January 2017

Detailed Analysis Of Convective Simulations From A WRF Microphysics Ensemble

Joshua Markel

DETAILED ANALYSIS OF CONVECTIVE SIMULATIONS FROM A WRF MICROPHYSICS ENSEMBLE

by

Joshua Luke Markel
Bachelor of Science, Pennsylvania State University, 2014

Thesis
Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

August
2017

This thesis, submitted by Joshua Luke Markel in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

.

_____
Dr. Matthew S. Gilmore

_____
Dr. Xiquan Dong

_____
Dr. Aaron Kennedy

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the Graduate School at the University of North Dakota and is hereby approved.

_____
Grant McGimpsey
Dean of the School of Graduate Studies

_____
Date

PERMISSION

Title Detailed Analysis of Convective Simulations from a WRF Microphysics Ensemble

Department Atmospheric Science

Degree Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of the University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor why supervised my thesis work or, in his absence, by the chairperson of the department of the dean of the Graduate School. It is understood that any copying or publication or other use of the thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Joshua Markel
7/27/2017

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES x

ACKNOWLEDGEMENTS xi

ABSTRACT xii

CHAPTER

I. INTRODUCTION 1

Motivation 1

Objectives 2

MCS Structure and Intensity 2

Convection Allowing Models 3

Microphysics Parameterization and Simulation within CAM 4

Past Studies: Sensitivity to Ice Species 6

Past Studies: Sensitivity to Number of Predicted Moments 9

Past Studies from the NSSL Spring Experiment 10

Observations: Stage IV Precipitation Data 12

Verification: MODE 15

II. METHODOLOGY 17

WRF Model Ensemble 17
Model Configuration and Domain Setup 18
Initialization / Selection of Cases 19
Object Tracking 22
MODE Convolution Thresholding and Sensitivity Test 24
Attributes of MODE Objects 25
III. RESULTS 27
Overview 27
Number of Detected Objects 28
Detected Object Size Distribution 31
Detected Object Size with Time 33
Initiation / Dissipation 35
Duration 37
Percentile [Rain] Intensity 38
Velocity 49
Analysis Summary 51
IV. DISCUSSION 54
Implications to Forecasting 54
Overprediction of Detected Objects 55
Microphysics Performance 57

Limitations / Weaknesses 59

Future Work 60

V. CONCLUSIONS 61

REFERENCES CITED 63

APPENDICES 71

LIST OF FIGURES

Figure Page

1. Example of the budget interpolation method. Solid lines and black triangles 14

2. Example of the NCAR DTC Method for Object-based Diagnostic Evaluation 16

3. Regions of interest (red boxes) that storms had to be located within to be 21

4. (a) Number of detected (b) percent of total detected objects ((# of objects at 30

5. Distribution of (a) the number or (b) percentage of detected objects ((# of 32

6. Distribution of (a) the number or (b) percentage of detected objects ((# of 33

7. Average size (square root of area, expressed in km) of detected objects as a 34

8. The number of tracked objects that were a) newly detected and b) discarded 36

9. Number of tracked objects (a) and percentage of tracked objects ((# of 38

10. Tracked object 10th percentile precipitation intensities sorted by object 41

11. Tracked object 25th percentile precipitation intensities sorted by object 41

12. Tracked object 50th percentile precipitation intensities sorted by object. 42

13. Tracked object 75th percentile precipitation intensities sorted by object 42

14. Tracked object 90th percentile precipitation intensities sorted by object 43

15. Tracked objects averaged precipitation percentile intensities from Figure 10 43

16. A MODE object for each microphysics ensemble member and Stage IV with 45

17. Tracked object 10th percentile precipitation intensities sorted by object 45

18. Tracked object 25th percentile precipitation intensities sorted by object..... 46

19. Tracked object 50th percentile precipitation intensities sorted by object..... 46

20. Tracked object 75th percentile precipitation intensities sorted by object..... 47

21. Tracked object 90th percentile precipitation intensities sorted by object..... 47

22. Tracked objects averaged velocity for the u-component (a), v-component 51

C1. (a) Number of detected (b) percent of total detected objects ((# of objects at 79

C2. Distribution of (a) the number or (b) percentage of detected objects ((# of 80

C3. Distribution of (a) the number or (b) percentage of detected objects ((# of 80

C4. Average size (square root of area, expressed in km) of detected objects as a 81

C5. The number of tracked objects that were a) newly detected and b) discarded 82

C6. Number of tracked objects (a) and percentage of tracked objects ((# of 83

C7. Tracked object 10th percentile precipitation intensities sorted by object..... 83

C8. Tracked object 25th percentile precipitation intensities sorted by object..... 84

C9. Tracked object 50th percentile precipitation intensities sorted by object..... 84

C10. Tracked object 75th percentile precipitation intensities sorted by object..... 85

C11. Tracked Object 90th percentile precipitation intensities sorted by object 85

D1. (a) Number of detected (b) percent of total detected objects ((# of objects at 86

D2. Distribution of (a) the number or (b) percentage of detected objects ((# of..... 86

D3. Distribution of (a) the number or (b) percentage of detected objects ((# of..... 87

D4. Average size (square root of area, expressed in km) of detected objects as a 87

D5. The number of tracked objects that were a) newly detected and b) discarded 88

D6. Number of tracked objects (a) and percentage of tracked objects ((# of 89

D7. Tracked object 10th percentile precipitation intensities sorted by object 89

D8. Tracked object 25th percentile precipitation intensities sorted by object 90

D9. Tracked object 50th percentile precipitation intensities sorted by object 90

D10. Tracked object 75th percentile precipitation intensities sorted by object 91

D11. Tracked Object 90th percentile precipitation intensities sorted by object 91

F1. Number of tracked objects as a function of object duration. Note that the y 94

G1. (a) Number of detected (b) percent of total detected objects ((# of objects at 95

G2. (a) Number of detected (b) percent of total detected objects ((# of objects at 95

G3. Average size (square root of area, expressed in km) of detected objects (only 96

G4. Average size (square root of area, expressed in km) of detected objects (only 96

LIST OF TABLES

Table Page

1. List of microphysics ensemble members, their predicted variables (including.....18

2. List of MODE-TD diagnosed detected objects across all forecast hours and case.....29

3. List of object averaged percentile precipitation intensities (mm h^{-1}) across all 39

4. List of object averaged percentile precipitation intensities (mm h^{-1}) across all 39

5. List of object averaged percentile precipitation intensities (mm h^{-}) across all.....40

6. List of object averaged percentile precipitation intensities (mm h^{-1}) across all 40

7. Best performing microphysics for each MODE-TD characteristic. Checkmarks 53

A1. List of the WRF namelist variables and the settings used for the retrospective.....72

B1. Chronological listing of the reduced retrospective case dates with one or 77

E1. Detected object number for Stage IV (St4), Thompson (Tho), Milbrandt (Mil).....92

E2. Detected object percent increase relative to Stage IV for Thompson (Tho) 93

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Matthew Gilmore, who has been extremely supportive and provided ample guidance throughout the entire thesis and graduate school process. Without the help of Dr. Gilmore, I would have not been able to reach the completion of my thesis. I would also like to thank my committee members, Xiquan Dong and Aaron Kennedy, for giving me the opportunity to work on funded research. Additionally, I much appreciate their feedback and constructive criticism throughout the process. I am also grateful for the all of the faculty and graduate students of the University of North Dakota Atmospheric Science Department for making my experience enjoyable and educational. Lastly, I want to thank my parents, Jim and Jan Markel, for supporting me throughout my life and providing me with the means needed to pursue my educational and career aspirations.

Funding for this work was provided by NOAA through the grant number NA15NWS4680004. The simulation runs were completed using the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation grant number ACI-1053575 (Towns et al. 2014). Stage IV observational data was provided by NCAR/EOL under sponsorship of the National Science Foundation. The MODE analysis tool used for analysis was included in the package of MET. Model Evaluation Tools (MET) was developed through a grant from the United States Air Force Weather Agency (AFWA).

ABSTRACT

This study performed a detailed analysis of convective storms across the Contiguous United States from 77 case dates using a 4-member microphysics Weather Research and Forecasting (WRF) ensemble and Stage IV gauge-adjusted radar derived precipitation. Dates included the 2016 NOAA Spring Forecast Experiment (SFE) with the remainder from 2010-2012. Quantitative attributes of precipitation objects in both simulations and observations were diagnosed using the Method of Object-Based Diagnostic Evaluation Time-Domain (MODE-TD). The microphysics schemes tested were WSM6, Thompson, Morrison, and Milbrandt.

Among all simulation case dates, compared to observations, the number of precipitation objects less than 90 km in length are overpredicted, with the WSM6 scheme greatest and Morrison scheme least. All simulation members also generally initiate and dissipate precipitation objects too early. For precipitation rates, the Morrison scheme predicts them best while the Milbrandt and WSM6 schemes overpredict the strongest rates. The microphysics biases found within this study should aid in the prediction of convective events.

CHAPTER I

INTRODUCTION

Motivation

Convective storms pose a major threat to American society including flooding, hail, damaging winds, and tornadoes. Each year billions of dollars in property damage and loss of life result. Just in 2013, eight out of nine weather-related billion-dollar losses resulted from convective storms and flooding (NOAA 2017). Many of these losses occur due to flash flooding - the largest weather-related cause of property damage (Rauber et al. 2008).

Squall lines are one type of mesoscale convective system (MCS) that can produce flooding rains. MCSs are a difficult weather phenomenon to forecast as they vary significantly in structure, time, and intensity. Further improvements in storm-scale forecasting are needed to better predict convective events and provide ample warning time to the public.

Within the last 10 years, the National Weather Service and Storm Prediction Center have been using numerical weather prediction (NWP) models to guide their convective forecasts. The precipitation processes within NWP models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008), are represented by so-called microphysics parameterization schemes (or just "microphysics schemes"), which come in varying levels of complexity. Approximately 20 microphysics schemes are available for use in WRF and most have not been extensively tested in their ability to represent important properties of convection. For those that have, testing has been limited such that scientists have not agreed

upon a best performing scheme to use operationally. To gain a better understanding of the microphysics scheme and its role in simulating convective properties, more than a few convective forecasts need to be tested and analyzed.

Objectives

The main objective of this study was to analyze the impacts microphysics parameterization, within high resolution models, have on simulated convective properties, focusing on multi-cellular thunderstorms and MCSs, in hopes of aiding operational forecasting. Specifically, the focuses included determining microphysics biases and a best performing microphysics scheme. Previous literature has investigated the role of various hydrometeors and predictive properties within microphysics parameterization, but has only done so for a limited amount of cases and atmospheric conditions. To further bridge the gap and possibly reiterate findings from past studies, this study performed an extensive analysis of a multitude of convective simulations spanning multiple warm seasons utilizing a microphysics ensemble. The analysis performed uses an object based performance metric comparing the microphysics ensemble simulations to radar derived precipitation observational data in hopes of determining biases within microphysics to aid the forecaster in prediction of convective events.

MCS Structure and Intensity

MCSs have been found to have two regions, convective and stratiform, which differ in their precipitation processes (Houze 1977; McAnelly and Cotton 1989; Houze et al. 1990). The convective region consists of strong cumulus-scale updrafts supporting development of a

diverse set of hydrometeors such as hail, graupel, rain and snow. Convective updrafts experience net condensational heating at all levels due to the freezing and condensing associated with hydrometeor formation and growth (Houze 2004). Air behind the convective region ascends more slowly in the stratiform region, and because of these weaker updrafts, precipitation there consists primarily of snow that melts into rain. Descent in both the convective and stratiform regions is aided by melting and evaporational cooling of falling hydrometeors (Houze 2004). Thus, the microphysical processes associated with hydrometeors play a significant role in updrafts and downdrafts, which in turn impact MCS structure and intensity.

Convection-Allowing Models

The WRF model utilized can be run as a "convection-allowing model" (CAM), meaning that the model can move air and associated precipitation within individual storm updrafts and downdrafts, albeit coarsely. As such, these models must use microphysics schemes. CAMs are commonly used in storm-scale forecasting, as they are capable of simulating a multitude of convective storms, such as supercells, multicells, and MCSs. Fowle et al. (2003) also found that CAMs are capable of accurately predicting the initiation and organizational mode of convective systems 36-48 h into the forecast. However, many past studies have shown that CAM forecasts are sensitive to the microphysics within the simulated convection.

When WRF is run at a grid spacing ($\Delta x < 4$ km), it is said to be a CAM because this grid spacing is fine enough to explicitly resolve deep convection to produce weather forecasts for large areas of the United States (Kain 2006). It is well known that to coarsely resolve a wave

within a model, at least five or six gridpoints are needed meaning that storm features being simulated need to be at least 20 km across. Thus, it is no surprise that Weisman et al. (1997) also suggested that grid spacing as coarse as 4 km is sufficient enough to declare a model to be a CAM as it would resolve basic mesoconvective circulations and net momentum and heat transports of convective systems in the mid-latitudes.

In recent history, CAMs have replaced non-CAMs that contained "convective parameterization schemes" (CPSs). Years ago when computational power lacked, WRF along with other NWP models were run at larger horizontal grid spacing and were unable to resolve individual storm updrafts. Thus, CPSs were used to parameterize the small scale updrafts associated with convective storms. CPSs adjust the environmental temperature and moisture profiles, create cloud, and precipitate out rainfall all within one model timestep in attempt to model the modifications convective storms have on the atmosphere (Straka 1994 and Stensrud 2007).

Microphysics Parameterization and Simulation within the CAM

Microphysics includes all processes that occur on the scale of hydrometeors (order of 10^{-3} to 10^{-2} m) located within clouds (Warner 2011, p. 121). However, because the forecast model grid spacing is, on order, 100 to 1000 times larger than individual precipitation particles, and because forecast models need to finish prior to the event occurring, simplifying assumptions are made to parameterize these sub-grid processed including that a smaller number of microphysical properties be predicted and diagnosed. More sophisticated schemes,

while generally slower to run, treat microphysical attributes or processes more realistically, such as including important species or predicting more moments of the size distribution for each species.

Microphysics representation is a primary factor affecting the skill of forecasted precipitation properties (Warner 2011, p. 121). As previously stated, microphysics schemes vary in their representation of hydrometeors and the affiliated microphysics processes. Specifically, schemes vary in the number of predicted hydrometeor species, number of moments predicted (e.g., mass mixing ratio and total number concentration), assumptions in parameterizing the microphysical processes, and in bulk particle properties. Hydrometeor species are the particle types involved in microphysical processes (e.g., cloud droplets, rain, cloud ice, snow, graupel, and hail). Bulk particle properties include mathematical descriptions of how the particle shapes, mass, and fall speeds vary across the size distribution. These relationships allow for deriving an analytical solution representing given microphysical processes that can be diagnosed and evaluated in one model time step for all particles in a model grid cell. Microphysical processes include: condensation, accretion (collision and coalescence in warm rain processes), evaporation or freezing of liquid drops, melting of ice, aggregation of ice crystals; accretion of smaller crystals by larger ice particles (snow aggregates or hail), and vapor deposition (Warner 2011, p. 123). In schemes where total number concentration is also predicted, the processes of droplet breakup and ice crystal shattering can also be parameterized. Microphysical parameterizations commonly used in operational models use the bulk method for size distributions and include either one or two prognostic categories for each hydrometeor represented in the microphysics scheme. The bulk method assumes a

functional form for the size distribution of each particle type. This allows for computational efficiency while maintaining an accurate assumption of the size distribution. Single-moment schemes have one prognostic variable for each hydrometeor species. The prognostic variable is commonly the mass mixing ratio. In addition to the mass mixing ratios, double-moment schemes also typically predict the species' total particle number concentration (Warner 2011, p. 125).

Past Studies: Sensitivity to Ice Species

Previous studies, such as Fovell and Ogura (1988), Gilmore et al. (2004a), Gilmore et al. (2004b), Morrison and Milbrandt (2011), and Wu et al. (2013) investigated the sensitivity of simulated convective simulations to the representation and inclusion of ice species. The inclusion of ice species into microphysical schemes began in the 1980s (Lin et al. 1983; Cotton et al. 1982, 1983; and Rutledge and Hobbs, 1984) and continues to this day. Studies followed in the late 1980s investigating the sensitivity of precipitation processes in convection to ice species (Fovell and Ogura 1988; Tao and Simpson, 1989; and others).

Wu et al. (2013) looked into the impact of ice hydrometeors on two WRF simulated squall lines across the Southern Great Plains. They found that the simulated differences from two ice and three ice schemes were prominent in convective/stratiform areal coverage. Schemes with only two ice species (hereafter 2-ice schemes) produced less convective coverage and more stratiform coverage. This resulted from 2-ice schemes containing more particles lower in density, allowing smaller particles to spread to the stratiform region rather than falling

out as precipitation in the convective region. Fovell and Ogura (1988) were among the first to discover the inclusion of ice species, particularly snow, in the microphysics resulted in a more realistic looking trailing portion of the storm and more widespread precipitation. They also found that melting hail/graupel was the largest source of rain for a simulated 2D squall line. In addition, Wu et al. (2013) found the scheme containing hail produced the most precipitation and highest peak reflectivity in the convective region, but underpredicted peak reflectivity in the stratiform region. This was attributed to a larger amount of denser ice hydrometeors and less overall in-cloud hydrometeors, resulting in less advection of hydrometeors into the stratiform region.

Gilmore et al. (2004a) tested the sensitivity of the inclusion of ice species in the microphysics scheme to storm morphology and precipitation characteristics. They showed the microphysics scheme that included ice species produced ~40% more total ground-accumulated precipitation after 2 h compared to liquid only microphysics. Gilmore et al. (2004b) was the first to explore the impact of particle species assumptions in a microphysics scheme. In their study, they performed idealized supercell simulations using a simple liquid-ice microphysics scheme similar to the scheme used in Lin et al. (1983). They varied the "large ice" concentration intercept and particle mass density to study model sensitivity to whether the "large ice" was more representative of graupel or hail. On one end of the parameter space was low density, numerous, and smaller "larger ice" particles (similar to graupel) and on the other end was high density, few, and giant particles (similar to large hailstones). For the small graupel cases (large hail intercept parameter and/or small particle density) hail mass mixing ratio (q_h) growth rates and updrafts were larger than the large hail cases (small hail intercept parameter

and/or large particle density). In addition, time and domain average q_h with height was larger at higher altitudes as one moved from large hail to small graupel. Despite the larger production rates aloft, Gilmore et al. (2004b) showed that large hail cases produced two to four times more accumulated rainfall at the surface compared to small graupel cases. This was attributed to the small graupel cases having low fallout velocities that led to less ground accumulation and more horizontal spread of graupel particles aloft.

Morrison and Milbrandt (2011) also studied representation of ice species in microphysics, but the focus was on comparing the two different double-moment schemes. Despite the similarities of the two parameterizations compared, the baseline simulations differed considerably in terms of storm dynamics, reflectivity structure, surface precipitation, and cold pool characteristics. Similar to Gilmore et al. (2004b), their study also showed the large sensitivities in the assumptions made about the single category (hail and graupel), rimed ice hydrometeor. Differences in simulated cold pool strengths and downdrafts were found to be primarily due to assumptions with fall speed parameters for different species. Also, cold pool and downdraft characteristics were sensitive to raindrop breakup for the Morrison scheme. In contrast, the Milbrandt scheme generally produced smaller raindrop sizes making the choice of drop breakup parameterization less significant. The large differences produced by Morrison and Milbrandt for a simulated thunderstorm, highlighted how the different assumptions in those schemes lead to forecast uncertainties. In particular, most of the differences were in ice particle properties (density, fall-out speeds, and rain-drop break-up or hail/graupel break-up). Morrison and Milbrandt (2011) argued that there is a critical need for

further observational studies investigating the ice-particle properties and PSDs to better understand microphysics and validate microphysics schemes.

Past Studies: Sensitivity to the Number of Predicted Moments

Previous studies have shown simulated convective properties are strongly sensitive to whether only a one moment (1M) of the species is predicted (typically mass mixing ratio) or whether two moments (2M) are predicted (typically total number concentration). Though the overall storm structure was similar in simple 2-D simulations, Morrison et al. (2009) argued that 2M schemes were better than 1M schemes in producing a squall line storm's stratiform region compared to actual squall line observations that are typically observed. The 1M scheme lacked hydrometeors and precipitation in trailing stratiform regions. The reason for differences was attributed to variability in the simulated updrafts between the schemes. Increased evaporation rates in the convective region occurred in the 2M compared to the 1M, which resulted in weaker updrafts and more detrainment from the convective to stratiform region.

Bryan and Morrison (2012) findings were similar to Morrison et al. (2009), but expanded to the investigation of convective sensitivities due to representation of ice species. They compared the Morrison 1M and 2M schemes for a simulated squall line, but also looked at differences when incorporating hail versus graupel in the microphysics scheme, similar to Gilmore et al (2004b). The 2M scheme again was claimed to represent the storm structure better than the 1M scheme because of its ability to represent a varying size distribution intercept parameter, as it would in actual storms. More evaporation in the convective region

and less evaporation in the stratiform region for the 2M scheme was argued to be more realistic. They also found, similar to Gilmore et al. (2004b) that surface precipitation and cold pool strength were sensitive to representation of graupel and hail in the microphysics scheme. The inclusion of graupel led to a wider, but weaker reflectivity in the convective region, as previously found in Gilmore et al. (2004b). When hail was included instead of graupel, storms were found to have larger reflectivity, but a narrower region of peak reflectivity. In addition, Bryan and Morrison (2012) performed additional sensitivity tests using the hail scheme, but with much slower graupel-like fall speeds. They found that assumptions in the fall speeds of large ice species have more of an impact over bulk density assumptions in terms of specification of rimed-ice characteristics. The modified bulk density simulation was similar to the original unmodified hail simulations in terms of reflectivity structure. The modified fall speed simulation showed noticeable differences in reflectivity structure and was similar to the scheme with graupel representation. These findings were also found in Morrison and Milbrandt (2011), but for surface precipitation and cold pool characteristics. Thus, hydrometeor precipitation loading, fallout rate, and evaporation rates from the microphysical processes have been shown to affect updrafts and downdrafts of a MCS, which result in differences in simulated reflectivity storm intensity and structure.

