January 2018

An Investigation Of Severe Weather Environments In Atmospheric Reanalyses

Austin Taylor King

AN INVESTIGATION OF SEVERE WEATHER ENVIRONMENTS IN ATMOSPHERIC REANALYSES

by

Austin Taylor King
Bachelor of Science, University of Oklahoma, 2016

A 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

May
2018

This thesis submitted by Austin Taylor King 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. Aaron Kennedy

Dr. Mark Askelson

Dr. Matthew Gilmore

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

Grant McGimpsey
Dean of the School of Graduate Studies

May 4, 2018
Date

PERMISSION

Title	An Investigation of Severe Weather Environments in Atmospheric Reanalyses
Department	Atmospheric Sciences
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 this 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 who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this 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.

Austin Taylor King
May 4, 2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to first thank my advisor Dr. Aaron Kennedy for his constant support and assistance while guiding me through the work presented in this thesis. Despite my near continual harassment, his door was always open and he was always willing to answer any questions I had. I could not have asked for a better advisor and friend during my time at UND. Secondly, I want to thank my loving wife, Kaley, for constantly supporting me and my goals. She selflessly transferred to UND despite being in the middle of her undergraduate career in Oklahoma. Not to mention she had to endure the winters here, which is worthy of my undying gratitude by itself. Lastly, I would like to thank all my friends at UND for their support and assistance in this process as well as the other members of my Master's committee, Drs. Mark Askelson and Matt Gilmore.

ABSTRACT

This thesis performs an intercomparison of reanalysis datasets with the goal of determining their respective proficiency in representing severe weather environments capable of producing phenomena such as strong wind, large hail, and tornadoes. A select reanalyses is then used to investigate the climatology and trends in pertinent severe weather parameters over a three-decade period from 1986-2015.

The intercomparison is performed by comparing a peer-reviewed dataset of Rapid Update Cycle 2 (RUC-2) proximity soundings to collocated soundings derived from the nearest grid point in six different modern reanalyses. These soundings are compared via various parameters related to severe weather such as: Convective available potential energy (CAPE), effective storm relative helicity (EFFSRH), and supercell composite parameter (SCP). Parameters are calculated using SHARPpy, which is an open source, peer reviewed python sounding analysis package modeled after the Storm Prediction Center's (SPC) Sounding and Hodograph Analysis and Research Program (SHARP).

Representation of severe weather environments varies across the reanalyses and the presented results have ramifications for climatological studies that use these datasets. In particular, thermodynamic parameters such as CAPE show the widest range in variations, and this property feeds back to other parameters that incorporate thermodynamic information directly or indirectly through the effective layer. As a result, better segregation of soundings by storm type is found for fixed-layer shear parameters.

Although no reanalysis can exactly reproduce the results of earlier RUC-2 studies, many of the reanalyses can broadly distinguish between environments that are significantly tornadic vs. nontornadic. Overall, the reanalyses found to have the most favorable error characteristics for severe weather environments are the North American Regional Reanalysis (NARR) and the Japanese 55-Year Reanalysis (JRA55).

Given the results of the first objective, NARR is used to understand the climatology and trends in severe weather parameters across the contiguous United States. A suite of severe weather parameters is calculated for the full domain of NARR by taking "pseudo-soundings" at each grid point. It is found that the spatial distribution of average severe weather climatologies are similar to prior studies but tend to have significantly larger magnitudes. It is also found that certain severe weather parameters may be increasing over select regions, while others have either a neutral trend or are decreasing over time.

The raw data used for this study, i.e. a suite of severe weather parameters for the full domain of NARR, will be made publicly available. This dataset is potentially useful to members of the climate science and atmospheric science communities. This is due in part to the large amount of computational resources and time that were required to produce this dataset.

CHAPTER 1

INTRODUCTION

Background

Severe Weather in Atmospheric Reanalyses

Atmospheric reanalyses combine historical observations with a fixed data assimilation scheme and underlying model to achieve a dynamically consistent, gridded representation of the atmosphere. This process fills data void regions, and, for this reason, reanalyses are commonly used to investigate various questions related to past weather events and climate. For example, reanalyses have shed light on the climatology and trends of severe weather environments given the numerous issues with observed reports associated with these events (e.g. Brooks et al. 2003, Brooks et al. 2007, Gensini and Ashley 2011, Romero et al. 2007, Blamey et al. 2016). Based off this climatological work, reanalyses have also been used to assess the fidelity of these environments in climate simulations (Marsh et al. 2007, Trapp et al. 2007). In summary, reanalyses are a key piece of the puzzle to determine how severe weather events may or may not change in a warming climate.

Since reanalyses are run at resolutions inadequate to resolve the processes that dictate severe weather events, alternative methodologies must be employed to relate reanalysis output to the potential for severe weather. One such way is to identify favorable environments by calculating diagnostic parameters such as Convective Available Potential Energy (CAPE) and 0-6km Bulk Wind Difference (BWD) from reanalyzed vertical profiles of the atmosphere. Parameter based evaluation of severe weather environments is rooted

in earlier forecasting studies, which utilized observed proximity soundings to discriminate between types of convection or the phenomenon produced (e.g. Davies 1993, Johns et al. 1993, Rasmussen and Blanchard 1998). Ideally, these soundings should be taken within the inflow environment of the storm, but this can significantly limit the number of samples for analysis. Further, there is risk of limiting the representativeness of the sounding due to interference from the storm itself (so-called convective contamination). Despite these issues, the viability of proximity soundings has been well documented in prior literature.

In lieu of observed soundings, the rise of mesoscale models has led to analyzed or forecast soundings being used as pseudo proximity soundings for convective events (Thompson et al. 2003, Davies 2004, Thompson et al. 2007, Reames 2017). Thompson et al. (2003) (hereafter T03) found that Rapid Update Cycle-2 (RUC-2, Benjamin et al. 2004) soundings were a good proxy for observed soundings and contained errors that were generally "…within 0.5° C for temperatures, 0.2 g kg^{-1} for mixing ratios, and 1 m s^{-1} for wind speed (all close to the ranges for radiosonde accuracy)". Additionally, T03 concluded that "Overall, the RUC-2 analysis soundings appear to be a reasonable proxy for observed soundings in supercell environments."

Given this information, it seems plausible that reanalyses can provide pseudo proximity soundings (hereafter proximity soundings) for convective events, especially for regions with ample observations. This concept was used by Brooks et al. (2003), who analyzed environments in the NCAR/NCEP (Kalney et al. 1996) reanalysis. Despite being coarse compared to the latest generation of reanalyses (~200km grid spacing with 6-hourly increments), this dataset allowed for proximity soundings to be taken within previously used definitions such as within three hours and 180 km of an event (T03). This was built

upon by Brooks et al. (2007) who examined the annual cycle of severe weather environments from the NCAR/NCEP reanalysis. Marsh et al. (2009) extended this framework to simulated soundings from the Community Climate System Model 3 (CCSM3) to investigate European severe weather environments in past and future climates.

Since these earlier studies, a number of reanalyses now exist, which leads to the question of whether severe weather environments are similar amongst these datasets. Overall, the capability of these reanalyses to represent severe weather environments is somewhat unclear. Kennedy et al. (2011) demonstrated that large differences in the vertical profiles of variables can exist between reanalyses, specifically the North American Regional Reanalysis (NARR, Mesinger et al. 2006) and the Modern-Era Retrospective Reanalysis for Research and Applications (MERRA, Rienecker et al. 2011). This would suggest that derived parameters, which are useful for investigating convective events, could differ significantly between reanalyses. This clouds the utility of these datasets for applied forecasting and climatological studies.

Although some work has been done to clarify this ambiguity in the representation of severe weather environments, most studies have focused on the NARR. Gensini et al. (2014) noted that kinematic variables were best represented by this reanalysis, while thermodynamic variables were generally overestimated and had regional biases. This study also demonstrated that sharp gradients in temperature associated with the elevated mixed layer (EML) were not adequately resolved by NARR in most soundings. In a study of high CAPE, low shear severe weather events, Sherburn et al. (2016) found that surface based CAPE (SBCAPE) calculated from the NARR did not strongly correlate with SBCAPE

values from the Storm Prediction Center's (SPC) Surface Objective Analysis (SFCOA). Additionally, the NARR overestimated SBCAPE in comparison to SFCOA.

Although the error characteristics of the NARR regarding severe weather parameters have been relatively well documented, a comprehensive study of all current generation reanalyses has never been completed. Many of these other reanalyses have been used in past studies to identify regional properties or trends in severe weather parameters. This has been done without knowledge of the capability of the reanalysis to represent severe weather environments, at least in comparison to other reanalysis datasets. For example, Blamey et al. (2016) used the Climate Forecast Systems Reanalysis (CFSR) to examine trends in severe weather over South Africa while Allen and Karoly (2013) performed a similar analysis with the European Interim Reanalysis (ERA-Interim, Dee et al. 2011) to determine the inter-annual variability and influence of El Niño-Southern Oscillation (ENSO) on severe thunderstorms in Australia. ERA-Interim has also been used to investigate the relationship of rainfall extremes to atmospheric properties (Lepore et al. 2014). It is possible that the use of a different reanalysis could have produced different results.

Trends in Severe Weather Environments

Some studies have suggested that occurrence of environments favorable for severe weather may be increasing due to a changing/warming climate (Marsh et al. 2007, Gensini and Brooks 2008, Hoogewind et al. 2017, etc.). However, there are numerous difficulties associated with accurately determining trends or climatologies regarding severe weather (e.g. Brooks et al. 2003, Brooks et al. 2007, Gensini and Ashley 2011, Romero et al. 2007,

Blamey et al. 2016). One of the primary issues is that severe weather phenomena occur on relatively small spatial and temporal scales. For example, tornadoes are typically on the order of 100s of meters wide and typically last for only a few minutes. This makes it impossible to resolve individual convective storms within the current generation of reanalyses or climate models, which typically have horizontal grid spacing on the order of tens to hundreds of kilometers.

One solution to this problem is the process of dynamical downscaling in which temporally and spatially coarse data from a Global Climate Model (GCM) or reanalysis is used as the boundary conditions for high-resolution regional model. For example, Hoogewind et al. (2017) used the Weather Research and Forecasting (WRF) model to perform dynamical downscaling of the Geophysical Fluid Dynamics Laboratory Climate Model, version 3 (GFDL CM3) for historical (1971–2000) and future (2071–2100) periods. They found that the proxy for severe weather occurrence (vertical velocity > 22 m s^{-1}) becomes more frequent by the end of the twenty-first century, primarily in the spring and summer months. However, dynamical downscaling is not without its drawbacks. It requires immense computational power to complete the simulations which makes it unfeasible for large areas or long time periods. There are also numerous other issues, including differences in physics packages between models, which are discussed in Hong and Kanamitsu (2014).

Tang et al. (2017) also utilized dynamical downscaling to study changes in the Great Plains low-level jet (GPLLJ). The GPLLJ is a wind speed maximum found within the lowest 2-3 km of the troposphere that typically develops in the evening hours of the Great Plains spring and summer season. This feature is important for convective storms

due to its ability to rapidly transport moisture from the Gulf of Mexico as well as to greatly increase the low level wind shear. This increase in wind shear is associated with an increase in tornado frequency during the evening hours. Tang et al. (2017) projected an increase in nocturnal GPLLJ frequency in the spring and summer, and persistence of current climatological patterns for daytime and cool season GPLLJs.

From a historical standpoint, the relative scarcity of severe weather also makes obtaining an entirely accurate record of events nearly impossible given past and current reporting methods. Many of these issues are examined in Brooks and Doswell (2001), Doswell et al. (2005), and Verbout et al. (2006). Since raw reports alone are not an adequate source of information, another route must be taken. As was previously discussed, a common solution to these issues is the use of diagnostic parameters such as CAPE, wind shear, and various combinations thereof (T03, T07, etc.). These parameters provide a view of the overall severe weather environment within which storms may exist within and serve as a proxy for severe weather occurrence. The relatively sparse network of upper air observations that typically only record data twice daily limits the usefulness of a purely observation-based approach. However, the relatively recent development of atmospheric reanalyses has allowed for the creation of a much larger dataset of severe weather parameters using "pseudo-soundings". This method was employed by Brooks et al. (2003) using the NCEP /NCAR Reanalysis (R1, Kalnay et al., 1996) to develop a severe weather climatology over North America. A similar approach was used by Gensini and Ashley (2011) using the newer North American Regional Reanalysis (NARR, Mesinger et al. 2006), which boasts a much higher spatial and temporal resolution (~32 km, 3 hrly respectively). It is important to note that the use of diagnostic parameters to evaluate severe

weather does not consider the actual occurrence of severe weather phenomena, but rather the *potential* for severe weather occurrence.

Purpose of Study

The primary goal of this thesis is to identify the characteristics of severe weather environments in the latest generation of atmospheric reanalyses. To do so, these datasets are compared to a database of RUC-2 proximity soundings from Thompson et al. (2007, hereafter T07). Aside from identifying characteristics of common severe weather parameters, this effort will explore whether reanalyses can reproduce earlier work from T03 and T07 that has identified differences in thermodynamic and kinematic environments by storm type.

This study offers several benefits. First, it will provide proper context to earlier literature that has focused on specific reanalyses. Second, the performance of these reanalyses will provide insight into which datasets should be used for future studies to examine the characteristics of severe weather phenomena. If a particular reanalysis can reproduce the results of T03 and T07, this may open the door to building a database of proximity soundings for rarer events such as long-track tornadoes that have a notoriously low sample size. Finally, these results will aid climatological studies, such as inter-comparisons between reanalyses and historical and future climate simulations.

The second goal of this thesis is to produce a dataset of severe weather parameters for the full domain of NARR at each grid point. NARR was chosen due to favorable error characteristics (King and Kennedy 2018) compared to other modern reanalyses, in addition to its relatively high spatial and temporal resolution (~32 km and 3 hourly, respectively).

Parameters are calculated using the Sounding and Hodograph Analysis and Research Program in Python (SHARPpy) package (Blumberg et al. 2017) at every grid point within NARR, with grid points over water and grid points with 2 m temperatures below freezing excluded to speed up the computational process. This dataset of severe weather parameters will be made publicly available. As this dataset was computationally expensive to produce, it will prove potentially useful to members of the climate science and atmospheric science communities needing a record of severe weather parameters for North America from 1979 through 2016.

To demonstrate the potential uses of this dataset, a climatology of severe weather parameters within NARR for a 30-year period from 1986-2015 is produced. Areas in the United States east of the Rocky Mountains are the primary focus. This serves to ensure that the average distribution of parameters within the dataset is reasonable as compared to prior studies that have conducted similar analyses. Finally, trends in various parameters are calculated at each grid point to determine any changes in environments favorable for severe weather over the time period of the dataset.

CHAPTER 2

REANALYSIS DATASETS

Overview

The reanalyses chosen for the intercomparison include the North American Regional Reanalysis (NARR), ECMWF interim Reanalysis (ERA-Interim), 2nd Modern-Era Retrospective Analysis for Research and Applications (MERRA2, Gellaro et al. 2017), Japanese 55-year Reanalysis (JRA55, Kobayashi et al. 2015), 20th Century Reanalysis (20CR, Compo et al. 2011), and the Climate Forecast System Reanalysis (CFSR, Saha et al. 2010). These datasets represent the latest generation of reanalyses, and as such, older datasets, such as NCAR/NCEP global reanalysis and first generation of MERRA, were omitted. It should be noted that the 5th generation of ERA (ERA5) is in production, but data were unavailable for analysis at the time of this thesis.

A unique aspect of reanalyses is the wide number of products generated. Due to space constraints and user purposes, reanalyses are often subsetted in space and/or time. A good example is NARR, which provides 29 isobaric levels vs. the original 45 model layers. For other reanalyses, multiple datastreams may exist. MERRA 2 includes datastreams that have atmospheric variables output in different time intervals, grid spacing, and on different vertical levels (model level vs. isobaric). Ideally, proxy soundings should have high temporal, spatial, and vertical resolutions to be most representative of the convective environment. To this end, datastreams were selected that maximized these properties for three-dimensional atmospheric variables. Grid spacing information along with datastream locations are also provided in Table 1. One important thing to note is the difference in

vertical resolution between reanalysis datasets. The difference in vertical resolution, especially at lower levels of the atmosphere, would theoretically have a large impact on parameters such as Convective Inhibition (CIN) that rely on having the lowest few hundred meters of the atmosphere adequately resolved. A synopsis of each reanalysis dataset is provided in the following section. For a more in-depth description of each dataset, the reader is referred to the primary publications that describe these datasets (Table 1).

Reanalysis Dataset Details

NARR

NARR is a regional reanalysis that was developed with the goal of improving upon the NCAR/NCEP global reanalysis (R1) by more accurately capturing the regional hydrological cycle, diurnal cycle, and other important features. This dataset was completed in 2004 following a 6-year development phase. NARR utilizes the NCEP Eta model and its Data Assimilation System (32-km, 45-layer resolution with 3-hourly output). It incorporates hourly assimilation of precipitation, a recent version of the Noah land surface model, as well as numerous other datasets that are additional or improved compared to R1. The currently available data for this reanalysis spans from 1979-2016 and is updated on a regular basis. Some of the reported highlights of the NARR dataset is a significantly more accurate analysis of precipitation, as well as 2 m temperatures and 10 m winds as compared to R1 (Kalney et al. 1996).

ERA-Interim

ERA-Interim is produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) partly in preparation for the ERA5 reanalysis that will replace the older ERA-40 reanalysis. The ERA-Interim covers the time period of 1979 onwards, and is updated in near real-time. The reanalysis utilizes a 12-hourly 4D variational analysis scheme (4D-Var) data assimilation scheme for upper-air atmospheric state. The 4D-Var analysis in ERA-Interim is obtained by successive linearizations of the model and observation operator (Dee et al. 2011, Courtier et al. 1994, Veersé and Thépaut, 1998). This is a vast improvement from the earlier ERA-40 that utilized a 3D variational analysis scheme (3D-Var; Courtier et al., 1998) scheme. The analysis itself is produced using the ECMWF IFS that incorporates a forecast model with coupled atmosphere, land surface, and ocean waves. Additionally, the observations being assimilated into the reanalysis increase from the start of the data availability to the finish, largely due to increasing satellite data (Dee et al. 2011).

MERRA2

The MERRA2, which is produced by NASA's Global Modeling and Assimilation Office (GMAO), represents an upgrade to MERRA. It includes assimilation of observations that were not available to its predecessor, such as aerosols and additional satellite data. MERRA2 also includes updates to the Goddard Earth Observing System (GEOS) model and analysis scheme. It is designed to be a stepping stone between the original MERRA and the next generation of integrated Earth system analysis (IESA) currently under development at GMAO. Additional improvements from MERRA include

a reduction of data gaps/jumps and reduced biases and imbalances in the water cycle (Gellaro et al. 2017).

JRA55

The JRA55 reanalysis dataset is produced by the Japan Meteorological Agency (JMA) and is the second global reanalysis produced by the agency. The reanalysis covers the time period of 1958 to present, which coincides with regular global radiosonde observations. The primary purpose behind the development of this dataset was to improve on the prior iteration, the Japanese 25-year Reanalysis (JRA25). As with other reanalyses, this second-generation reanalysis provides marked improvements relative to its predecessor. Most notably, the JRA55 includes newly available and improved observations as well as an updated data assimilation system (TL319 version of JMA's operational data assimilation system, Kobayashi et al. 2015). JRA55 also remedied a cold bias in the lower stratosphere that was present in JRA25 along with incorporating numerous other corrections.

20CR

The 20CR is an internationally developed dataset with the goal of producing a global atmospheric circulation dataset for the entirety of the 20th Century. The 20CR only assimilates surface pressure, as well as monthly sea surface temperatures and sea ice distributions. The primary purpose of this reanalysis is use in climate model validation. An Ensemble Kalman Filter data assimilation system (Whitaker and Hamill, 2002) is utilized in 20CR, as is a new version of the NCEP atmosphere-land model. The atmosphere-land

model is used to generate first-guess fields with interpolated sea-surface temperature and sea-ice concentrations being used as boundary conditions. This, in addition to assimilation of surface pressure leads to the generation of the reanalysis dataset.

