3-by-3 pixels. Basically, pixel i is determined by the mean of the rectangle centered on i. If the pixel i has no detecTable radar echo (NaN), then it will be preserved as NaN. At the edges of the VIL map, if the available pixels are not sufficient

to build a 3-by-3 rectangle, the smoothing program uses as many pixels available as possible. Two neighboring cells may appear to overlap but are still considered distinct as long as there are two VIL local maxima, which are separated by a valley at least 2 dB lower than the weaker peak of the two. If the valley between the two peaks is shallower than 2 dB, the weaker cell merges into the stronger one.

2.1.2 The MCIT algorithms

The purpose of the MCIT algorithm is to identify and track multiple cloud cells and save the time-height cross section of various radar variables and lightning properties throughout their lifecycle. Figure 2.1 is a flow chart of the MCIT algorithm. It includes four phases of the cell-tracking analysis. A detailed description of the four MCIT stages follows:

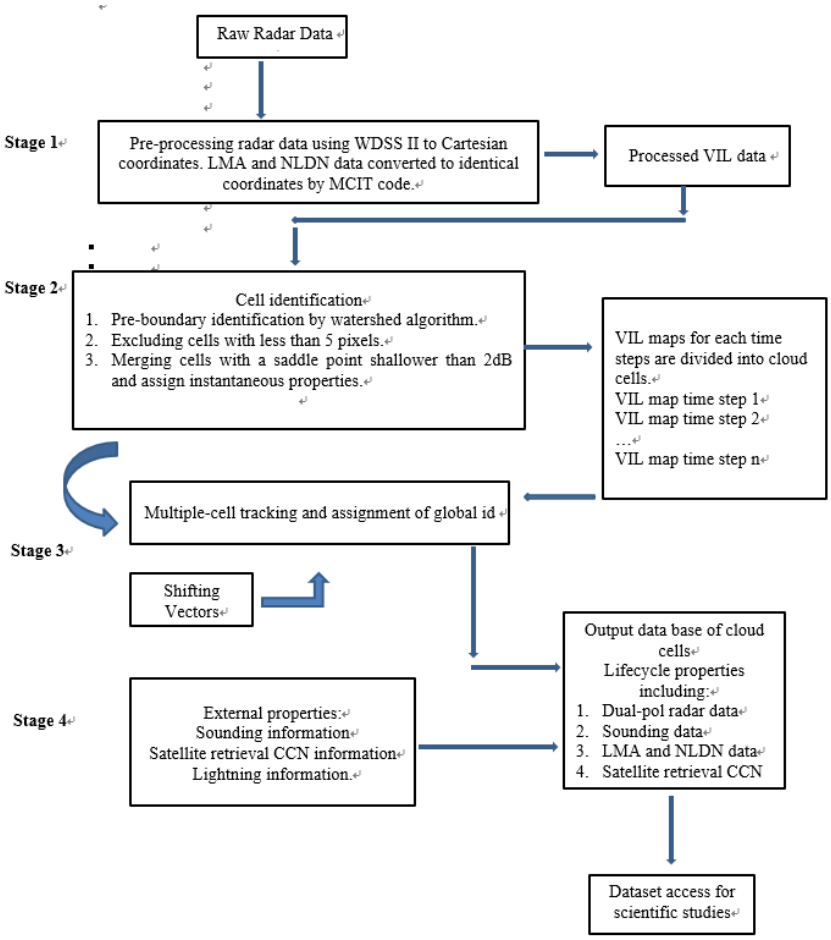

Figure 2.1 Flowchart of MCIT algorithm.

2.1.2.1 Stage 1: Pre-Processing Radar and Lightning Data

The first stage is to process Level 2 radar data, LMA and NLDN data sets into Cartesian coordinates. The radar data for each case is from a single WSR-88D radar closest to the storm.

The MCIT algorithm requires pre-processed 2-D VIL data and ERA interim 6-hourly wind field reanalysis as input. The WDSS II software package from the National Severe Storms Lab (NSSL) was used to pre-process raw Level 2 radar data into Cartesian

coordinates and calculate VIL. The variables that are converted to Cartesian coordinates are reflectivity (Z), differential reflectivity (Z_{DR}), specific differential phase (K_{DP}), copolar correlation coefficient (ρ_{hv}) and VIL from the WDSS II package. The temporal resolution equals the volume scan time, which is about 5 minutes, so that short-lived cells can be tracked. Vertical resolution better than 1 km (Rosenfeld 1987) is desirable. But this cannot be achieved everywhere within the radar umbrella because of the beam width and spacing of elevation angles. In this study data are spaced 300 m in the vertical. The horizontal resolution should be less than the meso-gamma scale (2 km) (Orlanski 1975) and this study data are spaced at 500 m. The range interval between 40 and 120 km from the radar is best for studies of cell properties.

The hail size discrimination algorithm (HSDA) developed by the NSSL is used to provide hydrometeors classifications based on Z, $Z_{DR,}$ and ρ_{hv} along with melting-level height within a fuzzy-logic scheme (Ortega et al., 2016; Park et al., 2009). The classification and abbreviations are as follows: No Echo (NE), Light/Moderate Rain (L/MR), HR (Heavy Rain), Rain/Hail (R/Ha), Big Drops (BD), Anomalous Propagation (AP), Biological Objects (BI), Unknown Hydrometeors Classification (UK), Dry Snow (DS). Wet Snow (WS), Ice Crystal (CY), Graupel (GR), Large Hail (LH), Giant Hail (GH). A detailed description of HSDA and each class of hydrometeors can be found in *(Ortega et al., 2016).*

LMA and NLDN data are also converted to Cartesian coordinates and put on the same grid as the radar data. Note that both LMA and NLDN data sets have much higher spatial and temporal resolution than the radar data.

2.1.2.2 Stage 2: Cell Identification and Instantaneous Property Assignment

The second phase of the MCIT algorithm uses the calculated VIL data, one scan at a time, as follows:

(i) The program uses the watershed algorithm (Meyer, 1994) to identify the boundary of local maximum VIL clusters with at least 5 pixels. The criterion for cells to be larger than 5 pixels (1.25 km^2) is to avoid noise and over-identification of cells. Each local maximum VIL cluster is assigned a local ID number.

(ii) Local maxima of VIL are considered as peaks of cells. All neighboring cells are tested for potential merging if the depth of the saddle point between them is shallower than 2 dB. The 2 dB threshold is a parameter in the MCIT algorithm and can be easily changed to fit different research needs.

(iii) The watershed algorithm ascribes pixels to each peak by locating watershed ridgelines in the VIL field. The relatively small VIL magnitude pixels represent low elevations and relatively large VIL magnitude pixels represent high elevations. The sign of the VIL is reversed before the watershed algorithm is applied to it. The reason for reversing the sign is that watershed algorithm assumes that water flows from high to low elevations. Therefore, the peaks get the lowest values when multiplied by -1.

(iv) Each watershed region is assigned a local ID number when it also satisfies size criteria from previous description (in this study the region must be with at least 5 pixels).

(v) For each local cell with local ID j, the program searches for peaks of other cells within a radius twice the cell's longest dimension. The local cell ID of the pixels along the lines between the peak of cell j and the peaks of the other cells around it ($k_1, k_2, \ldots k_n$) are examined. If there are pixels not assigned with local ID numbers of j or k_i between the peak of cell j and cell k_i, then cell j and cell k_j are not neighbors. Only when the pixels between the peak of cell j and cell k_i consist of local ID numbers of j and k_i exclusively, can cells j and k_i be treated as neighbors and then tested for merging by the 2 dB comparison. After searching all the neighboring peaks and after the merging process, all the pixels with valid VIL values have been assigned to cells with a unique local ID number.

(vi) For every local ID number that remained after merging, the program calculates and saves its instantaneous properties in a data structure for later use. The lightning properties that are added to the cell properties are:

1) LMA events count.

2) NLDN total event counts.

3) NLDN positive event counts.

4) NLDN negative event counts.

The program also saves the horizontal map of VIL and the identification local ID from the current time step. These VIL and ID maps are the input for PHASE 3.

2.1.2.3 Stage 3: Multi-Cell Tracking

The third phase of the MCIT algorithm is to track cells by comparing two consecutive VIL maps from PHASE 2. The temporal resolution is about 5 minutes for VIL maps, which is the time between radar volume scans for VIL maps. The assumption is that the duration of the targeted cells for tracking is longer than 5 minutes. A tracking error may occur only if a targeted cell dissipates between two tracking steps and a newly formed cell emerges around the location of the targeted cell between these two tracking steps. The false track is easy to tell by the sharp change of reflectivity and height in the cell area.

The cell tracking flow chart can be found in Rosenfeld (1987). The tracking algorithm summary and difference from Rosenfeld (1987) are as follows:

(i) Comparison of a pair of sequential VIL maps. This is done after shifting the VIL map at time step n+1 back to account for the mean cell motions between the time of maps n and n+1. The shifting vector is updated for each cell at each time step based on the remaining distance between the cell peaks in the two maps after the shift back. The detailed description of the shifting vector is discussed in STAGE 3. The program keeps track of all the overlapping cells. One cell from map n+1 can overlap with more than one cell from map n. This may be a result of merging of cells in map n. On the other hand, one cell in map n can overlap with more than one cell from map n+1. This can be a result of the cell splitting in the map n to several cells in the map n+1. The most important quantity for tracking is the overlapping or common integrated VIL. For each pair of overlapped cells between time step n and n+1, their common VIL values are integrated and saved.

(ii) This step does the actual tracking and is based on the logic in Rosenfeld (1987), with the difference that the present method uses the common VIL and VIL peaks instead of a common rain volume and reflectivity peaks. The program searches for all the overlapped cells at time step n+1 relative to targeted cell(i,n), i.e., cell i at time step n. The criteria to identify one of the cells at time step n+1 to be the continuation of cell(i,n) are:

1) The integrated common VIL is at least 50% of the total VIL of the smaller of the two compared cells.
2) The VIL peaks of both old and new cells are inside the common area of the two compared cells.
3) If the previous condition (peaks of old and new located within the common area of the two cells) is not fulfilled it is required that the integrated common VIL must be more than 75% of the total VIL of the smaller of the two compared cells.
4) If none of the conditions above is fulfilled, but the integrated common VIL is at least 25% of the total VIL of the smaller cell, then the cell from time step n can either be a merger or a split.

According to many tests and experiments, the new tracking code performs similarly the original (Rosenfeld 1987) in that the tracking results are not sensitive to the percentage limits of integrated common VIL. The percentage numbers are usually close to 100% or smaller than 50%.

There are four possible situations for candidate cells to emerge at time step n+1: simple continuation, simple birth, split and merger. A simple continuation status

occurs when cell(n,i) is the only source to cell(n+1,i) and cell(n+1,i) is the only one source for cell(n,i). A simple birth situation is when cell(n+1,i) has no source cell from the previous time step. If cell(n,i) has more than one continuation cell, only the one with maximum integrated common VIL gets the identity of cell(n,i) and all the rest of candidate continuation cells are labeled splits. A merger situation is similar to splits; cell(n+1,i) has more than one source cell and therefore is labeled as a merger.

At this stage of the MCIT algorithm, each cloud cell has been identified and tracked. A global ID number is assigned to each cell. Unlike the local ID numbers, the global ID number is a unique number assigned to each cell through its lifetime and is used as the identifier to assign each cell lifetime and instantaneous properties.

Two types of tracking examples are presented here to show the performance of the MCIT algorithm, i.e. isolated tracking and clustered tracking.

An isolated tracking example (Figure 2.2) is shown for June 8, 2013, KHGX (Houston) radar. The Figure contains both the VIL panels (Figure 2.2a - f) and the corresponding tracking panels (Figure 2.2g – l). The extracted tracking steps are from 19:51:04 UTC to 20:17:00 UTC in the northeast of Houston area. The cells included in Figure 2.2 are about 70 to 90 km from the KHGX radar site. Each color and the related black number in Figure 2.2g -l links to a unique cell. By examining each VIL map and comparing with the corresponding partition map, cells are well identified. The general direction of motion of these maps is from the west to east. Cell 1, 2, 3, 4 and 7 are relatively

stronger cells in these maps deduced through inspection of their VIL peak magnitudes. Cell 5 and 9 are relatively weaker cells and have shorter lifetime compare with other cells.

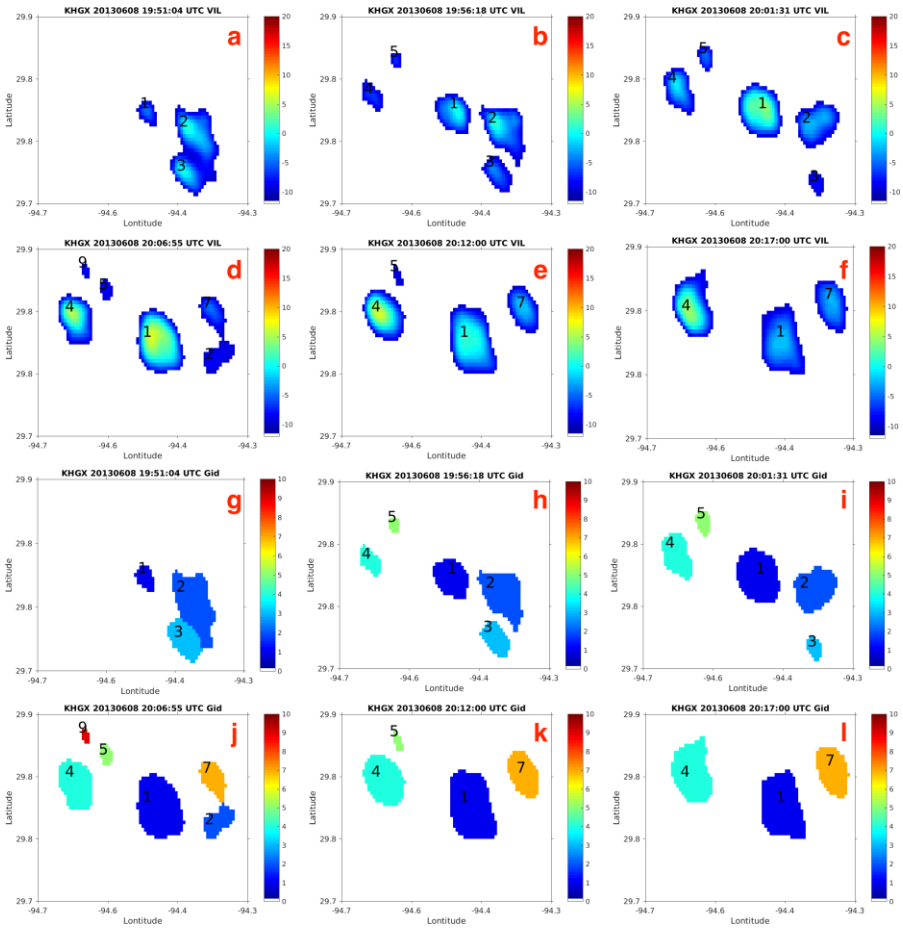

Figure 2.2 A Scattered cloud example with six identification and tracking steps for the KHGX on June 8, 2013, is shown. The first two rows show the horizontal VIL maps. Units are dB. Different colors in the lowest two rows divide the first two rows of VIL maps into cells. Each black number is a unique tracking id number for each cell with time.