Past Studies from the NSSL Spring Experiment

Each year during the peak of spring severe weather season (May – June) the Experimental Forecast Program (EFP) of the NOAA/Hazardous Weather Testbed (HWT) hosts a

Spring Forecasting Experiment (SFE) run by the National Severe Storms Laboratory (NSSL) and the Storm Prediction Center (SPC). The first official SFE was held in 2000 (Clark and Coauthors 2012b). Further information on the first few SFEs is detailed in Kain et al. (2003). For each SFE the specific subject of emphasis changes, but the general goal is to test and analyze new severe weather forecasting tools from research to operations. Since 2003, experiments have focused on the development and implementation of CAMs. Following the 2003-2005 SFE years, numerous studies were performed detailing the data from the first generation of CAMs (Kain et al. 2006, Kain et al. 2009, and Davis et al. 2009). The first real-time CAM ensemble was implemented during the 2007 SFE (Clark and Coauthors 2012b). Following the 2007 experiment, numerous studies were performed that investigated and analyzed data from the first real-time 10-member CAM ensemble. (Clark et al. 2009, Schwartz et al. 2009, and Schwartz et al. 2010).

In most recent years, emphasis has been put on the generation of severe probabilistic forecasts valid at better temporal and spatial resolutions than current SPC products utilizing CAMs. Testing of CAM forecasts to sensitivity of microphysics scheme started in the 2010 SFE where CAPS included four members in their Storm Scale Ensemble Forecast (SSEF) system configured identically, but with varying microphysics (Weiss et al. 2010). CAPS continued to reserve four to six members of their SSEF system to test microphysics sensitivity until the 2016 SFE (Weiss et al. 2010, Weiss et al. 2011, Clark et al. 2012, Coniglio et al. 2013, Clark et al. 2014, Clark et al. 2015). In the 2016 SFE, the University of North Dakota (UND) continued the effort of testing microphysics sensitivity by running a microphysics sub-ensemble, which was part of the 65-member Community Leveraged Unified Ensemble (CLUE).

Observations: Stage IV Precipitation Data

For verification of the microphysics ensemble simulations, Stage IV (Baldwin and Mitchell 1997) hourly precipitation data from National Centers for Environmental Prediction (NCEP) were used as observations. This multi-sensor precipitation analysis (MPE) merges precipitation radar estimates and rain gauge observations available across the contiguous United States. The National Weather Service (NWS) and 12 River Forecast Centers (RFC) across the United States produce the MPEs and then NCEP combines those into a national product. The processing of the raw radar data at each radar site to the national product of hourly precipitation contains four stages. The first step is the conversion of raw radar reflectivity data to accumulated precipitation. The second step referred to as Stage II, estimates an optimal hourly rainfall accumulation combining both radar and rain gauge observations. These first two steps are explained further below. For Stage III, each of the 12 RFCs interpolates mosaic Stage II data to their respective forecast areas. Manual control is performed separately by each RFC to create the Stage III dataset. The final step is Stage IV, where NCEP interpolates the Stage III data onto a national grid. Stage IV data are available in hourly, 6-hourly, and 24-hourly accumulated intervals. Fulton et al. (1998) explains the four stages of the rainfall processing in further detail.

Regarding the data sources and algorithms used in Stage I and II, rain rate observations are available from approximately 3000 automated rain gauge sites through the GOES Data Collections Platform (DCO) and Automated Surface Observing System (ASOS). Radar reflectivity

observations are from the National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) network of 160 S-band Weather Surveillance Radars (WSR-88D) across the United States. These observations are converted to an estimated rainfall accumulation using a standard z-R power law relationship (Battan 1973; Doviak and Zrnic 1984) within the Precipitation Processing System (PPS) algorithm. More about this algorithm can be found in Fulton et al. (1998).

The MPE is mapped onto a national grid referred to as the Hydrological Rainfall Analysis Project (HRAP). The HRAP grid is a 1221x881 polar stereographic grid spanning the contiguous United States and has horizontal grid spacing of 4 km. For the purpose of intercomparing simulated and observed storm behavior, the Stage IV data on the polar stereographic grid is regridded or interpolated to the WRF model grid. For this study, the budget interpolation method (Acadia et al. 2003) was used to remap the Stage IV data to the post-processing grid. That way, the precipitation accumulations from both the WRF simulation and the observations would be on the same grid for post processing.

The budget interpolation method was detailed along with bilinear interpolation in Acadia et al. (2003). In the following paragraph, the reader can think of the native grid as the HRAP grid. While bilinear interpolation would use a linear distance weighting between the center of the surrounding native gridcells and the center of the post-processing grid cell in two dimensions, the budget method instead uses an area weighting. The budget method otherwise known as the nearest-neighbor method or remapping, weights the interpolation by the approximate area from each native (HRAP) gridcell that overlaps the post-processing grid cell.

The area is estimated by subdividing the post-processing grid cells into smaller boxes. This is demonstrated in Figure 1 for a post-processing gridcell that has been subdivided into 5x5 smaller boxes. The smaller boxes are assigned the precipitation value associated with the native gridcell that the smaller box's center falls within. The average of these subgrid points within a post-processed grid cell are then used for the remapped value (Acadia et al. 2003). Using the example shown in Figure 1, the interpolated values assigned to the gridcell for post-processing would be weighted by 5/25 and 20/25 the values of the native (HRAP) gridcells #2 and #5 respectively. Although Figure 1 shows 5x5 subdivided grid boxes, the automated storm tracking analysis (explained below) utilizes 2x2 subdivided grid boxes to perform the interpolation.

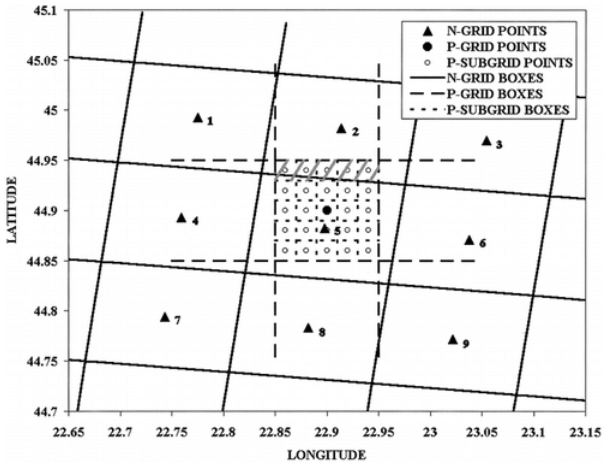

Figure 1. Example of the budget interpolation method. Solid lines and black triangles are native grid boxes (N grid) and native grid points, respectively. The dashed lines and black circles are the post processing grid box (P grid) and grid points, respectively. The dotted lines with the small white circles are the subgrid boxes. [Figure from Acadia et al. 2003]

Herein, the budget method is selected because Acadia et al. (2003) found the budget method resulted in less smoothing of the precipitation data compared to the bilinear interpolation method. Smoothing is not desirable since it acts to increase the minima and

decrease the maxima of the original precipitation values. Acadia et al. (2003) also compared the skill scores of both interpolation methods to the skill scores of the native grid and found that the budget method was superior. For these two reasons, the budget (area) interpolation method is used herein.

Verification: Method for Object-Based Diagnostic Evaluation (MODE)

The interpolation to the post processing grid is done within a tool called Method for Object-Based Diagnostic Evaluation (MODE; Davis et al. 2006a; Davis et al. 2009), which now comes as part of a suite of Model Evaluation Tools (MET)[1]. Davis et al. (2006a) introduced this object-based identification and comparison method to quantitatively and subjectively analyze model simulations to observations. After remapping to a common post-processing grid, Davis's method utilizes a smoothing, thresholding, and merging/matching process to create objects from model simulations and observations. Smoothing and thresholding operations can be adjusted to yield objects of different spatial and intensity scales. The smoothing performed is deliberate unlike the smoothing that results from regridding. Once objects are identified and matched the original raw data is applied to the objects and a series of performance metrics or attributes are applied for analysis. The object identifying method is illustrated in Figure 2. These performance metrics include area, number of objects, centroid location, orientation (axis angle), aspect ratio and additional attributes resulting from the added temporal dimension, such as duration, velocity, percentile intensities, and directional tracking.

[1] MET is managed by the Developmental Testbed Center (DTC) at the National Center for Atmospheric Research (NCAR).

MODE has been used in multiple studies within the past few years to verify and analyze data for convection-allowing forecasts. Specifically, it's been used to evaluate CAM data from the 2008, 2009 and 2010 SFEs (Kain et al. 2010a and Clark et al. 2012a). Multiple studies that used MODE to examine convective simulations noted a common trend where simulations tended to overpredict the amount of MODE precipitation objects (Davis et al. 2009, Johnson et al. 2013, Clark et al. 2014, and Goines 2017). Herein, MODE Time Domain (MODE-TD), an extension of MODE, was used to analyze simulated precipitation properties of convective storms (size, intensity, number, location, and duration) for the simulation suite to be described below, compared to NEXRAD Stage IV radar derived precipitation products. The details and settings regarding MODE-TD can be found in the methodology section below.

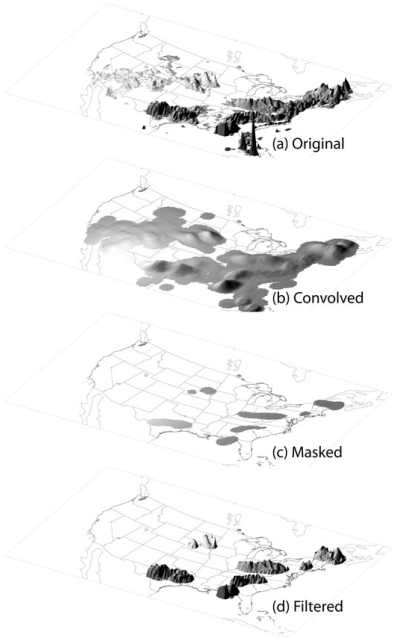

Figure 2. Example of the NCAR DTC Method for Object-based Diagnostic Evaluation (MODE). The object identifying process includes, (a) the original raw two-dimensional data field, (b) a convolution radius applied to smooth the field, (c) a threshold applied to create objects, and (d) original raw data applied inside the identified objects. [Figure from Davis et al. 2006]

CHAPTER II

METHODOLOGY

WRF Model Ensemble

The WRF model version 3.7.1 (Skamarock et al. 2008) was used to conduct the simulations for the microphysics ensemble. The WRF configuration for this study was the configuration used for the 2016 NOAA/HWT Spring Forecasting Experiment. This experiment included numerous CAM ensemble members, managed by various groups from around the country, referred to as CLUE. CLUE included sub-ensembles totaling 65 total members. All CLUE members used the same initialization time, horizontal grid spacing, domain, long- and shortwave radiation schemes, vertical levels, and model top (WRF configuration detailed in following text and Appendix A). For this study, the focus was on the microphysics ensemble provided by the University of North Dakota (UND) group, which included 5-members within CLUE. Four of the five microphysics schemes were analyzed in this study. The schemes making up each ensemble member, detailed in Table 1 include the WRF single-moment 6-class scheme (hereinafter WSM6, Hong and Lim 2006), Morrison double-moment scheme (Morrison et al. 2009), Milbrandt double-moment scheme (Milbrandt and Yau 2005), Thompson scheme (Thompson et al. 2008), and Predictive Particle Properties scheme (hereinafter P3, Morrison and Milbrandt 2015), which was excluded. The WSM6 scheme is the least sophisticated scheme as it is the only single-moment scheme in the ensemble. The remaining schemes are all double-moment with the Milbrandt scheme being the most complicated scheme as it predicts seven moments.

Table 1: List of microphysics ensemble members, their predicted variables (including mixing ratio Qx and total number concentration Nx), and the original reference that describes the scheme. Subscripts c, i, r, s, g, h refer to cloud water, cloud ice, rain water, snow, graupel, and hail, respectively. Q refers to a "free ice" category that is used instead of separate ice species.*

Microphysics Scheme	Variables predicted in addition to Qc, Qr, Qi, Qs, Qg	Original Reference
Ens 1. WSM6		Hong and Lim (2006)
Ens 2. Thompson	Ni, Nr	Thompson et al. (2008)
Ens 3. Morrison	Ni, Nr, Ns, Ng	Morrison et al. (2009)
Ens 4. Milbrandt	Qh, Ni, Nr, Ns, Ng, Nc, Nh	Milbrandt and Yau (2005)
Ens 5. P3	Qc, Qr, Nc, Nr, Q*	Morrison and Milbrandt (2015)

Model Configuration and Domain Setup

One domain covered the contiguous United States with a horizontal grid spacing of 3 km. There were 1681 grid points in the west/east direction and 1153 grid points in the north/south direction and 51 vertical levels. Since the focus domain was in the CONUS a Lambert Conformal map projection was used, which is often used for domains in the mid-latitudes. The 12 km North American Mesoscale (NAM; Rogers et al. 2009) model data was used for the initial and boundary conditions of the model, where the time-varying lateral boundary conditions were applied every 3 h during the simulation. The time step for the majority of the runs was 18 s. for each microphysics scheme. All simulations were initialized at 00z allowing for 12 h of spin up and the last 24 h to cover the full diurnal cycle. Certain runs had to be rerun using a 15 s timestep because of stability issues.

The default namelist settings for WRF were used except for a few settings that were changed for all CLUE members. Those changes are summarized here and the full name list is provided in Appendix A. The physical parameterizations used in this study included the Mellor_Yamada_Janic (MYJ; Mellor and Yamada 1982; Janjic 2002) boundary layer scheme, Rapid Radiative Transfer Model (RRTMG; Lacono et al. 2008) shortwave and longwave radiation schemes, and the Noah land-surface model (Chen and Dudhia 2011). The cumulus parameterization was turned off, due to the fine grid spacing. The time elapsed between radiation physics calls (model parameter, radt) was set to 15 min. For the dynamics, the vertical velocity damping (model parameter, w_damping) was turned on and the advection options for scalars and for moisture were both set to monotonic. Since a PBL scheme was used the K option (model, parameter, km_opt) was set to 2d deformation, which means horizontal diffusion is diagnosed from just horizontal deformation and the vertical diffusion is assumed to be done by the PBL scheme. Damping (model parameter, damp_opt) was set to use with Rayleigh damping, where an implicit gravity-wave damping layer is included near and above the tropopause. Again, these were the only changes made from the WRFV 3.7.1 default namelist.

Initialization / Selection of Cases

The WRF microphysics ensemble was run for case dates occurring during the 2016 CLUE time period, and for selected cases from a retrospective warm season period of April-September 2010-2012. WRF initial conditions for all cases originated from the 12 km NAM. WRF initial conditions for the real time runs were prepared by the Center of Analysis and

Prediction of Storms (CAPS) employees using the Advanced Regional Prediction System (ARPS) model. The retrospective cases were included to increase the sample size and thus the statistical significance of the results. Other than necessary differences in the initial conditions for the real time runs (to be described below), both the real time runs and retrospective runs used the same configuration. In total, there were 46 retrospective case dates simulated and 31 real time case dates during 2016 NOAA SFE simulated for a total of 77 cases. First, the retrospective case date selection criteria will be described.

The 46 retrospective case dates were chosen from a larger "case date" pool, where examination of hourly radar images revealed the presence of one or more MCSs having convective region reflectivity of at least 40 dBZ for at least three consecutive hours in their life cycle. MCSs were defined herein as storms that "evolve 3-6 h and longer, contain at some stage both convective and stratiform precipitation regions, and typically attain horizontal dimensions of at least 100 km" (Hilgendorf et al. 1998) and "include a wide variety of mesoscale phenomena ranging from short-lived indiscriminate aggregates of a few thunderstorms to well-organized squall lines" (Fritsch, 2001). This list of 307 case dates contained information on the start and end times, region formed, and region dissipated. This pool was reduced by finding 63 case dates that had one or more MCSs that were simultaneously located, or that passed through more than one region of the country (regions shown in Figure 3)[2]. Thus, from the master list of 307 case dates with MCSs, 63 case dates had MCSs that traveled through more than one specified region (in Figure 3), and of those, 46 cases dates were randomly selected for

[2] These regions are not related to this specific study, but are used for a follow-on study lead by Hagenhoff and Kennedy (personal communication, 2016).

simulation based upon available computational time. Appendix B shows the reduced list of 63 case dates and the 46 case dates within the reduced list that were simulated and included in the data analyzed.

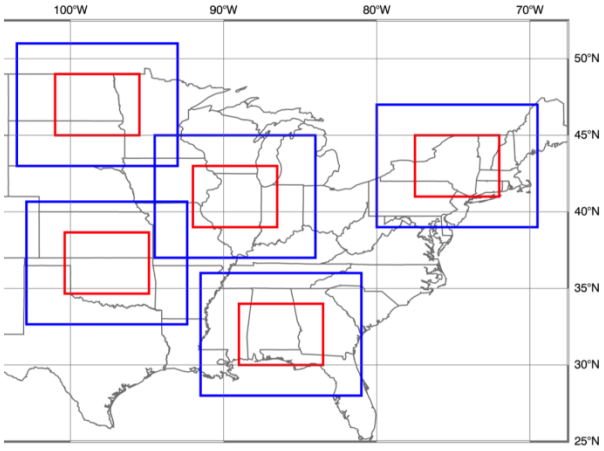

Figure 3. Regions of interest (red boxes) that storms had to be located within to be considered for inclusion.

In addition to the retrospective cases, 31 real time runs during 2016 NOAA SFE using the WRF microphysics ensemble were also completed and added to the dataset of case dates analyzed. This included case dates from April 20-22, 25-29, May 2-6, 9-10, 12-13, 16-20, 23-25,27, 30-31, and June 1-3, 2016 totaling 31 real time cases. The real time runs benefitted from improved initial conditions utilizing 3D variational assimilation (3DVAR; Xue et al. 2003; Gao et al. 2004) of radar observations and other high-resolution observations into 12 km NAM data using ARPS, which is known to give superior forecasts during the first 12 h of simulation (Kain et. al 2010a)- desirable for guiding the 2016 SFE forecasters. This may at first seem to be an important inconsistency, but data assimilation has been shown to have little impact on forecasts when lead times are greater than 12 h. Kain et al. (2010a) showed the “memory” of small convective features that are assimilated is lost between 6-12 h (compared to runs without

detailed assimilation), and mesoscale forcing mechanisms increasingly dominate the convection after 12 h into the simulation. Thus, since the work herein only analyzes model output between 12 and 36 h into the simulation, we believe that the output from the retrospective runs and the real time runs can be included together in the same analysis.

Object Tracking

MODE-TD was used for the analysis, which was released as an update made to the MODE toolkit on 26 October 2015. Clark et al. (2014) was one of the first to utilize a beta version of MODE-TD to analyze data from CAMs. The methodology herein in terms of the configurations and analysis of MODE-TD is similar to that of Clark et al. (2014) and Goines (2017). MODE objectively identifies objects within two-dimensional fields. Objects are classified from a single model variable, in this case hourly precipitation accumulation. The objects are identified through a convolution thresholding process, which applies a convolution radius and precipitation threshold to raw[3] hourly precipitation data. The convolution replaces the precipitation value at a grid point with the average of the surrounding grid points within a specified distance. For example, if a convolution radius of four is specified, gridpoints within a four grid point radius centered on a grid point are averaged, this averaged value then replaces the raw value at that centered grid point. Following the convolution step, a precipitation threshold is applied removing any grid point precipitation values less than the user specified value. This removes areas of weak precipitation and allows the object boundaries to be

[3] Raw refers to model output and Stage IV observational data.

detected. Lastly, the original raw precipitation values are applied back to the grid points within the identified objects (Process illustrated in Figure 2).

The aforementioned process details object identification at one time. To identify an object through time (*tracked object*), MODE-TD convolves in both time and space, matches precipitation regions forwards and backwards across all forecast times and then identifies them accordingly with an object identity (*object ID*). A single *object ID* is assigned to two or more precipitation regions that started as one precipitation region (common for supercell thunderstorms that split and move away from one another) or when two or more precipitation regions merge into a single precipitation region. Because the algorithm searches forward and backward in time, it is able to assign a single *object ID* and the precipitation region at any given time includes all of the combined precipitation regions that are encompassed by the *object ID*, even if they are temporarily separate.