CFSR

The CFSR is a high resolution, global reanalysis produced by NCEP-NCAR as an update to prior generation global reanalysis (R2). The CFSR dataset spans from 1979 through 2010 and utilizes a coupled atmosphere–ocean–land surface–sea ice system. The CFSR has a horizontal grid spacing of approximately 38km (T382) with 64 levels from the surface to 0.26 hPa. The reanalysis assimilates most available observations, including satellite data. According to Saha et al. (2010): "Satellite observations were used in radiance form, rather than retrieved values, and were bias corrected with 'spin up' runs at full resolution, taking into account variable CO2 concentrations. This procedure enabled the smooth transitions of the climate record resulting from evolutionary changes in the satellite".

Table 1. Properties of reanalysis datastreams used in this study. Primary references for these reanalyses are also provided.

Reanalysis	**Horizontal Grid Spacing**	**Vertical Levels**	**Temporal Spacing**	**Source**	**Reference**
NARR	~32 km	29	3-hourly	ESRL/PSD[1]	Mesinger et al. (2006)
ERA-Interim	~80 km (T255)	60	6-hourly	NCAR RDA[2]: ds627.0	Dee et al. (2011)
MERRA 2	~ 50 km (0.5°)	72	3-hourly	NASA GES DISC[3] inst6_3d_ana_Nv	Gelaro et al. (2011)
20CR (V2)	~200 km (T62)	28	6-hourly	ESRL/PSD[1]	Compo et al. (2011)
JRA55	~55 km (T319)	60	6-hourly	NCAR RDA[2]: ds628.0	Kobayashi et al. (2015)
CFSR	~38 km (T382)	64	6-hourly	NCAR RDA[2]: ds094.0	Saha et al. (2010)

[1] https://www.esrl.noaa.gov/psd/data/gridded/
[2] NCAR Research Data Archive: https://rda.ucar.edu/
[3] NASA Goddard Earth Sciences Data and Information Services Center: https://disc.sci.gsfc.nasa.gov/

CHAPTER 3

METHODS

Reanalysis Comparison

Reanalyses were compared to the RUC-2 proximity soundings from T07. This dataset contains 1185 proximity soundings that are separated into the following categories:

- "significantly tornadic" (F2-F5 tornadoes, denoted as sigtor),
- "weakly tornadic" (F0-F1 tornadoes, denoted as weaktor),
- "nontornadic" (discrete non-tornadic supercells, denoted as nontor),
- "marginal" (marginal supercellular structure, denoted as mrgl),
- "non-supercell" (discrete, non-supercells, denoted as nonsuper).

The nontornadic category was further split into surface-based and elevated storms (denoted as elevnt). This was performed by evaluating the effective inflow layer (T07). If the effective inflow layer was above ground level, the storm was classified as elevated.

RUC-2 model soundings were chosen in lieu of radiosonde data due to the temporal and spatial availability of the data, the availability of the T07 dataset with discrimination of soundings by storm type, and the performance of RUC-2 in representing the true atmospheric state (T03, Coniglio 2012). The RUC-2 soundings from T07 were compared to each reanalysis via numerous severe weather parameters that have been calculated from the proximity soundings. While some of these variables (such as CAPE and CIN) are provided by the reanalyses (and were presumably calculated using all model levels), documentation regarding these calculations is sparse. For example, there is no guarantee that all reanalyses use similar parcels (e.g. surface based) or the virtual temperature

correction (Doswell and Rasmussen 1994). For example, ERA-Interim uses an approximation for CAPE for the purposes of computational efficiency that leads to values ~20% higher than those calculated using virtual temperature (ECMWF, 2017). Because of these issues, severe weather parameters were calculated independently, using the Sounding and Hodograph Analysis and Research Program in Python (SHARPpy) package. This package has been developed to closely emulate methods developed by the SPC, which have been rigorously tested (Blumberg et al. 2017).

The proximity soundings for each reanalysis were selected by identifying the nearest grid-point to the T07 RUC-2 grid-point identifiers. Of the original 1187 samples within T07, twenty-two soundings, not associated with METAR locations, could not be identified and were not used. Two additional soundings did not have parcel traces completed due to erroneous values, and these soundings were also discarded.

Once the closest reanalysis grid-point was determined, the appropriate reanalysis time-step was selected. To understand how temporal variability may impact the analysis, as well as mitigate potential issues with convective contamination or passage of mesoscale features responsible for convective events, times were selected in two different ways:

1) Closest time to each RUC-2 sounding (e.g. 00 UTC for a 23 UTC RUC-2 analysis)
2) Closest time *prior* to each RUC-2 sounding (e.g. 21 UTC for the case above).

Once points were selected from each reanalysis, a vertical profile was extracted and required variables (temperature, dewpoint, u-component of wind, v-component of wind, height, and pressure) were passed to SHARPpy to compute desired parameters. This process was completed twice for each reanalysis: once while including surface variables (2

m temperature, 2 m dewpoint temperature, 10 m u-component of wind, 10 m v-component of wind, surface height, and surface pressure), and once starting from the first model layer. Additionally, for those reanalyses that did not utilize a hybrid-sigma pressure coordinate system, checks were performed to ensure that model layers below ground were not used. A full list of the severe weather parameters calculated using SHARPpy for this study is provided in Table 2. The net result of the process was four different calculations of parameters:

1) Closest spatial and temporal point with surface variables,

2) Closest spatial and prior temporal point with surface variables,

3) Closest spatial and temporal point without surface variables,

4) Closest spatial and prior temporal point without surface variables.

In most cases, the first calculation (closest in space and time, with surface variables) yielded the highest correlations with RUC-2 soundings. Using the prior time step (i.e. methods 2 and 4) yielded the lowest correlations, especially for kinematic variables. In the interest of time, calculations using the prior time step were only completed for NARR and JRA55. Table 3 shows correlation coefficients for each calculation method for these two reanlayses. The inclusion of surface variables typically had little effect on thermodynamic variables (e.g. correlations for NARR MLCAPE varied from 0.74 to 0.75 with the inclusion of surface variables, whereas JRA55 MLCAPE went from 0.68 to 0.69 for the calculations involving the closest temporal point). Correlation decreased slightly for NARR SBCAPE when including surface variables (0.70 to 0.66), which is likely due to the tendency of NARR to overestimate surface temperatures. Low-level shear parameters, parameters that include the effective layer, and composite parameters typically benefited from the inclusion

of surface parameters, especially for NARR. Broadly speaking, these results also hold true for other reanalyses. For example, not including surface variables (but still taking the closest temporal point) resulted in a correlation of 0.56 between ERA-INT and RUC-2 for SBCAPE, whereas including surface variables increased the correlation to 0.60 (Table 4). Overall, relatively small deviations can be noted across most reanalyses and variables (especially for surface-based parameters). As such, these results (from the first calculation) are presented herein.

Provided the objectives, SHARPpy was first used to reproduce the RUC-2 results of T03 and T07. Some differences should be expected between these studies. First, the sample size is slightly different. Second, there are some possible differences in the calculations of sounding parameters between SHARPpy and the Skew-T Hodograph Analysis and Research Program (NSHARP; Hart and Korotky 1991) used in T03 and T07. Although most parameters are nearly identical to NSHARP, those dependent on storm motion (e.g. ESRH) show more variability (Halbert et al. 2015). SHARPpy (by default) uses the ID method (Bunkers et al. 2000) for calculating storm motion vectors whereas T07 used radar identified centroids. Considering reanalyses may be used to study historical events that do not have radar data available or a priori knowledge of storm motion, identified storm motion vectors in SHARPpy were not modified.

Overall, results are similar to T03 and T07 (Fig. 1). Although the sample size increased from T03 and T07, box and whiskers plots of MLCAPE look similar to Fig. 6 in T03 (Fig. 1a). MLCAPE decreases from median values of 2220 to 1070 J kg^{-1} from sigtor to nontor supercells. This compares to the original values of 2152 to 952 J kg^{-1} in T03. Box and whiskers plots for effective SRH subject to the parcel constraints of CAPE $\geq$ 100 J kg^{-}

^{1} and CIN ≥ -250 J kg^{-1} are also shown (Fig. 1b, Fig 8 in T07). Similar to T07, ESRH decreases markedly across the categories with median values ranging from 225 to 19 m^{2}s^{-2}. The largest difference between this study and T07 occurs in the mrgl category. In this case, the storm motion is likely closer to the mean wind vector versus the vector derived from the ID method (Bunkers et al. 2000). As a result, storms with less deviant motion have reduced ESRH.

Climatology

Pseudo-soundings were taken at each grid point in NARR for the full time period of data availability (1979-2016). NARR was chosen for this analysis based upon prior results indicating favorable error characteristics compared to other modern reanalyses (King and Kennedy 2018) as well as a comparatively high spatial resolution (~32km). Additionally, NARR has been used in past research to determine severe weather climatologies (ex. Gensini and Ashley 2011), which makes a more direct comparison to past research possible. Given that 30 years is commonly used to determine a climatology, this thesis focused on the period 1986-2015. Soundings were taken at 00 UTC to allow for direct comparison to prior studies and observations. Additionally, only grid points over land and where temperatures were above freezing were considered. This was done to minimize the computational resources needed, and because parameters favorable for severe weather are unlikely to exist in sub-freezing temperatures. Parameters associated with the pseudo-soundings were computed using SHARPpy. A table of all computed parameters is provided in Table 2.

To enable a more direct comparison to prior studies (i.e. Brooks et al. 2003, hereafter B03; and Gensini and Ashley 2011, hereafter GA11) and ensure that the average distribution of parameters was reasonable, a climatology of severe weather parameters was developed. Similar to these earlier studies the annual mean number of days of parameters exceeding specific threshold values were calculated (e.g. days exceeding 2000 J kg^{-1} of CAPE).

While parameters can be explored individually, CAPE alone has a somewhat limited use in detecting environments that are potentially favorable for severe weather. CAPE can be useful in discriminating between environments of varying levels of severity, however, other parameters, especially those that include information about wind shear, tend to be better predictors (Rasmussen and Blanchard 1998). For example, many summer days are characterized by large CAPE values, but storms do not develop due to either too much CIN or lack of a forcing mechanism. Additionally, an environment may have large CAPE but little to no wind shear, which significantly limits the possibility of severe weather (Rasumussen and Blanchard 1997, T03, T07). Therefore, wind shear information should be included to produce a reasonable climatology of severe weather environments. A simple parameter was developed by B03 that is a combination of 0-6 km wind shear and CAPE. This technique was also used by GA11 in their expansion of the work done by B03 to examine the climatology of potentially severe environments from 1980-2009. The basic formula for this parameter from GA11 is:

$$0 - 6km\ BWD * MUCAPE$$

Values above 10,000 are typical for basic severe weather, whereas values of 20,000 and 30,000 represent environments capable of significant severe and significant tornadoes, respectively (GA11). This parameter is referred to as the composite C proximity parameter in GA11 and the same term will be used herein for the sake of consistency. Although the time periods differed between this thesis, B03, and GA11, the comparison of overall severe weather parameter distributions should still be valid.

Trends in various parameters were calculated by taking the number of days exceeding a predetermined threshold each year and performing a simple linear regression analysis at each grid point. The slope of the linear regression was used to determine whether the parameter in question has been increasing or decreasing over the 30 year period of this study. Statistical significance was determined using the Wald Test with t-distribution of the test statistic from the SciPy stats module in Python. This module outputs a p-value which represents the probability of obtaining the observed result given that the null hypothesis is true. In this case, the null hypothesis (the statement being tested), is that the slope is zero. In short, anywhere the p-value is less than 0.05, the null hypothesis can be rejected and it can be safely assumed that the slope is statistically significant. One of the primary assumptions that must be met for the statistical significance testing to be valid is that the errors (residuals) are normally distributed. This condition was tested using the Shapiro-Wilk test for normality. The test statistic is defined as:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2},$$

where $x_{(i)}$ is the ith order statistic, $\overline{x}$ is the sample mean, and the constants a_i are given by:

$$(a_i, \ldots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}.$$

Grid points where the p-value was less than 0.05 (meaning the null-hypothesis that the errors come from a normal distribution can be rejected) were discarded from the significance tests.

Table 2. List of calculated parameters provided by SHARPpy. Details for these parameters are found in Blumberg et al. (2017). Asterisks denote kinematic parameters that indirectly include thermodynamic information via the effective layer (Thompson et al. 2007). Mean values are taken from the surface to 400 hPa.

Kinematic Parameters	Thermodynamic Parameters	Composite Parameters
0-1km SRH	CAPE (SB, ML, MU)	K index
0-3km SRH	CIN (SB, ML, MU)	Fixed STP
*Effective SRH	LCL (SB, ML, MU)	CIN STP
0-1km bulk shear	LFC (SB, ML, MU)	SCP
*Effective bulk shear	EL (SB, ML, MU)	Total totals
0-1km helicity	LI (SB, ML, MU)	SWEAT index
0-3km helicity	BRN (SB, ML, MU)	
Critical angle	Max. profile temp	
	Mean mixing ratio	
	Mean theta	
	Mean theta e	
	Mean relative humidity	
	Perceptible water	

Table 3. Correlation coefficients for each method of parameter calculation for NARR and JRA55 as compared to RUC-2 for select parameters.

	mlcape	sbcape	shear1km	effsrh	scp
NARR					
Closest Time - With Sfc Vars	0.7542	0.6594	0.7809	0.6575	0.6524
Closest Time - No Sfc Vars	0.7438	0.6982	0.7193	0.5911	0.6086
Prior Time - With Sfc Vars	0.7594	0.6625	0.7689	0.653	0.6454
Prior Time - No Sfc Vars	0.7503	0.6996	0.7256	0.5891	0.5988
JRA55					
Closest Time - With Sfc Vars	0.6894	0.6371	0.7548	0.6484	0.6633
Closest Time - No Sfc Vars	0.6811	0.6167	0.7554	0.6485	0.6679
Prior Time - With Sfc Vars	0.6512	0.6543	0.7391	0.5896	0.624
Prior Time - No Sfc Vars	0.6367	0.5921	0.7269	0.6006	0.6387

Table 4. Correlation coefficients of popular parameters from the reanalyses compared to RUC-2. Cooler (warmer) colors represent stronger (weaker) correlations.

Variables	NARR	ERA	MERRA 2	20CR	JRA55	CFSR
sbcape	0.66	0.60	0.64	0.56	0.64	0.36
sbcin	0.37	0.33	0.21	0.17	0.37	0.17
mlcape	0.75	0.66	0.62	0.53	0.69	0.46
mlcin	0.54	0.39	0.28	0.15	0.39	0.24
mucape	0.70	0.62	0.66	0.57	0.68	0.42
mucin	0.38	0.25	0.33	0.19	0.34	0.23
srh1km	0.78	0.72	0.74	0.58	0.76	0.72
srh3km	0.78	0.73	0.76	0.57	0.80	0.76
effsrh	0.66	0.55	0.63	0.46	0.65	0.60
stpfix	0.68	0.63	0.68	0.37	0.70	0.54
scp	0.65	0.64	0.68	0.36	0.66	0.59
stpcin	0.57	0.56	0.62	0.27	0.61	0.46
hel1km	0.68	0.67	0.68	0.63	0.71	0.65
hel3km	0.79	0.73	0.75	0.68	0.79	0.77
blkshear1km	0.66	0.60	0.64	0.56	0.64	0.36

Table 5. As in Table 4 except for biases (reanalysis – RUC-2). Warmer (cooler) colors represent positive (negative) biases.

Variables	NARR	ERA	MERRA 2	20CR	JRA55	CFSR
sbcape	212.0	-1021.9	-1415.6	-896.8	-40.6	-1148.7
sbcin	34.9	37.1	11.5	34.6	14.8	-10.7
mlcape	-269.6	-1063.3	-1305.3	-955.3	-184.5	-942.7
mlcin	20.8	22.3	9.4	14.7	-11.4	-12.5
mucape	30.7	-1196.8	-1615.6	-1126.3	12.0	-1179.8
mucin	13.2	11.3	-5.6	4.5	0.7	-10.9
srh1km	-22.9	-8.1	21.1	-59.7	-16.9	-25.0
srh3km	-17.6	6.7	38.8	-89.7	-28.9	-19.8
effsrh	-31.4	-59.0	-40.1	-89.0	-35.2	-40.0
stpfix	-0.2	-0.7	-0.8	-0.9	-0.3	-0.8
scp	-1.3	-4.7	-5.0	-5.5	-2.3	-3.3
stpcin	-0.3	-0.7	-0.8	-0.9	-0.4	-0.7
hel1km	-29.9	-6.6	23.3	-23.8	-18.5	-17.0
hel3km	-30.9	-14.4	25.5	-77.0	-43.5	-20.4
blkshear1km	-1.6	1.5	2.8	-5.1	-0.00	-1.0

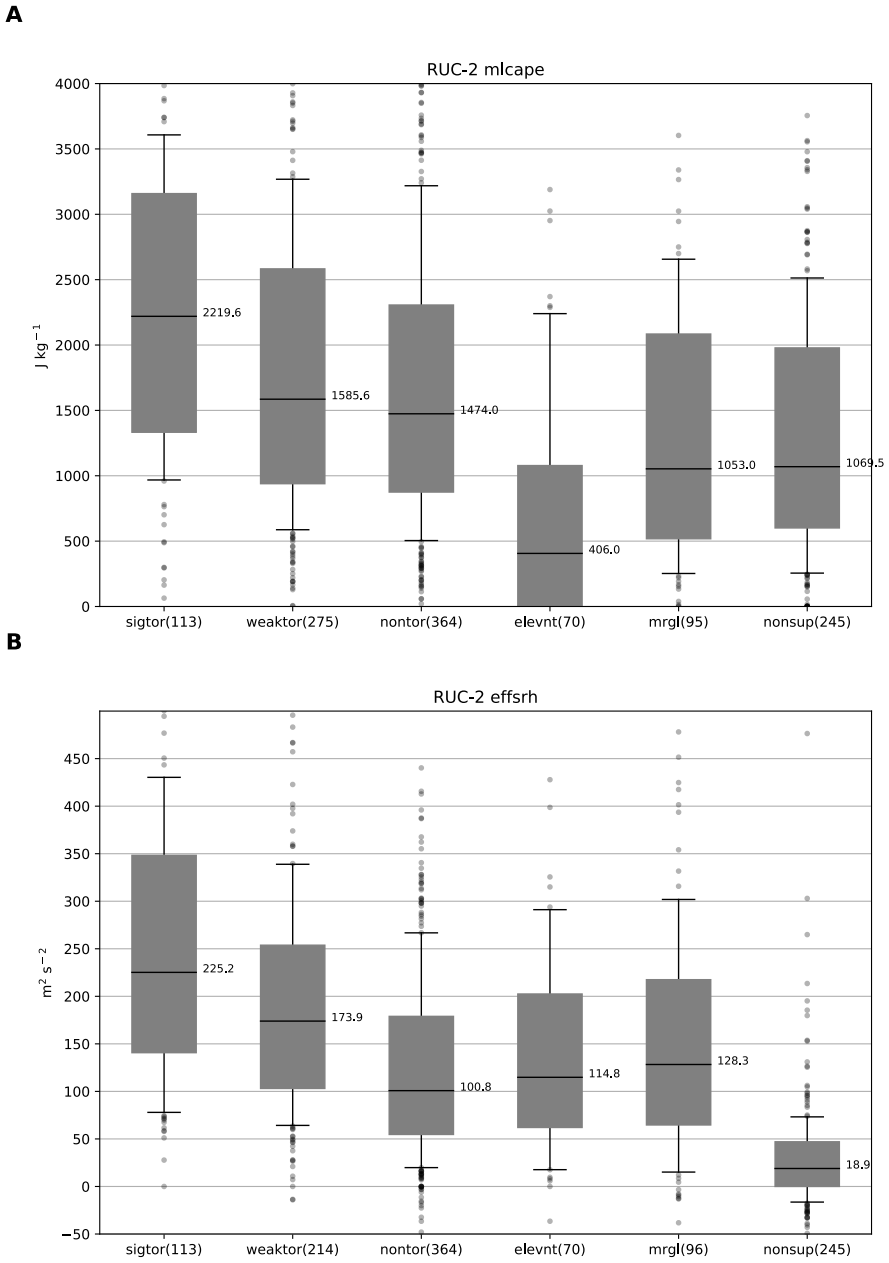

Figure 1. Box and whiskers plots of (A) MLCAPE and (B) ESRH for RUC-2 proximity soundings sorted by storm type. Shaded boxes enclose the 25th to 75th percentiles while black lines and numbers represent the medians. Whiskers extend to the lowest and highest values excluding outliers. Median values are printed.