A clustered tracking example (Figure 2.3.) is shown for August 8, 2016, KLBB (Lubbock) radar. Figure 2.3 has the same layout as Figure 2.2 for both VIL and the corresponding tracking partitions. The example is from 23:34:28 UTC to 00:00:33 UTC east of KLBB 100 to 120 km from the radar. These clustered cells are usually more difficult to identify and track, especially when convective cells are surrounded by stratiform clouds. By comparing pairs of the VIL panels with the partition panel, one can be confident in the multiple cell identification's capability of the MCIT algorithm. The general direction of motion of these clustered cells is from the west to the east. To validate the tracking capability of the MCIT algorithm, each pair of adjacent time steps is examined. The major cells are 2, 3 and 11 in these clustered cells. Cell 1 and 12 are mergers and merge to cell 2 and 11 respectively. Cells 4 and 8 are surrounded by stratiform clouds but uncoupled from other convective cells. Cell 5, 6, 7, 13, 14, 15, 20, 21 and 22 are clouds close to the coordinate boundary in this example and will not be discussed here.

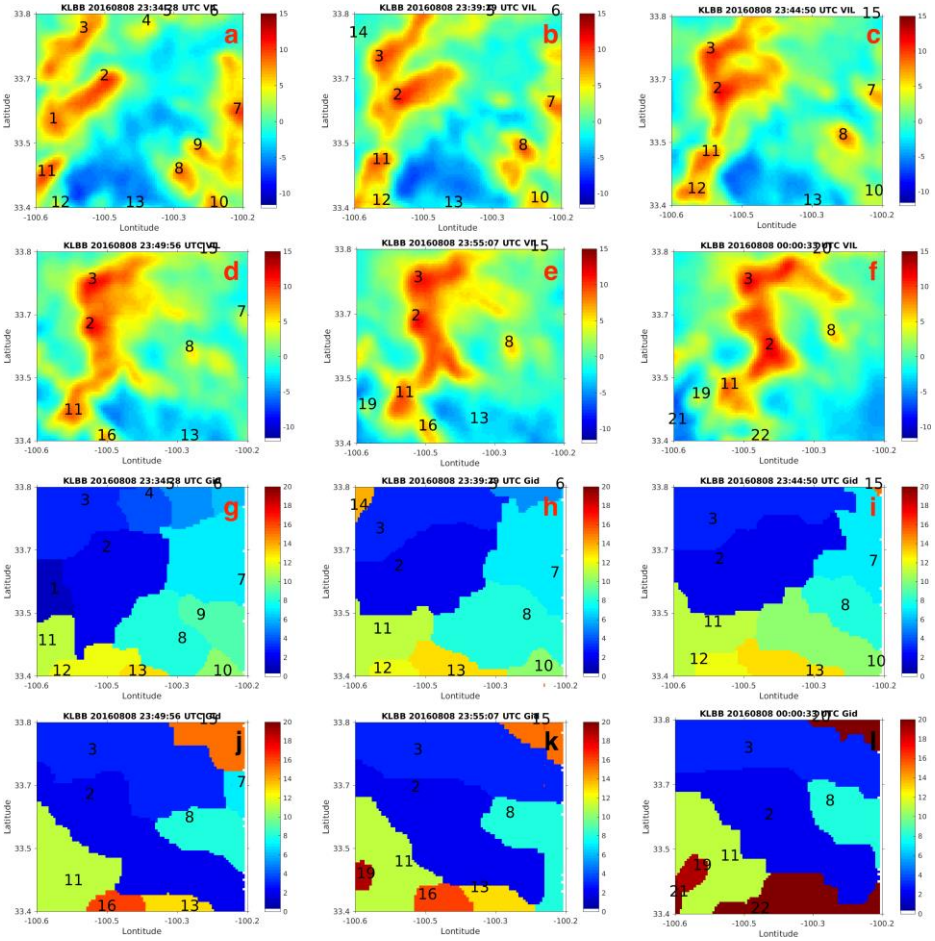

Figure 2.3 A squall line example with six identifications (a-f) and tracking steps (g-l) for the KLBB radar on August 1, 2015, are shown here. The first two rows show the horizontal VIL maps. Units are dB. Different colors in the last two rows divide the first two rows of VIL maps into cells. Each black number is a unique tracking id number for each cell with time.

(iii) The cell motion vector represents the movement of the cell peaks between time steps. According to Rosenfeld (1987), this motion is well correlated with the wind at the steering level (700 mb – 850 mb), so that cells in the same region usually have similar motion vectors. Therefore, shifting back the map by the average of

the cell motion vector in a given region works well for tracking of individual cells. The back shifting vector of the maps between time steps is a key part of the MCIT algorithm and is updated at each time step in the following way:

1) Based on the motion vector from previ*ous time steps*, each tracked pair of cells has a motion vector and the program saves all the shifting vectors. The initial shifting vector uses the mean wind at the steering level.
2) The full radar map is divided into 9 equal parts. The histogram of the u and v components of the cell motion vectors is constructed for each of the 9 sub domains. The median u and v components are used for constructing the shifting vector for all the cloud cells within the sub domain. The same is done for the full domain. All histograms and shifting vectors are saved. The statistic for each histogram needs to have a sample size of at least 25 and the median vector ±2 shifting vector samples must consist of at least 70% of each histogram's total shifting vector sample.

For each tracked pair of cells, the program first obtains the averaged shifting vector from the past hour for the sub domain in which the tracked cells reside. If the sample size of cell motion vectors for the previous hour is insufficient (i.e., <25), the program obtains the data from the previous two hours. If the statistics are still insufficient, the program uses the cell motions in the entire domain for the previous hour. If this is still insufficient, the program takes the cell motions from the entire domain during the previous two hours.

If none of the requirements above can be fulfilled, the shifting vector is taken as the mean wind from the steering level.

2.1.2.4 Stage 4: Assignment of Lifecycle Properties

The last stage of the MCIT algorithm is the construction of the time-height evolution of the observed polarimetric variables and LMA properties of the tracked cells. The peak reflectivity within the cell area at each height is taken as the vertical reflectivity profile. The vertical profiles of the other polarimetric variables are constructed by taking samples from the locations of peak reflectivities at the various heights. The lightning sources that occur at the various heights anywhere within the cell area are counted. These time height histories of the cells are saved in a database with a record for each cell.

The time-height evolution of the following cell properties is saved:

(i) Polarimetric data.

Z: The equivalent reflectivity factor, which is proportional to the 6^{th} moment of the raindrops size distribution.

Z_{DR}: The ratio of radar reflectivities at horizontal and vertical polarizations expressed in linear scale, which is an indicator of raindrop size. The combination of Z and Z_{DR} can give better assessment of hydrometeor types and rain rates. For example, large Z combined with low Z_{DR} values is a common indicator of hail.

K_{DP}: The specific differential phase is a good indicator for rain water content and is not affected by attenuation.

ρ_{hv}: The copolar correlation coefficient between horizontally and vertically polarized returns has low values for mixed-phase hydrometeors. For instance, rain and snow echoes

have high ρ_{hv} values, but in the melting layer with mixed-phase hydrometeors ρ_{hv} values can be as low as 0.8.

N_w and D_m: The normalized raindrop concentration and the mean volume diameter D_m are retrieved from Z and Z_{DR} as described by Illingworth and Thompson (2005), Tabary et al. (2011), and Ryzhkov et al. (2014)..

(ii) Lightning data. Two lightning networks are included – LMA (Lightning Mapping Array) and NLDN (National Lightning Detection Network). The LMA can provide both in cloud (IC) and cloud-to-ground (CG) lightning sources and radiated power strength in three dimensions. NLDN can provide only CG lightning locations and peak current. The merit of NLDN data is it can provide the polarity of each lightning event.

For example, cells 590, 3401, 2 and 1790 were detected by the MCIT algorithm on June 11, 2016 at KLBB; July 17, 2015 at KHGX; Aug 08, 2016 at KLBB and on June 11, 2016 at KLBB respectively. Selected properties of these cells as a function of time are plotted in Figure. 2.4. Both cell 3401 and 590 started as isolated cells. Cell 3401 is a warm rain cloud and cell 590 is a deep convective cloud with mixed phase processes. Cell 2 from Figure 2.3 is a squall line cell and cell 1790 is a supercell. It is found generally that the reflectivity peaks first, followed by echo top height, rain volume rate, and the cell's maximum area. The lightning peaked near the time of the echo top maximum height (Williams, 1985; Yoshida et al., 2009). Positive lightning flashes are much rarer than negative lightning flashes and total horizontal cloud to ground lightning flashes pattern looks similar to horizontal negative flashes pattern. This is mostly the case for the cells shown in Figure 2.4.

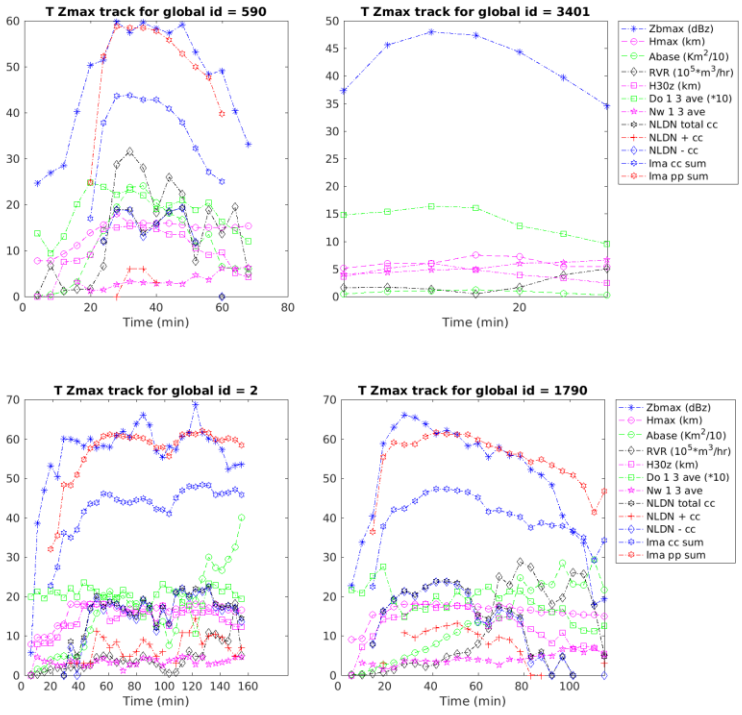

Figure 2.4 The properties of cell 590, 3401, 2 and 1790 as a function of time: Zbmax is peak reflectivity at cloud base in dBZ; Hmax is the cell's echo top height in 10^{-1} km above sea level; Abase is the area at cloud base in km^2; RVR is rain volume rate in 10^4 m^3 h^{-1}; H30z is the highest altitude reached by 30 dBZ in 10^{-1} km; D_m is the median raindrop diameter in mm. Nw is the raindrop concentration in log(Nw) m^{-3}; NLDN total cc is the total integrated NLDN CG flashes within each cells domain. NLDN + cc is similar to NLDN total cc but only consists of positive flashes. NLDN – cc is similar to NLDN total cc but only consists of negative flashes. LMA_CC is the integrated LMA source count in $10*\log_{10}(n)$. LMA_PP is the integrated LMA source power in $10*\log_{10}$ (P) [dBW].

With these data in hand, the fourth stage of the automatic procedure calculates and saves the lifecycle properties as follow:

1) Time-Height properties – Composites of the time-height evolution of each of the cell's radar variable (Z, Z_{DR}, K_{DP}, ρ_{hv}), results of hydrometeor classification, as well as the mean raindrop volume diameter D_m and the normalized concentration

N_w retrieved from the polarimetric radar variables (and LMA (counts and magnitudes of lightning sources). This is demonstrated here for four different kinds of clouds (Figures 2.5,2.6, 2.7 and 2.8). Each point in time and height in the different panels shows a polarimetric radar variable and the result of classification at the location of the reflectivity peak of the cell at that height and time. The values in the LMA panels show the integrated source counts within the cell area for each point in time and height.

2) Rain volume rate (RVR) [m^3/hr] – The integrated rain rate, which is calculated by the method from (Sachidananda and Zrnić, 1987) over the cell area at each time step is saved, as shown in Figure 2.4.

3) Maximum values – Maximum values throughout the entire life cycle of the cell for Z, Z_{DR}, K_{DP}, ρ_{hv} (minimum only), cell area, echo top height, rain volume rate and etc. The time of maximum during each cell's life cycle is also saved. Lifetime evolutions of these values during the cell lifetime are presented in Figure 2.4.

4) Location – The azimuth and range from the radar site of each cell at each time step. To ensure accuracy, all the cells in this paper are located between 40 and 120 km from the radar sites (Houston/Galveston, KHGX and Lubbock, KLBB in this paper).

5) Cell status and complexity – The four flags discussed in PHASE 3 (simple birth, simple continuation, merger and split) are saved for each cell. At each time step, the cluster in which each cell is embedded can have more than one cell. The

complexity keeps a record of the number of cells that make up the cluster at each time step during each cell's life cycle.

6) Thermodynamic parameters – The sounding data is from Eta Data Assimilation System (EADS). The temporal resolution is 3 hours and spatial resolution is 40 km. Melting level and Convective Available Potential Energy (CAPE) are saved for each cell. Since the temporal and spatial resolution is relatively coarse compared to the radar data, the program searches for the sounding data closest to each cell.

7） Satellite retrievals – By using the same method as in (Rosenfeld et al., 2016), for each cell a record is kept of its cloud base CCN concentration [mg^{-3}], cloud base supersaturation, cloud base temperature and cloud base vertical velocity. Although the satellite only has one data record for each case (around 13:30 solar time) and CCN concentration is slowly varying it is safest to use the retrieved CCNs for cells present at times close to the satellite scan time (+/- 2 hours of satellite overpass).

2.2 Four types of clouds

Here we present selected results that are intended to show the capabilities of the MCIT algorithm. Four types of convective clouds examples are shown: i) Isolated convective clouds with mixed phase processes; ii) Isolated convective clouds with mainly warm rain processes; iii) Convective clouds in a large cluster; iv) Supercell clouds. Examples of different types of convective clouds demonstrate some of the applications.

2.2.1 Isolated deep convective clouds with mixed phase

A common type of convective cloud is the isolated convective cloud with mixed phase processes and lightning events. Since these clouds are mainly isolated with little splits or mergers, they are relatively easy to identify and track, especially when the clouds are deep with mixed phase processes that indicate stronger updraft and more significant VIL signals compared to shallower clouds with little or no mixed phase processes.

An example is selected from June 6, 2016, near the Lubbock KLBB radar. The cloud cell id is 590 and was initiated at 32.82 °N 101.83 °W, 18:54 UTC, 93 km from the radar station at an azimuth angle of 181°. Figure 2.5a shows the time –height evolution of maximum reflectivity in dBZ. In this paper, the first echo height is defined as the height of the maximum reflectivity in the vertical at the first time when it exceeds 30 dBZ. The first echo height of cell 590 is detected at around 6.5 km (i.e., about 2 km above the environmental freezing level) and the cloud continues to grow for 25 minutes until the echo top reaches about 16 km. Dissipation in this paper is characterized as the reflectivity core reach its maximum and start to fall. The dissipation starts after 30 minutes of identification and lasts for another 55 minutes. The red numbers under each Z column are the complexity values and the magnitude is $\log_2(N)$, where N is the number of cells in the cluster in which the tracked cell resides.