The *number of detected objects* refers to the total number of times that MODE was able to track that region of precipitation, hour to hour, for the entire object lifespan (number of hour counts). This *number of detected objects* should not be confused with *object ID,* which is just an object numbering/naming system. This *number of detected objects* quantity is also potentially confusing since there can be more than one identified precipitation region being tracked by MODE-TD and each with a different lifespan. For example, if MODE-TD diagnosed and tracked a region of precipitation that initiated during the first hour and continued to the 24^{th} hour, the *number of detected objects* would equal 24 and the object lifespan would equal

24. If MODE-TD tracked a second region that lasted only 7 h, then the total *number of detected objects* would equal 31.

MODE Convolution Thresholding and Sensitivity Test

MODE-TD can be adjusted to identify precipitation regions of varying strengths. Certain thresholds within MODE-TD will remove areas of weak precipitation, allowing focus to be on precipitation regions convective in nature. Past studies have set length and intensity thresholds to identify MCSs while retaining small individual precipitation regions possibly related to that of individual thunderstorms. Davis et al. (2006a) expressed that the choice of convolution radius will vary depending on the interest of the study, but a minimum of four grid points is necessary as it relates to the scale resolved by the model. A convolution radius of four grid points was used in Davis et al. (2006a, 2006b), Clark et al. (2014), and Goines et al. (2017). Clark et al. (2014) also used a convolution radius of eight grid points. In addition, these studies used precipitation thresholds ranging from 2.5 – 15 mm/hr. Utilizing the studies mentioned, precipitation thresholds of 5 mm and 10 mm along with convolution radii of four grid points (12 km) and eight grid points (24 km) were tested for a case date (7/3/2010) to determine the most appropriate MODE settings for the purpose of this study.

The sensitivity test showed that increasing the convolution radius or increasing the precipitation threshold led to fewer detected objects, which was an expected outcome when using MODE. When the precipitation threshold was increased from 5 mm to 10 mm, while keeping the convolution radius at four grid points, the total number of detected objects was

approximately cut in half. In addition, the average size of the detected objects decreased with the smallest objects entailing the largest percentages. By increasing the precipitation threshold, the stratiform and lighter convective precipitation regions are smoothed leading to the detected objects being individual convective cells rather than larger mesoscale convective systems. When the convolution radius is increased from four grid points to eight grid points, while keeping the precipitation threshold at 5 mm, the number of detected objects was significantly reduced. The reduction in detected objects mainly occurred for the smallest objects with the larger objects being retained. This led to an increase in the average size of the detected objects over time as the focus of the detected objects shifted towards larger sizes. To focus more on large propagating storms rather than individual convective cells, a convolution radius of eight grid points (24 km) and a precipitation threshold of 5 mm was selected.

Attributes of MODE Objects

MODE-TD outputs multiple object attributes and statistics. Hourly detected object information includes centroid location and area. The addition of the temporal dimension adds information regarding the tracked object's velocity, start and end time (duration), and intensity percentile of the raw data within the object. For this study, the MODE-TD attributes analyzed were the *number of detected objects* and their *area/size*; along with attributes for tracked objects that include *initiation*, *dissipation*, *duration*, *velocity*, and *percentile [rain] intensities*.

For the detected objects, the *number of detected objects* over time is a useful attribute as it gives information about the total amount of precipitation regions summed over all times,

which should be proportional to the number of storm cells and their longevity. The *object area* of a detected object is a count of the grid squares a region of precipitation occupies at one time and gives vital information concerning spatial coverage of storms.

For each tracked object, *object duration* allows one to gain information on a storm's initiation and dissipation, which is of significant importance in forecasting. The *object velocity* is calculated by MODE-TD by computing the spatial difference of the precipitation region's centroid location over contiguous times, then averaging these distances over the duration of the precipitation region. The centroid is the geometric center of an identified precipitation region. Speed and direction are obtained from the *object velocity*, thus from *object velocity* information about the storm's propagation or track can be obtained. *Object [rain] intensity percentiles* (10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th}) mm h^{-1} were computed by sorting the raw (grid point) hourly precipitation accumulation data[4] within each tracked object for the entire *object duration*. Following traditional statistical definitions, the 50^{th} percentile is the median of the rain rate intensity whereas the 90^{th} percentile would be the rain rate intensity containing the lower 90% of rain rate values. These percentiles are computed automatically within MODE-TD. *Object [rain] intensity percentiles* are useful for comparing hourly rain rate distributions and intensities among storms.

[4] If 10 mm of rainfall accumulates at a grid point over 1 h period, then the associated rain rate is simply 10 mm h^{-1}.)

CHAPTER III

RESULTS

Overview

The following text within the results describes the findings from the MODE-TD analysis performed on the aforementioned case dates. The results section details MODE-TD characteristics in the following order: number of detected objects, detected object size distribution, detected object size with time, and tracked object characteristics including initiation/dissipation, duration, precipitation intensity, and velocity.

Various aforementioned characteristics of the precipitation objects to be evaluated herein were created by MODE-TD using a precipitation threshold of 5 mm/hr and a convolution radius of eight grid points (24 km). The analysis between Stage IV and forecasted ensemble members included 77 case dates. The results only evaluated the last 24 forecast hours, excluding the first 12 h (model spin-up), of each case date. Hereafter, “forecast times” only refers to the final 24 h period of the 36 h forecast. In addition, to help reduce areas where Stage IV lacks in coverage, the results only included objects that were within the 30 to 50 °N and 75 to 110 °W area. To determine if the results changed by region, a supplemental analyses was also performed for the southeastern United States (30 to 40 °N and 105 to 90 °W) and the south central United States (30 to 40 °N and 90 to 75 °W). The supplemental analyses for the Southeastern United States and the South Central United States are in Appendix C and Appendix D, respectively.

Number of Detected Objects

The total number of detected objects (i.e. the number of identified objects present over all forecast times and case dates), along with the corresponding percentage increase relative to Stage IV and model bias are shown in Table 2. All microphysics ensemble members produced 17 % or more detected objects compared to Stage IV. Thus, all microphysics members had a high bias for total number of detected precipitation objects. WSM6 and Milbrandt produced the most detected objects with 8,013 and 7,532, respectively. Therefore, they also had the highest percentage increase relative to Stage IV with 43.1 % for WSM6 and 34.5 % for Milbrandt. Thompson and Morrison had the least amount of detected objects out of the forecast ensemble members with 6,599 and 6,578 respectively. As a result, they had the lowest percentage increase relative to Stage IV with values of 17.8 % for Thompson and 17.4 % for Morrison. In terms of occurrence of detected precipitation objects, Morrison was more representative of the observations as it produced the amount of detected objects closest to Stage IV, although Morrison still significantly overpredicted the total number of detected objects. See Appendix E for number of detected objects and percentage increase relative to Stage IV separated by 4 °N by 5 °W regions across the entire analysis domain.

Table 2: List of MODE-TD diagnosed detected objects across all forecast hours and case dates. The increase relative to Stage IV and model bias are also listed

	Number of detected objects	Increase relative to Stage IV (%)	Model bias relative to Stage IV (number of detected objects)
Thompson	6,599	17.8%	+998
Milbrandt	7,532	34.5%	+1,931
Morrison	6,578	17.4%	+977
WSM6	8,013	43.1%	+2,412
Stage IV	5,601	0%	0

The number and percentage of detected objects at each forecast time for Stage IV and the forecast members are shown in Figure 4, respectively. Demonstrated in Figure 4a, simulations overpredicted detected object occurrence noticeably across forecast hours before the last six forecast hours. Although, Thompson predicted nearly the same way as Stage IV between 0-2 UTC. Consistent with the total number of detected objects summed over all forecast hours, WSM6 generally overpredicted the most with Thompson and Morison the least for the majority of the forecast hours, excluding the last 8 h of simulation time. In addition, the convective diurnal cycle was evident in the detected object curves with a minimum around forecast hours 15-17 (15-17 UTC) followed by a sharp increase in detected object occurrence until a peak was reached between forecast hours 22-24 (22-00 UTC). Following this peak, all datasets had a significant reduction in number of detected objects until another minimum was reached between forecast hours 32-34 (08-10 UTC). To mitigate issues with the high bias in detected object occurrence the percentage of detected objects as a function of forecast hour was illustrated (Figure 4b). Consistent with Figure 4a, simulations had a maximum and

minimum in detected object occurrence about 2 h earlier than Stage IV, excluding Milbrandt and Morrison. Stage IV's peak occurred at 25 h (01 UTC). Thompson reached detected object occurrence peak the earliest at 22 h (22 UTC), then WSM6 at 23 h (23 UTC) and Milbrandt and Morrison at hours 24 (00 UTC) and 25 (01 UTC), respectively. In general, the simulations detected object occurrence curve was shifted to the left indicating all microphysics initiate and dissipate precipitation objects earlier than observations. In terms of percentage of detected objects with time, Morrison was most similar to Stage IV as their shape curves are similar in both shape and percentage values, but Morrison was generally shifted 1 h earlier. Thompson appeared to resemble Stage IV the least as it peaks the earliest and its peak percentage value was furthest from Stage IV's peak. In addition, Thompson did not have a noticeable drop off in percentage of detected objects during the last 5 h (07-12 UTC) of forecast time. Bias issues aside, Morrison best resembled percentage of detected object occurrence and Thompson the worst.

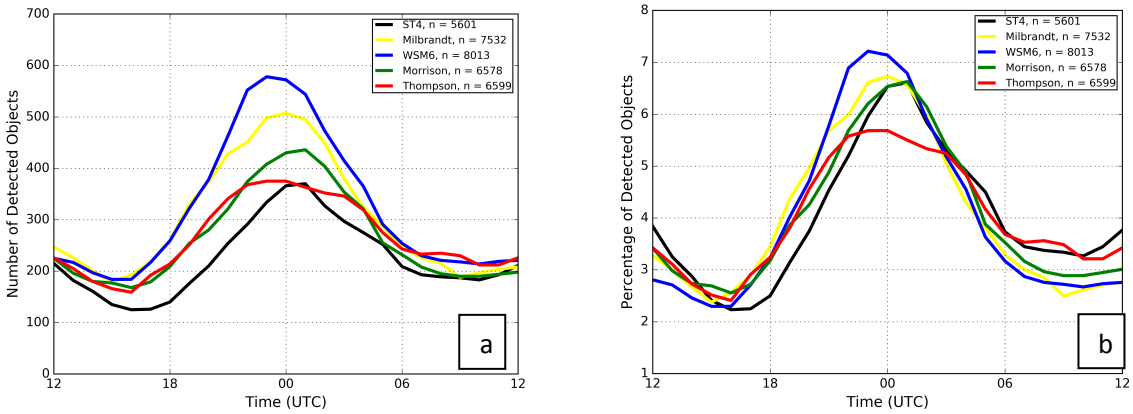

*Figure 4. (a) Number of detected (b) percent of total detected objects ((# of objects at time / n)*100) for the microphysics and Stage IV at each forecast hour (UTC time). The n refers to the total number of detected objects accumulated over all forecast hours and case dates.*

Detected Object Size Distribution

Size distributions of the detected objects for Stage IV and the microphysics forecast members are shown in Figure 5 and Figure 6. The "size" is defined herein as the square root of the detected object area giving a characteristic length of the detected object in km, assuming the detected object was a square. Recall, that the *object area* is a count of the grid squares a detected object occupies at a single time. Thus, to get the object size in km, the square root of the area was taken then multiplied by the horizontal grid spacing (3 km). The detected objects were binned into 30 km bins ranging from zero to 420 km.

Figure 5a shows the detected object count for particular object size bins over the forecast period (final 24 h) for every simulation. All simulations overpredicted the number of detected objects across all size intervals, excluding sizes larger than 240 km. Bins with the most detected objects occurred at sizes less than 120 km with the peak in detected objects occurring for the 30-60 km bin. The most likely reason that the smallest size (0-30 km) does not have the largest percentage of detected objects across the observations and forecasts was due to the selection of the larger convolution radius (8 grid points), which smoothed smaller objects. For bins with sizes less than 120 km, WSM6 overpredicted the most with Milbrandt, Thompson, and Morrison lesser amounts, in that order. Thus, Morrison in terms of number of detected objects was closest to observations at the smaller storm sizes (<120 km) and WSM6 had the highest bias for detected objects at smaller storm sizes. In terms of percentage of detected objects (Figure 5b), the simulations better resembled observations than the distribution in number of

detected objects. In general, Thompson and WSM6 had the highest biases for the smallest bin sizes (0-90 km) as they had the highest percentage of detected objects at these sizes. Otherwise, there was not a systematic bias among the smallest sizes.

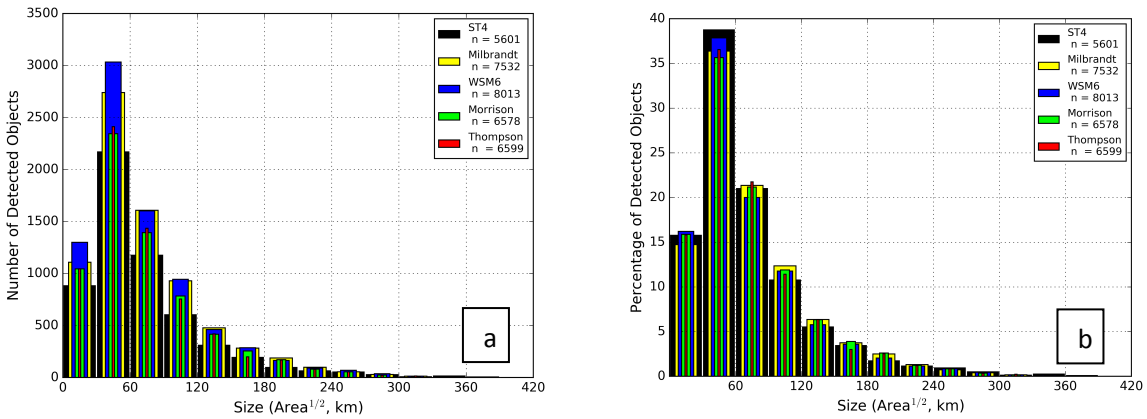

*Figure 5. Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size for all case dates*

The spread of the microphysics members and the observations appeared to be largest at the smallest detected object size intervals, with the spread decreasing towards the larger sizes. It was difficult to analyze the bins at the largest sizes (Figure 4), due to the large difference in detected objects at the smaller versus larger sizes, and thus, Figure 6 focuses on larger sizes.

Figure 6 like Figure 5 shows the size distribution of forecast and observed detected objects, but Figure 6 focuses on size bins larger than 120 km, which are likely more representative of MCSs. Like the smaller sizes, all microphysics had a high bias in number of detected objects, excluding sizes larger than 240 km. The differential in detected objects between the forecasts and observations seemingly decreased as the size increased. Among small to mid-size MCSs (120-240 km), Milbrandt had the most number of detected objects, where WSM6 had the most for sizes less than 120 km. Illustrated in Figure 6b, Stage IV for the

most part had higher percentages for number of detected objects at sizes greater than 240 km. Thus, detected object amount aside, the microphysics produced smaller percentages of large MCSs than observed. Although, the low bias was relatively small between the microphysics and Stage IV as the differences were less than 0.5%

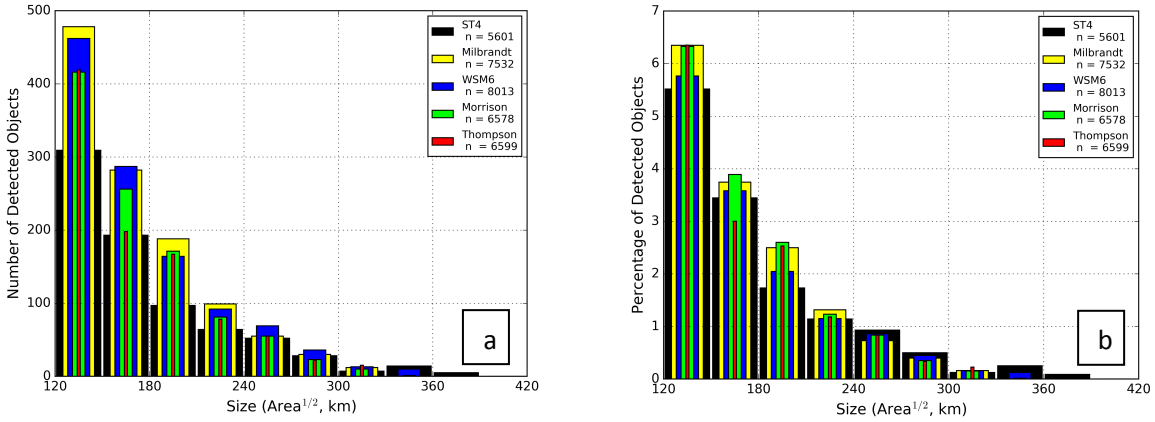

*Figure 6. Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size greater than 120 km for all case dates*

Overall, Morrison, Thompson, and Milbrandt consistently had the largest percentage of detected objects for small to mid-size MCSs (120-240 km) with WSM6 agreeing most with Stage IV. For larger MCSs (> 240 km), there was no noticeable trend among microphysics; just that Stage IV generally exceeded schemes in percentage values. Note sizes greater than 240 km occurred rather infrequently as all datasets had less than 1% of detected objects at each bin.

Detected Object Size with Time

Figure 7 expresses the average size of the detected objects at each hour for the observations and microphysics ensemble members. Like Figure 5 and Figure 6 the object size

used in Figure 7 is the square root of the object area, representing characteristic length of the object in km. In general, the microphysics represented the overall shape and trend of average size with time well, but it was evident that the simulations' peaks were shifted compared to Stage IV and the simulations failed to represent the smaller peak between forecast hours 3-6 (15-18 UTC). Thus, the microphysics prematurely reached their peak size 2-4 h before the observations and failed to simulate the observed smaller peak early in the forecast time. Milbrandt and Morrison produced the largest and similar average peak size, with WSM6 and Thompson having the smallest average peak size. Notice the last 3 h of the simulation times (09-12z), Thompson was the only scheme that did not have a significant drop off in average size. Subjectively, Milbrandt best represented truth as it was closest to observations in peak size and trend in average size with time. Meanwhile, Thompson was least representative of truth as it was among the lowest bias in peak size and did not undergo a significant drop-off in average size the last 3 h of simulation hours.

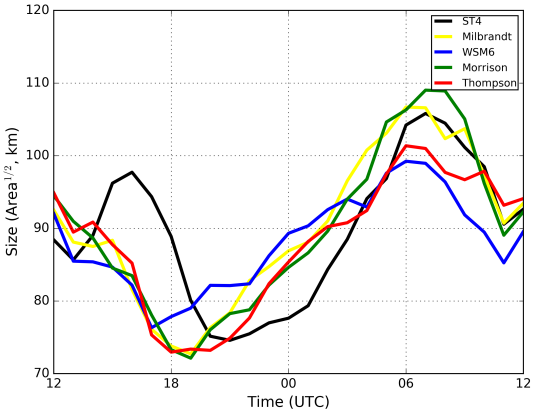

Figure 7. Average size (square root of area, expressed in km) of detected objects as a function of UTC time

Initiation / Dissipation

Figure 8 illustrates the number and percentage of newly tracked objects Figure 8a and c, respectively) and number and percentage of discarded tracked objects (Figure 8b and d, respectively) at each forecast hour in UTC time. First recognition is always associated with tracked object initiation while last recognition occurs when the tracked object is discarded by MODE-TD due to dissipation. Recall that object splitting does not initiate tracked objects. Also, because MODE-TD counts ongoing precipitation regions at the t=12 and t=36 h as "newly-detected" and "discarded", respectively, then those hours are omitted from Figure 8a and Figure 8b so that only actual object initiation and dissipation is shown. In the legend, N for each data set represents the sum of the number of tracked objects over all cases and is the number used when calculating the percentage. Thus, a tracked object that lasted multiple hours was counted only once rather than each hour. In Figure 8a, the microphysics all had a high bias in the amount of tracked objects initiated at each hour up to the forecast hour 11 (23 UTC) and the peak object initiation was again shifted so the microphysics peaked 1-3 h earlier than observed. This was consistent with the high bias in total number of detected objects with time (Figure 4), and a premature peak for some microphysics schemes. Figure 8c again shows the shift in initiation, but the newly tracked objects with time were normalized as percentages. With the large differential in number of tracked objects aside, Morrison resembled Stage IV the best as the percentage values were similar, but Morrison's initiation was shifted 1 h earlier.

Figure 8b and d express tracked object dissipation with time, in terms of number of tracked objects and percentage of tracked objects, respectively. Forecasts had a high bias in

dissipated tracked objects at most hours before forecast hour 12 (00 UTC). In terms of percentage of tracked objects (Figure 8d), Thompson least resembled observations as it reached peak dissipation the earliest and dissipated the highest percentage of tracked objects at the end of the simulations. In general, peak dissipation occurred 2-3 h earlier than observed and there did not seem to be a best performing microphysics scheme for dissipation.