CHAPTER 4

EXAMPLE CASES

To demonstrate how reanalyses depict severe weather environments, two example cases were chosen from T07: 3 May 1999 (Oklahoma City, OK) and 18 July 2004 (Grand Forks, ND). These events were selected based on varying meteorological environments as well as scales of impact.

3 May 1999 OKC

The Oklahoma event on 3 May 1999, was a significant severe weather outbreak that generated 69 tornadoes from 10 distinct supercell thunderstorms. Considering the primary purpose in this paper is to compare the suite of reanalyses to RUC-2, a full discussion of this event is not included. Rather, the reader is referred to Thompson and Edwards (2000), which provides an excellent overview of the complex surface and upper level features that contributed to the tornado outbreak. Besides this paper, Burgess et al. (2002) detailed a radar-based discussion of the event using WSR-88D and Doppler on Wheels (DOW) radars. Additional information on the event can be found in Markowski (2002), Roebber (2002), and Stensrud and Weiss (2001).

Surface analyses for this event at ~00 UTC 4 May 1999 are provided in Fig. 2. The primary features to note in this figure for all reanalyses are:

- The surface cyclone centered over northeastern Colorado and western Nebraska
- A broad region of lee troughing with low-level advection of moisture

- A dryline

There are a few notable differences from the RUC-2 (Fig. 2a). The extent of moisture and values of dewpoints varied across the warm sector (70-77 °F, 21-25 °C at OKC). Many of the reanalyses had a more diffuse dryline, while the 20CR (Fig. 2e) had a weaker, more northerly displaced cyclone. Many of these properties are expected due to reduced spatial and temporal resolution of the reanalyses and, in the case of 20CR, the ensemble nature of this reanalysis.

Upper air patterns at 500 hPa for this event are depicted in Fig. 3. Like many outbreaks, this event was associated with a shortwave trough and ~50kt jet streak embedded within larger-scale southwesterly flow. While the jet was present in the reanalyses, there were some subtle differences in both the positioning and magnitude of the jet between RUC-2 (Fig. 3a) and the reanalyses (Fig. 3b-g). Using the 50kt isopleth as a guideline, 20CR (Fig. 3e) appears to most poorly represent the event. The jet maxima over western Oklahoma, which is apparent in RUC-2 and other reanalyses, was absent for 20CR. Other differences include a slight low bias in NARR (Fig. 2b) for the local maxima just west of OKC, and a lack of higher magnitude winds over southern Texas in JRA55 (Fig. 2f). RAOB data from this time included a 65kt, 45kt, and 50-55kt observation over southern TX, the TX panhandle, and New Mexico, respectively. This suggests that, compared to observations, NARR may have had the best agreement aside from any possible observation errors.

To understand the vertical structure of the environment, soundings for this event are shown in Fig. 4. Variations in lapse rates, depth of moist and dry layers, effective inflow layers, hodographs, and surface temperature/dewpoint spreads are apparent in the

reanalyses. These factors contribute to a spread of values in derived parameters. Even the RUC-2 sounding from 23 UTC on 3 May 1999 at OKC (Fig. 4a) has noticeable differences with the 00 UTC OUN sounding on 4 May 1999 (Fig. 5). Although, given the proximity to ongoing convection, this should not be surprising.

Calculations of CAPE varied across the datasets. Observed and RUC-2 analyzed MLCAPE (SBCAPE) for this case were 2886 (3222) and 4199 (3891) J kg^{-1}, respectively. Reanalysis calculated CAPE fell largely into two camps with JRA55, NARR, and CFSR producing reasonable estimates (3681-2950 J kg^{-1} for MLCAPE) while 20CR, ERA-Interim, and MERRA 2 grossly under-reanalyzed this variable (1413-978 J kg^{-1}). The reasons for these differences are varied, but some of the more obvious factors include a dry bias in the boundary layer for ERA-Interim (Fig. 4c) and warmer mid-level temperatures for MERRA 2 (Fig. 4d). While NARR (Fig. 4b) and JRA55 (Fig. 4f) had the highest values of MLCAPE and a similar effective inflow layer to RUC-2, the higher values of CAPE can be attributed in part to a warm bias at the surface. This property led to higher LCL heights than observed.

Kinematic parameters also demonstrated a range in results. Variables such as SRH (effective or absolute height) were lower than observed for most of the reanalyses. For example, 0-1km SRH values ranged from 155-310 m^2s^{-2} compared to the observed and analyzed values of 255 and 309 m^2s^{-2}, respectively. Of these reanalyzed values, only MERRA 2 (Fig. 4d, 310 m^2s^{-2}) fell near this range, and this reanalysis was consistently higher than RUC-2 or OUN. Interestingly, this reanalysis arguably had one of the poorest representations of the sickle-shaped hodograph that was seen in most of the soundings, and this led to helicity values that were consistently higher than RUC-2 or OUN. Although

CFSR (Fig. 4g) had a 0-1km SRH lower than observed (due to lower 0-1km bulk shear), effective and 0-3km SRH values were also higher than RUC-2 and OUN. For ESRH, this is most likely related to an effective layer that was deeper than observed. Other than these two examples, the remainder of the reanalyses (NARR, JRA55, ERA Interim, and 20CR) were lower than RUC-2 and OUN for SRH. Despite this issue, these reanalyses were within a few knots of EBWD except for 20CR (Fig. 4e).

18 July 2004 GFK

The second case study was a more isolated severe weather event in eastern North Dakota that produced several tornadoes, one of which was an EF-4. Kellenbenz and Grafenauer (2007, hereafter KG07) examined this northwest-flow event and hypothesized that evapotranspiration processes were responsible for enhancing boundary layer moisture. While KG07 stressed the high MLLCL heights (~1800m), Edwards and Thompson (2009) pointed out that while corrected values (~1400m) were high, they still fell within upper percentile ranges for multiple studies (e.g. T03). Edwards and Thompson (2009) also provided other additional details and corrections to the study by KG07. Regardless of these issues, this is a reasonable case to explore due to the larger dewpoint depressions and weaker shear compared to 3 May 1999.

Surface analyses for 19 July 2004, 00 UTC are provided in Fig. 6. Unfortunately, RUC-2 analyses were unavailable for this case. Instead, comparisons are made to the surface analysis shown in Fig. 6 of KG07. Pertinent features from this figure included a surface pressure trough oriented from SW to NE across the region and surface temperature (dewpoints) around 83-89 (70-73) °F. While the wind shift is seen in many of the

reanalyses, the trough was much harder to discern. Furthermore, the reanalyses had varying biases for the surface thermodynamic fields. Although most of the reanalyses and the surface analysis had a moist axis in eastern ND, surface dewpoints were too dry (~0-3 °F). The reanalyses also had evidence of a thermal ridge that was oriented with the surface trough, but the position varied and this led to varying biases across the domain. NARR (Fig. 6a) had perhaps the most notable issue with surface temperatures, with values 5-10 °F too high across the domain. Gradients were too weak in 20CR (Fig. 2d), and the windshift was displaced well north of the observed location. Subjectively, ERA-Interim (Fig. 2b) appeared to have the best representation of the surface analysis.

Aloft, the region was under prevailing northwest flow at 500 mb (Fig. 7). Compared to Fig. 5 in KG07, reanalyses (aside from 20CR, Fig. 7d) captured the pattern quite well with a broad area of 30-40kt flow downstream of an upper-level ridge axis, which was located just west of the plotted domain. Minor nuances between the reanalyses included a weak shortwave trough in NARR (Fig. 7a) and separated maxima in the higher resolution reanalyses (e.g. MERRA 2 and CFSR).

Proximity soundings at 00 UTC 20 July 2004, were compared to the Grand Forks, ND 23 UTC RUC-2 sounding from T07 (Fig. 8). SBCAPE was ~4000 J kg^{-1} in RUC-2 (Fig. 8a), NARR (Fig. 8b), and JRA55 (Fig. 8c), while the other reanalyses had significantly less instability. This is similar to what was found for the 3 May 1999 case. The consideration of MLCAPE worsened the comparison; the closest reanalysis (NARR) had a deficit of ~1500 J kg^{-1}. In this case, the primary problem was a lack of moisture throughout the boundary layer for the reanalyses. In light of the earlier discussion about

this event, one can only speculate whether evapotranspiration (as hypothesized in KG07), problems with the boundary layer schemes, or some other cause was to blame.

Kinematically, this case was weaker than 3 May 1999; RUC-2 had 0-1km SRH (EFFSRH) of 105 (225) m^2s^{-2}, an EBWD of 22 kts, and 0-1km bulk shear of 13kts. With the exception of 20CR (Fig. 8e), which was, once again, too weak, the other reanalyses performed well for fixed height parameters, such as 0-1km SRH (85-130 m^2s^{-2}) and 0-1km bulk shear (11-15 kts). Larger variations were found when the effective layer was considered. For example, EFFSRH had a low bias with values from 96-179 m^2s^2, while EBWD ranged from 14-27 kts. Considering effective layers were similar in this case, most of this disparity/variation can be attributed to differences in the wind field, which may be expected provided the location of the soundings near a boundary.

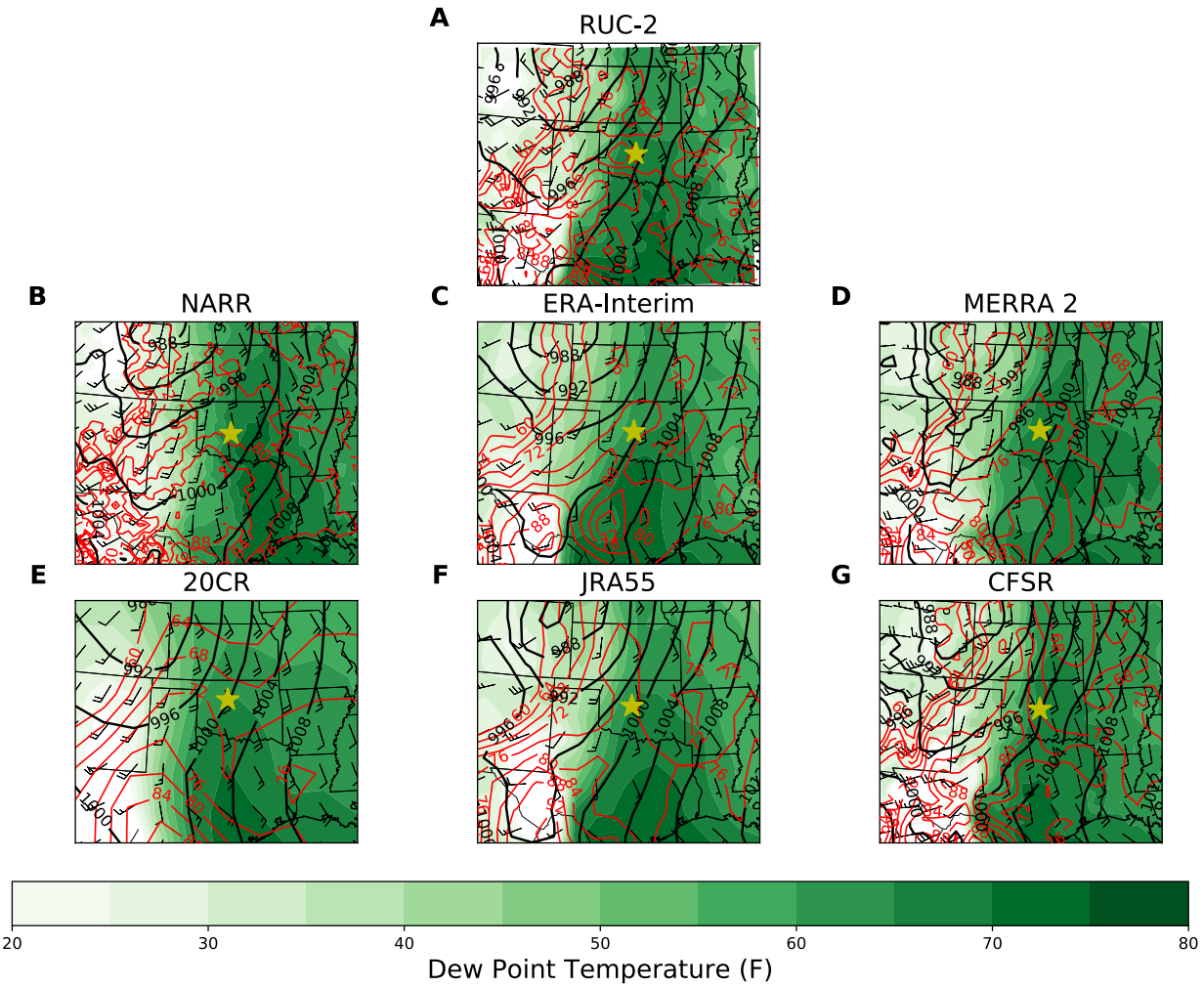

Figure 2. Surface analyses for (A) RUC-2 at 23 UTC 3 May 1999 and (B-G) reanalyses at 00 UTC 4 May 1999. Dew point temperatures are shaded, while MSLP and temperatures are contoured with black and red lines, respectively. Select wind barbs are shown while yellow stars represent the location for soundings shown in Figure 4.

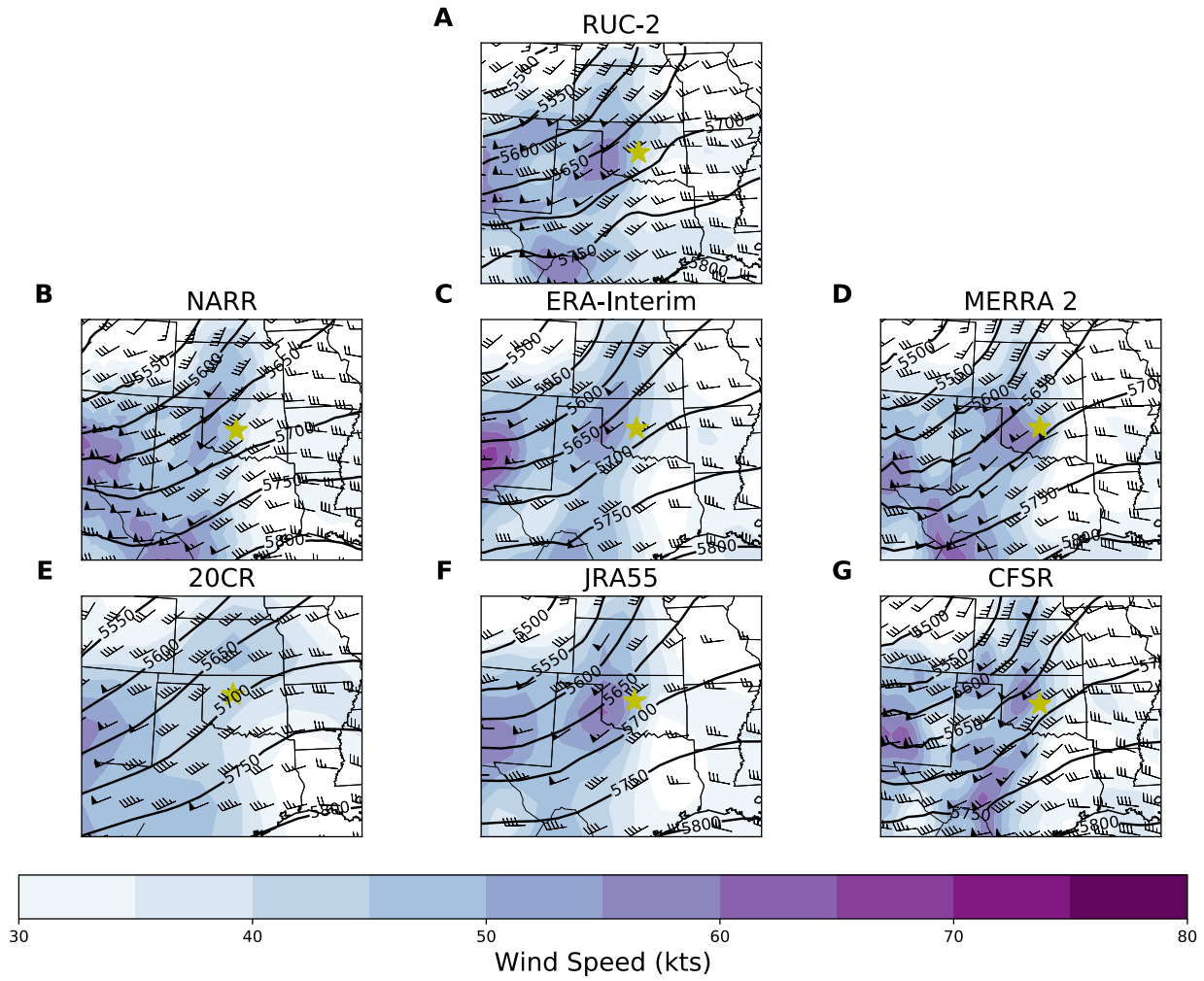

Figure 3. As in Fig. 2 except for 500 mb analyses. Wind magnitudes are shaded while geopotential heights are contoured with black lines.

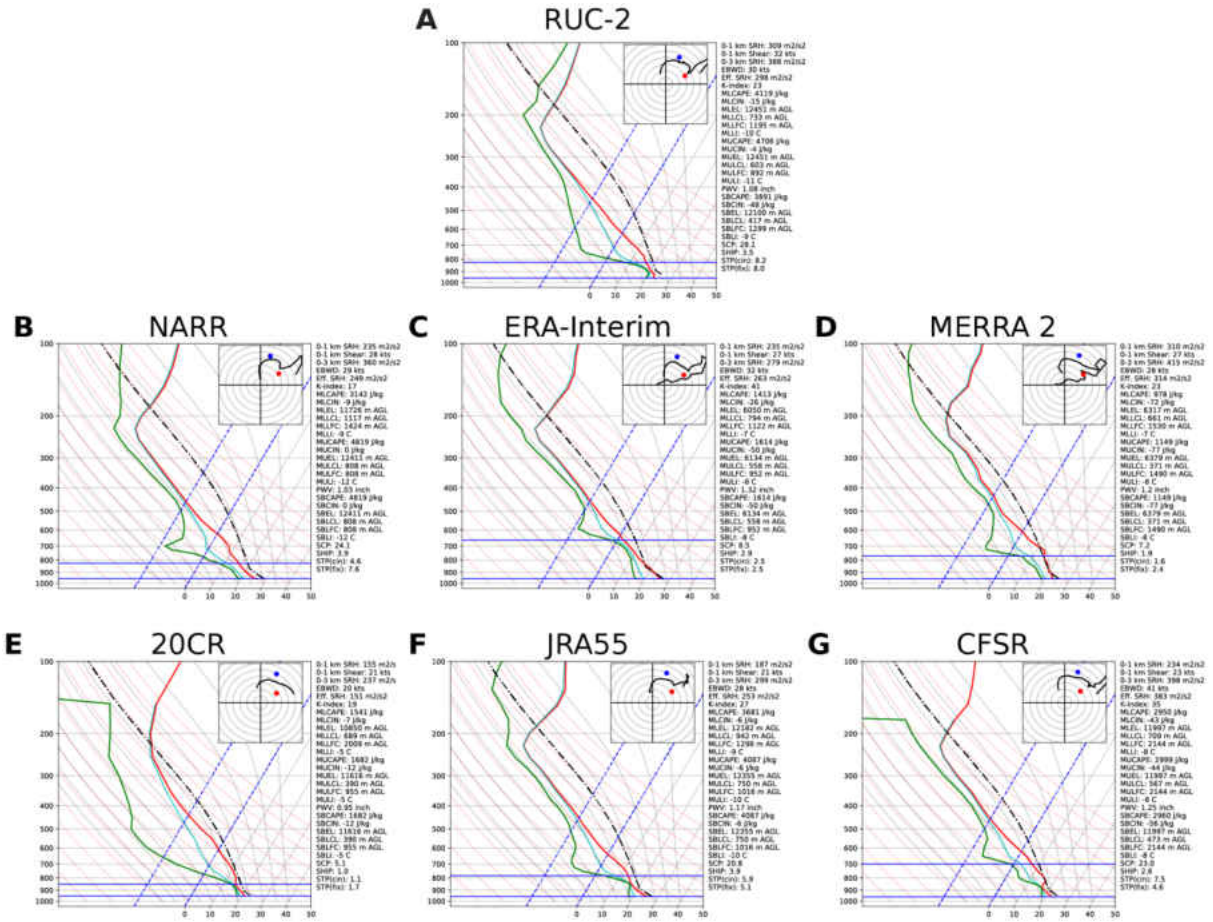

Figure 4. Proximity soundings near Oklahoma City, OK for (A) RUC-2 at 23 UTC 3 May 1999 and (B-G) reanalyses at 00 UTC 4 May 1999. Locations are marked by the yellow stars in Figs. 2-3.