The time-height cross section reveals some of the convective processes, which demonstrate the usefulness of this kind of presentation. Figure 2.5a shows a strong reflectivity core (> 60 dBZ) at t=30 min and H=8 km, which descends and reaches the surface after 5 to 10 minutes along the marked black line. The low value of Z_{DR} in the

core of this precipitation shaft marked by the black line indicates that this core is primarily composed of hailstones which start melting at the height of 3 km as indicated by the abrupt increase of Z_{DR} up to about 2 dB; this increase is caused by complete melting of smaller-size graupel / hail particles, partial melting of larger-size hailstones which acquire an external water coating, and emergence of very large raindrops having ice cores inside (Ryzhkov et al. 2013). The low values of ρ_{hv} (less than about 0.95, Figure 2.5d) through the precipitation shaft show that the hail is not completely melted all the way to the ground. The green line in Figure 2.5b marks a Z_{DR} column that represents the largest raindrops which form in the convective updraft, fall from the updraft area ahead of smaller-size raindrops, and reach the surface some 5 to 10 minutes before the main precipitation shaft (Kumjian and Ryzhkov, 2012). The high values (red, about 4 dB) of Z_{DR} extend up to 7 km and are caused by large supercooled raindrops. The low ρ_{hv} there indicates the existence of mixed phase rather than purely liquid supercooled raindrops. The low K_{DP} values within the Z_{DR} column (Figure 2.5c) along with the large reflectivities suggest that the column is composed of sparse large falling raindrops and lacks small drops which results in low rainwater content. The hail shaft transforms into a heavy rain shaft some 10 minutes later, as indicated by the large K_{DP} values, marked by the blue line. The K_{DP} cross section (Figure 2.5c) shows that hail or large graupel continues to fall and melt into rain at a height of 3 km between 45 and 60 minutes into the cell lifetime. The increase of the hydrometeor melting level from time 60 onward is evident in the Z_{DR} cross section by the increasing height of the top of the high Z_{DR} region. The large raindrop sizes and their respective low concentrations at low levels are shown in Figures 2.5e and 2.5f,

respectively during the times before onset of the major precipitation after 35 min. These are caused by differential sedimentation of large drops falling from the updraft. The indicated values (D_m and N_w) are not valid for frozen hydrometeors and only pixels classified as L/MR, HR and BD (Figure 2.5h) are plotted in color in the D_m and N_w panels (Figure 2.5e and 2.5f). It is possible to discern the locations of raindrops with various size distributions using the polarimetric variables Z_{DR} and K_{DP}. For instance, at lower levels (< 4 km) during the cell's dissipation stage (t = 60 to t = 90 minutes), a systematic pattern of decreasing D_m (Figure 2.5e) and increasing N_w (Figure 2.5f) caused by fallout of large raindrops and dominance of smaller raindrops before dissipation is obvious. The masked pixels in D_m and N_w panels (black dots in Figure 2.5e and 2.5f) due to mixed/solid phase processes are suggested in the ρ_{hv} (Figure 2.5d) panel (ρ_{hv} < 0.9 is a strong signal of hail and graupel) and DHCA (Figure 2.5h) panel.

The number of lightning sources (Figure 2.5g) starts peaking up 5 minutes after the collapse and freezing of the Z_{DR} column. The lightning flashes begin aloft (probably as intracloud flashes) and descend to the ground with the main precipitation hail shaft (Figure 2.5h). Flash rate can be calculated with MCIT algorithm package, but in this paper we are concentrating on the lightning source time-height patterns.

The DHCA panel (Figure 2.5h) shows the pattern of different hydrometeor habits. Most of hail and large graupel undergoes wet growth within the updraft at a height of above 4.5 km and starts melting below that level in the downdraft area containing a bulk of precipitation from t = 30 to t = 55 minutes after detection. Most solid phase hydrometeors in this cloud are small hail categorized as R/Ha, dry graupel (GR), dry snow

(DS) and crystals (CY). No large hail (LH), i.e., hail exceeding 2.5 cm is detected by DHCA. GR is the dominating hydrometeor above 6 km. During the mature stage (t = 30 to t = 50 minutes), most hydrometeor signals below 6 km are R/Ha type, suggesting consistent evidence of rain and small hail falling to ground during this time period. Heavy rain (HR) and light / moderate rain (L/MR) dominate the dissipation stage near the surface (z < 4.5 km) from t = 55 minutes to the end.

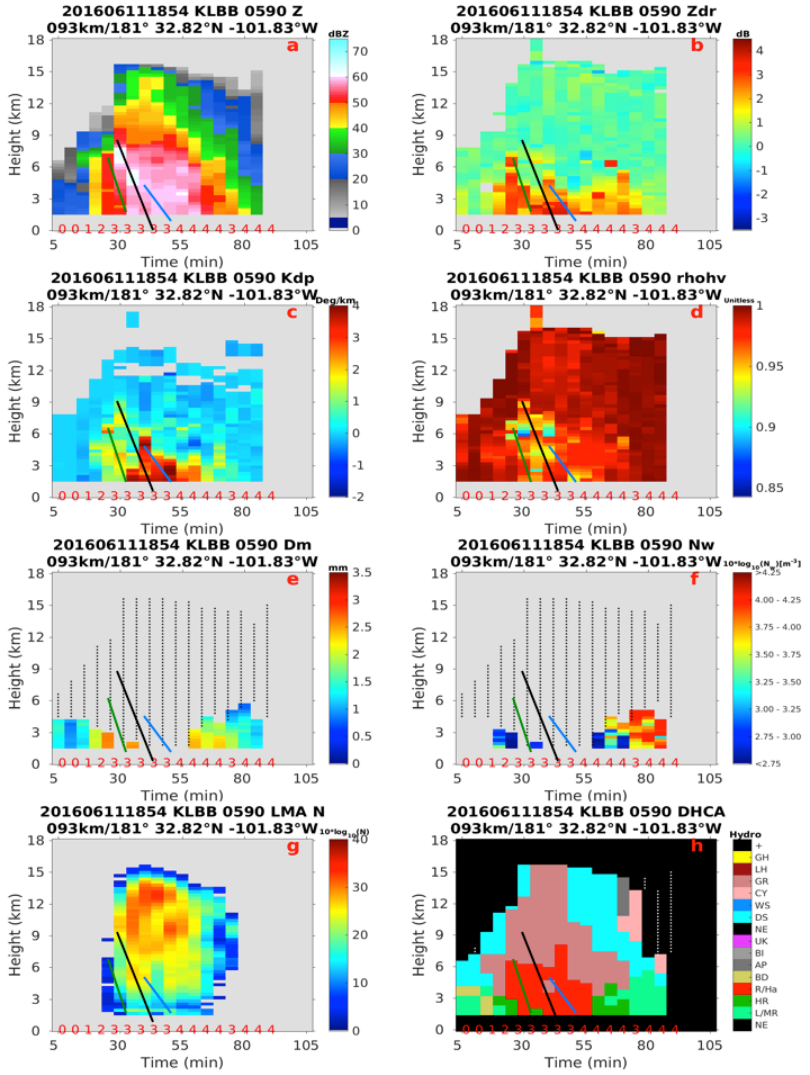

Figure 2.5 The time-height properties of cell 590 are shown here. Figure 2.5.a – h show the time-height evolution of Z, Z_{DR}, K_{DP}, $\boldsymbol{\rho_{hv}}$, D_m, N_w, LMA_CC and DHCA. For Figure 2.5.a – f, each pixel represents the magnitude of the variables at the same location of maximum Z of each time-height dimension for this cell. Figure 2.5.g pixels are the integrated VHF lightning sources for this cell at each time-height dimension. Figure 2.5.h pixels are hydrometeor classifications. Each panel title includes the specific UTC time of the cell shown in the Figure when the cell is first detected. For instance, 201606111854 means 18:54, June 11, 2016 and 093 km/181° 32.82°N -101.83°W means this cell is first detected 93 km away from KLBB site with an azimuth angle of 181 degree located at 32.82°N -101.83°W.

2.2.2 Isolated convective clouds with mainly warm rain processes

A typical isolated convective cloud with most precipitation formed through the warm rain process can occur in slightly unstable marine air mass. Warm rain clouds usually have relatively low echo top height that does not extend much above the freezing level, a slower growth rate of echo top height and produce less intense precipitation than mixed phase convective clouds.

An example is selected on July 17, 2015, near the Houston/Galveston KHGX radar. The cloud cell id is 3401 and was initiated at 28.85 °N 94.98 °W, 17:05 UTC, 48 km from the radar with an azimuth angle of 172°. Cloud cell 3401 was detected over the ocean. The first echo height of cell 3401 is detected at around 4 km (Figure 2.6a). The cell continued to grow for 20 minutes until the echo top reached around 7 km and subsequently dissipates in 15 minutes.

The reflectivity core is relatively weak (< 50 dBZ) compared to the previous cell example. The black line marks the reflectivity core descent to the ground in less than 10 minutes and represents the rain shaft of cell 3401. There are no clear K_{DP} and Z_{DR} enhancements in this cell. The low K_{DP} and moderate reflectivity magnitudes indicate that the cell 3401 has low rainwater content. Inspection of the Z, Z_{DR}, K_{DP}, and ρ_{hv} panels, shows no indication of mixed phase hydrometeors. Figure 2.6g shows no lightning activity (both LMA and NLDN) in this cell 3401. In the warm rain scenario, pre-existing rain embryos quickly grow by collision-coalescence to large raindrops which are sparser at the growing stage (t = 0 to t = 15 minutes) of the cloud. This is suggested by the relatively high Z_{DR} (Figure 2.6b) and D_m (Figure 2.6e) combined with low N_w (Figure 2.6f) values

below 3 km during this period. A closer look at the Z_{DR} descending pattern (Figure 2.6b, t = 10 to t = 15 minutes) shows that it is ahead of the main rain shaft from Z (Figure 2.6a, t = 15 to t = 25 minutes), which is due to the fallout of heavier and larger raindrops. At the dissipation stage (t = 15 to t = 35 minutes), most large raindrops have fallen out and more numerous smaller raindrops dominate the precipitation hydrometeor type. This transition is indicated by the relatively low Z_{DR} (Figure 2.6b), D_m (Figure 2.6e) and high Z (Figure 2.6a) and N_w (Figure 2.6f) values during the dissipation stage. The DHCA Figure (Figure 2.6h) is no surprise since this warm rain cloud consists mainly of L/MR and shows a consistent pattern of typical warm rain process as discussed above.

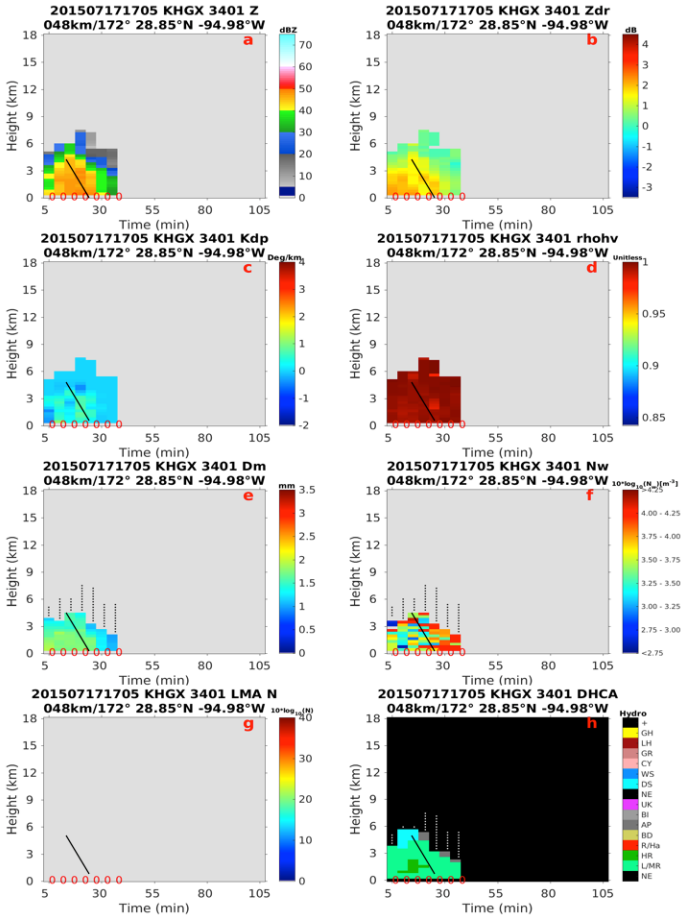

Figure 2.6 As in Fig 2.5, but for cell 3401 on July 17, 2015.

2.2.3 Multi-cell clouds

Multi-cell clouds usually refer to cloud clusters consisting of several clouds each of which is at a different stage of development. This type of cloud usually brings heavy precipitation, hail, frequent lightning, high wind, but rarely tornadoes. Multi-cell clouds tend to have longer lifetime and more intense precipitation than isolated convective clouds, except for supercells.

An example of a tracked cell from a squall line is selected on August 8, 2016 near the Lubbock KLBB radar. It is cell #2 in the horizontal tracks (Figure 2.3), which have been discussed in the methodology section. The cell was initiated at 34.25 °N 100.81 °W, 18:52 UTC, 114 km from the radar at the azimuth angle of 54°. The first echo height of cell 2 is detected at about 8 km and the cloud continues to grow for 30 minutes until the echo top reaches 18 km.

Three discernible cycles are visible visible in the Z panel (Figure 2.7a). The main precipitation shafts are marked by the black lines and each line represents the regeneration of this squall line. The combination of negative Z_{DR} (Figure 2.7b) and low ρ_{hv} (Figure 2.7d) coinciding with very high reflectivities (> 60 dBZ in Figure 2.7a) is a strong indication of large hail and graupel. The second and third black lines in each panel are well defined hail shafts that reach ground level. The strongest hail shaft belongs to the third cycle between 120 to 135 minutes, as indicated by the lowest ρ_{hv} region in Figure 2.7d. Figure 2.7c indicates an occasional heavy load of rainwater content at low levels, apparently as a result of the partial melting of the ice hydrometeors. The large (red) K_{DP} pixels combined with very low or even negative Z_{DR} and low ρ_{hv}pixels suggest heavy precipitation combined with wet hail and graupel. Between adjacent cycles delimited by the black lines, the regeneration of the squall line is most obvious in the Z_{DR} panel. For instance, within the time interval between t = 100 min to t = 120 min, the positive large (yellow to red) values in the Z_{DR} panel (Figure 2.7b) with relatively smaller reflectivity (Figure 2.7a) indicate that the liquid phase raindrops are dominating during this time slot. Before t = 115, when the hail shaft of the third cycle starts, large raindrops begin falling

at t = 110 min. The intensity of each regeneration is stronger than the previous cycle. This is indicated by the evolution of reflectivity, Z_{DR} and ρ_{hv} along the black lines. No obvious pattern can be found in D_m and N_w panels due to the overwhelming presence of ice or mixed-phase hydrometeors suggested by the DHCA product (Figure 2.7h) which inhibits the retrieval of D_m and N_w.

The LMA observations (Figure 2.7g) shows extremely high lightning activity (LMA source counts) from 30 minutes to 155 minutes after cell initiation. The lightning activity follows the pattern of the hail shafts. In accordance with the greatest intensity of the third hail shaft, the lightning activity reaches maximum source counts at the same time.

From the DHCA Figure (Figure 2.7h), the three major precipitation shafts marked by the black lines are dominated by R/Ha and LH (2nd and 3rd precipitation shafts) and reach ground level during at least two periods marked by two most intense precipitation shafts (t = 80 to t = 100 minutes and t = 120 to t = 135 minutes). LH reaches ground during the 3rd precipitation shaft in the Figure 2.7h is consistent with the ρ_{hv} panel (Figure 2.7d), during which ρ_{hv} value is less than 0.9. There is a gradual transition with increasing height from pure rain (L/MR, HR, and BD) to small hail mixed with rain (R/Ha), then to GR, DS and CY.