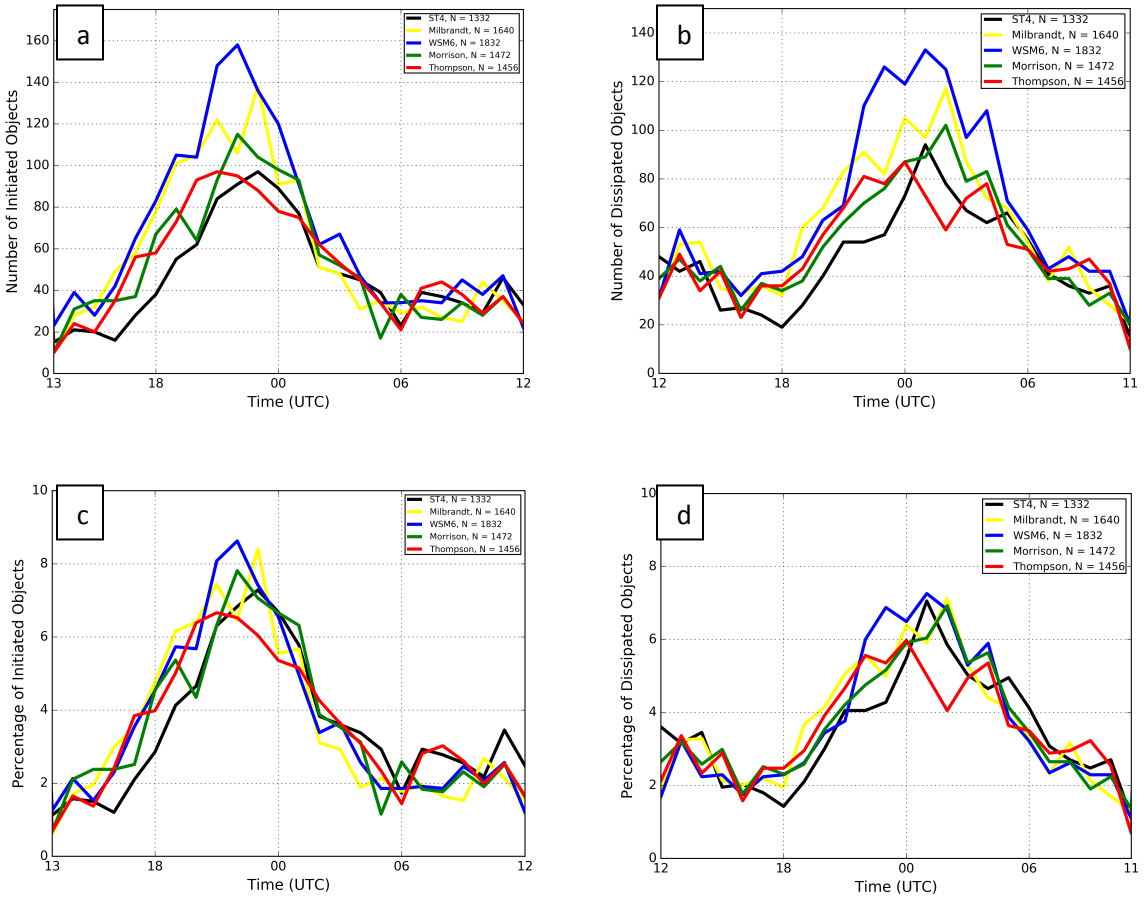

*Figure 8. The number of tracked objects that were a) newly detected and b) discarded as a function of UTC time. The percentage of detected objects ((# of objects at time / n)*100) that were c) newly detected and discarded d) as a function of UTC time. N represents the total number of tracked objects summed over all the case dates.*

Duration

Figure 9 expresses the count of tracked objects with a particular object duration (Figure 9a) and percentage of total tracked objects with a particular object duration (Figure 9b) per dataset. The total tracked objects of each dataset are listed in the legend, as in the start/end detection plot (Figure 8). Figure 9a shows the microphysics capture the trend and shape of the duration curve well, but the overprediction of number of tracked objects was apparent at object durations 6 h or less, especially for WSM6 and Milbrandt. All datasets had a peak in object duration at 2 h with a sharp drop-off for longer durations. To detail longer durations, which had significantly fewer objects than the shorter durations, the count of tracked objects with a particular object duration was plotted using a log scale for the y-axis shown in Appendix F. Figure 9b shows how similar object durations were between microphysics and observations when the datasets were normalized to percentages of total tracked objects. At 2 h, all simulations, excluding WSM6, had approximately 3-5 %, with respect to their total, less tracked objects than the Stage IV. Otherwise, there were little differences between simulations and observations.

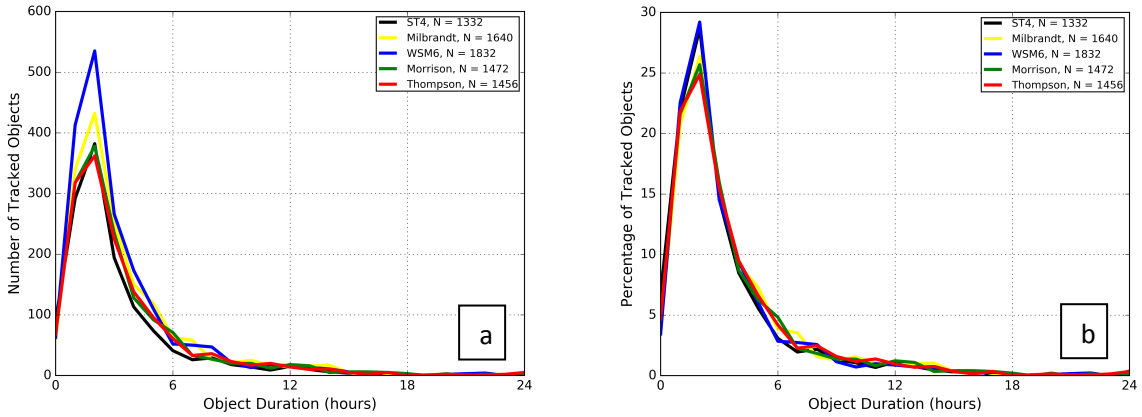

*Figure 9. Number of tracked objects (a) and percentage of tracked objects ((# of objects at time / n)*100) (b) as a function of object duration. N represents the total number of tracked objects summed over all the case dates*

Percentile [Rain] Intensity

Figure 10, Figure 11, Figure 12, Figure 13, Figure 14, and Figure 15 and Table 3, Table 4, Table 5, and Table 6 express rain intensities of the tracked objects through percentile [rain] intensities. Tracked objects are used as in the initiation, dissipation, and duration (Figure 8 and Figure 9) plots. The percentile [rain] intensities of a tracked object were determined by summing up all the grid-point precipitation values within the tracked object over its life span then calculating percentiles (10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th}). Figure 10, Figure 11, Figure 12, Figure 13, and Figure 14 represent the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentile [rain] intensities, respectively, of the tracked objects in terms of hourly precipitation rate and average size. The average size of a tracked object was calculated by summing the sizes of a tracked object at each time over its lifespan then dividing by the number of hours the tracked object was present. The four size ranges (<60 km; 60-120 km; 120-180 km; and >180 km) for each percentile allowed for separation of tracked objects representative of individual thunderstorms and MCSs and

maintain consistency with the size bins used in the size distribution plots (Figure 5 and Figure 6). Note the count of tracked objects for each microphysics dataset and size range differs. These values are shown on the right side of Figure 10 – Figure 14. Tracked objects that were only detected once were excluded from the percentile [rain] intensity analysis.

Table 3: List of object averaged percentile precipitation intensities (mm h^{-1}) across all cases for objects with an average size 60 km or less. Best performing forecast values are bolded.

	Size: < 60 km				
	10th	25th	50th	75th	90th
			(mm h^{-1})		
Thompson	1.71	3.65	7.69	**14.20**	22.97
Milbrandt	1.65	**3.86**	8.51	15.07	22.35
WSM6	1.24	3.18	**8.00**	15.52	23.50
Morrison	**2.20**	4.53	8.68	14.04	**19.87**
Stage IV	2.11	4.02	8.18	14.32	20.73

Table 4: List of object averaged percentile precipitation intensities (mm h^{-1}) across all cases for objects with an average size 60 – 120 km. Best performing forecast values are bolded

	Size: 60 -120 km				
	10th	25th	50th	75th	90th
			(mm h^{-1})		
Thompson	1.36	3.27	**6.89**	**12.64**	19.95
Milbrandt	0.92	2.83	7.31	14.04	21.78
WSM6	0.70	2.27	6.79	14.37	22.81
Morrison	**1.46**	**3.73**	7.79	13.10	**18.86**
Sage IV	1.71	3.51	7.03	12.67	18.94

Table 5: List of object averaged percentile precipitation intensities (mm h^{-}) across all cases for objects with an average size 120 – 180 km. Best performing forecast values are bolded.

	Size: 120-180 km				
	10th	25th	50th	75th	90th
			(mm h^{-1})		
Thompson	1.10	3.21	6.88	**12.96**	20.75
Milbrandt	0.86	2.99	7.41	14.14	21.60
WSM6	0.46	2.11	**6.97**	14.73	23.46
Morrison	**1.12**	**3.55**	7.90	13.63	**19.83**
Stage IV	1.72	3.75	7.07	12.19	18.35

Table 6: List of object averaged percentile precipitation intensities (mm h^{-1}) across all cases for objects with an average size greater than 180 km. Best performing forecast values are bolded.

	Size: > 180 km				
	10th	25th	50th	75th	90th
			(mm h^{-1})		
Thompson	**1.50**	3.79	**7.14**	**12.26**	**19.24**
Milbrandt	1.05	3.24	7.49	13.88	21.35
WSM6	0.63	2.48	7.21	14.35	22.74
Morrison	1.43	**3.73**	7.81	13.24	19.16
Stage IV	1.46	3.73	7.03	12.68	20.01

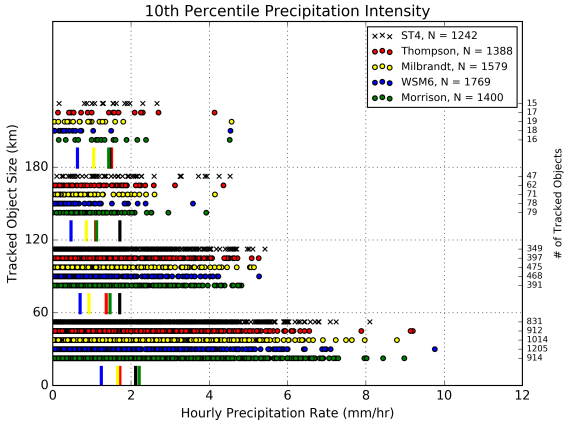

Figure 10. Tracked object 10th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 10th percentile precipitation intensities of all the tracked objects within the corresponding size range.

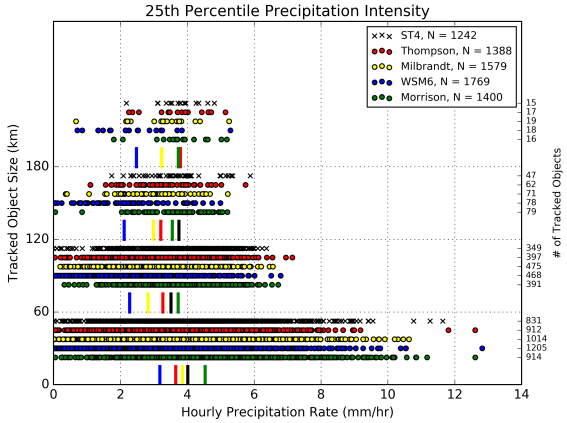

Figure 11. Tracked object 25th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 25th percentile precipitation intensities of all the tracked objects within the corresponding size range.

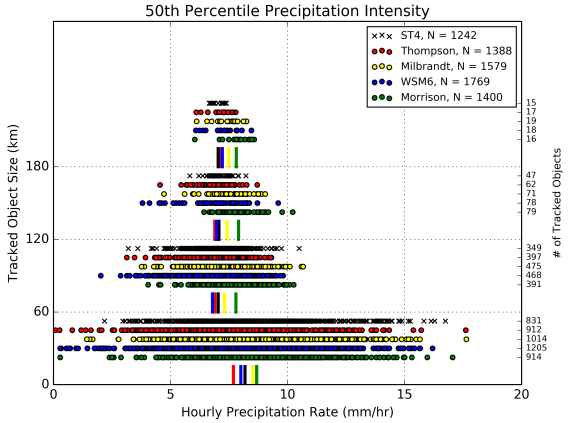

Figure 12. Tracked object 50th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 50th percentile precipitation intensities of all the tracked objects within the corresponding size range.

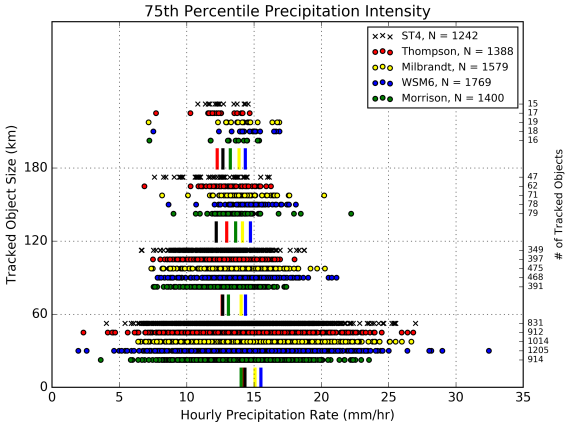

Figure 13. Tracked object 75th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 75th percentile precipitation intensities of all the tracked objects within the corresponding size range.

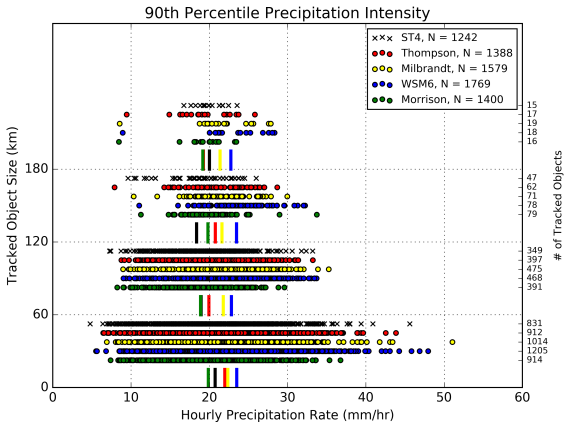

Figure 14. Tracked object 90th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 90th percentile precipitation intensities of all the tracked objects within the corresponding size range.

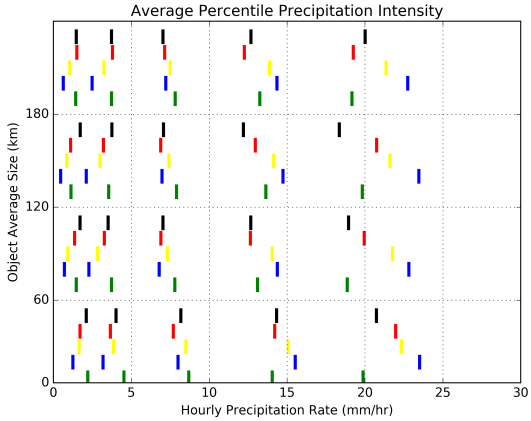

Figure 15. Tracked objects averaged precipitation percentile intensities from Figure 10, Figure, 11, Figure 12, Figure 13, and Figure 14 plotted together. Following the color convention of previous figures, black refers to observations, blue WSM6, red Thompson, green Morrison, and yellow Milbrandt.

To aid in the interpretation of the percentile [rain] intensities that were generated by grouping all case dates, Figure 16 was created for only a single case date. Figure 16 illustrates hourly precipitation values applied within the MODE defined objects. Figure 16 shows one forecast hour and one object for each microphysics and Stage IV from a retrospective case date (initialized 7/19/2010). This case date was selected from the 46 simulated retrospective case dates based off the criteria that there were a sufficient amount of detected objects across all

forecast times, and the case date was similar to the statistical analysis trend of detected object size with time and detected object number with time. The 17th forecast hour was selected because that forecast time had the largest average object size among the microphysics. Figure 17 – Figure 21 express the percentile [rain] intensities for the case date (initialized 7/19/2010). The percentile [rain] intensities that are outlined by boxes correspond to the objects in Figure 16. Note that these percentile [rain] intensities are calculated over the lifespan of the tracked object, not at one time.

Observing Figure 16, Thompson, Milbrandt, and WSM6 all appear to have stronger convective regions compared to Stage IV and Morrison as they have higher peak precipitation values. This is quantitatively expressed in Figure 20 and Figure 21 where WSM6, Milbrandt, and Thompson all have higher 75th and 90th percentile values than Stage IV and Morrison for the respective objects shown in Figure 16. Also, in Figure 16 it appears that there is more of an abrupt transition from the weaker to stronger precipitation values for WSM6 and Milbrandt compared to Stage IV. WSM6's and Milbrandt's weak intensity bias for the 10th and 25th percentiles is shown in Figure 17 and Figure 18 where their 10th and 25th percentiles values are the lowest. Thus, indicating that WSM6 and Milbrandt produced a weaker stratiform region and stronger convective region relative to Stage IV. As what will be shown in the following text, the results for this case study were consistent with the statistical analysis performed, excluding the large percentile [rain] intensity values produced by Thompson for the 75th and 90th percentiles.

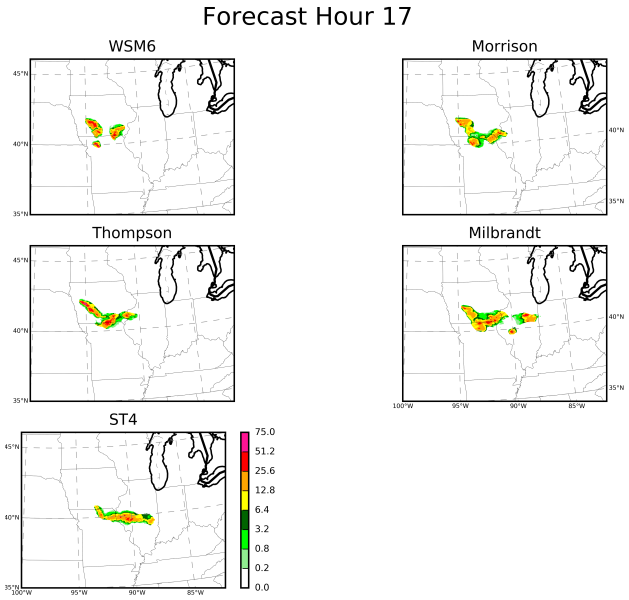

Figure 16. A MODE object for each microphysics ensemble member and Stage IV with the raw hourly precipitation values inside the object. Precipitation values are in mm. Valid for forecast hour 17 (12z 07/19/2010 – 12z 07/20/2010).

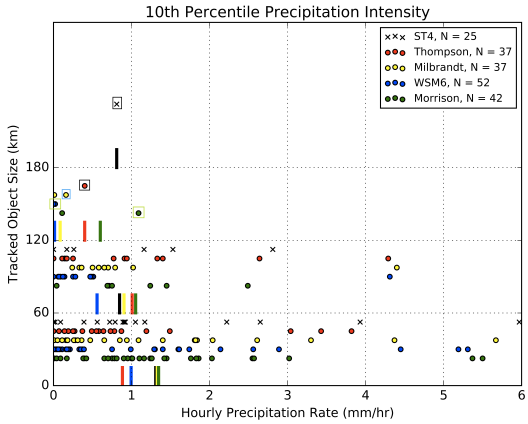

Figure 17. Tracked object 10th percentile precipitation intensities sorted by object average size (<60 km, 60-120 km, 120-180 km, and > 180 km) for tracked objects. Vertical lines are the averaged 90th percentile intensities of the tracked objects within the corresponding size range. Valid for one case (12z 07/19/10 – 12z 07/20/10). Boxes correspond to the objects in Figure 16.

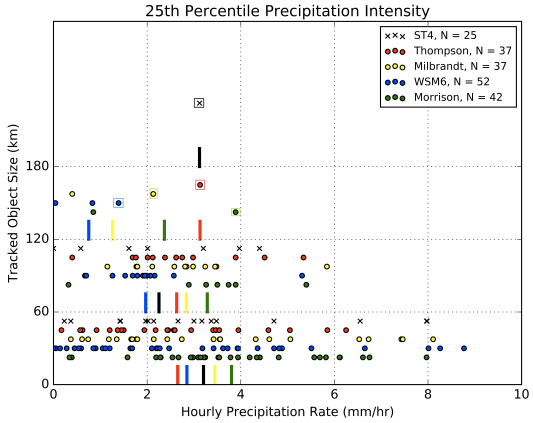

Figure 18. Tracked object 25th percentile precipitation intensities sorted by object average size (<60 km, 60-120 km, 120-180 km, and > 180 km) for tracked objects. Vertical lines are the averaged 90th percentile intensities of the tracked objects within the corresponding size range. Valid for one case (12z 07/19/10 – 12z 07/20/10). Boxes correspond to the objects in Figure 16.

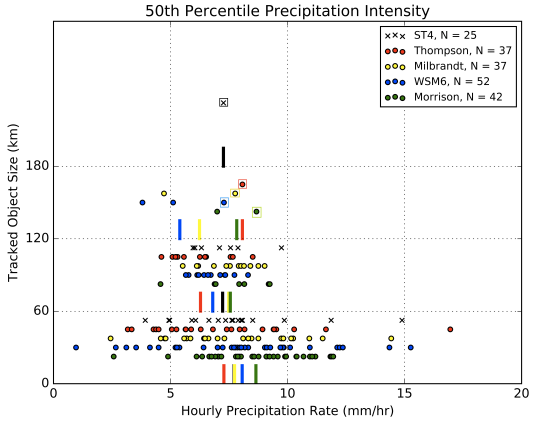

Figure 19. Tracked object 50th percentile precipitation intensities sorted by object average size (<60 km, 60-120 km, 120-180 km, and > 180 km) for tracked objects. Vertical lines are the averaged 90th percentile intensities of the tracked objects within the corresponding size range. Valid for one case (12z 07/19/10 – 12z 07/20/10). Boxes correspond to the objects in Figure 16.