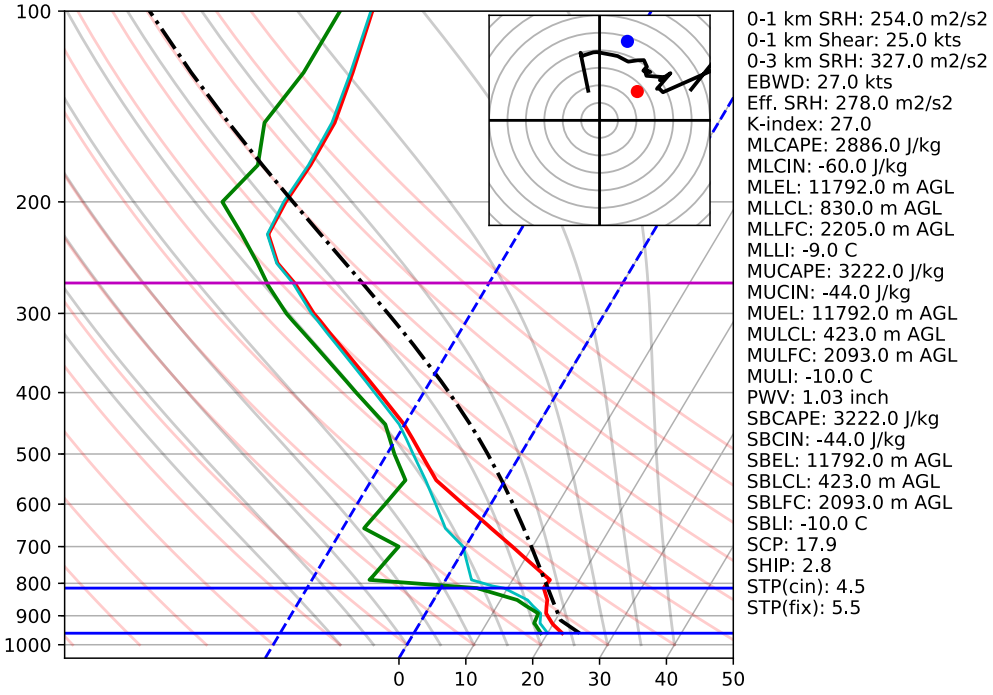

Figure 5. Observed sounding for Norman, OK (OUN) at 00 UTC 4 May 1999. Because the observed OUN sounding ended early (~275 mb, magenta line), the sounding was augmented with the RUC-2 proximity sounding above this level to allow for calculation of CAPE.

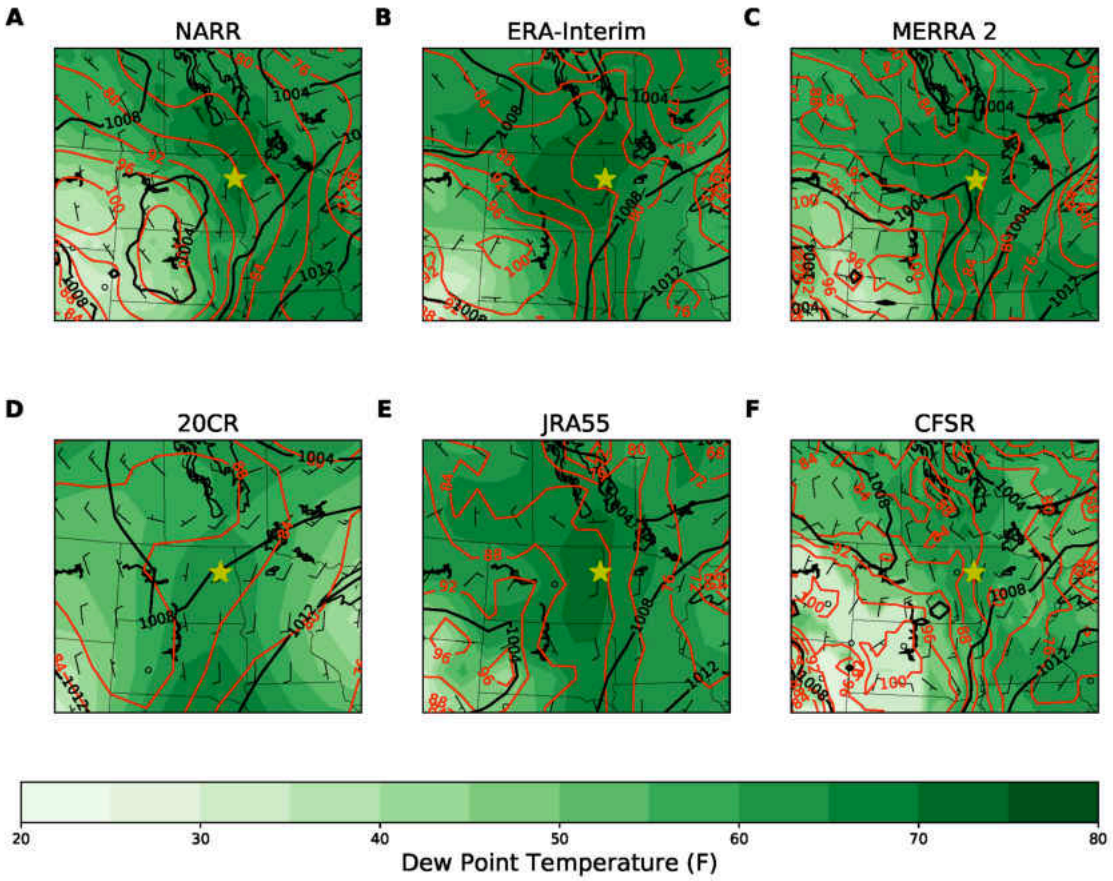

Figure 6. 00 UTC 19 July 2004 surface analyses for the reanalyses. Contoured variables are identical to Fig. 2.

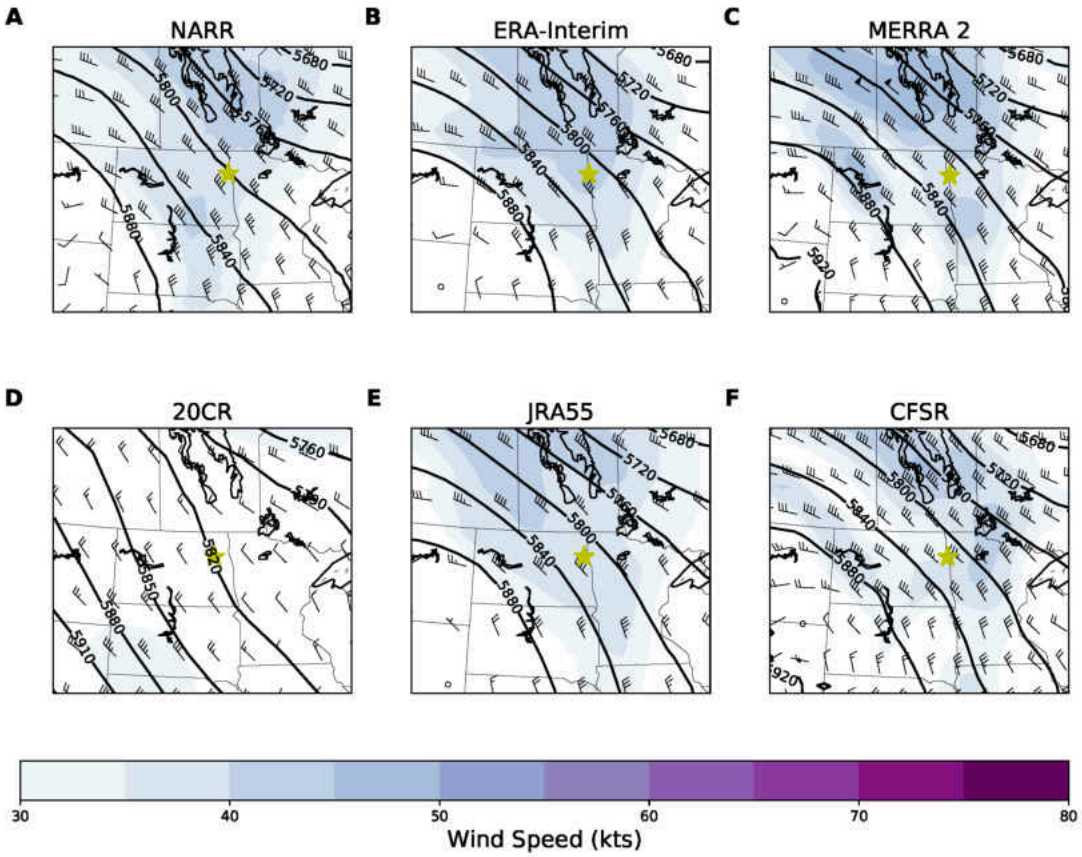

Figure 7. As in Fig. 6 except for 500 mb analyses. Contoured variables are identical to Fig. 4.

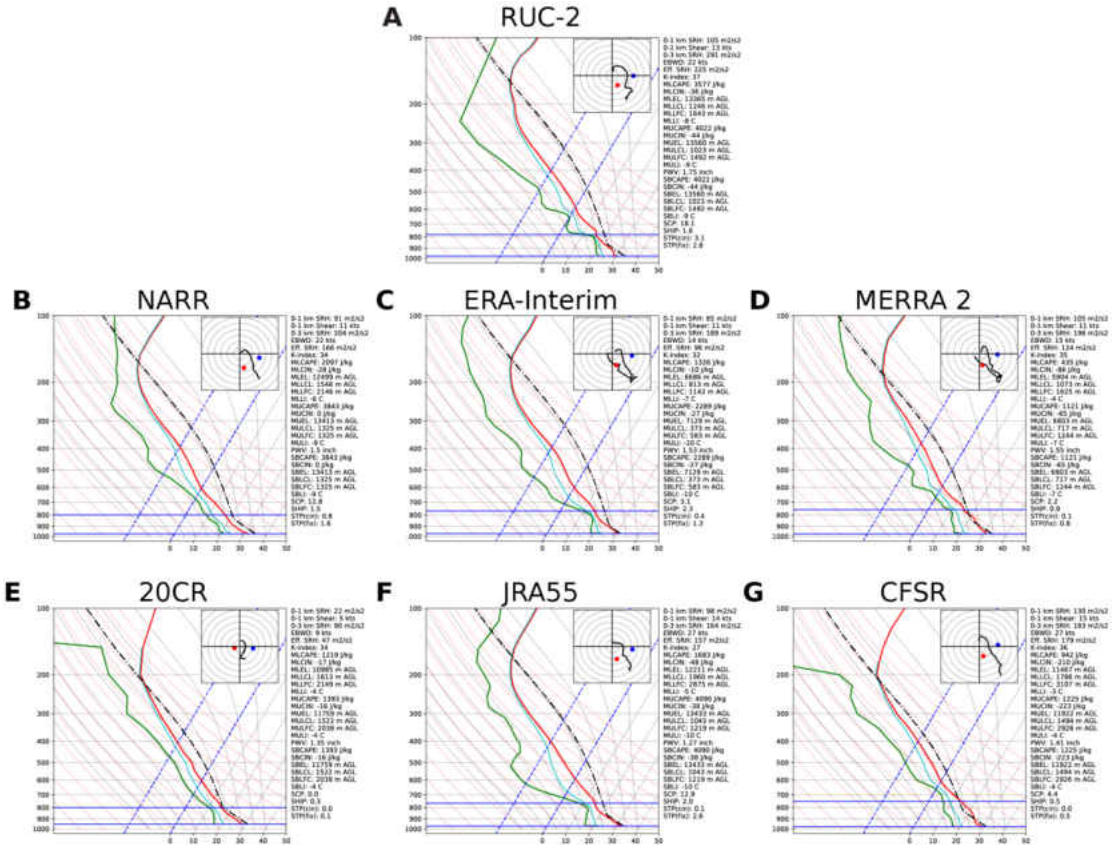

Figure 8. As in Fig. 4 except for locations closest to Grand Forks, ND at 00 UTC 19 July 2004.

CHAPTER 5

RESULTS

Reanalysis Performance and Comparison

Results for Part I are broken down into two sections. In section a, statistical properties calculated from all available soundings are presented. In section b, results are segregated by storm type to understand whether the reanalyses can reproduce the results of T07 and T03.

Statistical Results

Box and whiskers plots for the various convective parameters (Table 2) from each reanalysis were compared to those from RUC-2. Results are shown for select variables that are more commonly used (Fig. 9). Similar to what was found in the case studies, the reanalyses fall into two camps for MLCAPE (Fig. 9a), i.e. NARR and JRA55 have comparable values to RUC-2 and the remaining reanalyses are biased low. Compared to the median value of 1394 J kg^{-1} in RUC-2, NARR and JRA55 had values within 250 J kg^{-1}, and the box and whiskers were overlapped. The other reanalyses on the other hand suffered from a large negative bias with median values between 196-455 J kg^{-1}. This narrowed the distance between the quartiles, leading to the 3rd quartile of the reanalyzed distributions falling either below the median (ERA-Interim, CFSR, and 20CR) or the 1st quartile (MERRA 2) of RUC-2.

Kinematic variables such as 0-1km bulk shear (Fig. 9b) and EFFSRH (Fig. 9c) have smaller dichotomies than MLCAPE. In the former case, the reanalysis distributions are

similar in width (with exception of 20CR) with median values within +/- 4kts of the RUC-2 median (12.8 kts). When the effective layer and storm motion vector is considered to calculate EFFSRH, more separation is found between the reanalyses (Fig. 9c). While all reanalyses have a substantial negative bias compared to RUC-2, the inclusion of thermodynamic information (through the effective layer) causes a separation in the lower quartile for JRA55 and NARR vs. the other reanalyses. In short, the lower CAPE values (regardless of type of parcel) observed for ERA-Interim, MERRA 2, CFSR, and 20CR raise the probability that the criteria for inclusion within the effective layer (CAPE $\geq$ 100 J kg $^{-1}$ and CIN $\geq$ -250 J kg^{-1}) are not met. Therefore, these layers are narrower, lowering the values of EFFSRH. Given that SCP is also a function of CAPE, this property also significantly decreases the median and 3rd quartiles of this parameter (Fig. 9d). Although only four parameters are shown here (Fig. 9), most other parameters follow similar patterns. Overall, thermodynamic variables tend to have a strong negative bias for all reanalyses aside from NARR and JRA55. Kinematic variables are better represented by all reanalyses, while composite parameters such as SCP or STP are biased low for most reanalyses. This latter property is most pronounced in reanalyses that exhibited strong negative biases for thermodynamic parameters. These patterns can be further seen in the statistical calculations in the following section.

To provide a concise way to evaluate the ability of the reanalyses to reproduce RUC-2 calculated parameters, correlation coefficients and biases are calculated (Tables 3 and 4). These correlation values reinforce the notion that reanalyses are more likely to struggle with the thermodynamic vs. kinematic properties of the soundings. While moderate to strong correlations are found for CAPE depending on the type of parcel trace

(SB, ML, and MU), all reanalyses had difficulty accurately representing CIN regardless of how it is calculated. NARR, for example, has a maximum correlation of 0.54 for MLCIN with the other reanalyses lagging further behind (e.g. 0.39 for ERA-Interim and JRA55). Biases for the reanalyses are generally negative for CAPE and positive for CIN. Exceptions to this rule include: SBCAPE for NARR; ML and MUCIN for JRA55; and all CIN values for CFSR. Although surface-based parcels generally increase the CAPE and yield lower biases, this is at the expense of lower correlations. In most cases, this difference (e.g. 0.66 vs. 0.75 for NARR) is statistically significant.

Reanalyses perform more consistently for kinematic variables, such as fixed level SRH and bulk shear (Table 4). Excluding 20CR, correlations for SRH range between 0.72-0.80. While the reanalyses are biased low for SRH and helicity (with the exception of MERRA 2, Table 5), the strong correlations suggest that the reanalyses do a better job simulating the wind field versus the thermodynamic profile. However, given the strong dependence of storm type on low-level helicity (T07), low biases may mean it is more difficult to segregate between categories.

The final category of parameters is those that mix thermodynamic and kinematic properties. Compared to straight kinematic parameters, the inclusion of thermodynamic information leads to decreases in correlations (e.g. 0-1km SRH vs. EFFSRH, Table 4). Considering parameters that are a combination of CAPE and shear (e.g. SCP and STP), biases are largely controlled by the thermodynamic characteristics for a reanalysis. NARR and JRA55, which have higher CAPE values, contain the smallest biases for these parameters, whereas the other reanalyses are impacted by the low values of CAPE.

Comparison of reanalyses to T03 and T07

The results of T03 and T07 utilizing RUC-2 proximity soundings demonstrate certain parameters have skill in identifying environments supportive of supercells and tornadoes. In T03, relationships between storm categories and thermodynamic and kinematic parameters led to the development of SCP and STP, which are used operationally by the Storm Prediction Center (SPC) to help discriminate between severe weather environments. T07 defined the effective storm relative helicity (ESRH) and found it is more appropriate in discriminating between tornadic and non-tornadic environments than its counterpart, storm relative helicity (SRH), which does not take the effective inflow layer depth into account. The question remains whether these findings can be duplicated with proximity soundings from reanalyses.

Prior to presenting the results, it is important to consider what differences are needed to discriminate between environments. Similar to the methodology in Reames (2017), significant discriminators are defined as those in which the median of a defined distribution is below the lower quartile or above the upper quartile of another. To put T03 and T07 in context of this criterion, the following parameters are good discriminators between categories:

- MLCAPE: sigtor, elevnt, and mrgl / nonsup
- 0-1km Bulk Shear: sigtor, weaktor and nontor, and mrgl/ nonsup
- EFFSRH: sigtor, nontor, and nonsup
- SCP: sigtor, nontor, and mrgl/nonsup
- STP: sigtor, weaktor/nontor

To visualize both the quartiles and probability density functions for the various categories, violin plots are shown for select parameters (Figs. 10-14).

MLCAPE distributions in RUC-2 demonstrate a general decrease and broadening of MLCAPE as categories transition from sigtor, weaktor, to nontor categories (Fig. 10a). Although sigtor, elevnt, and margl/nonsup meet the criteria for being a good discriminator, significant overlap is still found among the categories. For the reanalyses, a quick comparison of median values demonstrates the same trend in magnitude, but with substantially more overlap. Of the presented datasets, only NARR (Fig. 10b) and CFSR (Fig. 10g) successfully discriminate between some of the categories (sigtor and mrgl). While JRA55 has MLCAPE on the same order of magnitude of NARR and RUC-2, inspection of the individual categories reveals a negative bias for sigtor offset by higher values for the other categories. This leads to additional overlap between the distributions, and no discrimination. For the other reanalyses, large negative biases for MLCAPE narrows the distributions making discrimination impossible.

As demonstrated by Fig. 11, ESRH is a reasonable discriminator between tornadic and non-tornadic events for RUC-2 soundings in T07. Compared to MLCAPE, comparisons between the reanalyses and RUC-2 are much better. Qualitatively, the shapes of the violins are similar, with a decrease and broadening of the distributions as storm type shifts from sigtor to nonsup. As a result, all of the reanalyses, except 20CR (Fig. 11e), discriminate between the sigtor and nontor categories. The predominant issue with 20CR is an overall negative bias in EFFSRH across all categories. RUC-2 and the other reanalyses have a probability maximum located between 150-250 m^2s^{-2}, whereas 20CR is located at ~25 m^2s^{-}

^{2}. The overall low bias in EFFSRH results in a mixed bag for discerning between nontor and nonsup categories.