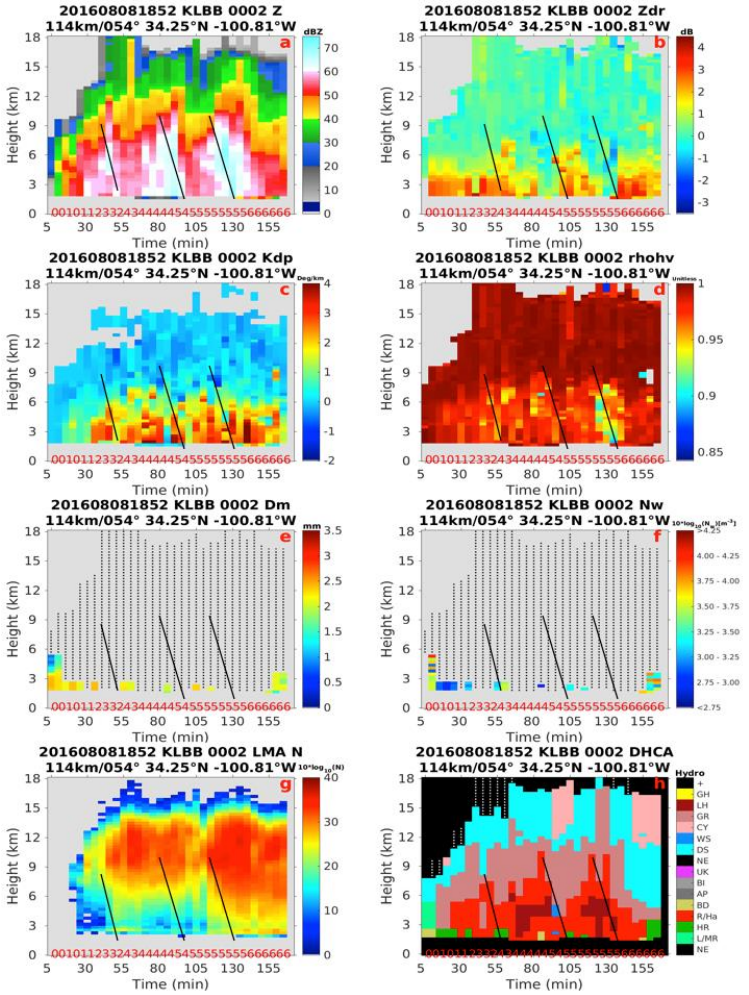

Figure 2.7 As in Fig 2.5, but for cell 2 on June 8, 2016.

2.2.4 Supercell

In the thunderstorm spectrum, supercells are the least common type of thunderstorm, but they have the highest propensity to produce severe weather, including damaging winds, very large hail, and sometimes tornadoes. What makes a supercell unique from all other thunderstorm types is that it contains a deep and persistent rotating

updraft called a mesocyclone. If the environment is favorable, supercell thunderstorms can last for several hours, and so are often easily tracked with algorithms of the kind considered here.

An example is selected on June 11, 2016, near Lubbock KLBB. The cloud cell id is 1790 and was initiated at 33.77 °N 101.03 °W, 20:18 UTC, 74 km from radar station with an azimuth angle of 80°. The first echo height of cell 1790 is detected at around 7 km and the cloud continues to grow for 30 minutes until the echo top reaches 18 km. The supercell echo top quickly develops to over 18 km in 30 minutes with a strong reflectivity core (>60 dBZ). The dissipation starts after 55 minutes of cell 1790's detection and lasts for 70 minutes. The supercell lifecycle is clearly longer (> 2 hours) than the lifecycle (< 1 hour) of an individual cell in a squall line as in the previous storm example (section iii). Of particular note in the present case is the long duration (30 min) of continuous high reflectivities and hail, which is a radar hallmark of a supercell.

The cross section of reflectivity shows that the major precipitation shaft starts around 25 minutes after the first echo. It is marked by the black lines in each panel. The weak Z_{DR} and K_{DP} at the core of high reflectivity indicate a strong signal from a hail shaft from 25 to 55 min. The high K_{DP} (red) pixels and low ρ_{hv} pixels at lower level (< 6km) indicate wet hail (Figure 2.8c and 2.8d). The ρ_{hv} cross section (Figure 2.8d) shows that the first detection of hail occurred at t = 20 min, as indicated by the low values of ρ_{hv}. The low Z_{DR} along the line of maximum Z indicates that the hailstones reach ground level. Hail production and fallout in this supercell lasts for 40 minutes and is much longer than that in the squall line example (< 20 min) as shown in Figure 2.7. The onset of the

dissipation stage is indicated by the rising of the hydrometeor melting level from t=50 min onward, which is shown in the Z_{DR} panel by the increasing height of the relatively large Z_{DR} (yellow to red pixels in Figure 2.8b). The decreasing raindrop size (Figure 2.8e) and increase of their concentration (Figure 2.8f) from 70 min also indicate the decay of Cell 1790.

The lightning panel (Figure 2.8g) shows that lightning is initiated right after the hail shaft started. During the mature stage of this supercell (t=25 min to t = 55 min), the top height of the lightning activity is close to the echo top height and the base of lightning activity practically reaches the surface. As the supercell dissipation starts, the top height of the lightning activity descends to about 10 km while echo top heights remains at 15 km. At the same time, the base of the lightning activity rises towards the level of 5 km, indicating a decrease of cloud-to-ground lightning activity and more stratiform cloud-to-cloud lightning during the supercell's dissipation stage. The lightning height top descends with the 25 dBZ reflectivity contour line at nearly the same rate (Figure 2.8a and 2.8g), which is likely due to the descent of large hydrometeors as a result of the weaker vertical velocity. This pattern of lightning reaching the radar echo top during the active convective phase of a cell and descending as the cell starts to decay was also observed in other thunderstorms that were tracked through their decaying stages.

From the DHCA Figure (Figure 2.8h), a clear hail shaft is revealed from t = 25 to t = 55 minutes after detection of "big drops" (BD) and its transition (t = 5 to 10 minutes) to L/MR and HR (t = 5 to t = 20 minutes) and then to solid phase hydrometeors (R/Ha, LH and GR from t = 15 minutes) is obvious. HR at low levels (z < 4 km) is associated

with GR and R/Ha at z > 4 km. Comparing Figure 2.8g and 8h, the core of LMA sources counts descends after the core of GR falls downward. After surface rain became weaker (L/MR from t = 105 to the end), upper level (z > 4.5 km) GR is replaced by DS.

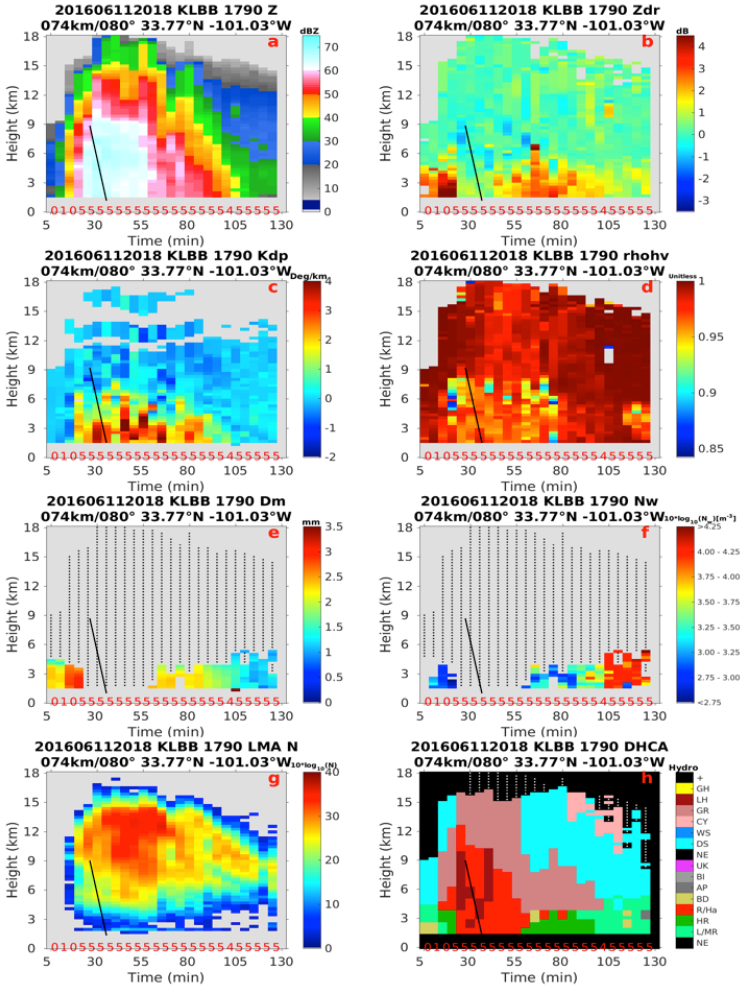

Figure 2.8 As in Fig 2.5, but for cell 1790 on June 11, 2016

3. CONVECTIVE CELL TRACKING APPLICATION

The MCIT algorithm package (Hu et al., 2018) has been used as the primary tool to produce data set with cloud cell life information in this study. A cell is defined as a local maximum of vertically integrated liquid (VIL), where VIL is calculated with a fixed Z-R relationship from the WDSS II software package developed by the National Severe Storm Laboratory (Lakshmanan et al., 2007). The cells are tracked by the maximum common VIL in consecutive radar scans. Various parameters are collected for each cell. For instance, in this study we use maximum Z and the corresponding Z_{DR} at each step of tracking at all elevation angles to construct time-height evolutions of cloud cell life cycle. To isolate the effect of CCN on the tracked cells' properties, each tracked cell was assigned to a class of CCN concentration, while constraining the echo top height vertical growth rates within bounds of up to 3 km within the first 30 minutes (Fig 3.1 and 3.2). The integrated LMA source counts for each cell at each step of tracking at all elevations are also presented in the results section under different cell selection criteria.

The CCN number concentration and super saturation (SS) are retrieved from the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi National Polar-Orbiting Partnership (NPP) satellite by using convective clouds as natural CCN chambers (Rosenfeld et al., 2016). The CCN number concentration is equal to the satellite-retrieved cloud base drop concentration, which is calculated by a VIIRS-retrieved T-r_e relationship and cloud base updraft (W_b). Although the satellite only has one overpass (around 13:30 solar time), we apply the retrieved CCN for ±3 hours, This time window is short with

respect to time scale of changes in air masses, while still allowing for a useful sample size. The CCN, SS and W_b are mapped into the data records of the tracked cells by MCIT.

Cells were tracked under various conditions of CAPE and CCN conditions and were further classified as occurring over land, coast or ocean and as occurring over rural or urban areas. CAPE is retrieved from Eta Data Assimilation System (Project, 2004) with temporal resolution of 3 hours and spatial resolution of 40 km. Events occurring during and around 19 satellite overpasses were analyzed, resulting in a total of 2859 tracked cells. The cases are listed in Table 3.1.

Table 3.1 Cases dates and time period in UTC time.

Dates	Time (UTC)
20130619	16:00-03:00
20130722	17:00-21:00
20140527	17:00-22:00
20140703	16:00-23:59
20140704	19:00-23:59
20140714	15:00-23:59
20140724	15:00-23:00
20140906	14:00-23:59
20140911	16:00-23:00
20150626	12:00-2:00
20150717	12:00-23:00
20150813	19:00-4:00
20150818	10:00-23:59
20160521	12:00-5:00
20160527	15:00-5:00
20160601	13:00-5:00
20160628	16:00-23:59
20160730	18:00-23:59
20160926	14:00-23:59

Table 3.2 Definitions of the selection criteria for cell selection.

Sample	Cloud sample size for selected scenarios.
HT0	The height range of samples' highest pixel that exceeds 1.5 dB at time T0.
H1GW	The growth interval of the first growth period in km.
Gt	The growth time length interval for the first growth in minutes.
NC	Max number of cells in selected cloud clusters.
CCN	CCN concentration range in cm^{-3}.
Rdis	Distance range between NEXRAD radar site (KHGX Houston/Galveston in this study) to sample clouds' initial detection location
Udis	Distance range between urban center (Houston city center in this study) to sample clouds' initial detection location
CAPE	CAPE range in J/kg
Hour	The time range of samples' detection time in UTC.
Reg	Region of cloud samples, 1 = ocean, 2 = costal, 3 = continental.
T0	The time range of samples' first and highest column top Z_{DR} value exceeds 1.5 dB in minutes after detection.
Mean CCN	Mean CCN concentration in cm^{-3}.
Mean SS	Mean super saturation. (%)
Mean CAPE	Mean CAPE in J/kg.
dBZ H 1	Initial echo top height limit in km.

It is crucial to define selection criteria to divide cloud cell samples into different categories based on their various CCN, CAPE and regional attributes in order to compare

the effect these attributes have on hydrometeor and lightning properties. A complete Table of selection criteria can be found in SI Table S3.2. Here we discuss a few important criteria. To isolate the microphysical from thermodynamic effects, the initial (within the first 30 minutes) growth rate of each cell's echo top height is restricted to be less than 10 m/s. The analyses are done with and without this restriction.

The cells that pass the selection criteria are composited by averaging them after shifting their times to a common time basis. Two kinds of cell compositions are used here, as defined by different reference time (t_0). To study the CCN effect on cloud microstructure and invigoration, t_0 was defined as the time step where Z_{DR} > 1.5 dB reached a peak height for the first time. This is the time when the largest raindrops in the updraft start to fall or freeze. Indeed, the composites in Fig. 3.1 show clearly patterns of rain shafts initiated aloft at t_0.

Superposition of the cells' lightning properties was done by defining the time of first lighting as t_0. This allows assessing the time evolution of lightning as a function of different CCN, CAPE and location/region. Cells that had no LMA-detected lightning sources could not be included.

3.1 CCN Effect on Cloud Microphysics

First, we test the role of CCN concentrations on the hydrometeors of convective cells while excluding the aerosol invigoration effect. This is done by selecting cells that have initial vertical growth rates (as seen in their echo top heights) less than 10 m s^{-1}. The full classification criteria are shown in Table 3.3. The t_0 for superpositioning of the cells was their time of rain initiation.

Table 3.3 Criteria tables for all scenarios in this study for figure 2-4. The most left column shows the acronym of all the criteria. The top row shows the corresponding scenario and figure number of each table. Detailed criteria acronym is described in Table 3.2.

Criterion	CCN effect with growth rate control, Figs 2-4			
Fig panel	a	b	c	d
N of cells	10	40	18	15
HT0	[4 10]	[4 10]	[4 10]	[4 10]
H1GW	[0.1 3]	[0.1 3]	[0.1 3]	[0.1 3]
Gt	[4 31]	[4 31]	[4 31]	[4 31]
NC	[1 10]	[1 10]	[1 10]	[1 10]
CCN	[1 300]	[301 600]	[601 1000]	[1001 9999]
Rdis	[40 120]	[40 120]	[40 120]	[40 120]
Udis	[1 9999]	[1 9999]	[1 9999]	[1 9999]
CAPE	[0 9999]	[0 9999]	[0 9999]	[0 9999]
Hour	[16 22]	[16 22]	[16 22]	[16 22]
Reg	[1 3]	[1 3]	[1 3]	[1 3]
T0	[4 31]	[4 31]	[4 31]	[4 31]
Mean CCN	204	451.9	723.3	1379
Mean SS	0.282	0.218	0.161	0.114
Mean CAPE	3054	3004	2903	2403
Mean CBH	914.1	939.3	922.6	1071
dBZ H 1	9	9	9	9

Fig. 3.2 shows time-height evolution of the reflectivity patterns for 4 intervals of CCN concentration with the values of <300, 301-600, 601-1000 and >1000 CCN cm^{-3}, in panels a to d, respectively. CAPE is limited between 2000 J/kg to 3700 J/kg so that the cloud samples in each CCN interval have similar mean CAPE values. The similarity in CAPE and CBH (Table 3.3) minimize their effects. The descending maxima of Z that are evident in Fig. 3.2 indicate the descending cores of the first precipitation shafts. The core reaches the ground ~10 minutes after its initiation aloft for average CCN<300 cm^{-3} (Fig. 3.2a) but increases to ~20 minutes for the highest CCN concentrations of > 1000 cm^{-3}.

This increase in time of precipitation core reaching the ground has high statistical significance (Fig. 3.2c and Table 3.8). These patterns behave according to expectations that large concentrations of CCN inhibit coalescence and delay the initiation of rain. The initial location of cell samples can be found in Fig. 3.2.