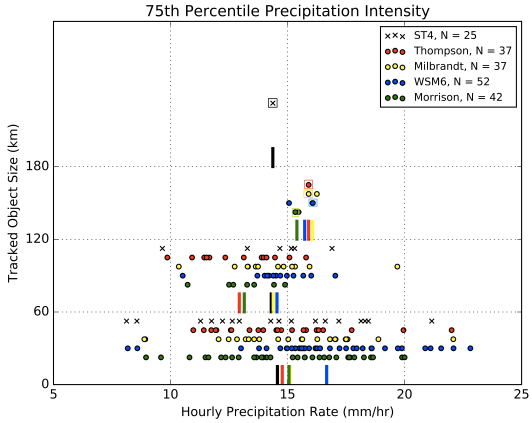

Figure 20. Tracked object 75th percentile precipitation intensities sorted by object average size (<60 km, 60-120 km, 120-180 km, and > 180 km) for tracked objects. Vertical lines are the averaged 90th percentile intensities of the tracked objects within the corresponding size range. Valid for one case (12z 07/19/10 – 12z 07/20/10). Boxes correspond to the objects in Figure 16.

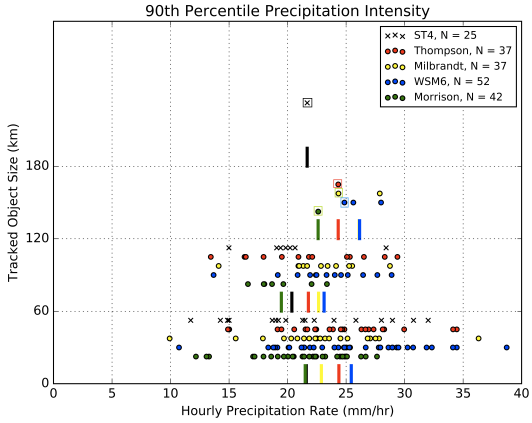

Figure 21. Tracked object 90th percentile precipitation intensities sorted by object average size (<60 km, 60-120 km, 120-180 km, and > 180 km) for tracked objects. Vertical lines are the averaged 90th percentile intensities of the tracked objects within the corresponding size range. Valid for one case (12z 07/19/10 – 12z 07/20/10). Boxes correspond to the objects in Figure 16.

Figure 10 shows that across all size ranges the averaged 10th percentile [rain] intensities were consistent between microphysics with Thompson and Morrison having the largest and WSM6 the smallest. All microphysics had low intensity biases relative to observations, excluding Morrison at sizes less than 60 km. Morrison performed closest to observations with

Thompson a close second. WSM6 and Milbrandt had the largest magnitude of low negative average 10th percentile biases across all sizes. WSM6's 10th percentile averages were approximately 1 mm/hr less than Stage IV across most sizes (Tables 3-6). Milbrandt's low bias for the 10th percentile average was not as significant as WSM6's bias, but still larger in magnitude than Morrison and Thompson. The relative magnitude of the biases between schemes in the 25th percentile averages (Figure 11) was similar to the 10th percentile averages. In other words, the microphysics schemes were weaker than observations, excluding Morrison at sizes less than 120 km, and were ordered similarly, relative to each other. Overall, Morrison and Thompson best represented truth with Morrison closest to observed. The WSM6 and Milbrandt schemes were least representative of Stage IV with WSM6 performing the worst. Thus, for lighter rain (10th and 25th percentile intensities) Morrison and Thompson tracked objects produced precipitation rates closest to Stage IV, and WSM6 objects produced rates furthest from observed.

The 50th percentile (a.k.a. median) [rain] intensities and the corresponding averages are illustrated in Figure 12. Unlike the 10th and 25th averaged percentiles, there was no longer a low intensity bias across the majority of the microphysics and sizes. Unlike other percentiles, Morrison did not perform well. In general, WSM6 tracked objects best-represented median precipitation intensity values within simulated storms.

Figure 13 and Figure 14 show the 75th and 90th percentile [rain] intensities, respectively. Unlike the low percentile intensities (10th and 25th), the high percentile intensities (75th and 90th) for the forecasts had more of a high bias in rain rates. . WSM6 performed the worst when

predicting intense storm rain rates by producing the highest biases and largest differentials from Stage IV. Meanwhile, Thompson captured rain intensities quite well for the 75th percentile and Morrison for the 90th percentile. Thompson and Morrison for the 75th and 90th percentiles have the smallest differentials from Stage IV. Thus, when forecasting convective precipitation, Thompson and Morrison were most likely to produce intensities similar to observations.

Figure 15 is a summary plot that displays the tracked object-averaged percentile intensities from all percentiles for the four size ranges. Plotting the 10th, 25th, 50th, 75th, and 90th percentiles together made it is easier to compare the range of precipitation rates between the microphysics and Stage IV. In general, across all sizes the WSM6 had the largest relative bias in average rain percentiles, thus having the largest spread from the 10th to 90th percentiles making WSM6 the furthest from observations. Morrison performed best as its averaged rain rates across all percentiles and sizes closely resemble observations. Thompson also performed well, as its averaged rain rates did not deviate from observations nearly as much as Milbrandt and WSM6. Therefore, despite having overpredicted the number of detected objects, the Morrison microphysics scheme best matched the observed rain rate intensities and distributions.

Velocity

Figure 22. illustrates the tracked object-averaged velocity components and magnitude as a function of tracked object duration. The object velocity is the average velocity for all

tracked objects of the specified duration. Thus, all tracked objects across all cases were separated according to their duration then the average tracked object's u- and v-components and magnitude of the velocity were calculated for each duration. Only durations with at least ten total tracked objects were considered; thus durations of more than 14 h were not considered. Note there were significantly more tracked objects at shorter durations than longer durations (see Figure 9), thus the averages at shorter durations were more heavily weighted.

In general, as the object duration increased in simulations or observations, the u-component of the velocity also increased (Figure 22a). In addition, for all durations, the average u-velocity was positive, indicating the simulated and observed tracked objects, on average, propagate to the east. The average v-velocity (Figure 22b) for microphysics and observations at shorter durations had slight positive values for duration, indicating that storms with northward movement dominated over storms that had southward movement. As the duration increased, the simulations and observations did not appear to have a noticeable trend. In general, datasets had a much stronger u-velocity than v-velocity component. This was affirmed with the magnitude of the velocity (Figure 22c), which looks significantly similar to the u-velocity (Figure 22b).

The simulations tended to have a slow bias across durations greater than 5 h for u-velocity and durations of 5-7 h for the magnitude of velocity. Simulations had a slow bias in magnitude for durations 5-7 h with a bias as low as approximately 8 km/hr. Thompson performed slightly better at durations less than 6 h and Milbrandt performed better at the

longest durations (> 10 h) as it also had a spike in u-velocity at duration hours 12 and 13. In summary, the simulations all had a slight slow bias in storm propagation between durations 5-7 h with better representation at durations less than 10 h and more variation for longer duration objects (~12 km/hr difference).

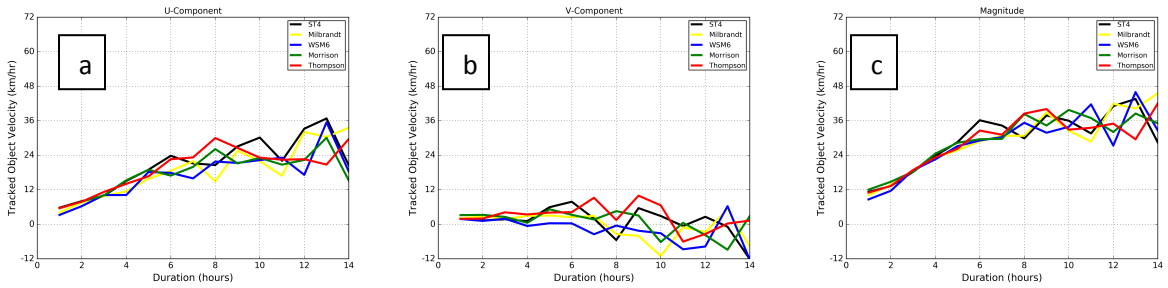

Figure 22. Tracked objects averaged velocity for the u-component (a), v-component (b), and magnitude(c) as a function of object duration. Only objects with durations of 14 h or less are plotted

The tracked object-averaged velocity components and magnitude as a function of tracked object duration, but only for objects greater than 120 km were also plotted to focus on large tracked objects similar in size of an MCS. But, this was not included herein because there were no noticeable trends of significance.

Analysis Summary

Table 7 summarizes the best performing microphysics scheme for the MODE-TD characteristics analyzed in the results. The MODE-TD characteristics in the table include: total number of detected objects (Table 2), number of detected objects with time (Figure 4a), percentage of objects with time (Figure 4b), size distribution of detected object count (Figure 5a), size distribution of percentage of detected objects (Figure 5b), detected object size with

time (Figure 7), number of newly tracked objects (Figure 8a), percentage of newly tracked objects (Figure 8c), number of discarded tracked objects (Figure 8b), percentage of discarded objects (Figure 8d), the 10^{th} – 90^{th} percentile [rain] intensities (Figure 10, Figure 11, Figure 12, Figure 13, and Figure 14) and tacked object velocity (Figure 22). The determination of the best performing microphysics for each attribute was chosen based off the plots in the analysis. The affiliated plot of each MODE-TD characteristic was examined independently and weighted equally. Note that MODE-TD characteristics that were analyzed in terms of both number and percentage of objects were scored separately. Morrison was the best performing microphysics scheme with nine and a half points and Milbrandt was the worst performing scheme with only three points.

Table 7: Best performing microphysics for each MODE-TD characteristic. Checkmarks (✓) represent a clear winner and are worth 1.0 point. Hyphens (-) represent no clear best performing microphysics or tie for best performing microphysics and are worth 0.5 points.

	Thompson	Milbrandt	Morrison	WSM6
Total # of detected objects	-		-	
# of detected objects with time			✓	
% of detected objects with time			✓	
Size distribution (#)	-		-	
Size distribution (%)	-	-	-	-
Size with time		✓		
Initiated (#)	-		-	
Initiated (%)		-	-	
Dissipated (#)	-		-	
Dissipated (%)		-	-	-
Duration (#)	-		-	
Duration (%)				✓
10th Percentile			✓	
25th Percentile			✓	
50th Percentile				✓
75th Percentile	✓			
90th Percentile			✓	
Velocity	-	-	-	-
Total:	**4.5**	**3.0**	**9.5**	**3.5**

CHAPTER IV

DISCUSSION

Implications to Forecasting

As mentioned earlier, previous studies have shown that the choice of microphysics parameterization significantly affects simulated convective storm structure and intensity. The study herein examined sensitivities of simulated convective events due to the choice of microphysics parameterization over a large number of cases that span a variety of spring and summer conditions, in hopes of obtaining results and systematic biases that are more significant than examining a single case study. Therefore, obtaining results that could aid forecaster when interpreting model guidance. In addition, the results could help indicate a best performing microphysics scheme that could be implemented more widely among operational models. The utilization of MODE-TD to analyze the results was deliberately used as it provides many performance metrics useful to an operational forecaster. Model behavior and bias was established by analyzing precipitation regions with time using MODE-TD, which does not rely on point-to-point, verification methods, but rather uses an object-based approach that replicates more of the forecaster's interpretation of skill of model guidance. MODE-TD distinguishes individual storms and gives quantitative metrics, such as size, intensity, propagation speed that common verification methods are not capable of doing. Thus, by utilizing MODE-TD for the analysis and analyzing a multitude of cases, the study herein was able to gather results that are transferable to operational forecasting.

In general, the results revealed previously-established and well-known conclusions for CAM forecasts and lesser-known conclusions specific to microphysics scheme choice, all of which have forecast implications. The well-documented conclusions include CAMs tendency to overforecast surface rainfall and over simulate the occurrence of convective storms. Thus, the high precipitation bias reaffirmed in this study reiterates a CAM bias forecasters should continue to take into consideration. As mentioned earlier, Milbrandt and the sole single-moment scheme in the study, WSM6, were the two worst performing schemes, with Milbrandt performing the worst. It is surprising that Milbrandt performed the worst as it is the most complex scheme within the microphysics ensemble used in this study. On the contrary is not that surprising that WSM6 was not among the best performers, as many studies have shown double-moment schemes outperform single-moment schemes when simulating convective storm structure and intensity. Although Morrison was declared the best performing microphysics scheme based off the ranking criteria used in this study, there was no clear best performer in the schemes tested as Morrison was not consistently the best performer across all of the MODE-TD characteristics. The results from this study also support the use of probabilistic forecasting utilizing ensembles rather than a single deterministic solution as there is no clear best performer among microphysics.

Overprediction of Detected Objects

The overforecast of detected objects by all microphysics is one of the most noticeable biases evident from the results herein. WSM6 and Milbrandt had the largest biases with

percentage increases relative to Stage IV of 43.1% and 34.5%, respectively; and Morrison and Thompson had the smallest biases of 17.4% and 17.8%, respectively. On the contrary, Clark et al. (2014) found Thompson to have the largest bias in detected objects when comparing simulated forecasts from WSM6, WDM6, Morrison, and Thompson to Stage IV. The discrepancy is most likely due to the difference in choice of precipitation thresholds and convolution radii between studies. Clark et. al (2014) used smaller thresholds and radii retaining the weaker smaller precipitation areas, which Thompson seems to generate the most of. By increasing the precipitation threshold and convolution radius the smaller, weaker precipitation regions are smoothed Yet, the simulations and observations herein still produced many small objects (30 – 60 km); more than any other size range.

The high precipitation bias found is not abnormal as CAMs typically overpredict precipitation based off the findings from previous studies (Davis et al. 2009, Johnson et al. 2013, Clark et al. 2014, and Goines 2017). The exact reason why CAMs tend to have a high bias in precipitation remains unknown. Johnson et al. (2013) suggested it is due to errors within the model dynamics and physics, which is a broad reason and does not further clarify the problem. Davis et al. (2009) suggested the source of the problem might be from the misrepresentation of numerical dissipation leading to excessive small-scale convection. While Goines et al. (2017) hypothesized the model bias may be from the choice of microphysics scheme; the work herein indicates that all microphysics schemes significantly overproduce rain regions, indicating the choice of microphysics is not the sole reason for the high precipitation bias. Although, some schemes are more likely to produce a higher occurrence of rain regions. Still one cannot completely rule out microphysics, because all microphysics may treat a process or make an

assumption in a similar manner causing the precipitation bias. One possibility that can be ruled out is the choice of CPS since that was turned off in these CAM simulations.

Microphysics Performance

Based of the microphysics performance (Table 7) and interpretation of the plots displaying the objective and quantitative MODE-TD attributes, Morrison performed the best overall and Milbrandt the worst. Note that the interpretation of the plots and the weights applied to each MODE-TD characteristic could affect the ranking scores. In addition, Morrison did not did not outperform the other microphysics in all the MODE-TD characteristics.

A noteworthy result is the microphysics performance for the percentile [rain] intensities. Morrison best represented percentile [rain] intensities in simulated storms, while WSM6 performed the worst. WSM6 had the largest spread and Milbrandt the second largest spread in rain rate distributions. They produced the lowest rain rates for the smallest percentiles (10^{th} and 25^{th}) and the highest rain rates for the largest percentiles (75^{th} and 90^{th}), possibly indicating a poor representation of typical progressive transition of rain rates between the convective and stratiform regions.

The differences in percentile [rain] intensities and distributions can possibly be attributed to the differences in predicted moments and species among microphysics schemes. WSM6 is the only single-moment scheme used in the microphysics ensemble; therefore, it does not include number concentration as a prognostic variable. Double-moment schemes have

been shown to better represent the convective and stratiform regions and associated transition between regions within a storm, for case studies and idealized 2D studies. Single-moment schemes, such as WSM6 herein, tend to overpredict peak convective rainfall and underpredict light stratiform rainfall, suggesting why WSM6 produced the lowest rain rates for the 10th and 25th percentiles and the highest rates for the 75th and 90th percentiles. These findings are supported in Morrison et al. (2009) and Bryan and Morrison (2012). The Milbrandt scheme had the largest spread of rain intensities among double-moment schemes in the microphysics ensemble herein. The differences in performance among double-moment schemes (Thompson, Morrison, and Milbrandt) are most likely due to the representation of rimed ice species between the schemes. Other microphysical properties, such as the representation of particle species and assumptions that may differ between schemes could also be reason for the differences. Milbrandt includes hail as a prognostic species in contrast to Morrison and Thompson, which only predict graupel. Schemes containing hail tend to produce the most precipitation and highest peak reflectivities in the convective region, but underpredict peak reflectivity and precipitation in the stratiform region (e.g., Gilmore et al. 2004b and Wu et al. 2013). Thus, this could explain why Milbrandts' 10th and 25th percentile [rain] intensities were the lowest and the 75th and 90th percentiles were the highest, making it the worst performing microphysics among the double-moment schemes for the percentile [precipitation] intensity characteristic.

Limitations / Weaknesses

As in all studies, some limitations exist within this study. This study only analyzed the microphysics ensemble simulations using one combination of a precipitation threshold and convolution radius. Although, sensitivities of CAM analysis do arise depending on choice of precipitation threshold and convolution radius as explained in past studies (Clark et al. 2014 and Goines et al. 2017). After deliberation, it was decided that there was little significance in performing the analysis using multiple MODE thresholds and radii because past studies have done so and our focus was on large propagating convective systems. Upon further analysis, it maybe helpful to adjust to a larger convolution radius to further the focus on MCSs and eliminate small scale convection that might impact the results. Although, this may not be the best option as this could smooth too much and possibly remove convective precipitation regions important to the formation of MCSs. Another option would be to discard short-lived objects from the statistical analysis.

There are also a few limitations, assumptions, or nuances to consider when interpreting the results. These include the size assumption, object identification process, initiation/dissipation of tracked objects at first/last lead times, and poor model initialization. Recall that when determining object size, the size was calculated by taking the square root of the area, which is a count of the grid squares a region of precipitation occupies. The resulting length, as calculated herein, corresponds to a side of a square – often quite unlike the original storm's shape. A nuance to be considered arises when MODE-TD identifies a tracked object. MODE-TD also identifies individual precipitation regions that merge into one precipitation

region or split into separate precipitation regions as one tracked object overtime. Thus, this needs to be considered when analyzing the MODE-TD characteristics of tracked objects (initiation, dissipation, duration, precipitation percentile intensities, and velocity). Also, MODE-TD inherently identifies tracked objects present at the first analysis time (12 UTC) as newly-detected and precipitation regions remaining during the last forecast time as discarded. Storms may, in fact, be ongoing at both of those times. Thus, this may affect the tracked object attribute duration as MODE-TD could have identified objects as being first initiated or dissipated when they actually were not. Lastly, the results do not consider simulations that initialized poorly where all microphysics schemes performed poorly from the start and never capture the main convective activity. To investigate further the affect of model initialization, selected MODE-TD attributes were analyzed, which were separated by real-time and retrospective cases (Appendix G).

Future Work

For future work, further analysis could be performed separating simulations that captured the presence of the MCS or main convective activity versus those that did not. It is possible that there were some case dates where all simulations performed poorly and there was little resemblance to the Stage IV observations. Thus, the intercomparison of microphysics for these cases would be of little use because the main convective activity or MCS was not present. In addition, separating the small-scale convection and MCSs as to focus the analysis more on large-scale convection would provide further insight on the impact of microphysics on

MCS behavior or structure. One could possibly address this by applying a larger convolution radius when utilizing MODE-TD to help eliminate small scale convection, although as mentioned previously, this may smooth/remove precipitation regions pertinent to MCS formation. A more robust study that considers the impact of the boundary layer scheme, the roles of particle processes and assumptions within the microphysics, and the model dynamics would also be of use to determine the reason for the high precipitation bias. Lastly, analyzing the microphysics ensemble simulations according to synoptic regime would reveal information regarding the performance of microphysics across varying atmospheric conditions.

CHAPTER V

CONCLUSIONS

This study performed a detailed analysis of convective simulations from a 4-member microphysics WRF ensemble. These convective simulations included 77 total case dates from retrospective simulations spanning from 2010-2012 and real-time simulations performed during the 2016 NOAA Spring Forecast Experiment (SFE). The microphysics schemes in the ensemble included WSM6, Thompson, Morrison, and Milbrandt. The convective simulations were compared to Stage IV gauge-adjusted radar derived precipitation and analyzed using MODE-TD. The main objective of this study was to analyze the impacts microphysics parameterization, within high-resolution models, have on simulated convective properties, mainly MCSs, in hopes of aiding operational forecasting. Key results from this study included:

- There was an overprediction of total detected objects by all microphysics schemes with WSM6 the most and Morrison the least.
- For all microphysics schemes and Stage IV, the 30 – 60 km size range had the most detected objects.
- The simulations represented temporal trends in detected object size well, but the simulations were shifted earlier than Stage IV. Milbrandt and Morrison had a large bias in detected object size, while Thompson and Morrison had a small bias.
- All microphysics schemes all captured trend of object duration well with Morrison and Thompson performing best for duration, in terms of number tracked objects, and WSM6 performing best for duration, in terms of percentage of tracked objects. All datasets had a peak object duration of 2 h with a sharp drop off for longer durations.
- Morrison matched Stage IV the best and WSM6 the worst for percentile [precipitation] intensities.
- Based off the ranking criteria herein, Morrison was the best performing microphysics scheme overall and Milbrandt the worst.

REFERENCES CITED

Accadia, C., S. Mariani, M. Casaioli, A. Lavagini, and A. Speranza, 2003: Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high resolution verification grids. *Wea. Forecasting,* **18**, 918-932.