A better kinematic discriminator for these environments in reanalyses is 0-1km bulk shear (Fig. 12). While RUC-2 discriminates between categories in a similar fashion using this parameter and EFFSRH, a more clearly defined partitioning is found between categories for the reanalyses. Even 20CR, which has the largest negative bias for kinematic parameters, has significant separation between categories (Fig. 12e).

To discriminate between environments supportive of long-lived, rotating updrafts, SCP is used (T03). Values > 1 are strongly supportive of supercells as noted by the differences between nontor and mrgl/nonsup categories in RUC-2 (Fig. 13a). Due to SCP being a product of kinematic and thermodynamic parameters, performance is impacted by biases in CAPE. While the distributions in reanalyses are concentrated on a value of 0 for the mrgl and nonsup categories, SCP values for stronger categories are biased low. As a result, this makes discrimination of supercell and non-supercell categories difficult. Even for reanalyses like NARR (Fig. 13b) and JRA55 (Fig. 13f), which have reasonable performance for MLCAPE, formal discrimination is limited to non-supercell categories vs. strongly and weakly tornadic storms. CFSR (Fig. 13g) is also capable of making this distinction, while ERA-Interim (Fig. 13c) and MERRA 2 (Fig. 13d) can only distinguish between sigtor and the non-supercell categories.

STP is used to distinguish between strongly tornadic and nontornadic supercell environments with values > 1 maximizing forecast skill between these categories (Thompson et al. 2002, T03). SHARPpy provides two versions of this parameter: one based on the effective inflow layer and the other with a fixed layer. Based on prior discussion

(EFFSRH vs. 0-1km bulk shear), the fixed layer STP is shown in Fig. 14 and discussed herein. As shown in T03 and Fig. 14a, RUC-2 sigtor can be discriminated from all other categories as the median of weaktor (0.85) is lower than the 1st quartile of sigtor (0.97). Although STP is lower for all reanalyses, NARR (Fig. 14b), ERA-Interim (Fig. 14c) and MERRA 2 (Fig. 14d) successfully discriminate between the sigtor and weaktor categories. That said, the distance between the 1st quartile of sigtor and median of weaktor is much narrower than RUC-2, and more overlap occurs. JRA55 (Fig. 14f) and CFSR (Fig. 14g) have more overlap, but can discriminate between sigtor and nontor categories. 20CR has the worst overlap and discrimination can only occur between sigtor and mrgl/nonsup cases.

Climatology and Trends in Severe Weather Parameters

Results for Part II are broken into two sections. The first section consists of a brief climatological evaluation of severe weather variables. The focus of the second section is to examine any trends in severe weather parameters during the period of 1986-2015. Throughout Part II, various parameters and combinations thereof are examined for varying thresholds. In particular, the impact of calculation method choice for CAPE is examined (i.e. MLCAPE, SBCAPE, MLCAPE). Since prior studies have listed surface temperature and associated moisture increases as possible effects of a changing climate, it would be reasonable to assume that those changes may affect surface-based CAPE (SBCAPE) more strongly than mixed-level CAPE (MLCAPE). Thresholds for parameters are chosen somewhat arbitrarily, albeit with results of parameter based studies (i.e. Rasumussen and Blanchard 1997, T03, T07), taken into consideration.

Climatology

One of the simplest ways to examine environments favorable for severe weather is to look at a single thermodynamic variable such as CAPE. Fig. 15a shows the number of days exceeding a threshold of 2000 J/kg of most unstable CAPE (MUCAPE) calculated from NARR for the time period of 1986-2015. Areas with the maximum number of days exceeding the chosen threshold are focused near the Gulf Coast region. However, this is not historically where one would expect to see a maximum in severe weather occurrence. This result also holds true for MLCAPE (Fig. 15b) and SBCAPE (not shown). As was previously discussed, CAPE alone is not the best proxy for severe weather occurrence (e.g. there are many days with very high CAPE in the plains regions during the summertime that are not always associated with severe weather). Therefore, inclusion of wind shear information is vital to getting a better idea of the overall capability of an environment to produce severe weather. One such way to do this is via the use of the composite C parameter (as in Gensini and Ashley 2011, hereafter GA11) which combines 0-6 km shear and MUCAPE. An example of the average climatology of this parameter exceeding 20,000 (i.e. environments favorable for significant severe weather) can be seen in Fig. 16. Although there is perhaps marginal improvement from using CAPE alone, the areas of maximum days exceeding the threshold are still biased south from reality. Perhaps a more useful parameter to examine is SCP (Fig. 17). The spatial distribution of days exceeding the chosen threshold (2.0) for this parameter is far more reasonable, albeit with some likely overestimation along the east coast.

Trends in Environments Favorable for Severe Weather

The second portion of this section is an analysis of trends in severe weather parameters for the historical period of 1986-2015. The first parameter that is examined is CAPE. Overall, days with MLCAPE greater than 2000 J/kg (Fig. 18) have decreased during Summer and Fall months (Fig. 18c, 18d). The strongest negative trends (-0.5 to -1.25 days per year) for this parameter occur over the Central and Southern Plains regions as well as areas of the Southeast. Trends for this parameter during Spring months are somewhat mixed with areas of statistically significant increase (decrease) over Arkansas (Texas coastal region, Fig. 18b). Similarly, days with MUCAPE greater than 2000 J/kg (Fig. 19) are primarily decreasing during summer and fall months over portions of the Southeast and Southern Plains (-0.25 to -1.6 days per year, Fig. 19c, 19d) and increasing during Spring months over the Midwest and portions of the Southeast (0.15 to 0.45 days per year, Fig. 19b). SBCAPE (not shown) displays a similar pattern to MUCAPE.

Of equal importance to severe weather environments is CIN, which is defined as the amount of energy required to lift a parcel to the Level of Free Convection (LFC). Even if other atmospheric parameters are favorable for severe convection, if the magnitude of CIN is too large, convective initiation will be inhibited. An analysis of the trend in mixed layer CIN (MLCIN) exceeding -75 J/kg can be seen in Fig. 20. The most significant trend exists for the summer months which shows a statistically significant decrease in the number of days with favorable magnitudes of MLCIN over the Southern, Central, and Northern Plains regions (Fig. 20c).

The next parameter examined is wind shear, which represents the other fundamental component of severe weather environments. An example of deep layer (0-6

km) shear can be seen in Fig. 21. The slope of the trend line for this parameter is positive nearly everywhere in the domain for winter and spring months (Fig 20a, 20b). The largest increases (0.3 to 0.75 days per year) are located over portions of the Central and Southern Plains. For summer and fall months, the trend is primarily neutral with areas of statistically significant decrease over portion of the Northeast (Fig. 21c, 21d). Note that as with CAPE, wind shear alone is not a particularly useful way to identify severe weather environments (ex. there are many days where there is little to no instability, but large values of wind shear).

To incorporate both instability and wind shear information into the analysis, SCP was investigated (Fig. 22). The trends in SCP are primarily positive across the southern portions of the Midwest as well parts of the Southeast for winter and spring months (0.2 to 0.48 days per year, Fig. 22a, 20b). The slope of the linear regression of this parameter for summer and fall months is varied, with negative slopes over the Southern Plains and positive slopes over portions of the Southeast. To gain a better understanding of why these trends exist for SCP, each component of the parameter is examined. The equation for SCP is given by:

$$SCP = \left(\frac{MUCAPE}{1000\frac{J}{kg}}\right) * \left(\frac{EFFSRH}{50\frac{m^2}{s^2}}\right) * \left(\frac{EBWD}{20\frac{m}{s}}\right)$$

ESRH exceeding 100 m^2/s^2 is discussed first (Fig. 23). This cutoff was chosen as it is a typical low-end value for significant severe weather. Note that the areas of positive trends for ESRH are over areas where SCP also have a positive trend. Days with EBWD > 15 m/s

(Fig. 24) show a pattern similar to that of ESRH, although the magnitudes of trends are lower. As was previously discussed, trends in days with MUCAPE exceeding 2000 J/kg (Fig. 19) largely match the trends of SCP.

In summary, these results indicate that parameters favorable for severe weather have been primarily increasing during spring months, and decreasing during the summer. The strongest trend is for MLCAPE during the summer (Fig. 18). Additionally, it seems likely that the magnitude of CIN is increasing during summer months, which would further suppress convection (Fig. 20c). The areas where positive trends for both shear and instability parameters overlap are most likely to have experienced an increase in favorable severe weather environments. SCP (Fig. 22) is a decent representation of this overlap, which shows the positive trends are confined to portion of the Midwest and Southeast during spring, while negative trends are focused over the Southern Plains regions during summer.

To gain better understanding of reanalyzed trends and to perform a "sanity check", select locations were also examined and compared to parameters calculated from observed soundings. The first examined location is OUN (Norman, OK). The time period for this site was limited to 1990-2015 since 1990 was the lower limit of data availability for observed soundings at OUN. Fig. 25 shows time series plots of days per year with MLCAPE > 2000 J/kg for both observations and NARR. Note that the negative trend in MLCAPE for summer months (Fig. 25c) exists in both observations and NARR and the two datasets have a strong correlation of 0.83. This gives credence to the trend depicted over this region as seen in Fig. 18. However, of the two datasets, only the trend depicted by NARR is statistically significant (p-value of 0.006).

The next location analyzed is JAN (Jacksonville, MS). This location was chosen due to Fig. 22, which shows that SCP should be increasing for most months. This trend is positive for both NARR and observations for all time periods (Fig. 26). The trend in observations is statistically significant for winter and spring seasons (p-values of 0.05 and 0.03, respectively).

Trends for other sites and variables were less pronounced. Tables of slopes for all months for observations and NARR can be found in Table 6 and Table 7, respectively while cutoffs used to calculate these slopes are listed in Table 8. The correlations between the two datasets for each site and parameter combination can be seen in Table 9. Most of the parameters have strong correlations between the datasets, especially for MLCAPE. SBCAPE and MUCAPE had lower correlations (XXX to XXX), at least at select locations (ex. correlation of 0.16 for MUCAPE at JAN).

Since the trends presented above are for time accumulated number of days exceeding a threshold, it is unclear how often observations and NARR soundings agree on a day-to-day basis. To better understand this property of the reanalyses, percentages of days exceeding a threshold were calculated (Table 10). These percentages were calculated by dividing the number of days where the threshold was exceeded in both NARR and observations on the same day by the total number of days exceeding the threshold in either dataset. As would be expected given prior results, CAPE parameters (MLCAPE, SBCAPE, MUCAPE) tended to have a higher percentage of days than other parameters.

Table 6. Linear regression slopes (days per year) for parameters calculated from observations at select locations. Bolded values are statistically significant (p-value <= 0.05).

	mlcape	sbcape	mucape	effsrh	scp
OUN	-0.2	**0.86**	0.59	0.05	**0.34**
TOP	-0.23	**0.57**	0.12	0.16	0.21
BIS	-0.1	-0.04	-0.11	0.02	-0.02
JAN	**-0.49**	-0.59	**-0.71**	**0.36**	**0.4**
DVN	-0.05	-0.04	-0.06	**0.41**	**0.38**
ILX	-0.11	-0.14	-0.22	**0.23**	**0.4**

Table 7. As in Table 8, but for NARR.

	mlcape	sbcape	mucape	effsrh	scp
OUN	**-0.43**	-0.4	-0.35	**-0.28**	**-0.4**
TOP	-0.29	-0.21	-0.15	-0.07	-0.15
BIS	-0.12	0.06	0.01	-0.02	-0.06
JAN	-0.46	-0.08	-0.09	0.18	**0.28**
DVN	-0.11	0.14	0.20	0.06	0.05
ILX	-0.01	0.14	0.19	0.11	0.15

Table 8. List of variables and associated thresholds used for parameters calculated for sites in Tables 7-9.

Variable	Threshold
mlcape	2000 J/kg
sbcape	2000 J/kg
mucape	2000 J/kg
effsrh	100 m^2/s^2
scp	2

Table 9. Correlation coefficients for parameters calculated from NARR compared to parameters calculated from observations at select locations.

	mlcape	sbcape	mucape	effsrh	scp
OUN	0.75	0.64	0.68	0.57	0.38
TOP	0.81	0.4	0.57	0.78	0.68
BIS	0.96	0.78	0.79	0.53	0.58
JAN	0.69	0.2	0.16	0.58	0.59
DVN	0.83	0.55	0.6	0.67	0.59
ILX	0.89	0.45	0.52	0.49	0.47

Table 10. Percentage of days threshold values were exceeded in both observations and NARR.

	mlcape	sbcape	mucape	effsrh	scp	stpcin
OUN	0.473	0.626	0.625	0.299	0.397	0.183
TOP	0.438	0.577	0.588	0.324	0.406	0.208
BIS	0.304	0.307	0.352	0.123	0.183	0.048
JAN	0.444	0.779	0.772	0.266	0.31	0.216
DVN	0.276	0.497	0.526	0.157	0.241	0.072
ILX	0.314	0.567	0.563	0.21	0.302	0.094

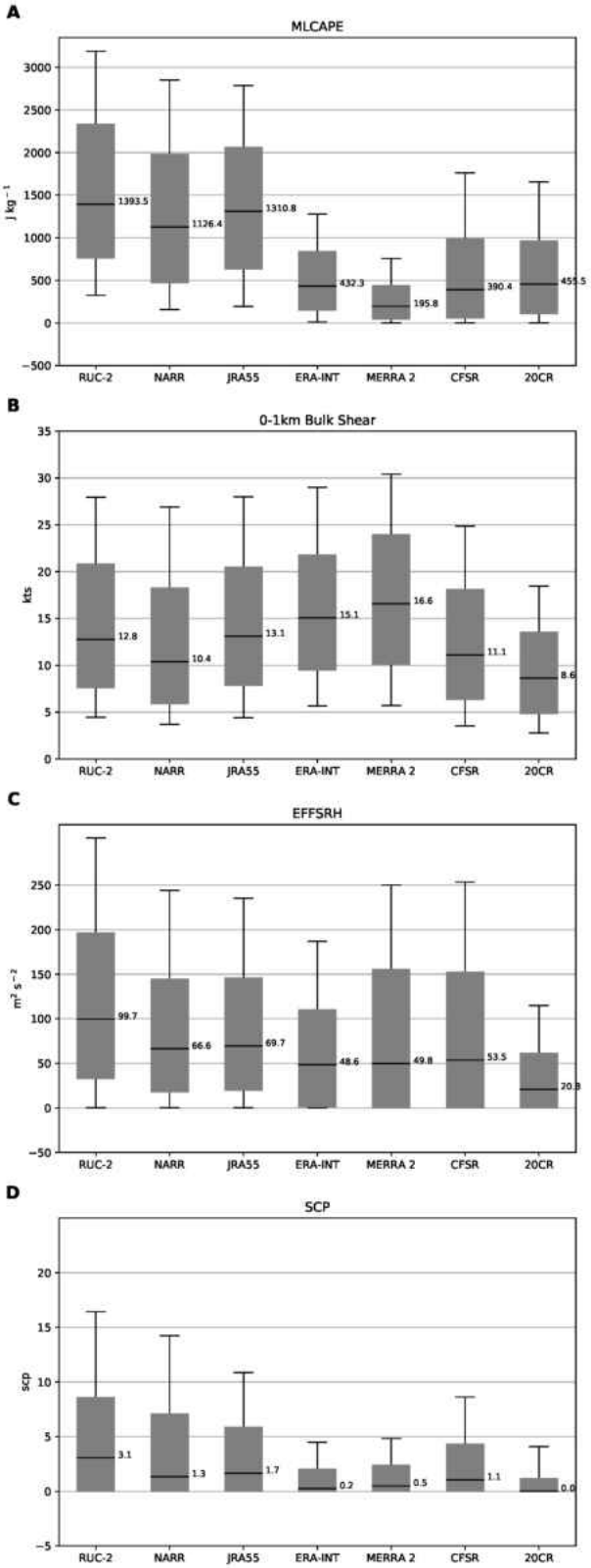

Figure 9. Box and whiskers plots of (A) MLCAPE, (B) EFFSRH, (C) SCP, and (D) 0-1km bulk shear for RUC-2 and the reanalyses.

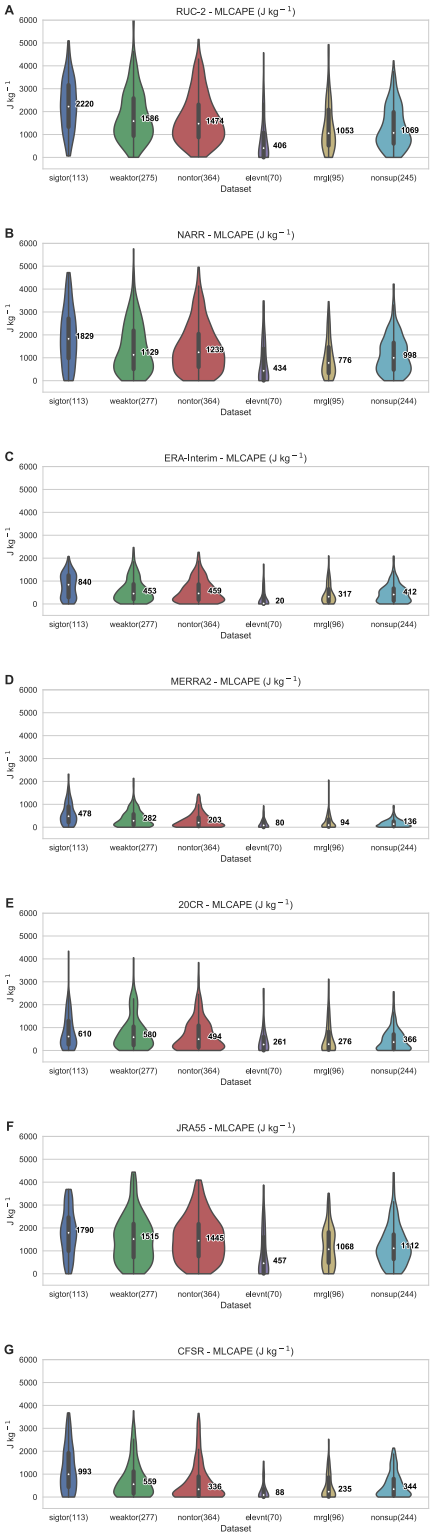

Figure 10. Violin plots of MLCAPE for (A) RUC-2 and (B-G) the reanalyses. Besides displaying traditional box and whiskers with the median values listed, the area represents a kernel density function for the distributions. Wider (narrower) sections of a violin plot represent regions where more (less) members of the distributions reside. The width of the violins are scaled by the count of values within each bin.

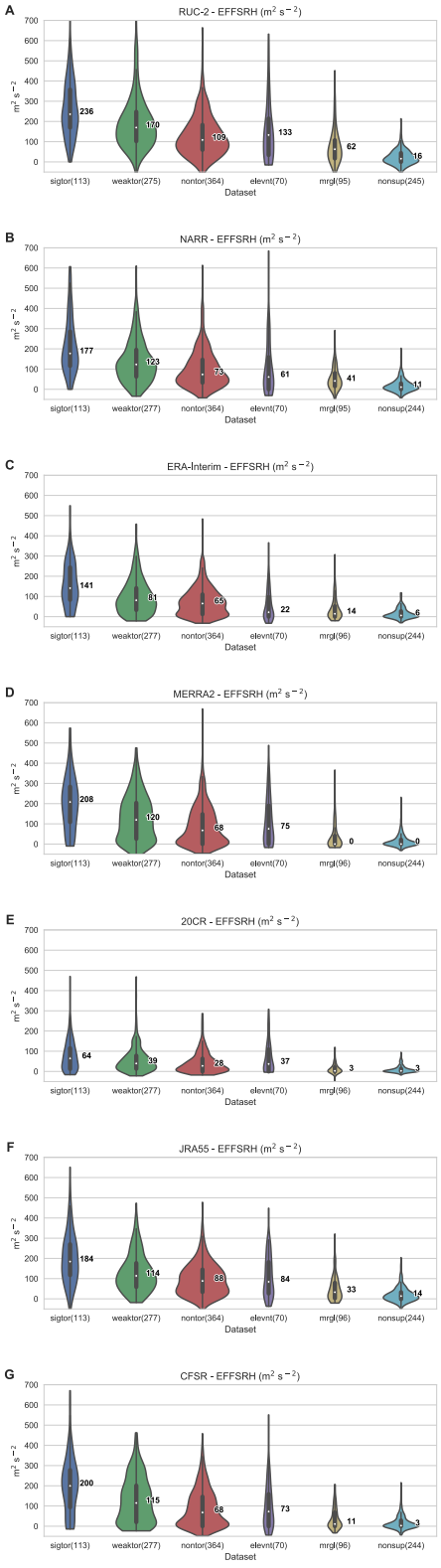

Figure 11. As in Fig. 10 except for EFFSRH.