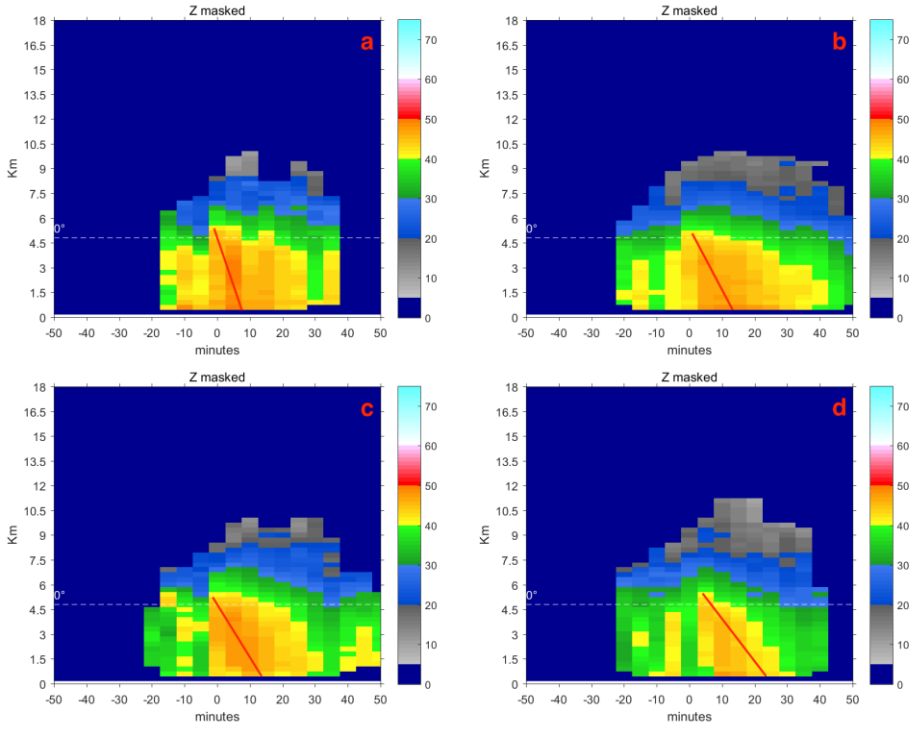

Figure 3.1 Time-height reflectivity composites of tracked cells. Time zero is defined as the time of first peak height of Z_{DR}>1.5 dB. The composites are for CCN concentrations of four intervals, a: 1-300 cm^{-3}, b: 301-600 cm^{-3}, c: 601-1000 cm^{-3}, d: 1001-9999 cm^{-3}. The sample clouds include clouds from three regions, i.e. land, coastal area and ocean. The first detected echo top height is limited to be less than 9 km in order to filter out cells that were not tracked from their initiation. To isolate the microphysical effect of CCN from the invigoration feedbacks, the cloud vertical growth rate is up to 3 km within the first 30 minutes. Detailed criteria is in Table 3.3.

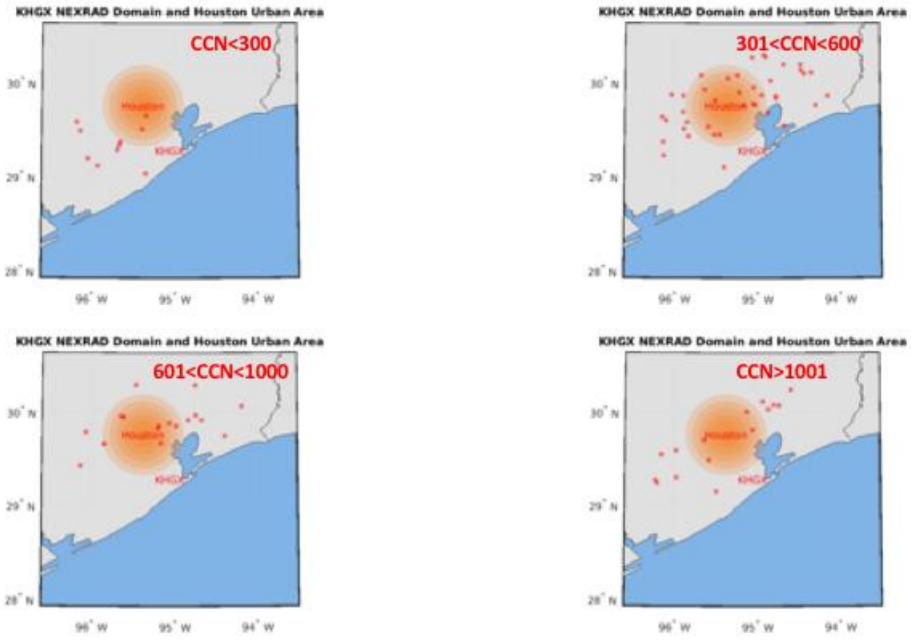

Figure 3.2 Location of selected cells for Fig. 3.1, 3.3 to 3.4.

Fig 3.3 shows time-height evolution of Z_{DR} of the same cloud samples and CCN intervals as in Fig 3.1. The core of peak Z_{DR} reaches the ground some 5 to 15 minutes before the reflectivity cores do, with statistically significant longer time for higher CCN (Fig. 3.4b and Table 3.8). This indicates that the initial relatively sparse large raindrops that form in the clouds fall before the main precipitation shaft due to their larger terminal fall velocity. It also shows that high CCN delays the formation of the initial large raindrops.

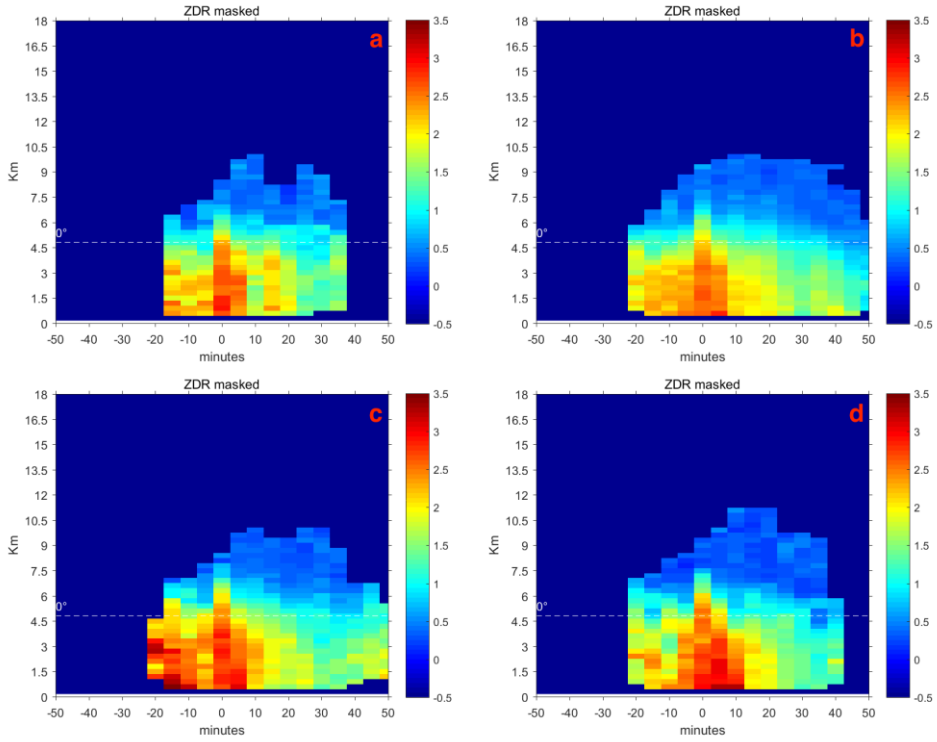

Figure 3.3 Same as Fig 3.1, except for Z_{DR}. Detailed criteria is in Table 3.3.

The boxplot of the time of main precipitation shaft reaching ground is shown in Fig 3.4a and the time of Z_{DR} shaft reaching ground is shown in Fig 3.4b. Z and Z_{DR} shaft are calculated by weighted linear regression function (Pozo, 1997) and their times to reach ground are the intercept to the x-axis calculated from the function. The box plots of the times for rain shaft and Z_{DR} shaft reaching ground are using the same samples and CCN intervals as Fig 3.1 and Fig 3.3. The mean time of rain shaft reaching ground, marked by the red star, for the four CCN intervals are 14.5, 20, 23.3 and 29.7 minutes respectively. The mean times for Z_{DR} shaft reaching ground for the four CCN intervals are 9, 13.7, 15

and 22 minutes respectively. These patterns of delay rain initiation are consistent with the reflectivity patterns in Fig 3.1.

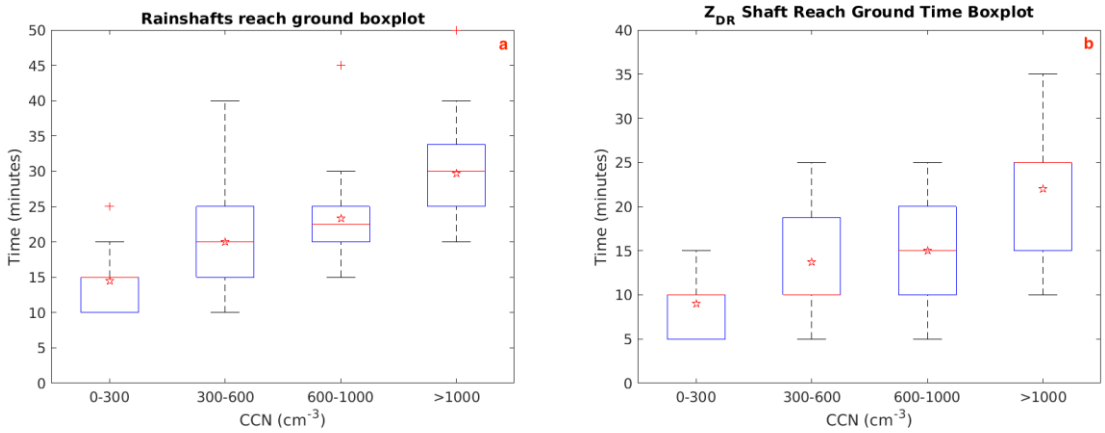

Figure 3.4 Box plots of the time difference from rainshafts initiation aloft until they reach ground (Fig 3.4a) and the same for Z_{DR} shafts, for the same CCN interval as Figs 3.1 and 3.3. Rain shaft and ZDR shaft are calculated by weighted linear regression function (Pozo, 1997) between height and maximum Z or Z_{DR} in the vertical. The stars represent the mean values. Detailed criteria is in Table 3.3.

The mean volume diameter (D_m) of raindrop is retrieved by Eq. 1 (Ryzhkov et al., 2014; Testud et al., 2001):

$$D_m(mm) = 1.53[Z_{DR}(dB)]^{0.47} \qquad (1)$$

The magnitude of D_m increases with more CCN, indicating larger hydrometeors with more CCN (Fig 3.5). This is ascribed to the CCN-induced suppressed coalescence, which leaves more cloud water available for faster accretion and resultant growth of raindrops or

hailstones. (The statistical significance of the dependence of D_m on CCN is shown in Fig 3.6).

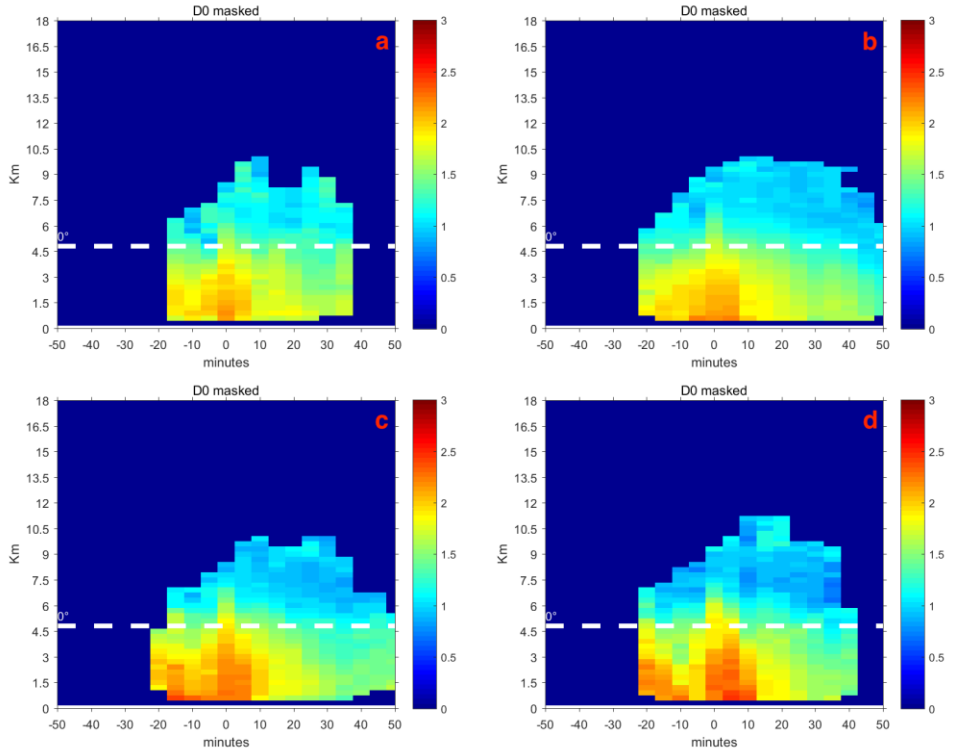

Figure 3.5 Same as Fig 3.1, except for cloud drop mean volume diameter, D_m. The retrieved D_m is valid only for liquid raindrops.

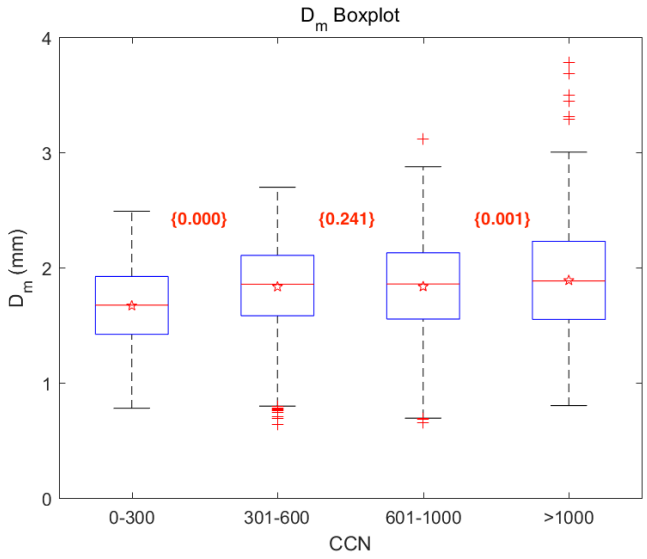

Figure 3.6 Boxplot of D_m under 4km from 0 to 25 minutes as in Fig 3.5. The height maximum is set to 4 km is due to the calculation of D_m is only valid in liquid phase.

Fig 3.7 shows the CCN convection invigoration effect on cell's lifetime maximum echo top heights (H_{max}) [km]. The box plots of H_{max} in Fig. 3.7 are made for the same CCN intervals as in Figs. 3.1 and 3.3. Therefore, CAPE and CBH for the CCN groups are also similar. The means of H_{max} increase with CCN concentration from 10.5 km (CCN<300 cm^{-3}) to 12.9 km (600-1000 cm^{-3}) and become saturated for higher CCN. The indicated increase in H_{max} is statistically significant (Table S3.8) and serves as a manifestation of the aerosol invigoration effect. The initial location of cell samples can be found in Fig. 3.8. Fan et al. (2009) showed that wind shear > 10 ms^{-2} suppresses the invigoration. The wind shear is calculated as vector difference between 925 and 500 hPa winds. However, the wind is not likely to suppress the CCN effects, because its magnitude for the various CCN groups was weak and similar.