Baldwin, M.E., and K.E. Mitchell, 1997: The NCEP hourly multisensory U.S. precipitation analysis for operations and GCIP research. Preprints, 13th Conf. on Hydrology, Long Beach, CA, *Amer. Meteor. Soc.,* 54-55.

Battan, L., 1973: Radar Observation of the Atmosphere. University of Chicago Press, 324 pp.

Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev*., **140**, 202-225.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev*., **129**, 569–585.

Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting*, **24**, 1121–1140,

Clark, A. J., and Coauthors, 2012a: An Overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc*., **93**, 55-74.

Clark, A.J., 2012: Program Overview and Operations Plan. Spring Forecasting Experiment 2012 Conducted by the Experimental Forecast Program of the NOAA/Hazardous Weather Testbed, 18.

Clark, A.J., 2014: Program Overview and Operations Plan. Spring Forecasting Experiment 2014

Conducted by the Experimental Forecast Program of the NOAA/Hazardous Weather Testbed, 24.

Clark, A.J., R. G. Bullock, T. L. Jensen, M. Xue, and F. Kong, 2014: Application of object-based domain diagnostics for tracking precipitation systems in convection allowing models. *Wea. Forecasting,* ***29***, 517-542.

Clark, A.J., 2015: Program Overview and Operations Plan. Spring Forecasting Experiment 2015 Conducted by the Experimental Forecast Program of the NOAA/Hazardous Weather Testbed, 23.

Coniglio, M., 2013: Program Overview and Operations Plan. Spring Forecasting Experiment 2013 Conducted by the Experimental Forecast Program of the NOAA/Hazardous Weather Testbed, 21.

Cotton, W.R., M. A. Stephens, T. Nehrkorn, and G.J Tripoli, 1982: The Colorado State University three-dimensional cloud-mesoscale model-1982. Part II: An ice-phase parameterization. *J Rech. Atmos*., **16**, 295-320.

Davis, C. A., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: methodology and application to mesoscale rain areas. *Mon. Wea. Rev*., **134**, 1772-1784.

Davis, C.A., B. Brown, and R. Bullock, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev*., **134**, 1785-1795.

Davis, C.A., B. Brown, and R. Bullock, J. Halley-Gotway, 2009: The Method for Object-Based

Diagnostic Evaluation (MODE) applied to numerical forecast from the 2005 NSSL/SPC Spring Program. *Wea. Forecasting*, **24**, 1252-1267.

Doviak, R., and D. Zrnić, 1984: Doppler Radar and Weather Observations. Academic Press, 458 pp.

Fovell, R.G., and Y. Ogura, 1988: Numerical simulation of a mid-latitude squall line in two dimensions. *J. Atmos. Sci*., **45**, 3846-3879.

Fowle, M. A., and P. J. Roebber, 2003: Short-range (0–48 h) numerical prediction of convective occurrence, mode, and location. *Wea. Forecasting*, **18**, 782–794.

Fritsch, J. M., and G. S. Forbes, 2001: Mesoscale convective systems. *Meteor. Monogr.,* **28**, 323-358.

Fulton, R.A., J.P. Breidenbach, D.J. Seo, D.A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. and Fore*., **13**, 377-395.

Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J Atmos. Oceanic Technol*., **21**, 457-469.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004a: Precipitation and evolution sensitivity in simulated deep convective storms: comparisons between liquid-only and simple ice and liquid phase microphysics. *Mon. Wea. Rev*., **132**, 1897–1916.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004b: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev*., **132**, 2610–2627.

Goines, D., 2017: Precipitation properties in a multi-year database of convection-allowing WRF simulations. M.S. thesis, Dept. of Atmospheric Sciences, University of North Dakota, 35 pp.

Hong, S.–Y., and J.–O. J. Lim, 2006: The WRF single–moment 6–class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Houze, R. A. Jr., 1977: Structure and dynamics of a tropical squall-line system. *Mon. Wea. Rev*., **105**, 1540–1567.

Houze, R.A. Jr., B. F. Smull, and P. Dodge, 1990: Mesoscale organization of springtime rainstorms in Oklahoma. *Mon. Wea. Rev*., **118**, 613–654.

Houze, R. A. Jr., 2004: Mesoscale convective systems, *Rev. Geophys*., **42**, RG4003.

Janjic, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.

Johnson, A., and X. Wang, 2013: Object-based evaluation of a storm-scale ensemble during the 2009 NOAA Hazardous Weather Testbed Spring Experiment. *Mon. Wea. Rev.,* **141,** 1079-1098.

Kain, J. S., P. R. Janish, S. J. Weiss, M. E. Baldwin, R. S. Schneider, and H. E. Brooks, 2003: Collaboration between forecasters and research scientists at the NSSL and SPC: The spring program. Bull. *Amer. Meteor. Soc*., **84**, 1797–1806.

Kain, J. S., J. Weiss, J. Levit, M. E. Baldwin, and D.R. Bright, 2006: Examination of convection-allowing configurations of the WRF model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Wea. Forecasting*, **21**, 167–181.

Kain, J. S., J. Weiss, D.R. Bright, M.E. Baldwin, J. J. Levit, G. W. Carbin, and K.W. Thomas, 2009. Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23(5)**, 931-952.

Kain, J. S., and Coauthors, 2010a: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting–research environment. *Wea. Forecasting*, **25**, 1510–1521.

Hilgendorf, E.R., and R.H. Johnson, 1998: A study of the evolution of mesoscale convective systems using WSR-88D data. *Wea. Forecasting*, **13**, 437-452.

Lin, Y.-L., R.D. Farlet, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor*., **22**, 1065-1092.

McAnelly, R. L., and W. R. Cotton (1989), The precipitation life cycle of mesoscale convective Complexes. *Mon. Wea. Rev*., **117**, 784–808.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys*., **20**, 851–875.

Milbrandt, J. A., and M. K. Yau (2005), A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description, *J. Atmos. Sci*., **62**, 3065_3081.

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two- moment schemes. *Mon. Wea. Rev*., **137**, 991-1007.

Morrison, H., and J. Milbrandt, 2011: Comparison of two-moment bulk microphysics schemes in idealized supercell thunderstorm simulations. *Mon. Wea. Rev*, **139**, 1103-1130.

NOAA, 2017: Billion-Dollar Weather and Climate Disasters: Overview. Accessed on 17 June 2017. [Available online at http://www. ncdc.noaa.gov/billions/.]

Rauber, R.M., J.H. Walsh. and D.H Charlevoix, 2008. Severe and Hazardous Weather, 3rd ed., 642 pp., Kendall/Hunt Publishing Company, Dubuque, Iowa.

Rogers, E., and Coauthors, 2009: The NCEP North American Mesoscale modeling system: Recent changes and future plans. Preprints, 23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction, Omaha, NE, Amer. Meteor. Soc., 2A.4.

Rutledge, S. A., and P.V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in mid-latitude clouds. VIII: A model for the "seeder-feeder" process in warm-frontal rainbands. *J. Atmos. Sci*., **40**, 1185-1206.

Rutledge, S. A., and P.V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in mid-latitude clouds. Part XII: A diagnostic modeling study of precipitation development in narrow cold frontal rainbands. *J. Atmos. Sci*., **41**, 2949-2972.

Schwartz, C. S., J.S. Kain, J. Weiss, M. Xue, D.R. Brigh, F. Kong, and M.C Coniglio, 2009: Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev*., **137(10)**, 3351-3372.

Schwartz, C. S., J.S. Kain, J. Weiss, M. Xue, D.R. Bright, F. Kong, and M.S. Wandishin, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25(1)**, 263-280.

Skamarock W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech Note NCAR/TN-475+STR, 113 pp.

Stensrud, David, 2007: Parameterization Schemes: Keys to Understanding Numerical Weather Prediction. 1st ed. Cambridge University Press, 478 pp.

Straka, Jerry M., 1994: Representing Moisture Processes in Mesoscale Numerical Models. *Mesoscale Modeling of the Atmosphere, Meteor. Monogr.,* No. 47, Amer. Meteor. Soc, 29-39.

Tao, W.-K., and J. Simpson, 1989: Modeling study of a tropical squall-type convective line. *J. Atmos. Sci*., **46**, 177-202.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev*., **136**, 5095-5115.

Towns J., T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, N. Wilkins-Diehr, "XSEDE: Accelerating Scientific Discovery", Computing in Science & Engineering, vol. 16, no. 5, pp. 62 – 74, Sept.-Oct. 2014.

Xue, M.D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: the Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Metero. Atmos. Phys*., **82**, 139-170.

Warner, Thomas, 2011: Numerical Weather and Climate Prediction. 1st ed. Cambridge University Press, 522 pp.

Weisman Morris L., Skamarock William C., and Klemp Joseph B., 1997: The Resolution Dependence of Explicitly Modeled Convective Systems. *Mon. Wea. Rev.,* **125**, 527–548.

Weiss, S., 2010: Program Overview and Operations Plan. NOAA Hazardous Weather Testbed Experimental Forecast Program Spring Experiment 2010, 62.

Weiss, S., 2011: Program Overview and Operations Plan. NOAA Hazardous Weather Testbed Experimental Forecast Program Spring Experiment 2011, 64.

Wu, D., X. Dong, B. Xi, Z. Feng, A. Kennedy, G. Mullendore, M. Gilmore, and W.-K. Tao, 2013: Impacts of microphysical scheme on convective and stratiform characteristics in two high precipitation squall line events, *J. Geophys. Res. Atmos*., **118**, 11, 119–11.

APPENDICES

APPENDIX A
WRF Namelist Variables

Table A1: List of the WRF namelist variables and the settings used for the retrospective and real-time runs.

Namelist Variable	**Value**
&time_control	
run_days	1,
run_hours	12,
run_minutes	0,
run_seconds	0,
start_year	2016,
start_month	05,
start_day	24,
start_hour	00,
start_minute	00,
start_second	00,
end_year	2016,
end_month	05,
end_day	25,
end_hour	12,
end_minute	00,
end_second	00,
interval_seconds	10800,
input_from_file	.true.,
history_interval	60,
frames_per_outfile	1,
restart	.false.,
restart_interval	5000,
io_form_history	2,
io_form_restart	2,
io_form_input	2,
io_form_boundary	2,
debug_level	0,
history_outname	./wrfout_d<domain>_<date>',
nwp_diagnostics	1,
output_ready_flag	.true.,
/	
&domains	
time_step	15,
time_step_fract_num	0,

time_step_fract_den	1,
max_dom	1,
s_we	1,
e_we	1681,
s_sn	1,
e_sn	1153,
s_vert	1,
e_vert	51,
num_metgrid_levels	40,
dx	3000,
dy	3000,
grid_id	1,
parent_id	0,
i_parent_start	1,
j_parent_start	1,
parent_grid_ratio	1,
parent_time_step_ratio	1,
feedback	1,
smooth_option	0,
eta_levels	1.0000, 0.9980, 0.9940, 0.9870, 0.9750, 0.9590, 0.9390, 0.9160, 0.8920, 0.8650, 0.8350, 0.8020, 0.7660, 0.7270, 0.6850, 0.6400, 0.5920, 0.5420, 0.4970, 0.4565, 0.4205, 0.3877, 0.3582, 0.3317, 0.3078, 0.2863, 0.2670, 0.2496, 0.2329, 0.2188, 0.2047, 0.1906, 0.1765, 0.1624, 0.1483, 0.1342, 0.1201, 0.1060, 0.0919, 0.0778, 0.0657, 0.0568, 0.0486, 0.0409, 0.0337, 0.0271, 0.0209, 0.0151, 0.0097, 0.0047, 0.0000,
/	
&physics	
mp_physics	8,
ra_lw_physics	4,
ra_sw_physics	4,
radt	15,

sf_sfclay_physics	2,
sf_surface_physics	2,
bl_pbl_physics	2,
bldt	0,
cu_physics	0,
cudt	5,
isfflx	1,
ifsnow	0,
icloud	1,
surface_input_source	1,
num_soil_layers	4,
sf_urban_physics	0,
iz0tlnd	0,
mp_zero_out	0,
mp_zero_out_thresh	1.e-8,
do_radar_ref	1,
cugd_avedx	3,
maxiens	1,
maxens	3,
maxens2	3,
maxens3	16,
ensdim	144,
lightning_option	0,
lightning_dt	18,
lightning_start_seconds	0,
flashrate_factor	1.0,
cellcount_method	1,
iccg_method	2,
num_land_cat	24,
prec_acc_dt	60.,
hailcast_opt	1,
/	
&fdda	
/	
&dynamics	
w_damping	1,
diff_opt	1,
km_opt	4,
diff_6th_opt	0,
diff_6th_factor	0.12,
base_temp	290.00,

damp_opt	3,
zdamp	5000.,
dampcoef	0.2,
khdif	0,
kvdif	0,
non_hydrostatic	.true.,
moist_adv_opt	2,
scalar_adv_opt	2,
use_input_w	.false.,
iso_temp	0,
/	
&bdy_control	
spec_bdy_width	5,
spec_zone	1,
relax_zone	4,
specified	.true.,
nested	.false.,
/	
&grib2	
/	
&namelist_quilt	
nio_tasks_per_group	0,
nio_groups	1,
/	
&dfi_control	
/	
&afwa	
afwa_diag_opt	1,
afwa_severe_opt	1,
afwa_ptype_opt	0,
afwa_buoy_opt	0,
afwa_therm_opt	0,
afwa_turb_opt	0,
afwa_radar_opt	0,
afwa_vil_opt	0,
afwa_icing_opt	0,
afwa_vis_opt	0,
afwa_cloud_opt	0,
/	
&stoch	
stoch_force_opt	0,

stoch_vertstruc_opt	0,
tot_backscat_psi	1.E-05,
tot_backscat_t	1.E-06,
nens	1,

Appendix B
Retrospective Case Dates

Table B1: Chronological listing of the reduced retrospective case dates with one or more MCSs traveling through the specified regions in Figure 3. The regions are southern Great Plains (SGP), northern Great Plains (NGP), Midwest (MW), Northeast (NE), and Gulf Coast (GC). The 'x' in the simulated column, represents the retrospective case dates simulated and included in the statistical analysis.

Simulated	Year	Month	Day	SGP	NGP	MW	NE	GC	Total
x	2010	4	7			1		1	2
x	2010	4	23			1		1	2
x	2010	4	24			1		1	2
x	2010	5	12	1		1			2
x	2010	5	15	1				1	2
x	2010	5	24	1	1				2
x	2010	5	29		1			1	2
x	2010	5	30	1				1	2
x	2010	5	31			1		1	2
x	2010	6	1			1	1		2
x	2010	6	4			1	1	1	3
x	2010	6	5			1		1	2
x	2010	6	8	1	1	1			3
x	2010	6	10		1		1		2
x	2010	6	12	1		1			2
x	2010	6	13	1		1			2
x	2010	6	14	1		1	1	1	4
x	2010	6	15			1		1	2
x	2010	6	17		1	1			2
x	2010	6	18			2			2
x	2010	6	22			1	1		2
	2010	6	26		1	1			2
x	2010	6	28				1	1	2
x	2010	7	3	1	1				2
x	2010	7	5	1				1	2
x	2010	7	12	1				1	2
x	2010	7	16				1	1	2
x	2010	7	19			1	1		2
x	2010	7	22		1	1	1		3

x	2010	7	23			1		1	2
x	2010	7	24	1		1		1	3
x	2010	8	4			1	1		2
x	2010	8	10		1	1			2
x	2010	8	13		1	1			2
	2010	8	15				1	1	2
	2010	8	16	1			1		2
	2010	8	21				1	1	2
	2010	8	31	1		1			2
x	2010	9	1	1	1	1			3
	2010	9	2	1		1			2
	2011	4	3			2			2
	2011	4	4				1	1	2
	2011	4	11				1	1	2
	2011	4	15			1		1	2
	2011	4	19			1	1		2
	2011	4	24						
x	2011	4	25						
	2011	4	26				1	1	2
x	2011	5	13			2		1	3
x	2011	5	18	1			1		2
x	2011	5	20						
x	2011	5	21		1	1	1		3
	2011	5	22		1	1			2
	2011	5	23	1			1		2
	2011	5	24						
x	2011	6	10	1		1	1		3
	2011	6	17		1	2	1	1	5
x	2011	6	28	1			1	1	3
x	2011	7	14		2			1	3
x	2011	7	26		1		1	1	3
x	2011	8	6	1			1	1	3
x	2012	5	6	1		1		1	3
x	2012	8	9			1	1	1	3

Appendix C
South Central US Region (40° to 30° N and 105° to °90 W)

In general, the results for the south central region were similar to the results from the analysis domain used for the main analysis of the paper. A high bias in detected objects across all microphysics was still evident in the south central region, with WSM6 and Milbrandt overpredicting the most. In addition, the performance of the microphysics schemes for the MODE-TD characteristics of initiation/dissipation, duration, and percentile [precipitation] intensities had little difference than the results from the main analysis. One noticeably difference was that all microphysics underpredicted the detected object size across most forecast hours (Figure C4), which was not the case in the main analysis.

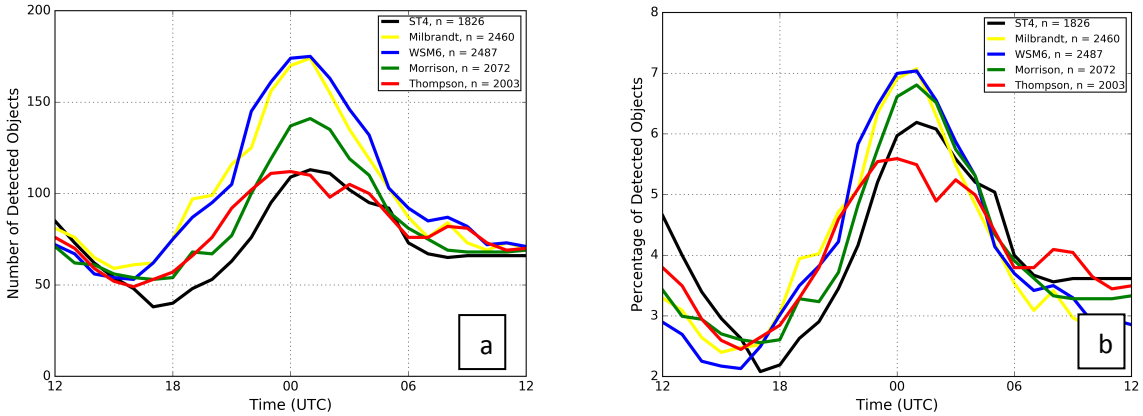

*Figure C1: (a) Number of detected (b) percent of total detected objects ((# of objects at time / n)*100) for the microphysics and Stage IV at each forecast hour (UTC time). The n refers to the total number of detected objects accumulated over all forecast hours and case dates.*

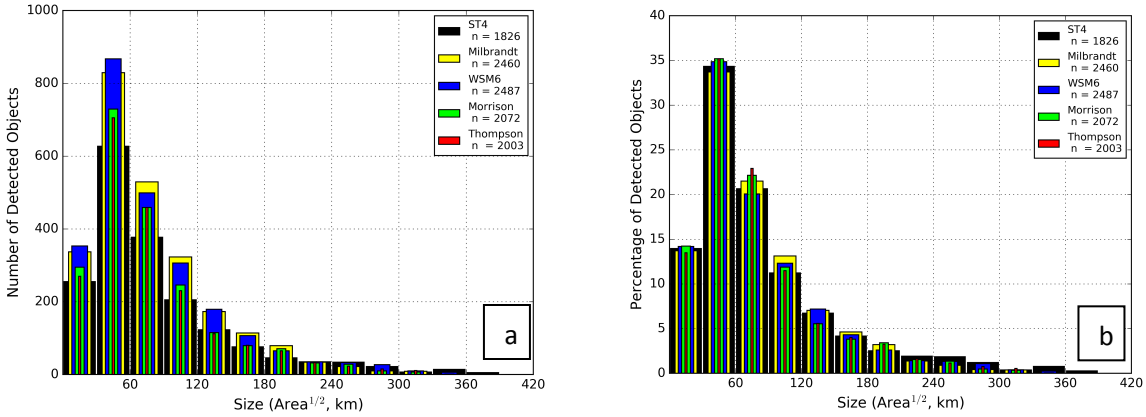

*Figure C2: Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size for all case dates*

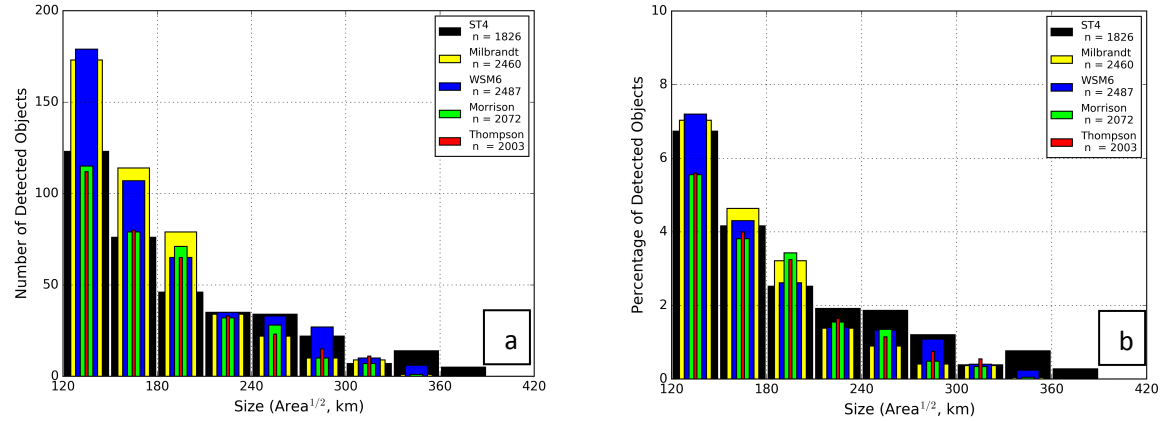

*Figure C3: Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size greater than 120 km for all case dates*