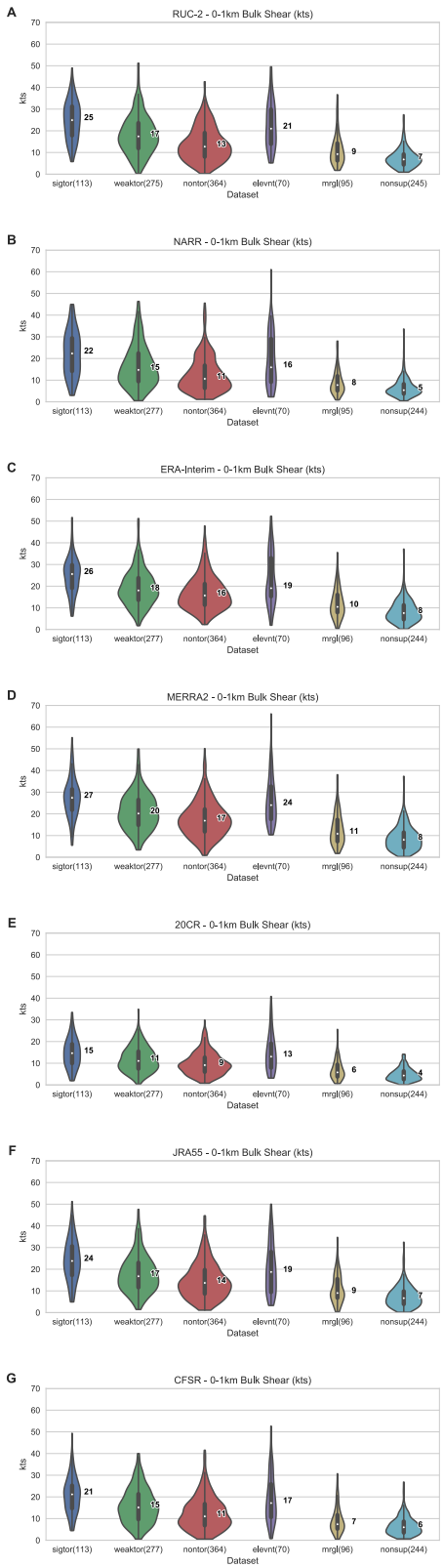

Figure 12. As in Fig. 10 except for 0-1km bulk shear.

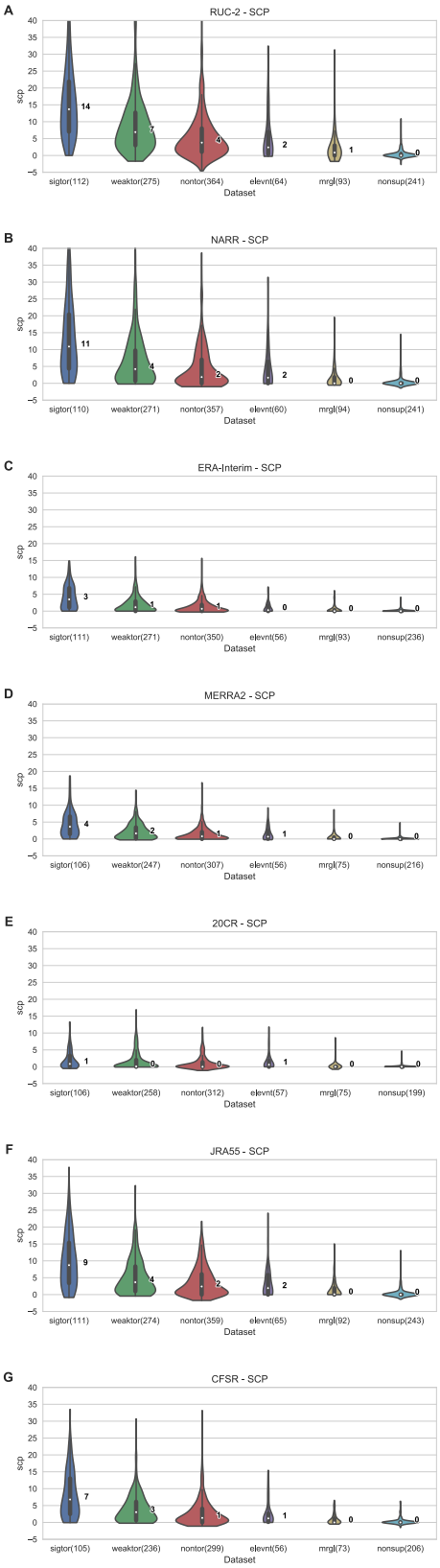

Figure 13. As in Fig. 10 except for SCP.

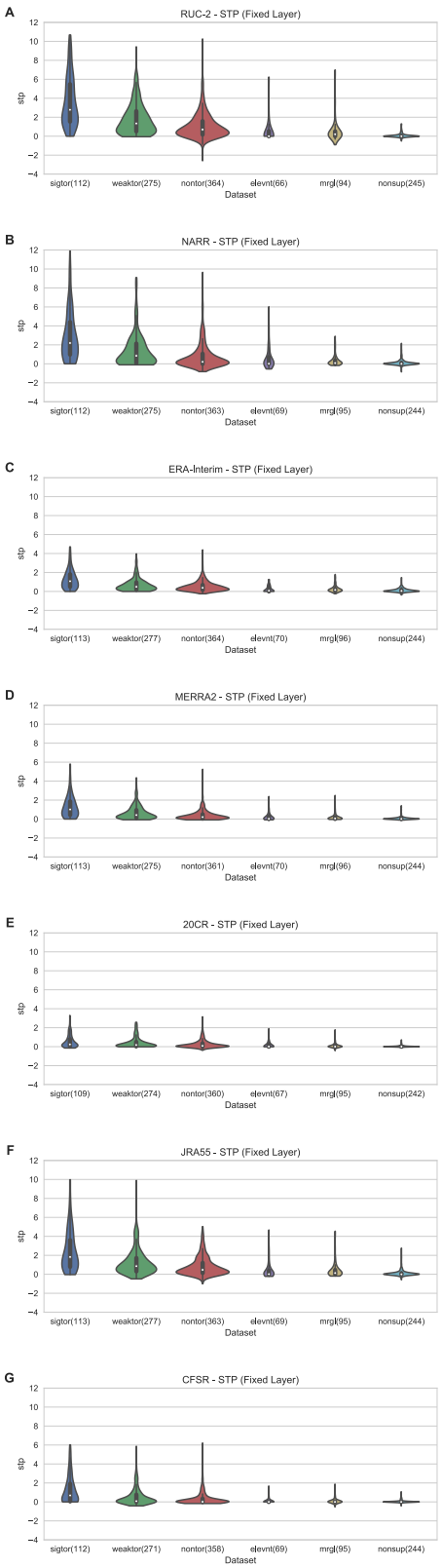

Figure 14. As in Fig. 10 except for STP.

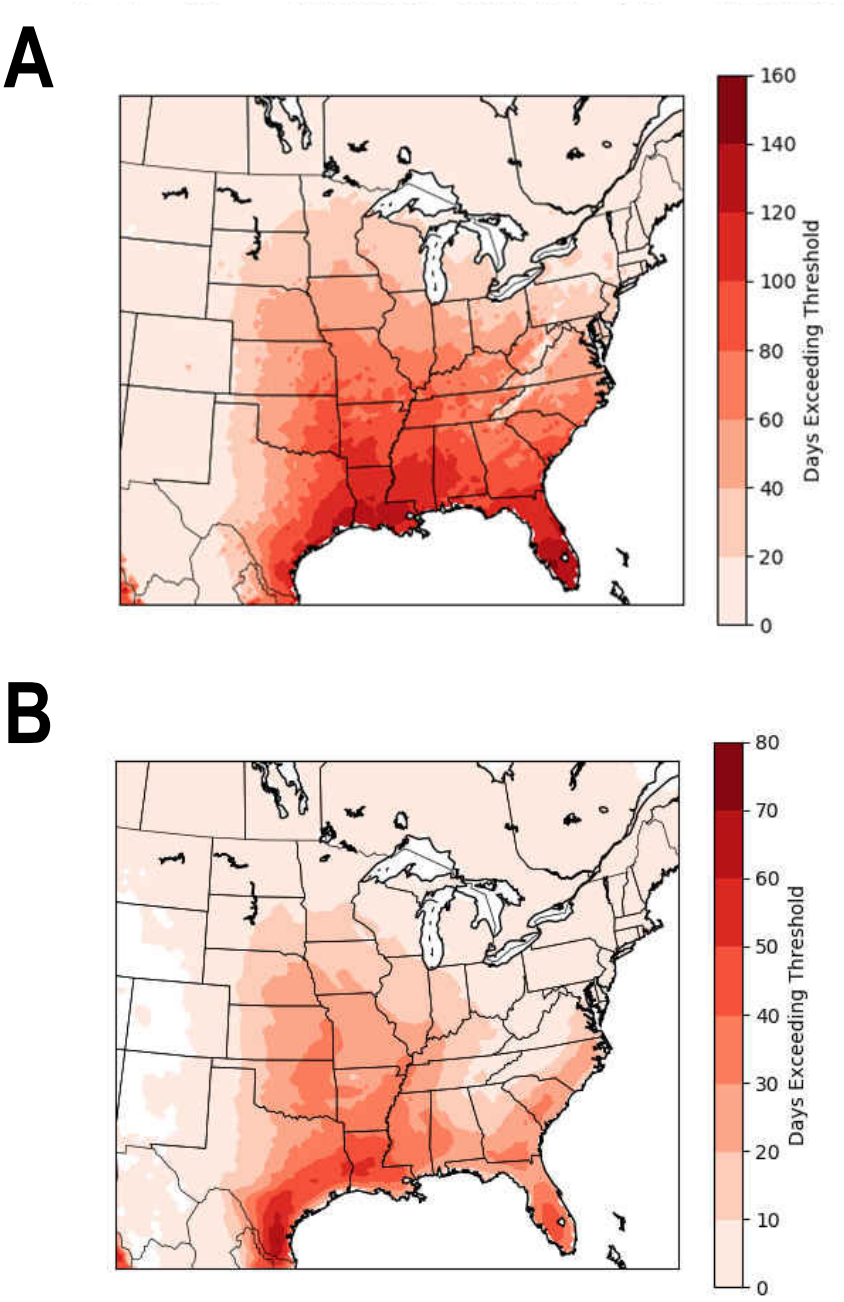

Figure 15: Average number of days during the period from 1986-2015 with 00 UTC NARR (A) MUCAPE values exceeding 2000 J kg^{-1}

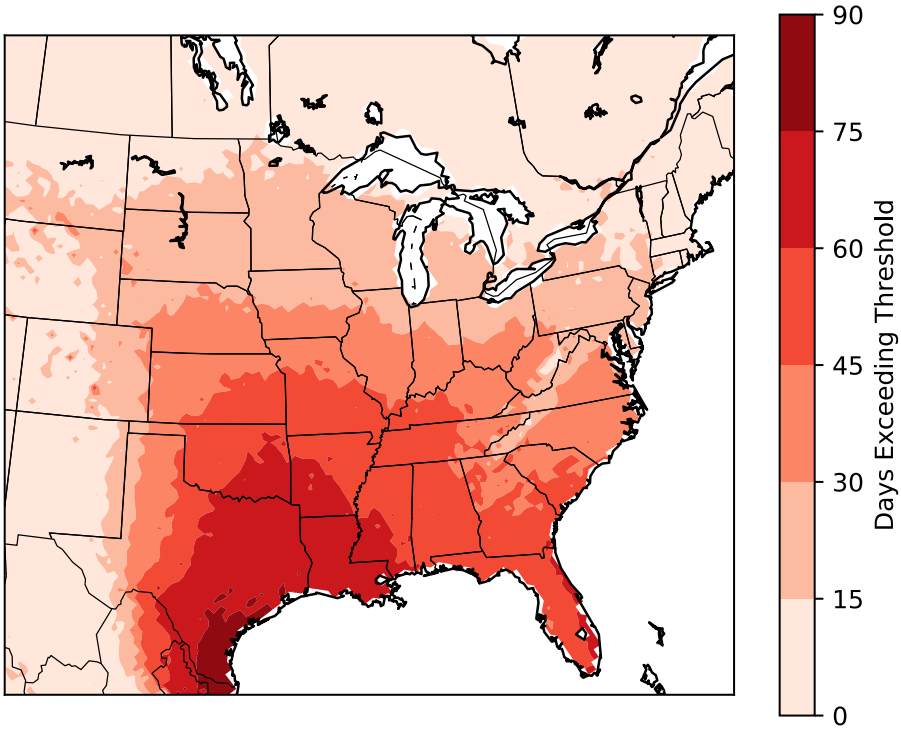

Figure 16: As in Fig.15 but for the product of 0-6km shear and MUCAPE exceeding 20000

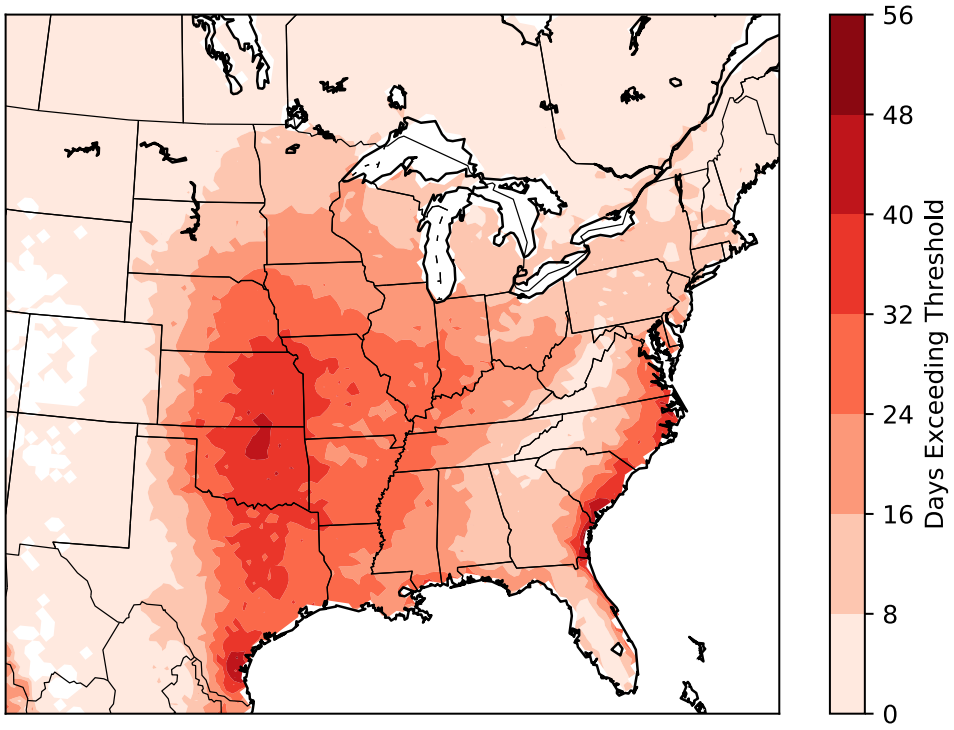

Figure 17: As in Fig.15 but for SCP exceeding 2.

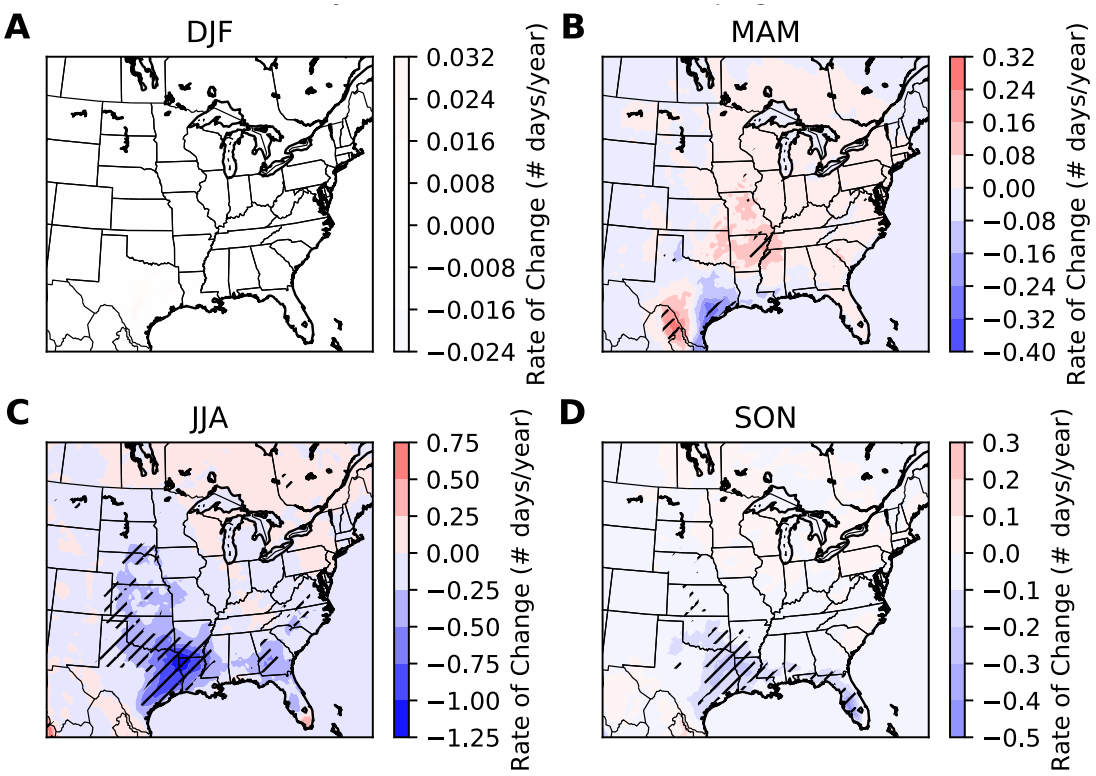

Figure 18: Linear trend in days with MLCAPE exceeding 2000 J/kg for (A) Winter (December, January, February), (B) Spring (March, April, May), (C) Summer (June, July, August), and (D) Fall (September, October, November) for 1986-2015. Only 00 UTC "pseudo-soundings" are considered. Hashing indicates statistically significant slopes with the p-value ≤ 0.05.