Table 3.4 Criteria Tables for all scenarios in this study for figure 5. The most left column shows the acronym of all the criteria. The top row shows the corresponding scenario and Figure number of each Table. Detailed criteria acronym is described in Table 3.2.

Criterion	CCN effect without growth rate control, Fig 5			
Fig panel	a	b	c	d
N of cells	40	138	73	17
HT0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
H1GW	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Gt	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
NC	[1 10]	[1 10]	[1 10]	[1 10]
CCN	[1 300]	[301 600]	[601 1000]	[1001 9999]
Rdis	[40 120]	[40 120]	[40 120]	[40 120]
Udis	[1 9999]	[1 9999]	[1 9999]	[1 9999]
CAPE	[0 9999]	[0 9999]	[0 9999]	[0 9999]
Hour	[16 22]	[16 22]	[16 22]	[16 22]
Reg	[1 3]	[1 3]	[1 3]	[1 3]
T0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Mean CCN	212.4	450.3	753.8	1348
Mean SS	0.273	0.209	0.167	0.118
Mean CAPE	3032	3170	3057	2417
Mean CBH	842.9	928.1	913	998.3
dBZ H 1	9	9	9	9

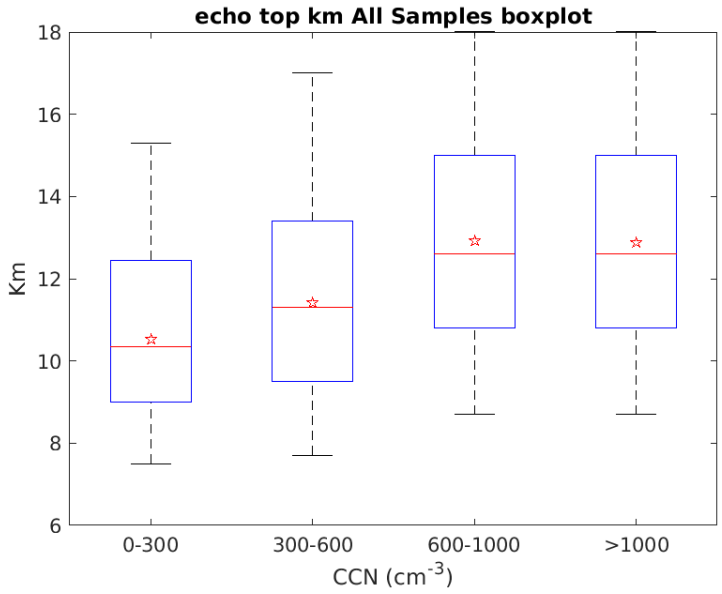

Figure 3.7 Box plots of cloud samples max echo top height (km) using the same CCN interval as Fig 3.1 and 3.3. The first growth height and time limitations in Fig 3.1 and 3.3 are released in order to show the CCN invigoration effect. Other criteria stay the same as Fig 3.1 and 3.3. Detailed criteria is in Table 3.4.

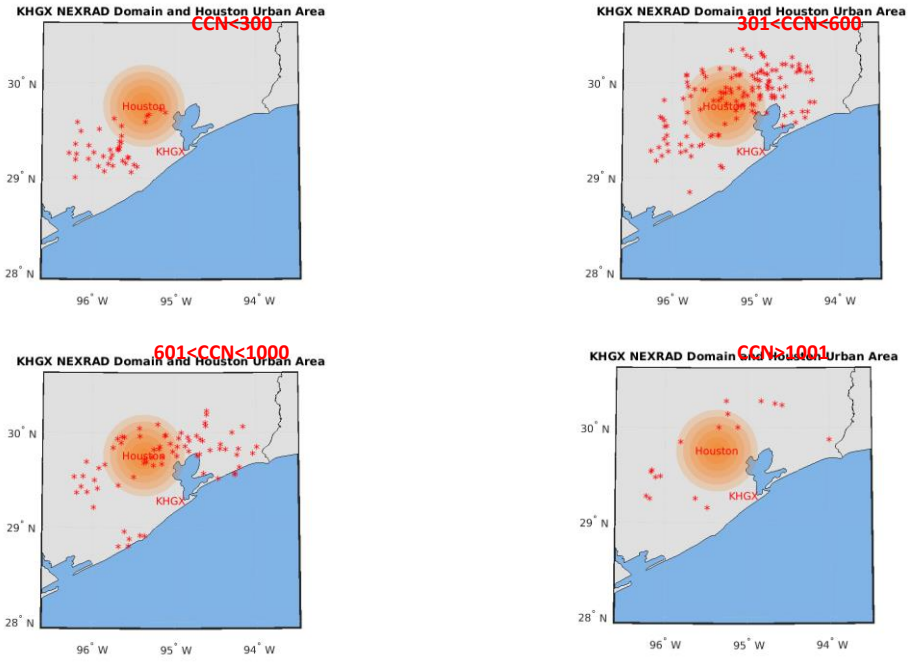

Figure 3.8 Location of selected cells for Fig. 3.7.

3.2 CCN Effect on Lightning

To study the CCN effect on lightning, we first select cloud cells that have detected LMA sources. Then the cells are composited with t_0 defined as the time of detection of the first LMA source. Fig 3.9 shows the time-height composition of the LMA source count for 4 CCN concentration intervals of <400, 400-800, 800-1200, and >1200 cm^{-3}, as shown in Figs 3.9a to 3.9d, respectively. The CCN intervals in Fig. 3.9 are different than those in Figs. 3.1-3.8 because there is a scarcity of cells with lightning at low CCN concentrations. CAPE is limited to between 2000 to 4000 J/kg and CBH is limited to between 900 to 1200m in order to keep all CCN scenarios within similar thermodynamic conditions. According to Figs. 3.9a-d, LMA source count increases with increasing CCN concentration to an optimal range of 800-1200 cm^{-3} (the average within this range is 955 cm^{-3}) and decreases when the CCN concentration exceeds 1200 cm^{-3}; additional significant results can be found in Fig 3.10 and Table 3.5. The enhanced lightning activity seen in Fig. 3.9a-c is likely due to the delayed warm rain process associated with higher CCN concentrations, which allows cloud drops to reach higher altitudes, enhancing mixed-phase processes and resulting in stronger updrafts (Rosenfeld et al., 2008). This creates favorable conditions for charge separation as non-neutral charge buildup occurs as hail/graupel collide with ice crystals in a supercooled cloud that has a strong updraft. There is a limit, however, where increasing CCN becomes associated with decreasing lightning activity as seen in the highest CCN concentration (Fig 3.9d, mean CCN = 1354 cm^{-3}). With CCN concentration exceeding optimal range (N_d >1200 cm^{-3} in this study), the

initiation of ice crystals is delayed to higher altitudes (Braga et al., 2017; Rosenfeld et al., 2011). Since the top of the mixed phase layer cannot exceed the -38°C isotherm, delayed ice crystal initiation decreases the depth of the layer where robust cloud electrification can take place. The initial location of cell samples can be found in Fig. 3.11. Variability in wind shear is not a likely alternative explanation, due to the low and similar wind shears (2.9 ms^{-1} to 3.9 ms^{-1}) for the CCN groups.

Table 3.5 Criteria Tables for all scenarios in this study for figure 6. The most left column shows the acronym of all the criteria. The top row shows the corresponding scenario and Figure number of each Table. Detailed criteria acronym is described in Table 3.2.

Criterion	CCN effect on LMA, Figs 6			
Fig panel	a	b	c	d
N of cells	19	73	33	24
HT0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
H1GW	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Gt	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
NC	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
CCN	[1 400]	[401 800]	[801 1200]	[1201 9999]
Rdis	[40 120]	[40 120]	[40 120]	[40 120]
Udis	[1 9999]	[1 9999]	[1 9999]	[1 9999]
CAPE	[2000 3700]	[2000 3700]	[2000 3700]	[2000 3700]
Hour	[16 22]	[16 22]	[16 22]	[16 22]
Reg	[1 3]	[1 3]	[1 3]	[1 3]
T0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Mean CCN	274.8	566.6	956	1426
Mean SS	0.25	0.18	0.142	0.121
Mean CAPE	3068	2978	2776	2814
Mean CBH	1037	1009	1051	1092
dBZ H 1	9	9	9	9

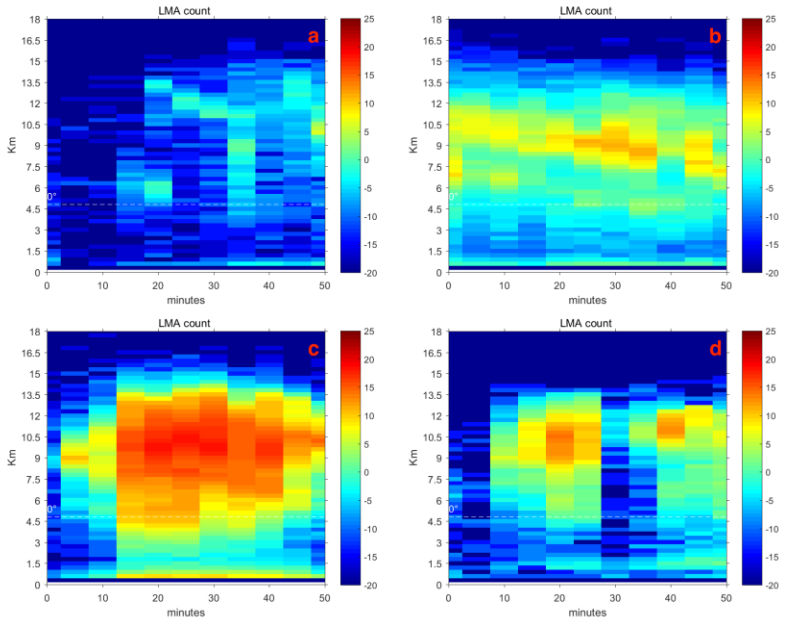

Figure 3.9 Composite time-height evolution LMA source count in radar tracked convective cells under different CCN concentrations. The color scale is ($10*\log_{10}$ (N)), where N is number of LMA sources per 5 minutes at a 300 m height interval. The composites are shown for CCN concentrations of, a: <400 cm^{-3}, b: 400-800 cm^{-3}, c: 800-1200 cm^{-3}, d: >1200 cm^{-3}., The growth-related criteria are released in all the Figures of LMA source counts. The rest criteria are identical as Fig 3.1 and 3.3. Detailed criteria is in Table 3.5.

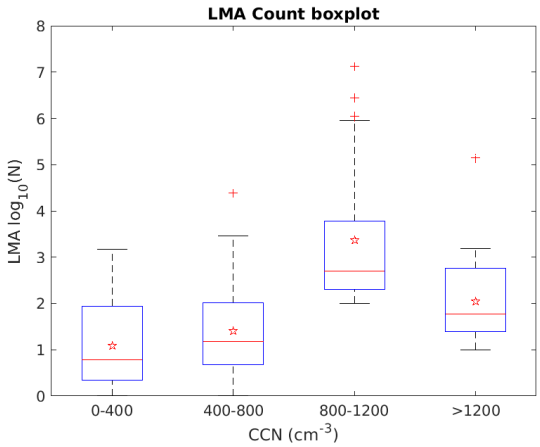

Figure 3.10 Box plot of the integrated LMA source count of sample clouds for the CCN interval samples as in Fig 3.9. P-values can be found in Table 3.5.

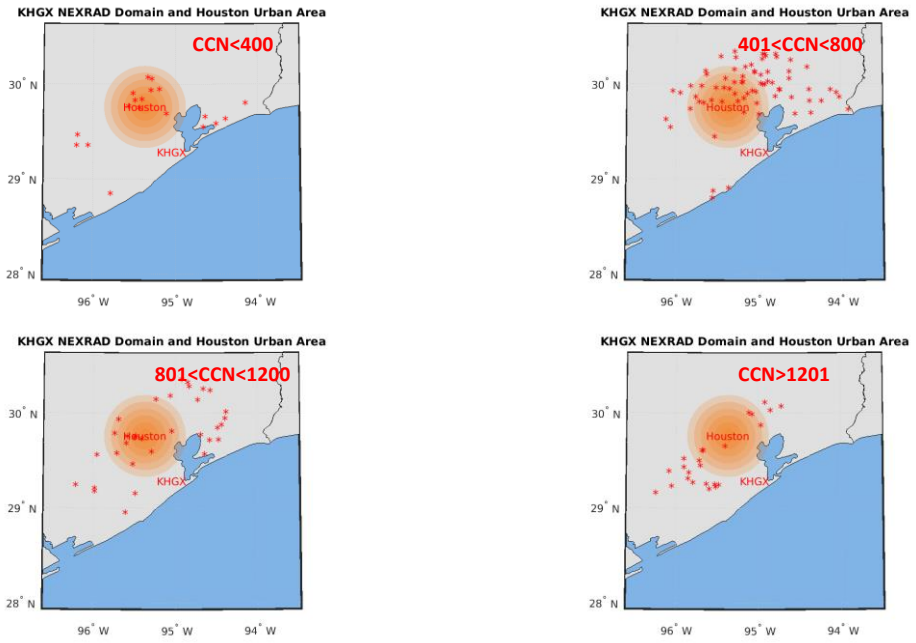

Figure 3.11 Location of selected cells for Fig. 3.9.

3.3 CCN/CAPE Effect on Lightning

The ambiguity between the roles of CCN and CAPE (i.e., the thermodynamic impact) in lightning enhancement was addressed in this study by classifying the tracked cells into low and high CCN groups occurring in low and high CAPE environments. This allows us to better separate the role that CCN and CAPE have on lightning occurrence. Fig 3.12 shows time-height composites of the LMA source counts under different CCN and CAPE combinations. On the left-hand side of Fig 3.12 (LHS, Fig 3.12a and 3.12c), the CAPE values of sample clouds are limited to less than 3000 J/kg with either low CCN (i.e., < 500 cm^{-3}; Fig. 3.12a) or high CCN (i.e., > 500 cm^{-3}; Fig. 3.12c) concentrations.

The right-hand side (RHS) of Fig 3.12 (Fig3.12b and 3.12d) has the same CCN partition but is valid for cells occurring in environments of CAPE > 3000 J/kg. By comparing the LHS and RHS of Fig 3.12 (Fig 3.12a vs 3.12c or Fig 3.12b vs 3.12d) vertically, the enhanced lightning associated with increasing CCN concentration (mean CCN increases from 360 cm^{-3} to 924 cm^{-3} from Fig 3.12a to 3.12c for example) is apparent, though it should be noted that mean CAPE decreased slightly from 2604 J/kg (Fig 3.12a) to 2376 J/kg. These observations support the idea that increasing CCN concentration enhances mixed-phase precipitation and affects lightning as discussed above.

Table 3.6 Criteria Tables for all scenarios in this study for figure 7. The most left column shows the acronym of all the criteria. The top row shows the corresponding scenario and Figure number of each Table. Detailed criteria acronym is described in Table 3.2.