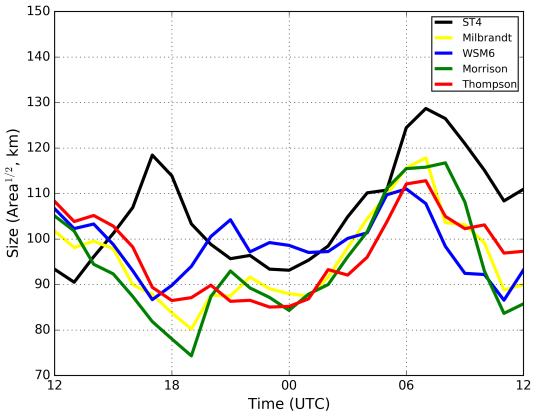

Figure C4: Average size (square root of area, expressed in km) of detected objects as a function of UTC time

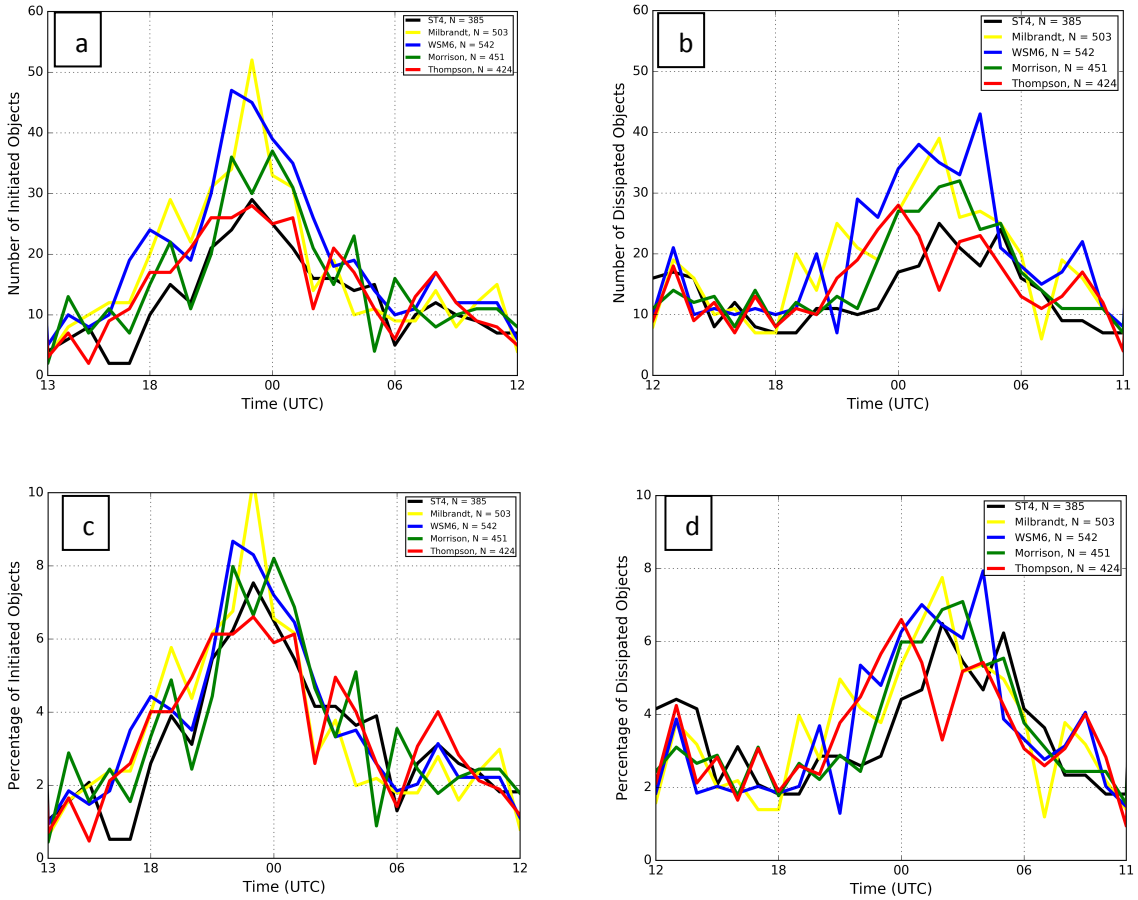

*Figure C5: The number of tracked objects that were a) newly detected and b) discarded as a function of UTC time. The percentage of detected objects ((# of objects at time / n)*100) that were c) newly detected and discarded d) as a function of UTC time. N represents the total number of tracked objects summed over all the case dates.*

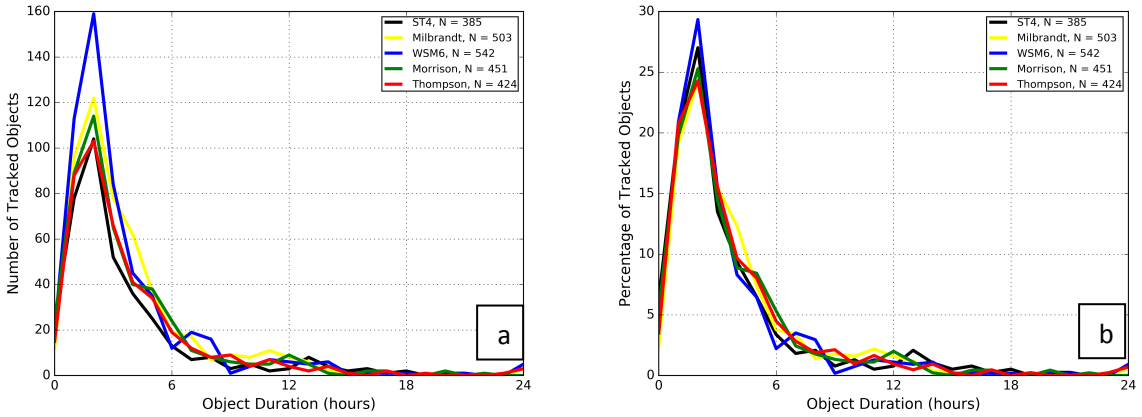

*Figure C6: Number of tracked objects (a) and percentage of tracked objects ((# of objects at time / n)*100) (b) as a function of object duration. N represents the total number of tracked objects summed over all the case dates*

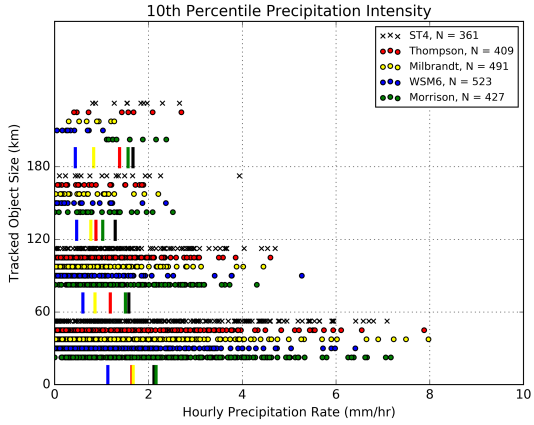

Figure C7: Tracked object 10th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 10th percentile precipitation intensities of all the tracked objects within the corresponding size range.

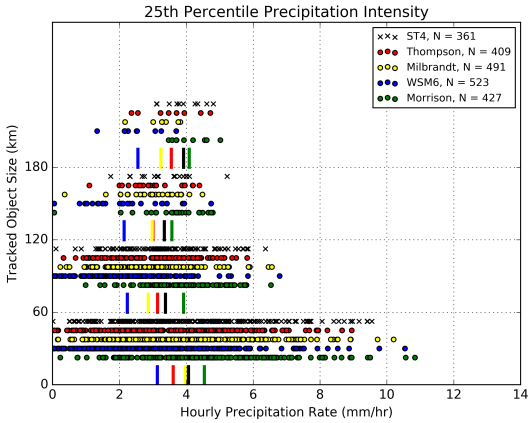

Figure C8: Tracked object 25th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 25th percentile precipitation intensities of all the tracked objects within the corresponding size range.

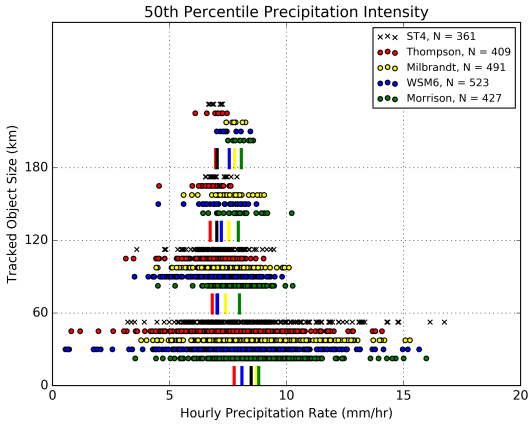

Figure C9: Tracked object 50th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 50th percentile precipitation intensities of all the tracked objects within the corresponding size range.

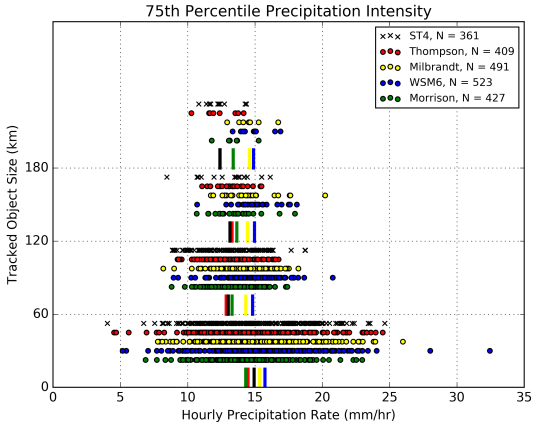

Figure C10: Tracked object 75th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 75th percentile precipitation intensities of all the tracked objects within the corresponding size range.

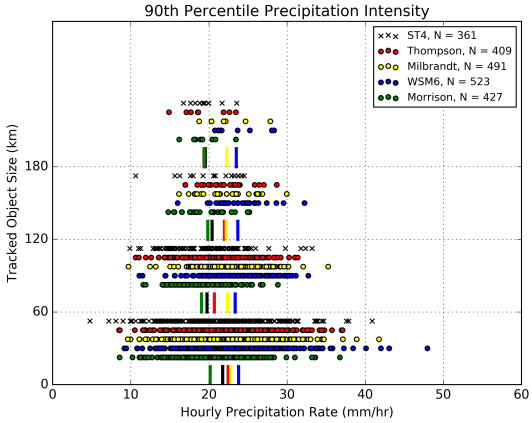

Figure C11: Tracked object 90th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 90th percentile precipitation intensities of all the tracked objects within the corresponding size range.

Appendix D
Southeastern US Region (40° to 30° N and 90° to 75° W)

Similar to the south central region (Appendix C), the results for the southeastern region were similar to the main analysis results. A noticeable difference from the main analysis was with object size at each forecast hour (Figure D4). All microphysics overpredicted the sizes at each forecast hour, excluding, the first 6 h, which was not the case in the main analysis.

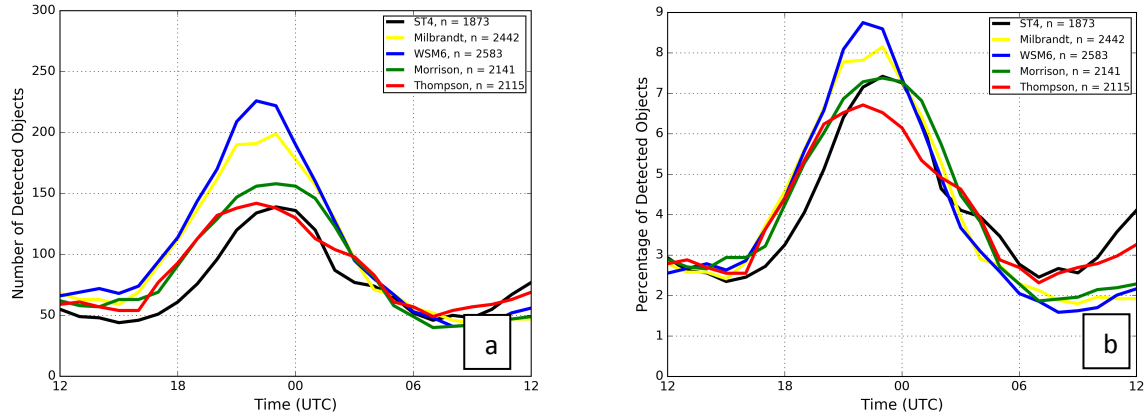

*Figure D1: (a) Number of detected (b) percent of total detected objects ((# of objects at time / n)*100) for the microphysics and Stage IV at each forecast hour (UTC time). The n refers to the total number of detected objects accumulated over all forecast hours and case dates.*

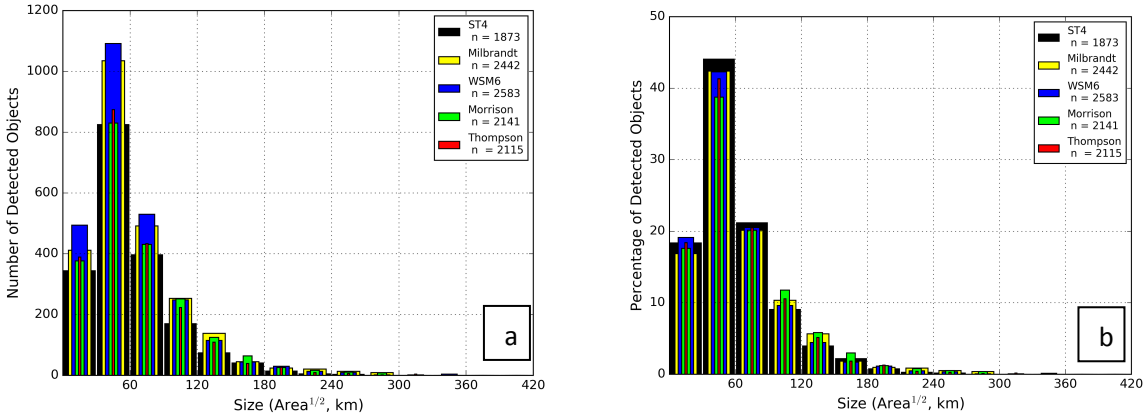

*Figure D2: Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size for all case dates*

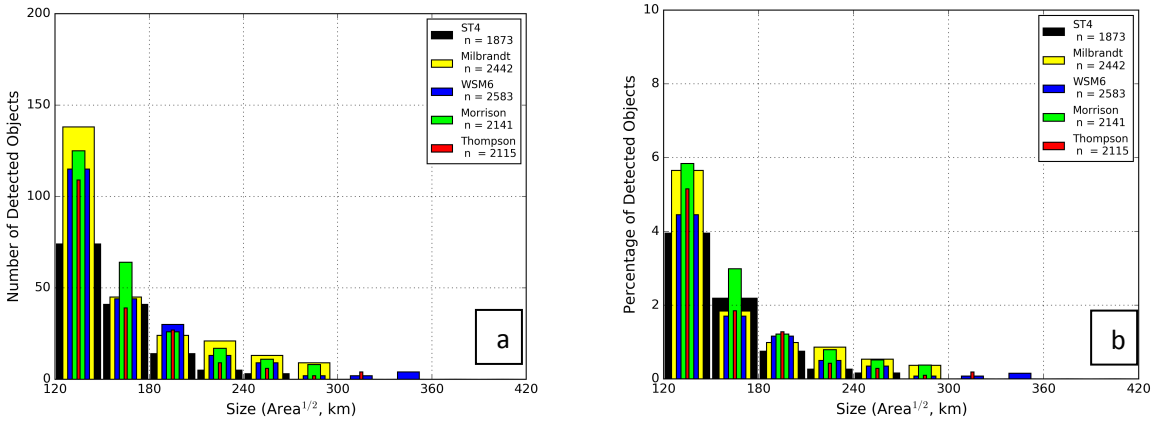

*Figure D3: Distribution of (a) the number or (b) percentage of detected objects ((# of objects at size / n)*100) for each object size greater than 120 km for all case dates*

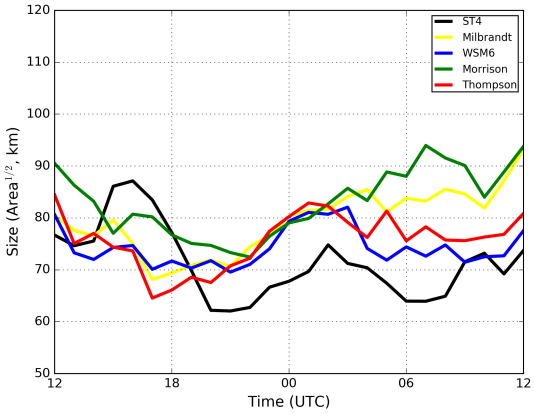

Figure D4: Average size (square root of area, expressed in km) of detected objects as a function of UTC time

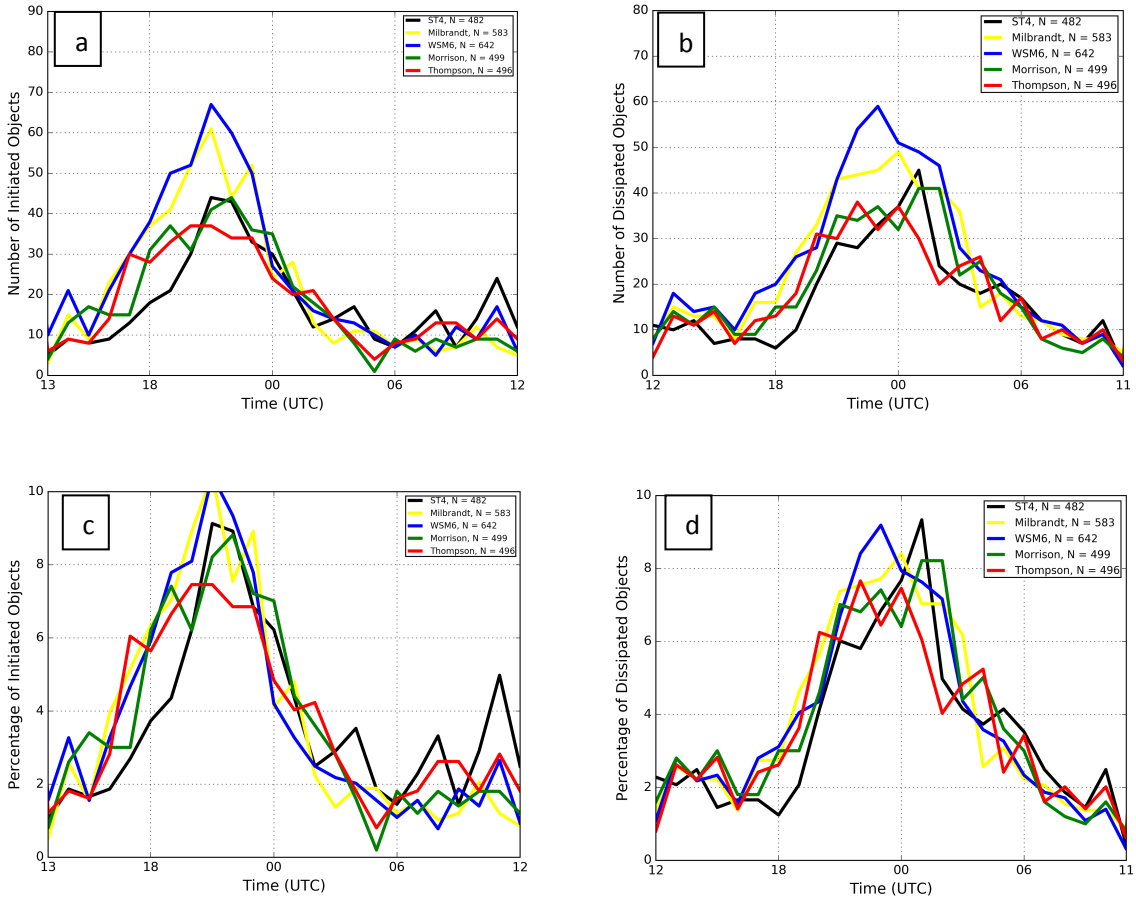

*Figure D5: The number of tracked objects that were a) newly detected and b) discarded as a function of UTC time. The percentage of detected objects ((# of objects at time / n)*100) that were c) newly detected and discarded d) as a function of UTC time. N represents the total number of tracked objects summed over all the case dates.*

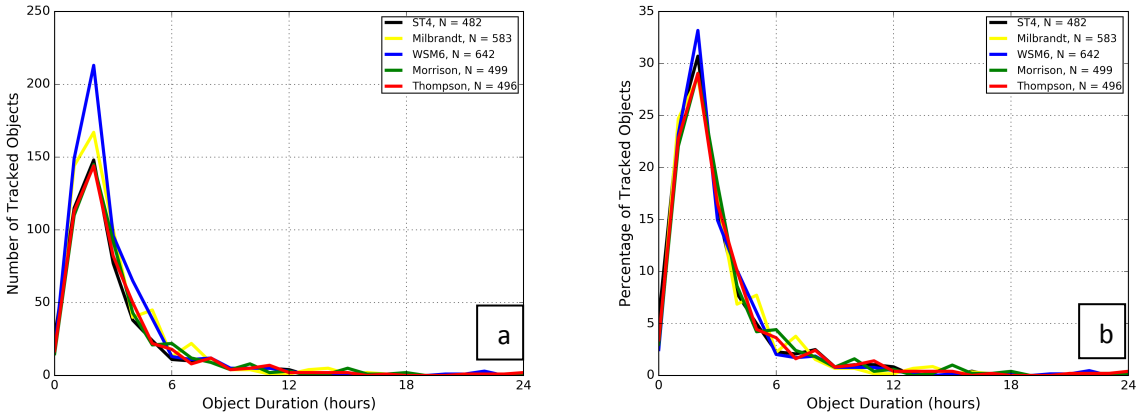

*Figure D6: Number of tracked objects (a) and percentage of tracked objects ((# of objects at time / n)*100) (b) as a function of object duration. N represents the total number of tracked objects summed over all the case dates*

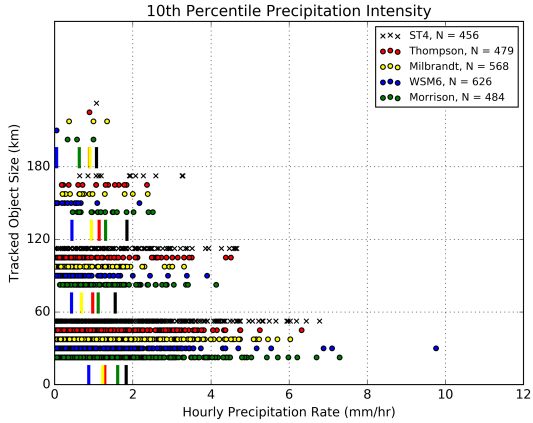

Figure D7: Tracked object 10th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 10th percentile precipitation intensities of all the tracked objects within the corresponding size range.