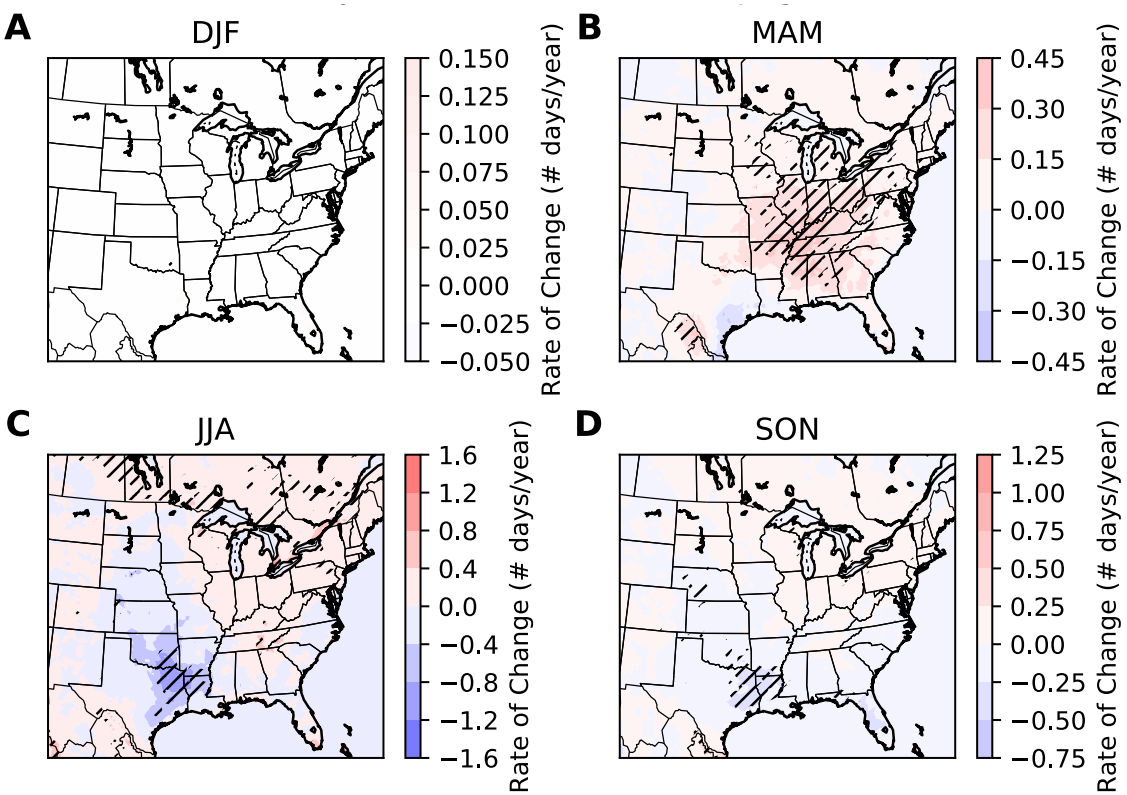

Figure 19: As in Fig. 18 but for MUCAPE exceeding 2000 J/kg.

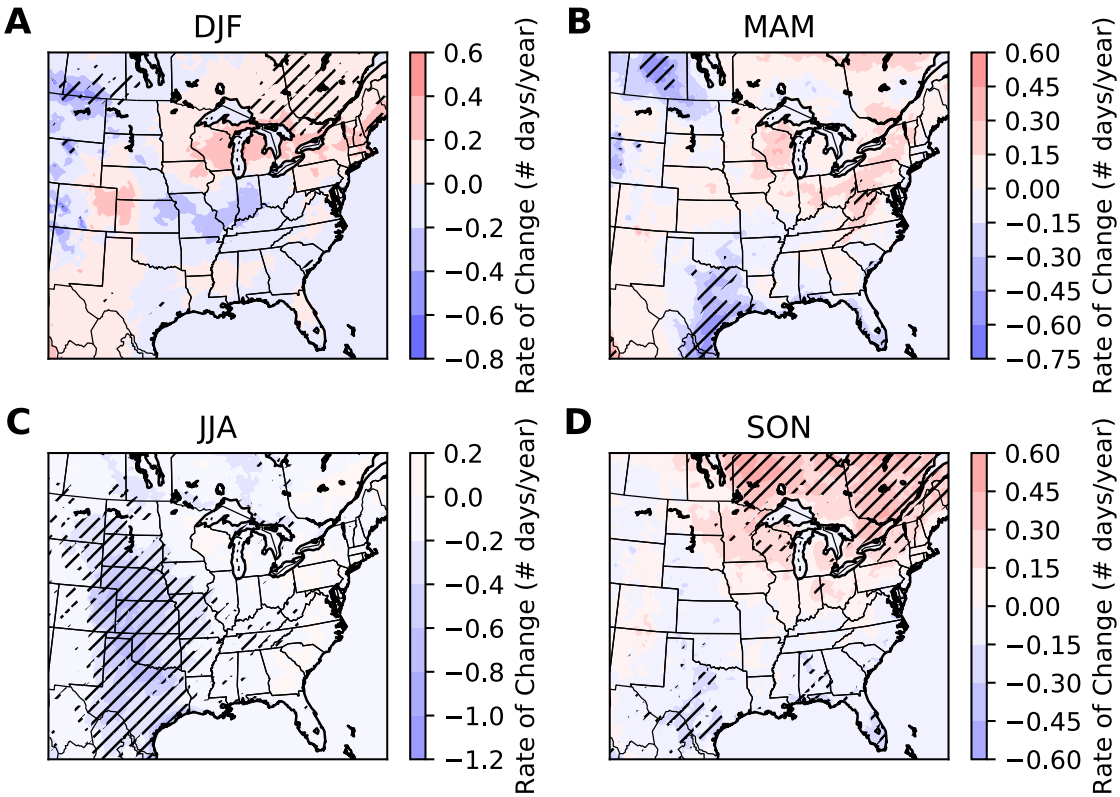

Figure 20: As in Fig. 18 but for MLCIN exceeding -75 J/kg.

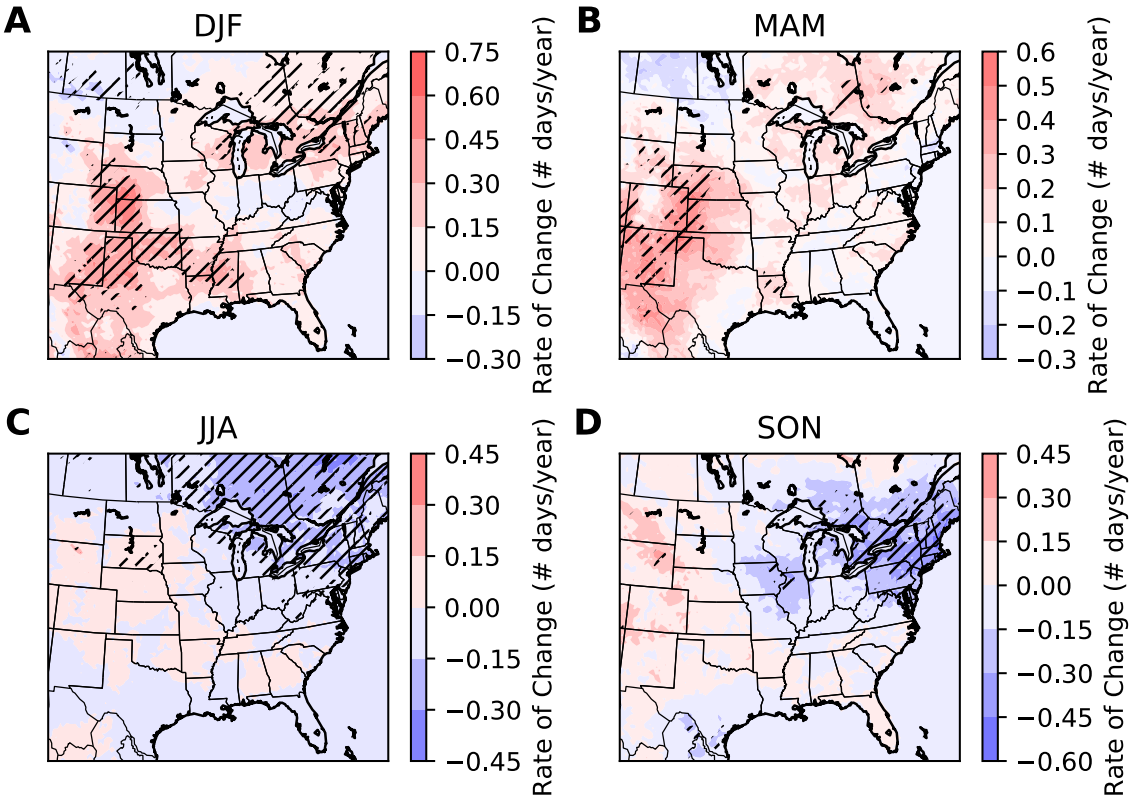

Figure 21: As in Fig. 18 but for 0-6 km shear exceeding 40 kts. Only days with nonzero MLCAPE are considered.

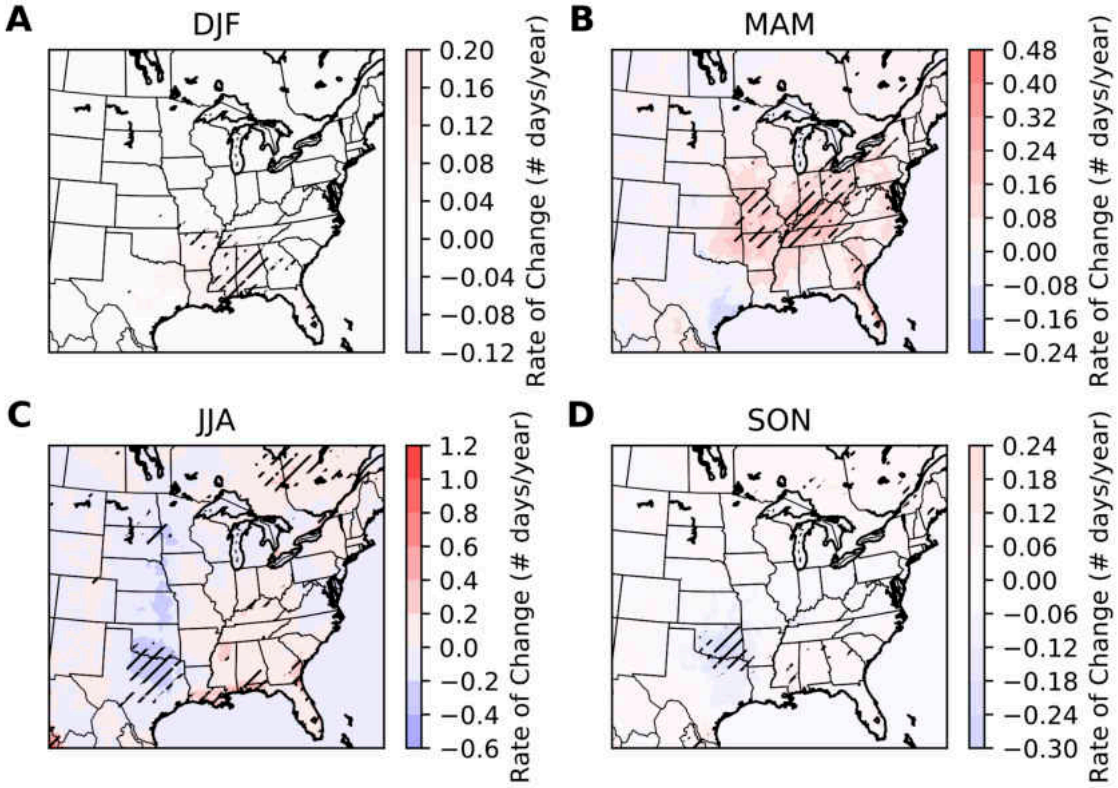

Figure 22: As in Fig. 18 but for SCP greater than 2

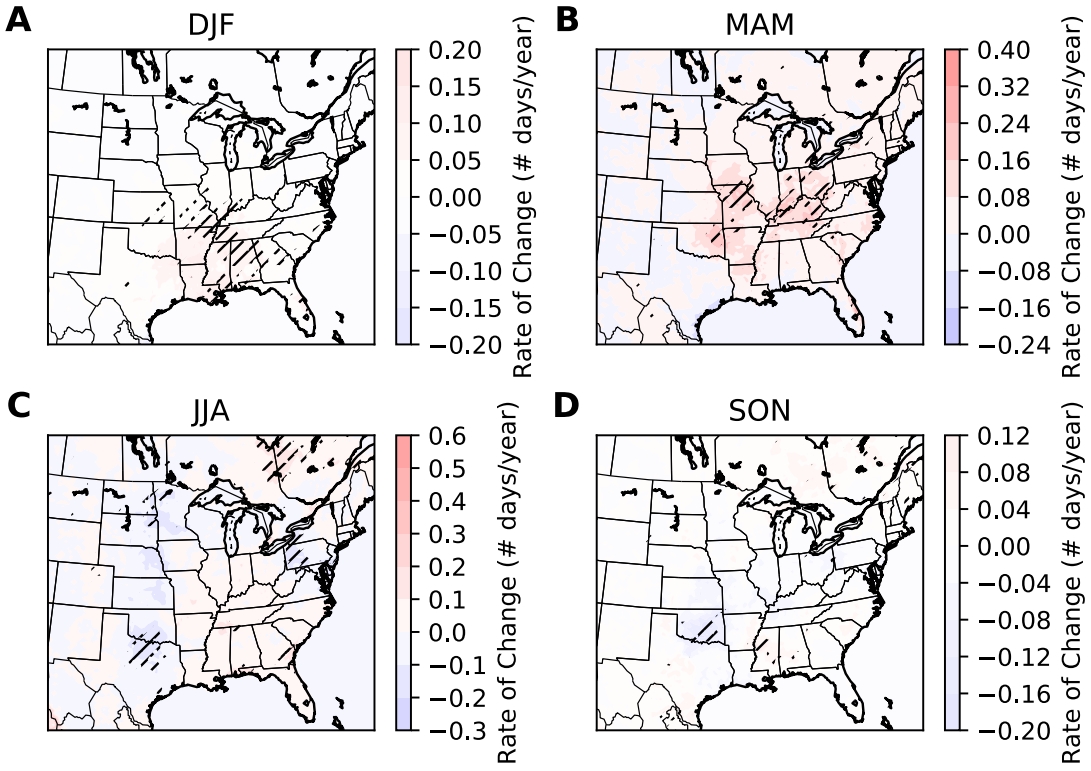

Figure 23: As in Fig. 18 but for EFFSRH greater than 100 m^2s^{-2}

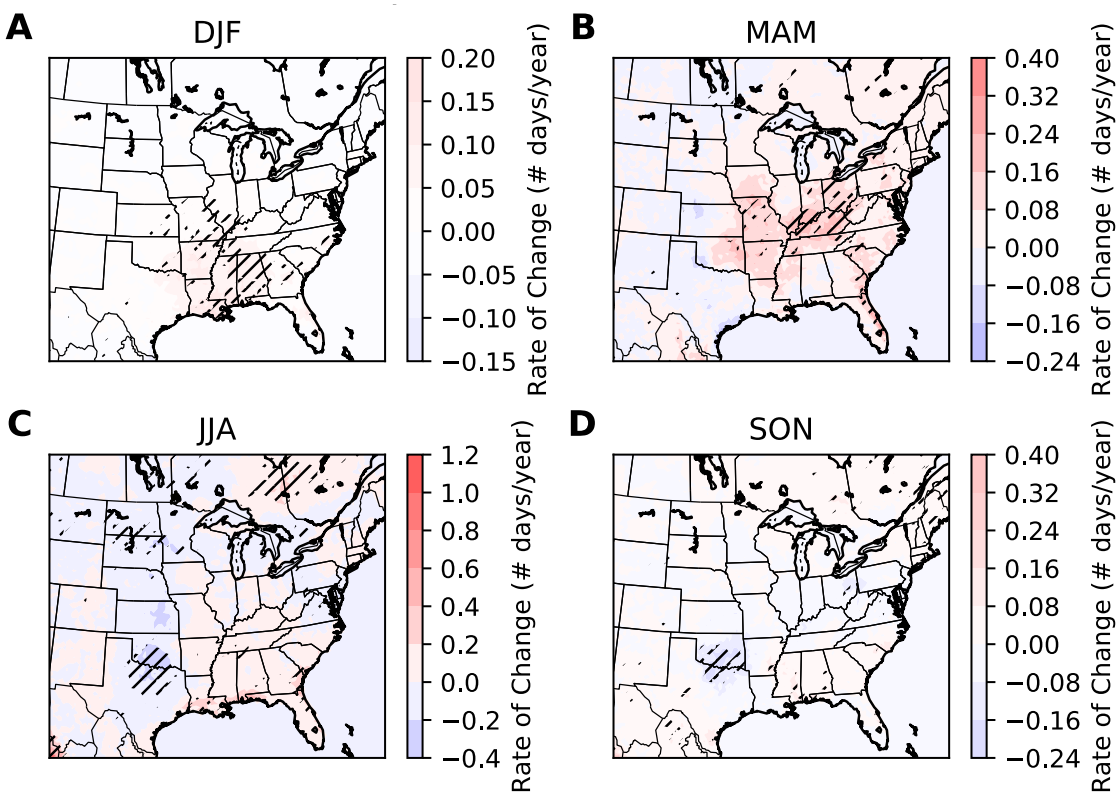

Figure 24: As in Fig. 18 but for EBWD greater than 15 ms^{-1}

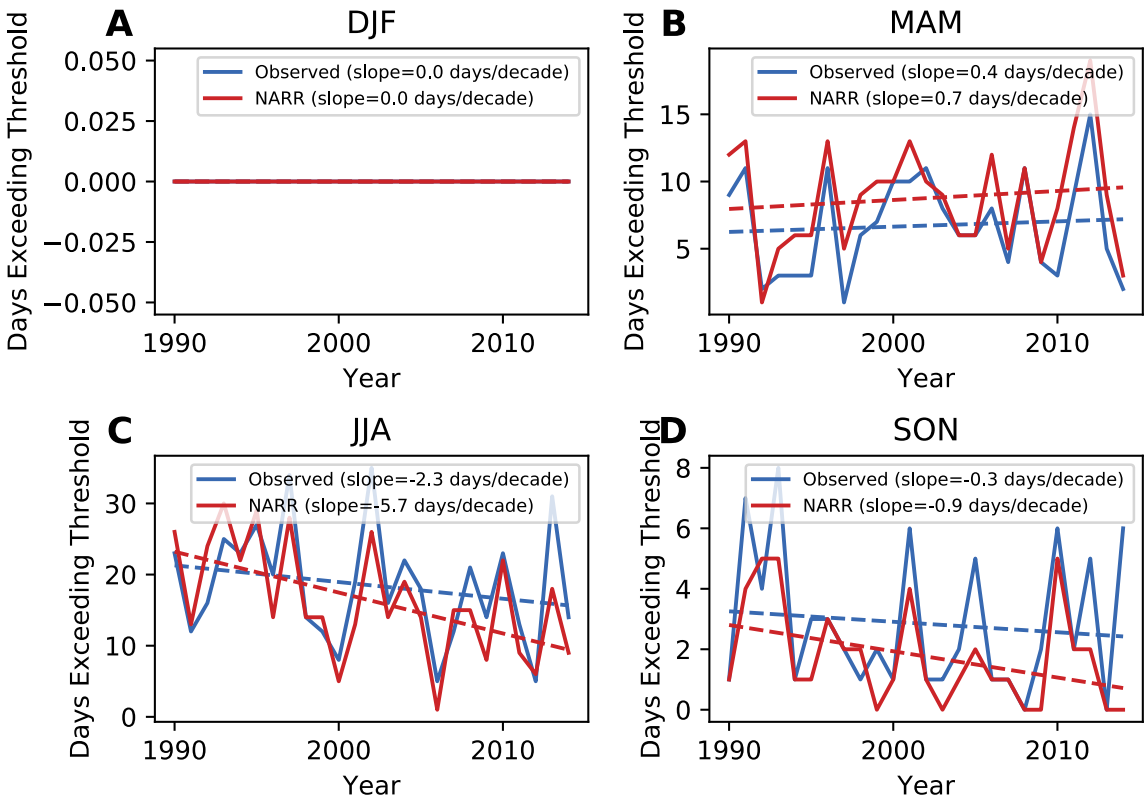

Figure 25: Linear trend in days with MLCAPE exceeding 2000 J/kg for both NARR and observations at OUN (Norman, OK) from 1990-2015.