Criterion	CAPE effect on LMA, Fig. 7			
Fig panel	a	b	c	d
N of cells	38	60	94	132
HT0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
H1GW	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Gt	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
NC	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
CCN	[1 500]	[1 500]	[501 9999]	[501 9999]
Rdis	[40 120]	[40 120]	[40 120]	[40 120]
Udis	[0 9999]	[0 9999]	[0 9999]	[0 9999]
CAPE	[1 3000]	[3001 9999]	[1 3000]	[3001 9999]
Hour	[16 22]	[16 22]	[16 22]	[16 22]
Reg	[1 3]	[1 3]	[1 3]	[1 3]
T0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Mean CCN	342.2	355.8	977	760.3
Mean SS	0.225	0.231	0.142	0.15
Mean CAPE	2517	3500	2197	3694
Mean CBH	902.3	969.1	936.2	887.5
dBZ H 1	9	9	9	9

On the other hand, comparing Fig 3.12 panels horizontally (Fig 3.12a vs 3.12b or Fig 3.12c vs 3.12d) elucidates the CAPE effect on lightning. For instance, Fig 3.12a and 3.12b have similar mean CCN concentration (360 cm^{-3} for Fig 3.12a and 373 cm^{-3} for Fig 3.12b), but Fig 3.12b has about 1000 J/kg higher mean CAPE. The lightning enhancement is not obvious with higher CAPE when keeping the CCN concentration similar; the hypothesis that higher CAPE enhances mixed phase process and the resultant charge separation and lightning is not seen in these data. However, Fig 3.12c and 3.12d did not show such lightning enhancement for the high CCN. Although Fig 3.12d samples have a much higher mean CAPE (3698 J/kg) than Fig 3.12c (2197 J/kg), the lightning is not enhanced, and it even decreased to some extent. This result suggests that CAPE is not important to enhancing lightning. Once CAPE is sufficient to support deep convective clouds, CCN appears to dominate the lightning activity, and these results are statistically significant, as shown in Table 3.8. The initial location of cell samples can be found in Fig. 3.13. Here again, variability in wind shear is a not likely alternative explanation, due to the low and similar wind shears (2.9 ms^{-1} to 3.9 ms^{-1}) for the four CCN and CAPE groups.

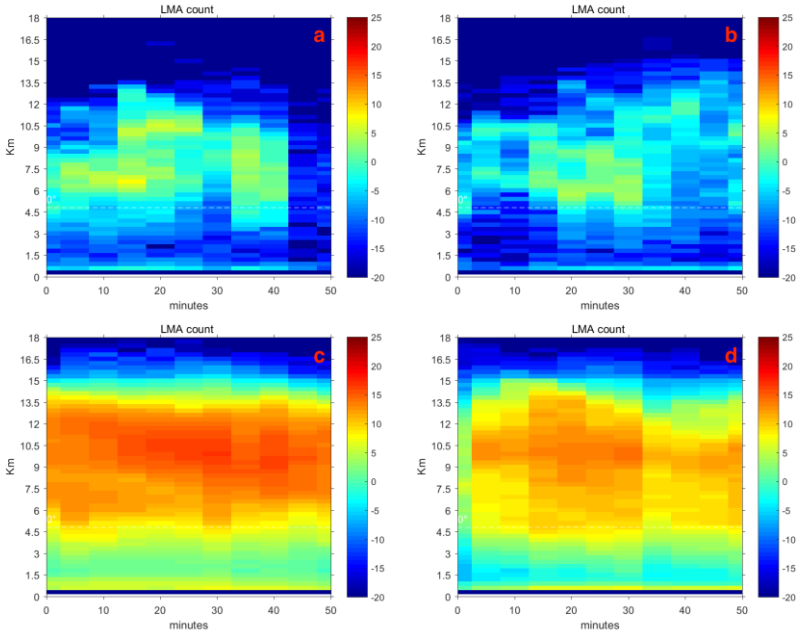

Figure 3.12 Same as Fig. 3.9, but for combined classification of CAPE and CCN. The CCN classes are - a & b: <500 cm^{-3}, c & d: >500 cm^{-3}. The CAPE classes are - a & c: 1500-3000 J/kg, b & d: >3000 J/kg. The sample clouds CAPE start with 1500 J/kg is due to the lack of sample with lightning associated with CAPE < 1500 J/kg. The rest of the criteria are as in Fig 3.9. Detailed criteria is in Table 3.6.

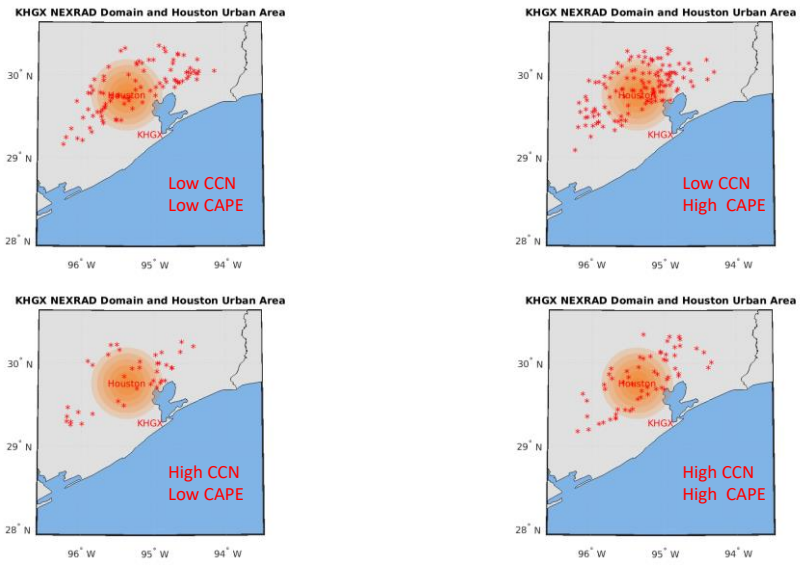

Figure 3.13 Location of selected cells for Fig. 3.12.

3.4 CCN/Urban Heat Island Effect on Lightning

Another outstanding question is the ambiguity between the impacts of urban aerosol emissions and urban island effect on lightning. Previous studies (Bourscheidt et al., 2016; Orville et al., 2001; Soriano and de Pablo, 2002; Stallins, 2004) have pointed out that urban areas tend to enhance lightning with respect to nearby rural areas because urban areas have, among differences, higher CAPE, more CCN, and more complicated surface roughness conditions. In this study, an urban area is defined as a 50-km radius circle centered at Houston city center (29.7604° N, 95.3698° W). Rural area includes the rest of the search domain centered at KHGX (Houston/Galveston, TX radar) radar site (Fig 1.1). The cloud cells over the Gulf of Mexico have been excluded for this selection because of possible additional difference between land and sea surface properties.

Table 3.7 Criteria tables for all scenarios in this study for figure 8. The most left column shows the acronym of all the criteria. The top row shows the corresponding scenario and figure number of each table. Detailed criteria acronym is described in Table 3.2.

Criterion	CCN effect between urban and rural area, Fig. 8			
Fig panel	a	b	c	d
N of cells	47	50	97	126
HT0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
H1GW	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Gt	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
NC	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
CCN	[1 500]	[1 500]	[501 9999]	[501 9999]
Rdis	[40 120]	[40 120]	[40 120]	[40 120]
Udis	[51 9999]	[1 50]	[51 9999]	[1 50]
CAPE	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Hour	[16 22]	[16 22]	[16 22]	[16 22]
Reg	[3 3]	[3 3]	[3 3]	[3 3]
T0	[-9999 9999]	[-9999 9999]	[-9999 9999]	[-9999 9999]
Mean CCN	383.7	355.1	892.6	771.3
Mean SS	0.224	0.251	0.147	0.153
Mean CAPE	3055	3220	3156	3219
Mean CBH	907.5	975.3	906.8	905.9
Mean W	0.8168	0.8778	0.8161	0.8153

The time-height evolutions of LMA source counts in rural and urban areas are compared in Fig. 3.14. The LHS panels of Fig 3.14 show the lightning in the rural area with low (Fig 3.14a, CCN < 500 cm^{-3}) and high (Fig 3.14c, CCN > 500 cm^{-3}) CCN classes. The RHS panels show the lightning source counts in the urban area with the same CCN classes. By comparing the panels vertically, i.e. Fig 3.14a vs 8c and Fig 3.14b vs 3.14d, increasing CCN concentration induced significantly more numerous lightning sources as expected from previous discussion both in the rural and urban areas. On the other hand, comparing panels horizontally (Fig 3.14a vs 3.14b and Fig 3.14c vs 3.14d) indicates that

urban area has significantly more lightning than rural area under similar CCN and CAPE conditions primarily when CCN concentration was low. The increase is small and statistically insignificant under high CCN concentrations (the significance levels can be found in Table 3.8). Lightning enhancement over urban areas has been previously attributed to the urban topography (Bornstein et al., 1990) and to urban heat island effect (Osmar et al., 2013). The urban heat island effect works by increasing the temperature, but not the absolute humidity, thus increasing cloud base heights (owing to lower mean relative humidity) and updraft strength (owing to higher temperatures). This hypothesis can be tested by comparing the cloud base heights in the urban and rural areas (Fig 3.15). The cloud base height between rural and urban areas under similar CCN and thermodynamic conditions are similar and show no statistically significant differences, which means urban heat island probably does not play a significant role here. Note that, since the urban heat island effect impacts CAPE, we reduce some of the impact of the urban heat island by grouping the data by similar CAPE. The initial location of cell samples can be found in Fig. 3.16. Wind shear is not likely an alternative explanation due to the weak and similar values for the four groups.

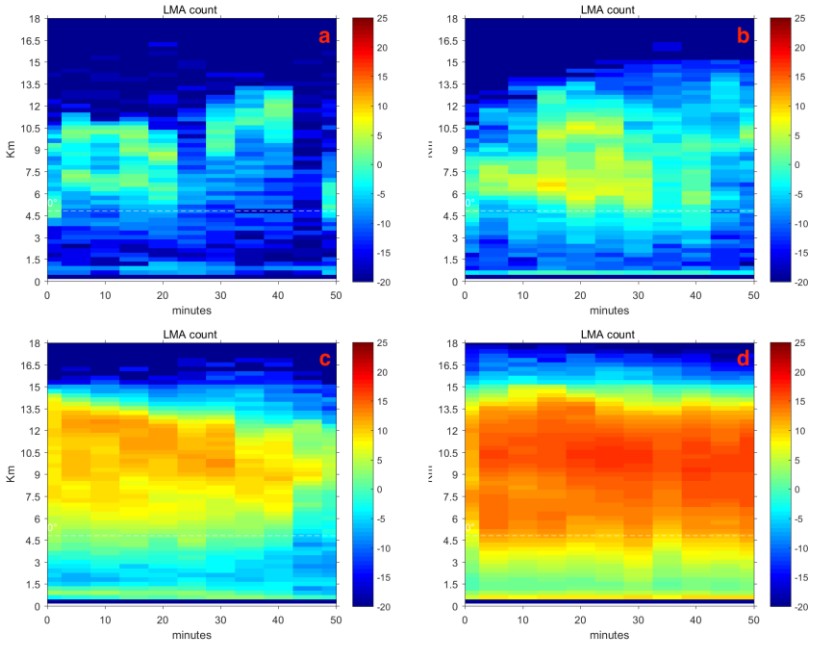

Figure 3.14 Same as Fig. 3.12, but for urban and land rural areas under low and high CCN. Panels a and b have CCN<500 cm^{-3}, and panels c and d have CCN>500 cm^{-3}. Panels a and c are over rural areas, whereas panels b and d are over urban areas. Detailed criteria is in Table 3.7.

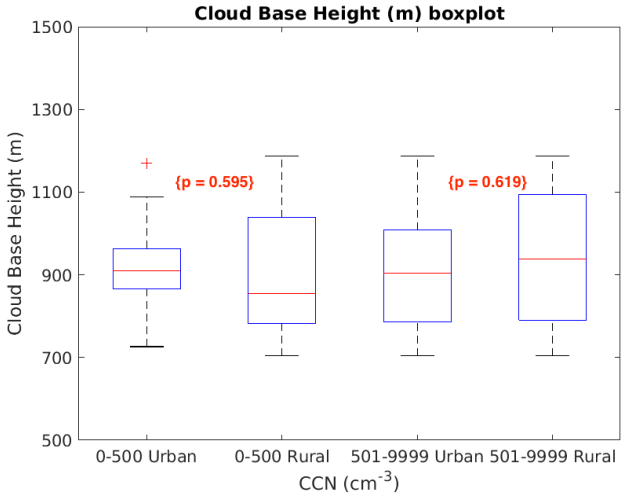

Figure 3.15 Boxplot of cloud base height of the cloud samples from Fig. 3.14. The red numbers is the P-Values calculated by the two-paired t-test of the adjacent boxplots samples assuming unequal variance.

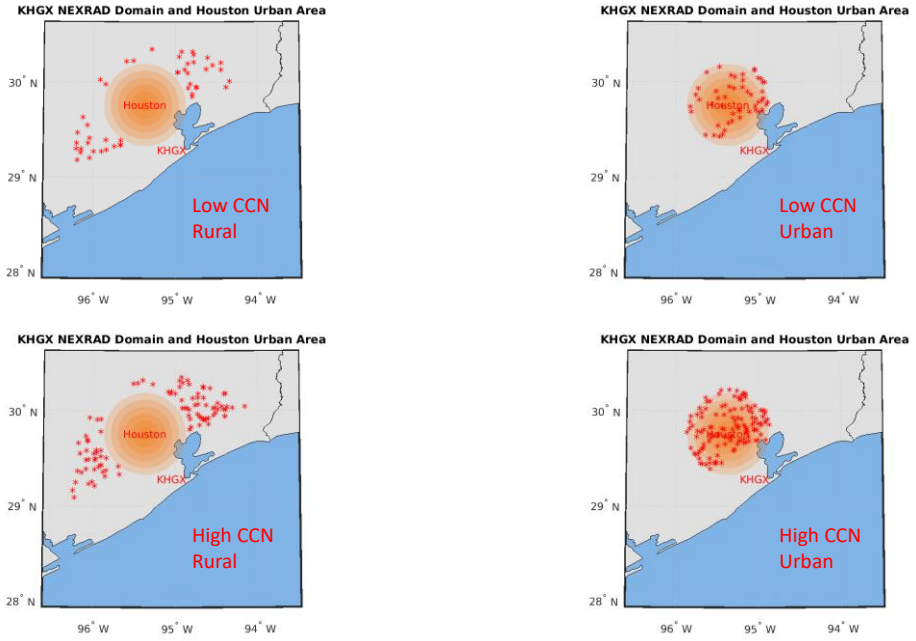

Figure 3.16 Location of selected cells for Fig. 3.14.

Low CCN concentrations over the Houston urban area are possible despite the large emission source because most of the newly formed aerosols over urban areas are ultra-fine aerosol particles (UAPs) that are too small to serve as CCN (Almeida et al., 2016; Pikridas et al., 2015). However, it was recently shown that CN can invigorate deep tropical convective clouds with low CCN concentrations (Khain et al., 2012, Fan et al., 2018). The low CCN leads to fast coalescence that reduces the integrated drop surface area and increases the vapor supersaturation (S) to more than 10% in the absence of large numbers of UAP. The added UAP (of urban origin in this case) can activate additional cloud droplets at an S of few % and limits S to lower values. This means additional

production of cloud droplets aloft along with condensation and respective latent heat release that invigorate the convection and enhance the lightning. This could be a major contributor to this phenomenon and should draw attention for future studies.

Previous studies indicated that CBH and SHEAR are important parameters in regulating lightning activities (Stolz et al., 2017; Williams et al., 2005). In this study, CBH is used as a control parameter in each above scenario to assure that all panels have similar CBH conditions. The mean CBH magnitudes for each classification can be found in Table S2. Fan et al. (2009) showed that wind shear > 10 ms^{-2} suppresses the invigoration. The wind shear is calculated as vector difference between 925 and 500 hPa winds. The SHEAR distributions are shown in Figure S4, S6, S9, S11 and S14. The median values of SHEAR for the various classifications (Table S6) are around or less than 5m/s and are considered in the weak SHEAR category. Therefore, SHEAR has little effect on the overall conclusions in this study. CBH, SHEAR and other thermodynamic parameters do affect lightning activity and polarity, but other environment than Houston provide a much larger range of these properties. The relatively small variability in this study allow focusing on the effect of CCN.

Table 3.8 Tables of P-values calculated with 2-pair t-test assuming unequal variance for Fig 3.2, 3.3, 3.4, 3.6, 3.7 and 3.8. There are 4x4 matrices that provide the comparisons between all the four panels in each figure. Fig 6 does not require P-values since the general Z_{DR} collapsing pattern are similar and expected.