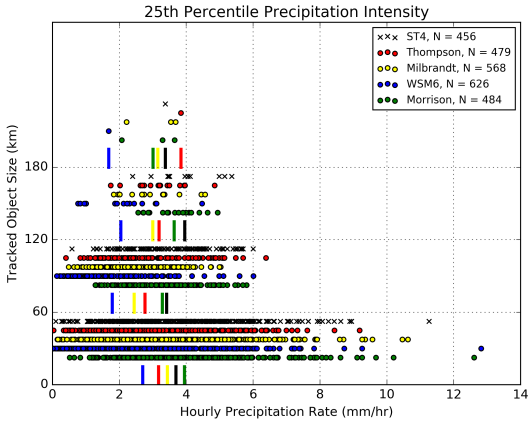

Figure D8: Tracked object 25th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 25th percentile precipitation intensities of all the tracked objects within the corresponding size range.

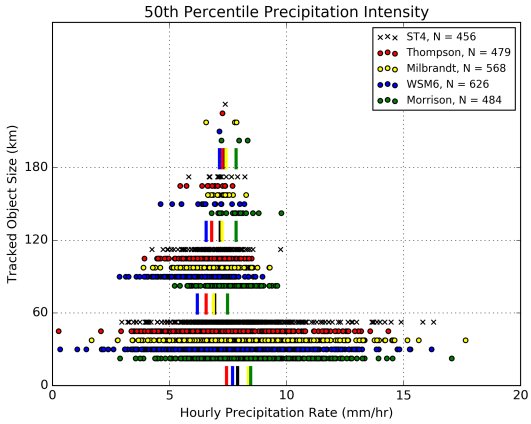

Figure D9: Tracked object 50th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 50th percentile precipitation intensities of all the tracked objects within the corresponding size range.

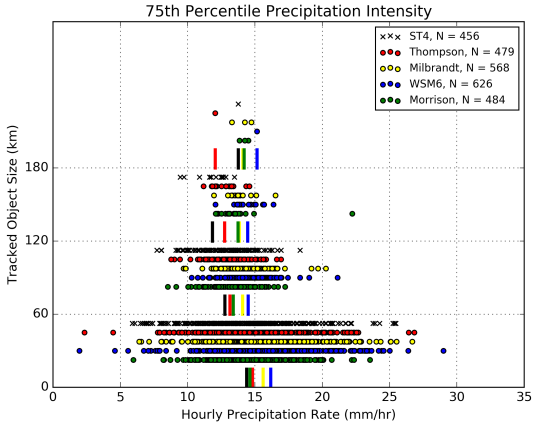

Figure D10: Tracked object 75th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 75th percentile precipitation intensities of all the tracked objects within the corresponding size range

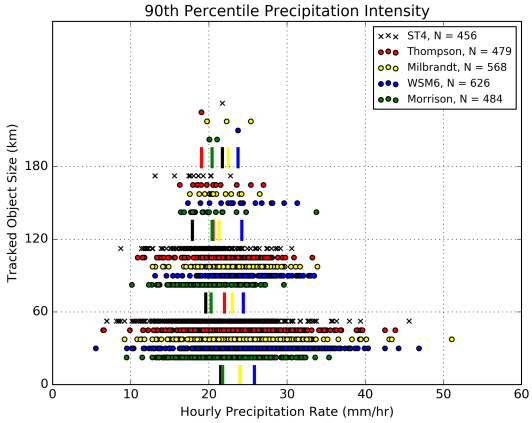

Figure D11: Tracked object 90th percentile precipitation intensities sorted by object average size (<60km, 60-120-km, 120-180km, and > 180km) for all tracked objects. Vertical lines are the averaged 90th percentile precipitation intensities of all the tracked objects within the corresponding size range

Appendix E
Detected Object Count Per Region

To determine if there were any biases in number of detected objects specific to regions of the analysis domain, the analysis domain was separated into 4 °N by 5 °W. Tables E1 and Table E2 show the number of detected objects and percentage increase relative to Stage IV, respectively. Notice in the 50-46 °N by 85-80 °W and 80-75 °W regions there was a significant overprediction of detected objects. This was due to the lack of Stage IV coverage (north of the Great Lakes). At about 1 % of all detected objects they had marginal affect on the main analysis. Also, note the underprediction of detected objects in the 38-34 °N and 34-30 °N by 105-100 °W regions (eastern Colorado and west Texas).

Table E1: Detected object number for Stage IV (St4), Thompson (Tho), Milbrandt (Mil), Morrison (Mor), and WSM6 (Wsm) across all case dates and forecast hours. The detected object count is separated by regions (4 °N by 5°W) covering the entire analysis domain.

	110-105 °W	105-100 °W	100-95 °W	95-90 °W	90-85 °W	85-80 °W	80-75 °W
50-46 °N	**Tho:** 46 **Mil:** 40 **Mor:** 42 **Wsm**: 73 **St4:** 34	**Tho:** 91 **Mil:** 111 **Mor:** 117 **Wsm**: 135 **St4:** 68	**Tho:** 172 **Mil:** 145 **Mor:** 147 **Wsm**: 156 **St4:** 94	**Tho:** 112 **Mil:** 129 **Mor:** 102 **Wsm**: 114 **St4:** 80	**Tho:** 38 **Mil:** 44 **Mor:** 28 **Wsm**: 68 **St4:** 34	**Tho:** 46 **Mil:** 48 **Mor:** 55 **Wsm**: 47 **St4:** 3	**Tho:** 46 **Mil:** 64 **Mor:** 67 **Wsm**: 83 **St4:** 0
46-42 °N	**Tho:** 22 **Mil:** 32 **Mor:** 32 **Wsm**: 47 **St4:** 19	**Tho:** 90 **Mil:** 148 **Mor:** 91 **Wsm**: 129 **St4:** 82	**Tho:** 250 **Mil:** 247 **Mor:** 236 **Wsm**: 271 **St4:** 176	**Tho:** 232 **Mil:** 297 **Mor:** 283 **Wsm**: 307 **St4:** 225	**Tho:** 195 **Mil:** 205 **Mor:** 177 **Wsm**: 224 **St4:** 156	**Tho:** 134 **Mil:** 108 **Mor:** 118 **Wsm**: 114 **St4:** 81	**Tho:** 109 **Mil:** 119 **Mor:** 94 **Wsm**: 124 **St4:** 49
42-38 °N	**Tho:** 2 **Mil:** 2 **Mor:** 0 **Wsm**: 7 **St4:** 5	**Tho:** 102 **Mil:** 116 **Mor:** 85 **Wsm**: 141 **St4:** 93	**Tho:** 323 **Mil:** 328 **Mor:** 274 **Wsm**: 336 **St4:** 293	**Tho:** 403 **Mil:** 466 **Mor:** 376 **Wsm**: 451 **St4:** 342	**Tho:** 338 **Mil:** 319 **Mor:** 315 **Wsm**: 287 **St4:** 277	**Tho:** 223 **Mil:** 217 **Mor:** 185 **Wsm**: 220 **St4:** 214	**Tho:** 191 **Mil:** 205 **Mor:** 166 **Wsm**: 210 **St4:** 143
38-34 °N	**Tho:** 8 **Mil:** 5 **Mor:** 6 **Wsm**: 13 **St4:** 3	**Tho:** 147 **Mil:** 166 **Mor:** 123 **Wsm**: 151 **St4:** 176	**Tho:** 264 **Mil:** 324 **Mor:** 270 **Wsm**: 382 **St4:** 303	**Tho:** 390 **Mil:** 479 **Mor:** 417 **Wsm**: 474 **St4:** 272	**Tho:** 300 **Mil:** 302 **Mor:** 286 **Wsm**: 379 **St4:** 286	**Tho:** 174 **Mil:** 226 **Mor:** 179 **Wsm**: 243 **St4:** 209	**Tho:** 315 **Mil:** 382 **Mor:** 329 **Wsm**: 363 **St4:** 353
34-30 °N	**Tho:** 26 **Mil:** 34 **Mor:** 37 **Wsm**: 40 **St4:** 35	**Tho:** 146 **Mil:** 201 **Mor:** 183 **Wsm**: 176 **St4:** 187	**Tho:** 293 **Mil:** 386 **Mor:** 332 **Wsm**: 366 **St4:** 253	**Tho:** 369 **Mil:** 454 **Mor:** 392 **Wsm**: 473 **St4:** 295	**Tho:** 396 **Mil:** 538 **Mor:** 428 **Wsm**: 529 **St4:** 339	**Tho:** 277 **Mil:** 359 **Mor:** 308 **Wsm**: 408 **St4:** 248	**Tho:** 291 **Mil:** 306 **Mor:** 278 **Wsm**: 340 **St4:** 163

Table E2: Detected object percent increase relative to Stage IV for Thompson (Tho), Milbrandt (Mil), Morrison (Mor), and WSM (Wsm) across all case dates and forecast times. As in Table E1, the statistics are separated into regions (4 °N by 5 °W) covering the entire analysis domain.

	110-105 °W	105-100 °W	100-95 °W	95-90 °W	90-85 °W	85-80 °W	80-75 °W
50-46 °N	**Tho:** 35.3 **Mil:** 17.6 **Mor:** 23.5 **Wsm:** 114.7	**Tho:** 33. 8 **Mil:** 63.2 **Mor:** 72.1 **Wsm:** 98.5	**Tho:** 83.0 **Mil:** 54.3 **Mor:** 56.4 **Wsm:** 66.0	**Tho:** 40.0 **Mil:** 61.3 **Mor:** 27.5 **Wsm:** 42.5	**Tho:** 11.8 **Mil:** 29.4 **Mor:** -17.6 **Wsm:** 100.0	**Tho:** 1433.3 **Mil:** 1500.0 **Mor:** 1733.3 **Wsm:** 1466.7	**Tho:** N/A **Mil:** N/A **Mor:** N/A **Wsm:** N/A
46-42 °N	**Tho:** 15.8 **Mil:** 68.4 **Mor:** 68.4 **Wsm:** 147.4	**Tho:** 9.8 **Mil:** 80.5 **Mor:** 11.0 **Wsm:** 57.3	**Tho:** 42.0 **Mil:** 40.3 **Mor:** 34.1 **Wsm:** 54.0	**Tho:** 3.1 **Mil:** 32.0 **Mor:** 25.8 **Wsm:** 36.4	**Tho:** 25.0 **Mil:** 31.4 **Mor:** 13.5 **Wsm:** 43.6	**Tho:** 65.4 **Mil:** 33.3 **Mor:** 45.7 **Wsm:** 40.7	**Tho:** 122.4 **Mil:** 142.9 **Mor:** 91.8 **Wsm:** 153.1
42-38 °N	**Tho:** -60.0 **Mil:** -60.0 **Mor:** -100.0 **Wsm:** 40.0	**Tho:** 9.7 **Mil:** 24.7 **Mor:** -8.6 **Wsm:** 51.6	**Tho:** 10.2 **Mil:** 11.9 **Mor:** -6.5 **Wsm:** 14.7	**Tho:** 17.8 **Mil:** 36.3 **Mor:** 9.9 **Wsm:** 31.9	**Tho:** 22.0 **Mil:** 15.2 **Mor:** 13.7 **Wsm:** 39.7	**Tho:** 4.7 **Mil:** 1.9 **Mor:** -13.1 **Wsm:** 3.3	**Tho:** 33.6 **Mil:** 43.4 **Mor:** 16.1 **Wsm:** 46.9
38-34 °N	**Tho:** 166.7 **Mil:** 66.7 **Mor:** 100.0 **Wsm:** 333.3	**Tho:** -16.5 **Mil:** -5.7 **Mor:** -30.1 **Wsm:** -14.2	**Tho:** -12.9 **Mil:** 6.9 **Mor:** -10.9 **Wsm:** 26.1	**Tho:** 43.4 **Mil:** 76.1 **Mor:** 53.5 **Wsm:** 74.3	**Tho:** 4.9 **Mil:** 5.6 **Mor:** 0.0 **Wsm:** 32.5	**Tho:** -16.7 **Mil:** 8.1 **Mor:** -14.4 **Wsm:** 16.3	**Tho:** -10.8 **Mil:** 8.2 **Mor:** -6.8 **Wsm:** 2.8
34-30 °N	**Tho:** -25.7 **Mil:** -2.9 **Mor:** 5.7 **Wsm:** 14.3	**Tho:** -21.9 **Mil:** 7.5 **Mor:** -2.1 **Wsm:** -5.9	**Tho:** 15.8 **Mil:** 52.6 **Mor:** 31.2 **Wsm:** 44.7	**Tho:** 25.1 **Mil:** 53.9 **Mor:** 32.9 **Wsm:** 60.3	**Tho:** 16.8 **Mil:** 55.8 **Mor:** 26.3 **Wsm:** 56.0	**Tho:** 11.7 **Mil:** 44.8 **Mor:** 24.2 **Wsm:** 64.5	**Tho:** 78.5 **Mil:** 87.7 **Mor:** 70.6 **Wsm:** 108.6

Appendix F
Duration Plot using Log Y-Scale

To help determine any trends for tracked object duration at longer durations where there were fewer objects, Figure 9 was re-plotted using a log scale for the y-axis as shown in Figure F1. At durations greater than 17 h there was no noticeable trend and appeared random.

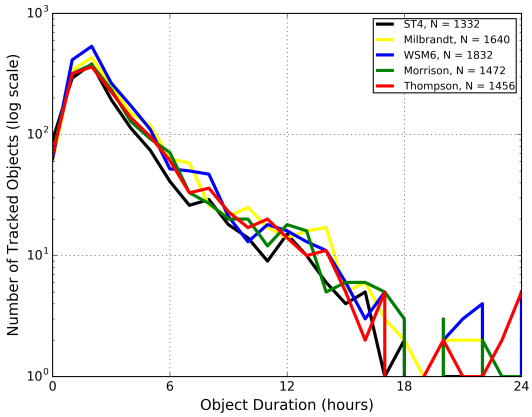

Figure F1: Number of tracked objects as a function of object duration. Note that the y-axis is a log scale. N represents the total number of tracked objects summed over all the case dates

Appendix G
Separation of Retrospective and Real-Time Case Dates

The MODE-TD attributes of number of detected objects and size of detected objects were analyzed where retrospective and real-time case dates were separated. The real-time simulations (Figure G4) did contain a peak in average size the first few hours, although the simulation peaks were earlier and smaller than observed. The retrospective simulations (Figure G3) did not have any indication of a peak in size during the first few forecast hours, even though the observations had a peak.

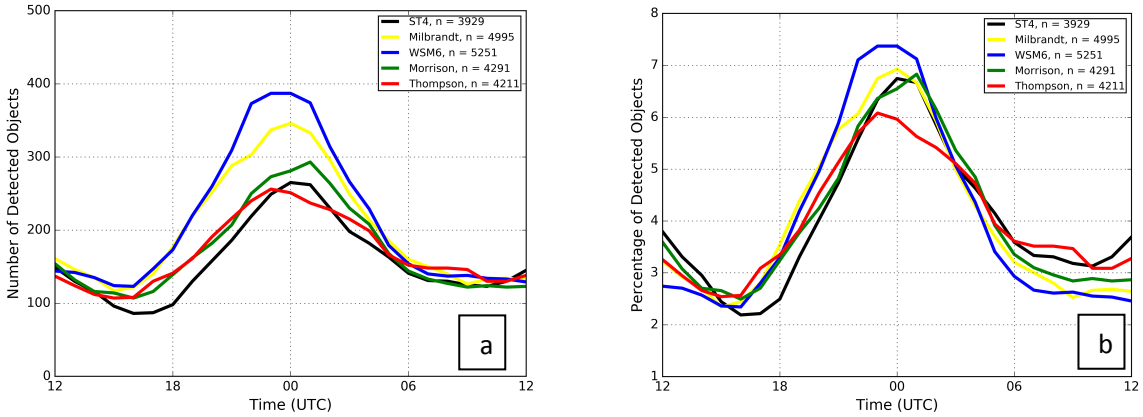

*Figure G1: (a) Number of detected (b) percent of total detected objects ((# of objects at time / n)*100) for the microphysics and Stage IV as a function of UTC time. The n refers to the total number of detected objects accumulated over all forecast hours and retrospective case dates.*

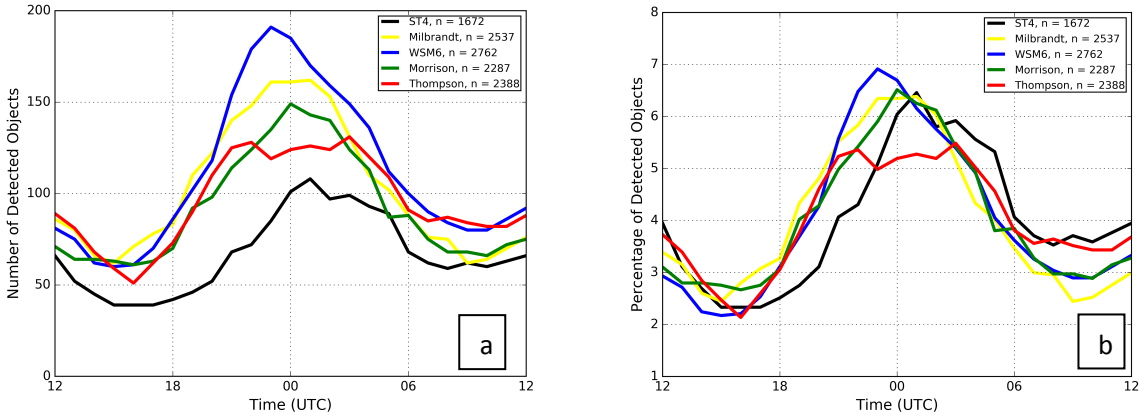

*Figure G2: (a) Number of detected (b) percent of total detected objects ((# of objects at time / n)*100) with time for the ensemble forecasts and Stage IV as a function of UTC time. The n refers to the total number of detected objects accumulated over all forecast hours and real-time case dates.*

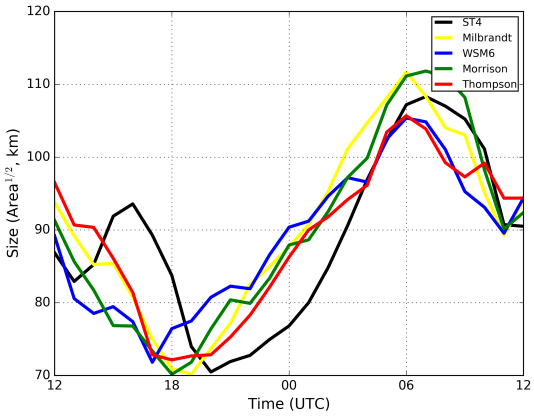

Figure G3: Average size (square root of area, expressed in km) of detected objects (only retrospective case dates) as a function of UTC time.

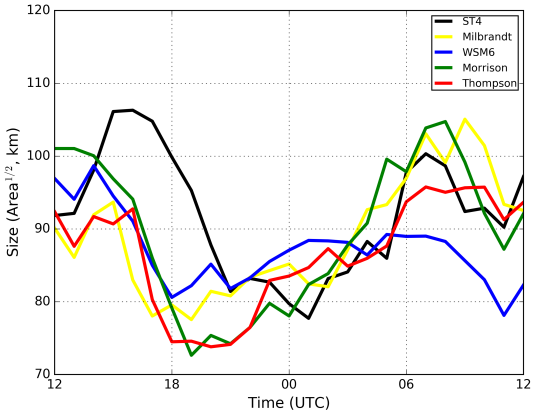

Figure G4: Average size (square root of area, expressed in km) of detected objects (only real-time case dates) as a function of UTC time.