Fig 2			
1.000	0.002	0.048	0.004
0.002	1.000	0.157	0.598
0.048	0.157	1.000	0.088
0.004	0.598	0.088	1.000
Fig 3			
1.000	0.001	0.007	0.127
0.001	1.000	0.823	0.928
0.007	0.823	1.000	0.970
0.127	0.928	0.970	1.000
Fig 4			
1.000	0.012	0.000	0.000
0.012	1.000	0.016	0.020
0.000	0.016	1.000	0.929
0.000	0.020	0.929	1.000
Fig 5			
1.000	0.029	0.000	0.000
0.029	1.000	0.000	0.009
0.000	0.000	1.000	0.752
0.000	0.009	0.752	1.000
Fig 6			
1.000	0.360	0.000	0.041
0.360	1.000	0.001	0.279
0.000	0.001	1.000	0.011
0.041	0.279	0.011	1.000
Fig 7			
1.000	0.000	0.000	0.000
0.000	1.000	0.580	0.088
0.000	0.580	1.000	0.244
0.000	0.088	0.244	1.000

4. CONCLUSION

The MCIT algorithm uses 2-dimensional VIL maps as input for cloud cell identification. A watershed algorithm divides VIL maps into individual cloud cells. The tracking of cloud cells between time steps is implemented by comparing the maximum common VIL, which is determined by the shifting vectors calculated from the median shifting vector from previous time steps (detailed discussion in methodology phase 3). Each cloud cell is given a unique global id number and its life cycle information from polarimetric radar, lightning, satellite retrieval and reanalysis sounding data is saved in the output data structure.

The dataset output from the MCIT algorithm is extremely useful and handy to study different convective storm cases. For instance, a combination of different convective storm cases can be used to study and validate the CCN effect by selecting convective clouds with specific CCN concentration at cloud base height and then studying their differences in cloud echo top height, radar vertical development, lightning activity etc. Thermodynamic influences may also be investigated.

The MCIT package was applied as a demonstration of different types of convective storm and several cloud examples have been analyzed with the MCIT algorithm. MCIT algorithm shows good performance in identifying and tracking all examples of convective systems, isolated or clustered. The output from the MCIT algorithm is a comprehensive and expandable dataset of cloud cell microphysical properties, which can be used to study storm structure and electrification. The time-height evolution of rain, hail and electrification was clearly evident in the tracked cells, with some potentially new insights.

The Houston area serves as a natural laboratory for identifying and examining the effects of aerosols on cloud microstructure, precipitation-forming processes and cloud electrification. The setting of abundant deep tropical clouds occurring over a major metropolis with many industrial emission sources on the shore of tropical ocean with onshore flow allows one to examine the differences between polluted and clean clouds over land and to study the large variability of aerosols from marine, rural and urban origins. A unique combination of polarimetric radar, LMA and newly developed satellite retrievals of CCN makes it possible, for the first time, to evaluate quantitatively the effects of CCN on cloud microphysical, dynamic and electrification properties. The main results show that:

- Added CCN delays the initiation of precipitation (Fig 2 to Fig 4). The physics behind this is that more CCN leads to more numerous and smaller cloud droplets and weakened coalescence, which delays the initiation of warm rain.
- Added CCN invigorate the convection as indicated by increased H_{max}. The effect saturates near CCN = 1000 cm^{-3}, as shown in Fig 5. Since CCN is completely decorrelated from CAPE in the study area, these observations are in line with the aerosol convective invigoration hypothesis (Rosenfeld et al., 2008).
- Increasing CCN from <400 cm^{-3} to the range of 800-1200 cm^{-3} is associated with an increase of lightning activity by one order of magnitude, while holding CAPE constant (Figs 6c to 6c). This is in line with the aerosol convective invigoration effect which brings more cloud water and larger updrafts to the temperature range were mixed phase precipitation and cloud electrification can occur.

- Increasing CCN further beyond 1200 cm^{-3} has the opposite effect, i.e., reducing lightning activity (Figs 6c and 6d). It is hypothesized that this is caused by increasing height of ice initiation when supercooled cloud droplets become smaller.
- Adding CAPE in high CCN environment (CCN>500 cm^{-3}) does no longer enhance further the already high lightning activity (Figs 7c and 7d). The enhanced updrafts probably push the cloud electrification beyond the optimal point by delaying the initiation of mixed phase processes to lower temperatures.
- The urban area has considerably more lightning in low CCN environment (CCN<500 cm^{-3}), but the effect fades in higher CCN concentrations (Fig. 8). There is no indication for urban heat island effect on cloud base heights. It is hypothesized that the urban ultrafine aerosols that are too small to serve as CCN at cloud base are activated into additional cloud droplets aloft and invigorate the convection by the latent heating during the added condensation. This mechanism can add also supercooled water and further enhance the cloud electrification.

CAPE is essential for the initiation of deep convection. But when CAPE is high, which means that deep convection is already in progress, aerosols dominate the lightning activity. The low magnitude of the wind shear for the cases in this study excludes it as a possible alternative explanation. The aerosols include both CCN and smaller particles, which may activate cloud droplets and invigorate the convection well above cloud base. These observations and inferences serve as an impetus for additional in situ observations of

aerosols, cloud and thermodynamic properties around Houston, for substantiating the emerging conclusions and hypotheses of this study.

REFERENCES

Almeida, G.P., Santana, M.E., Sousa, R., Godoi, R., Aircraft observations of the interactions of the manaus plume with aerosols forest during rainy season: a case study, *Ciencia Natura* 38(2016), p. 7.

Altaratz, O., Koren, I., Remer, L.A., Hirsch, E., Review: Cloud invigoration by aerosols-Coupling between microphysics and dynamics, *Atmospheric Research* 140(2014), pp. 38-60.

Andreae, M.O. *et al.*, Smoking rain clouds over the Amazon, *Science* 303(2004), pp. 1337-1342.

Bellon, A., Lovejoy, S., Austin, G.L., Combining Satellite and Radar Data for the Short-Range Forecasting of Precipitation, *Monthly Weather Review* 108(1980), pp. 1554-1566.

Bjerkaas, C.L., Forsyth, D.E., Real-Time Automated Tracking of Severe Thunderstorms Using Doppler Weather Radar, *Bulletin of the American Meteorological Society* 60(1979), pp. 533-533.

Bang, S.D., Zipser, E.J., Seeking reasons for the differences in size spectra of electrified storms over land and ocean, *Journal of Geophysical Research-Atmospheres* 121(2016), pp. 9048-9068.

Blanchard, D.O., Assessing the vertical distribution of convective available potential energy, *Weather Forecast* 13(1998), pp. 870-877.

Bourscheidt, V., Pinto, O., Naccarato, K.P., The effects of Sao Paulo urban heat island on lightning activity: Decadal analysis (1999-2009), *Journal of Geophysical Research-Atmospheres* 121(2016), pp. 4429-4442.

Braga, R.C. *et al.*, Further evidence for CCN aerosol concentrations determining the height of warm rain and ice initiation in convective clouds over the Amazon basin, *Atmospheric Chemistry and Physics* 17(2017), pp. 14433-14456.

Corporation, T.E.D., Petroleum Refining & Chemical Products. (2018).

Crane, R.K., Automatic Cell Detection and Tracking, *Ieee Transactions on Geoscience and Remote Sensing* 17(1979), pp. 250-262.

Dixon, M., Wiener, G., Titan - Thunderstorm Identification, Tracking, Analysis, and Nowcasting - a Radar-Based Methodology, *Journal of Atmospheric and Oceanic Technology* 10(1993), pp. 785-797.

Emanuel, K.A., Overview of atmospheric convection, *Nato Adv Sci I C-Mat* 505(1997), pp. 1-28.

Gunn, R., Phillips, B.B., An Experimental Investigation of the Effect of Air Pollution on the Initiation of Rain, *J Meteorol* 14(1957), pp. 272-280.

Greene, D.R., Clark, R.A., Vertically Integrated Liquid Water - New Analysis Tool, *Monthly Weather Review* 100(1972), pp. 548-&.

Handwerker, J., Cell tracking with TRACE3D - a new algorithm, *Atmospheric Research* 61(2002), pp. 15-34.

Joe, P. *et al.*, The S2K severe weather detection algorithms and their performance, *Weather and Forecasting* 19(2004), pp. 43-63.

Johnson, J.T. *et al.*, The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm, *Weather and Forecasting* 13(1998), pp. 263-276.

Khain, A., Rosenfeld, D., Pokrovsky, A., Aerosol impact on the dynamics and microphysics of deep convective clouds, *Quarterly Journal of the Royal Meteorological Society* 131(2005), pp. 2639-2663.

Khain, A.P., BenMoshe, N., Pokrovsky, A., Factors determining the impact of aerosols on surface precipitation from clouds: An attempt at classification, *J Atmos Sci* 65(2008), pp. 1721-1748.

Kumjian, M.R., Ryzhkov, A.V., The Impact of Size Sorting on the Polarimetric Radar Variables, *Journal of the Atmospheric Sciences* 69(2012), pp. 2042-2060.

Lakshmanan, V., Smith, T., Stumpf, G., Hondl, K., The Warning Decision Support System-Integrated Information, *Weather and Forecasting* 22(2007), pp. 596-612.

Latham, J., The Electrification of Thunderstorms, *Quarterly Journal of the Royal Meteorological Society* 107(1981), pp. 277-298.

Li, L., Schmid, W., Joss, J., Nowcasting of Motion and Growth of Precipitation with Radar over a Complex Orography, *Journal of Applied Meteorology* 34(1995), pp. 1286-1300.

Ortega, K.L., Krause, J.M., Ryzhkov, A.V., Polarimetric Radar Characteristics of Melting Hail. Part III: Validation of the Algorithm for Hail Size Discrimination, *Journal of Applied Meteorology and Climatology* 55(2016), pp. 829-848.

Orville, R.E. *et al.*, Enhancement of cloud-to-ground lightning over Houston, Texas, *Geophysical Research Letters* 28(2001), pp. 2597-2600.

Park, H., Ryzhkov, A.V., Zrnic, D.S., Kim, K.E., The Hydrometeor Classification Algorithm for the Polarimetric WSR-88D: Description and Application to an MCS, *Weather and Forecasting* 24(2009), pp. 730-748.

Pikridas, M. *et al.*, In situ formation and spatial variability of particle number concentration in a European megacity, *Atmospheric Chemistry and Physics* 15(2015), pp. 10219-10237.

Pozo, R., Template numerical toolkit for linear algebra: High performance programming with C++ and the standard template library, *Int J Supercomput Ap* 11(1997), pp. 251-263.

Project, N.C.G.C., NAM Data Assimilation System. (2004).

Rinehart, R.E., Garvey, E.T., 3-Dimensional Storm Motion Detection by Conventional Weather Radar, *Nature* 273(1978), pp. 287-289.

Reynolds, S.E., Brook, M., Gourley, M.F., Thunderstorm Charge Separation, *J Meteorol* 14(1957), pp. 426-436.

Rosenfeld, D. *et al.*, Global observations of aerosol-cloud-precipitation-climate interactions, *Reviews of Geophysics* 52(2014), pp. 750-808.

Rosenfeld, D., Lensky, I.M., Satellite-based insights into precipitation formation processes in continental and maritime convective clouds, *Bulletin of the American Meteorological Society* 79(1998), pp. 2457-2476.

Rosenfeld, D. *et al.*, Flood or drought: How do aerosols affect precipitation?, *Science* 321(2008), pp. 1309-1313.

Rosenfeld, D. *et al.*, Glaciation temperatures of convective clouds ingesting desert dust, air pollution and smoke from forest fires, *Geophysical Research Letters* 38(2011).

Rosenfeld, D. *et al.*, Satellite retrieval of cloud condensation nuclei concentrations by using clouds as CCN chambers, *Proceedings of the National Academy of Sciences of the United States of America* 113(2016), pp. 5828-5834.

Rosenfeld, D., Objective Method for Analysis and Tracking of Convective Cells as Seen by Radar, *Journal of Atmospheric and Oceanic Technology* 4(1987), pp. 422-434.

Rosenfeld, D. *et al.*, Satellite retrieval of cloud condensation nuclei concentrations by using clouds as CCN chambers, *Proceedings of the National Academy of Sciences of the United States of America* 113(2016), pp. 5828-5834.

Ryzhkov, A., Diederich, M., Zhang, P.F., Simmer, C., Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking, *J Atmos Ocean Tech* 31(2014), pp. 599-619.

Sachidananda, M., Zrnić, D.S., Rain Rate Estimates from Differential Polarization Measurements, *Journal of Atmospheric and Oceanic Technology* 4(1987), pp. 588-598.

Soriano, L.R., de Pablo, F., Effect of small urban areas in central Spain on the enhancement of cloud-to-ground lightning activity, *Atmospheric Environment* 36(2002), pp. 2809-2816.

Stallins, J.A., Characteristics of urban lightning hazards for Atlanta, Georgia, *Climatic Change* 66(2004), pp. 137-150.

Stolz, D.C., Rutledge, S.A., Pierce, J.R., Simultaneous influences of thermodynamics and aerosols on deep convection and lightning in the tropics, *Journal of Geophysical Research-Atmospheres* 120(2015), pp. 6207-6231.

Takahashi, T., Riming Electrification as a Charge Generation Mechanism in Thunderstorms, *J Atmos Sci* 35(1978), pp. 1536-1548.

Tao, W.K., Chen, J.P., Li, Z.Q., Wang, C., Zhang, C.D., Impact of Aerosols on Convective Clouds and Precipitation, *Reviews of Geophysics* 50(2012).

Tao, W.K. *et al.*, Role of atmospheric aerosol concentration on deep convective precipitation: Cloud-resolving model simulations, *Journal of Geophysical Research-Atmospheres* 112(2007).

Testud, J., Oury, S., Black, R.A., Amayenc, P., Dou, X.K., The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing, *J Appl Meteorol* 40(2001), pp. 1118-1140.

Tuttle, J.D., Foote, G.B., Determination of the Boundary-Layer Air-Flow from a Single Doppler Radar, *Journal of Atmospheric and Oceanic Technology* 7(1990), pp. 218-232.

van den Heever, S.C., Carrio, G.G., Cotton, W.R., DeMott, P.J., Prenni, A.J., Impacts of nucleating aerosol on Florida storms. Part I: Mesoscale simulations, *J Atmos Sci* 63(2006), pp. 1752-1775.

van den Heever, S.C., Stephens, G.L., Wood, N.B., Aerosol Indirect Effects on Tropical Convection Characteristics under Conditions of Radiative-Convective Equilibrium, *J Atmos Sci* 68(2011), pp. 699-718.

Williams, E.R., Large-Scale Charge Separation in Thunderclouds, *Journal of Geophysical Research-Atmospheres* 90(1985), pp. 6013-6025.

Williams, E. *et al.*, Contrasting convective regimes over the Amazon: Implications for cloud electrification, *Journal of Geophysical Research-Atmospheres* 107(2002).

Williams, E.R., The Electrification of Thunderstorms, *Sci Am* 259(1988), pp. 88-&.

Williams, E.R. *et al.*, A Radar and Electrical Study of Tropical Hot Towers, *J Atmos Sci* 49(1992), pp. 1386-1395.

Williams, E.R., Satori, G., Lightning, thermodynamic and hydrological comparison of the two tropical continental chimneys, *J Atmos Sol-Terr Phy* 66(2004), pp. 1213-1231.

Wilson, J.W. *et al.*, Sydney 2000 Forecast Demonstration Project: Convective storm nowcasting, *Weather and Forecasting* 19(2004), pp. 131-150.

Yoshida, S., Morimoto, T., Ushio, T., Kawasaki, Z., A fifth-power relationship for lightning activity from Tropical Rainfall Measuring Mission satellite observations, *Journal of Geophysical Research-Atmospheres* 114(2009).