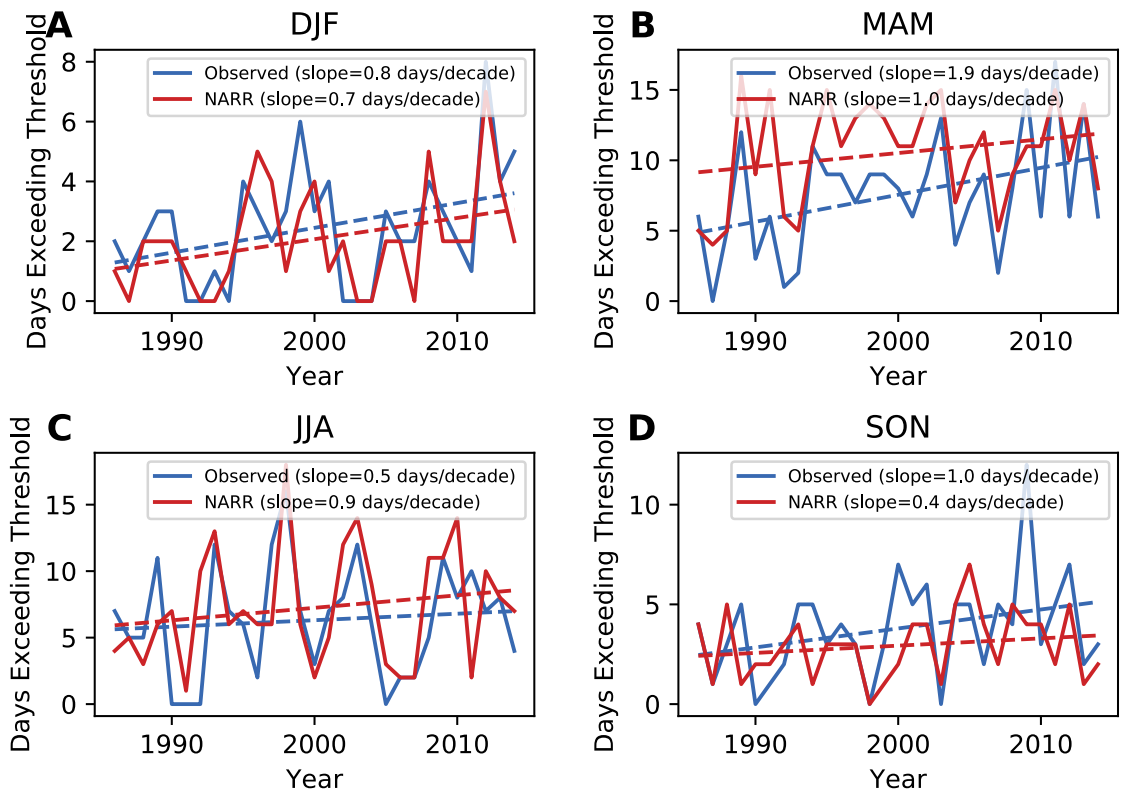

Figure 26: As in Fig. 25 but for SCP exceeding 2 at JAN (Jackson, MS).

CHAPTER 6

DISCUSSION

Reanalysis Performance and Comparison

Overall performance of each reanalysis is linked to the ability of each dataset to reproduce the thermodynamic environment. NARR and JRA55 had values of MLCAPE within the error bars of RUC-2 (< 500 J kg^{-1}, T03 and Coniglio 2012), while all other reanalyses had significant negative biases. Since more recently developed convective parameters, such as EFFSRH, SCP, and STP incorporate thermodynamic information, these parameters were suppressed for many of the reanalyses. By focusing on parameters that relied purely on the wind field, better separation was found between storm types (e.g. 0-1km bulk shear vs. EFFSRH).

Discrepancies in the thermodynamic environment may be explained by a number of factors. While the quantity of reanalyses and variety of causes precludes an extensive analysis herein, an attempt was made to investigate several possible causes: resolution; surface and mid-level thermodynamic biases; and convective contamination (which may in part cause said biases).

It is a reasonable assumption that the superior spatial and temporal grid spacing for NARR may have led to its improved representation of thermodynamic parameters. To test this, NARR was upscaled to a lower resolution by averaging grid-points and using only synoptic times (12Z, 18Z, 00Z, 06Z). No statistically significant differences were found for almost all of the parameters. This implies that resolution differences are not the primary

factor impacting the ability of a reanalysis to accurately portray the thermodynamic environment.

Because small variations in moisture and temperature can have a large impact on CAPE, biases in the thermodynamic properties were explored by generating composite soundings for the various storm types (not shown). Although composite hodographs were similar amongst the datasets, thermodynamic biases were evident for some of the reanalyses. To demonstrate this, box and whiskers plots of biases for surface properties and 400 mb temperatures are shown in Fig. 27. Many of the biases noted in the case studies are also apparent when considering the full dataset. For surface temperatures (Fig. 27a), notable reanalyses with biases exceeding the range of RUC-2 errors include NARR (+1.5°C) and 20CR (-2.9°C). Smaller biases are found for surface moisture. While all reanalyses have slightly negative surface mixing ratios compared to RUC-2, only CFSR has a median that falls outside errors quantified in T03 (Fig. 27b). Later work by Coniglio (2012) documented surface dewpoint biases for RUC-2 that were 1-2°C too moist. Median values for most reanalyses are < 1° C drier than RUC, which is an improvement from RUC-2 for surface moisture. The lone exception is CFSR, which is biased 2.5° C drier than RUC-2.

Mid-level temperature biases also impact calculations of CAPE for select reanalyses (Fig. 27c). Notably, MERRA 2 and 20CR have warm biases of 3°C and 1.5°C, respectively, which reduce values for the lifted index and CAPE. Given that these reanalyses are also too cool at the surface (Fig. 27b), this is most likely an impact of convection occurring. While a full exploration of this issue is beyond the scope of this paper, it is worth noting that both of these reanalyses use the Relaxed Arakawa-Schubert

(RAS, Moorthi and Suarez 1992) convective parameterization. Common in the climate modeling community, a by-product of this choice is a diurnal cycle of convection that is closely tied to diurnal heating and production of CAPE (Lee et al. 2007). For these reanalyses, CAPE effectively triggers convection earlier in the day which: a) reduces the CAPE (Fig. 9a) and b) introduces thermodynamic biases (Fig. 27c). In MERRA-2, for example, 70% of the cases had convective precipitation recorded within six hours prior to the proximity soundings.

In summary, the reliance of reanalyses on parameterizations leads to a variety of solutions, many of which have consequences regarding thermodynamic and kinematic parameters which are frequently used in convective studies. This was also noted in studies such as Gensini et al. (2014). This leads one to ponder what impact this may have on climatological studies that may use threshold values (e.g. MLCAPE $\geq$ 1000 J kg^{-1}) to diagnose properties, such as the frequency or trends in environments favorable for severe convection. Further complicating matters is the question of whether a user should calculate parameters such as CAPE or utilize the values provided by the reanalysis.

Consider, for example, a user of ERA-Interim; the results of this study demonstrate that MLCAPE is biased low by ~1000 J kg^{-1} when calculated independently with SHARPpy. As mentioned in Section 3, however, ERA-Interim uses an approximate form of CAPE that results in a ~20% positive bias (ECMWF, 2016). Utilizing the included values for CAPE in ERA-Interim results in median values 547, 193, and 113 J kg^{-1} higher than the values calculated in this study for mixed-layer, surface-based, and most-unstable parcels, respectively. In conclusion, there are a number of choices a user can make and it is not immediately clear what the impact may be for climatological studies. What should

be evident is that users *should* use a standard methodology so that results from different studies (reanalyses) can be directly compared.

Climatology and Trends in Severe Weather Parameters

It is important to validate the climatology portion of this thesis by comparing results to prior studies that have performed a similar analysis. Fig. 15a, which shows the number of days exceeding a threshold of 2000 J/kg of most unstable CAPE (MUCAPE), corresponds to similar figures such as Fig. 2 from GA11 and Fig. 6 from B03. Overall, the general spatial distribution of days exceeding the threshold value shows good agreement between the studies. The primary differentiating factor is the magnitude of days exceeding the threshold. B03 had a maximum around 12 days per year, whereas GA11 had a maximum over 40 (the actual maximum is unclear, as the scale is maxed out). Historical records of days with severe weather suggest that the numbers from B03 are too low. The large discrepancy between this study and the earlier work done by B03 can be partially explained by the course resolution of the R1 reanalysis used in that study (~200km). The discrepancy between GA11 and this study is potentially due to differences in time period and CAPE calculation methods. Interestingly, although Fig. 2 from GA11 is listed as MUCAPE, the overall distribution (and magnitude) have better geographical agreement with MLCAPE in this study (Fig. 15b).

Another comparison that can be made between studies is the composite C parameter which combines 0-6 BWD and MUCAPE (Fig. 16). Overall, this field resembles Fig. 4 from GA11. That said, the areas associated with the largest counts are wider and do not extend as far northward as they do in GA11.

The historical analysis of severe weather parameters/environments contained within this thesis meshes well with prior studies that have found certain parameters to have identifiable trends (e.g. Marsh et al. 2007, Gensini and Brooks 2008, Hoogewind et al. 2017). In general, days with favorable thermodynamic parameters may be increasing during spring months (Figs. 18b, 19b) but decreasing during the summer (Figs. 18c, 19c). The overall decrease in severe weather potential during summer months is also supported by a statistically significant decrease in days with favorable CIN during for the Plains regions (Fig. 20c). Other studies such as Hoogewind et al. (2017) have made similar claims regarding this trend in CIN over time. This decrease in favorable thermodynamic parameters (i.e. decrease in CAPE, increase in CIN) during the summer months, can likely be attributed to increasing surface temperatures and decreasing surface moisture over time due to a changing climate. Figures 28 and 29 show that these results are consistent with the trends in 2m temperature and dewpoint at OUN (Norman, Ok) during the summer months.

Days with parameters that take the effective layer into account (either directly or indirectly, e.g. Figs. 21-23) have shown a tendency to increase over primarily the Midwest during the spring and decrease in the Southern Plains during the summer. One potential reason for the increase in these parameters during the spring is an increase in the strength of the LLJ This ties into studies such as Tang et al. (2017) that have shown that the strength of the LLJ may be increasing with a changing climate. An increase in LLJ frequency or magnitude would theoretically have a large (positive) impact on parameters that rely heavily on the lowest layers of the atmosphere.

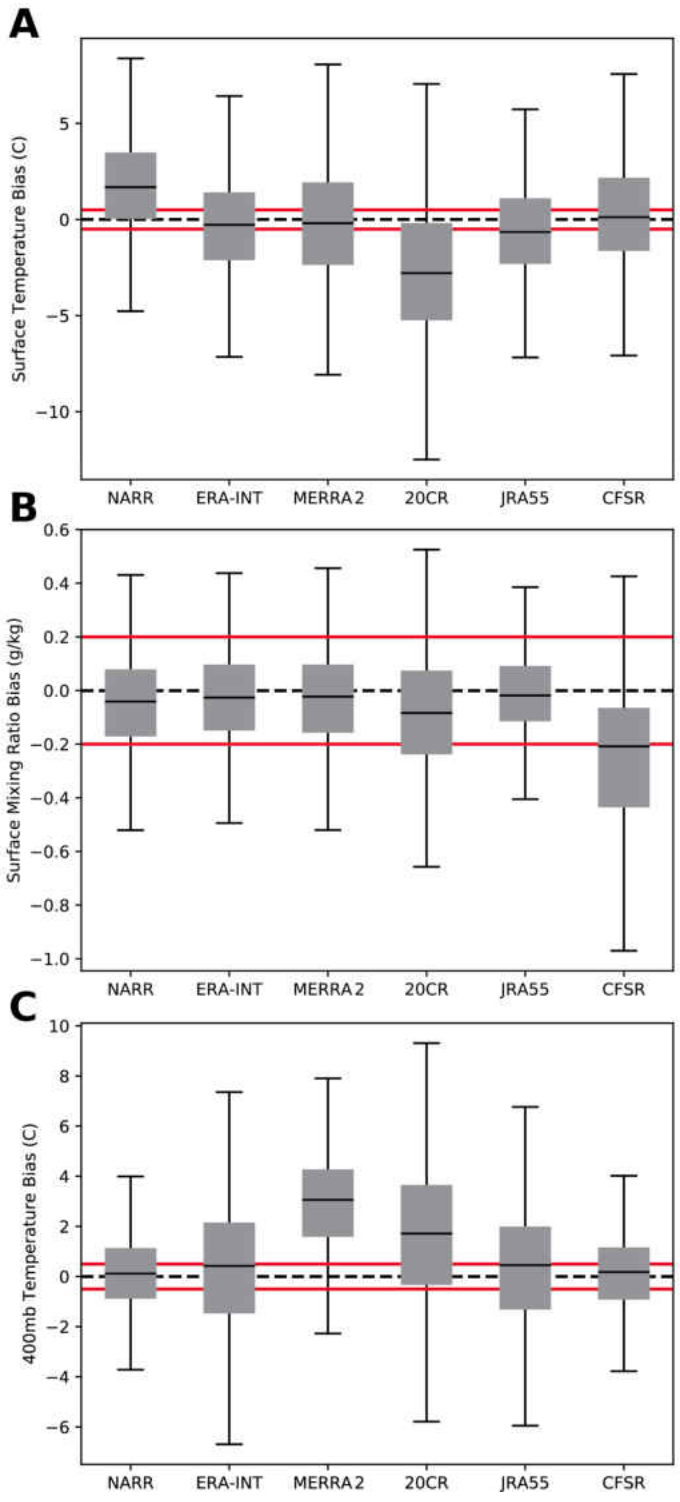

Figure 27. Box and whiskers plots of biases (reanalysis – RUC-2) for (A) surface temperatures, (B) surface mixing ratios, (C) and 400 mb temperatures.

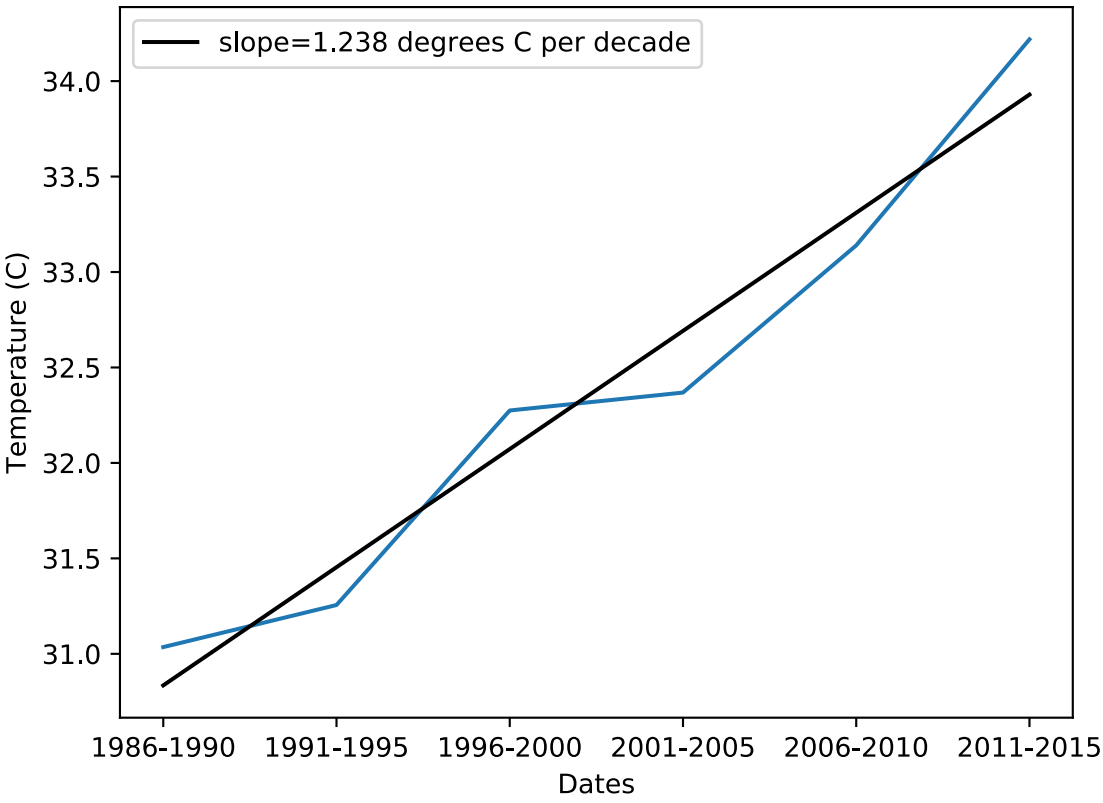

Figure 28. Trend in 2m temperature from NARR at OUN averaged over 5 year periods for Summer (June, July, August) months.

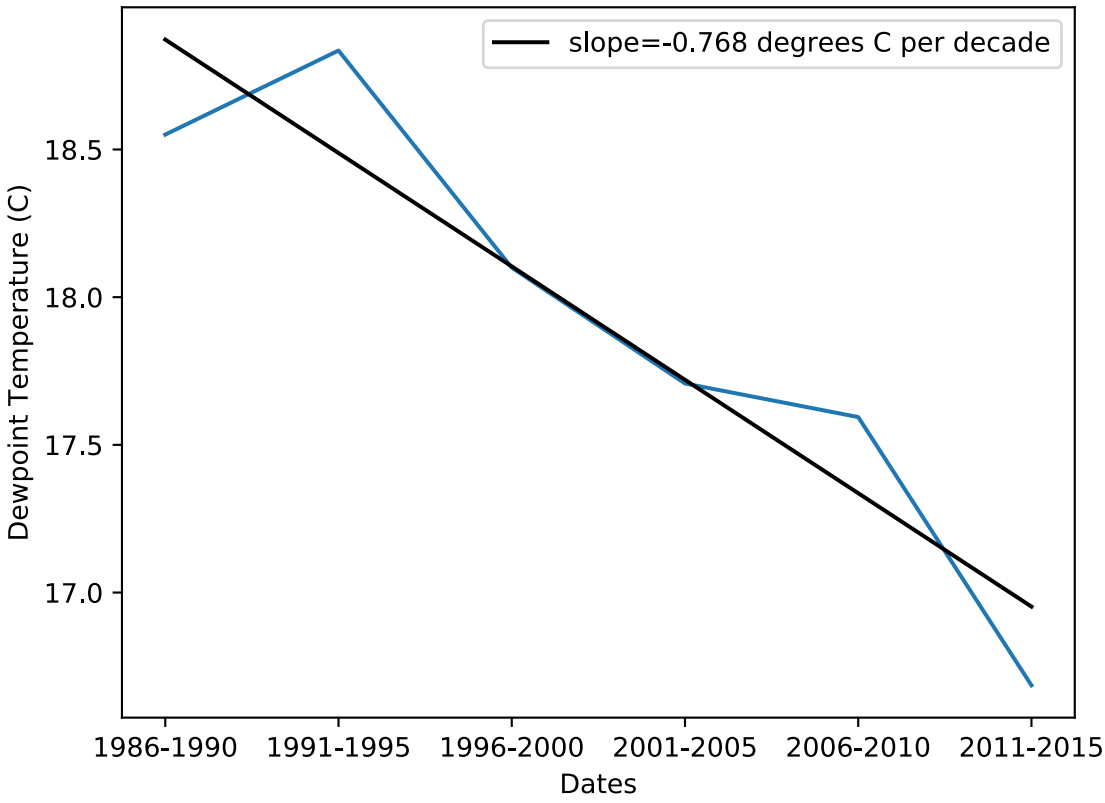

Figure 29. As in Fig. 28 but for 2m dewpoint temperature.

CHAPTER 7

SUMMARY AND CONCLUSIONS

Reanalysis Performance and Comparison

One of the primary objectives of this thesis was to compare a suite of modern reanalyses to the T07 dataset of RUC-2 proximity soundings. Besides investigating biases in commonly used convective indices, the study explored whether the reanalyses could reproduce the results of T03 and T07. To that end, a common sounding analysis package (SHARPpy) was used so that direct intercomparisons could be made. Provided below is a bulleted list of findings:

- Thermodynamic performance varied substantially across the reanalyses. Only NARR and JRA55 had CAPE values within the error bars of RUC-2. Although NARR and JRA55 were technically biased low for MLCAPE, these values may be more representable considering documented high biases for RUC-2.
- Biases in CAPE can be directly related to issues in the thermodynamic fields. Some reanalyses, such as MERRA 2 and 20CR, are biased low due to mid-level temperatures that are too warm, presumably due to convection contamination, while CFSR was impacted by negative biases for surface moisture.
- Thermodynamic performance negatively impacts reanalyses for composite and kinematic parameters that indirectly incorporate thermodynamic information via the effective layer. As a result, all reanalyses, except NARR and JRA55, were biased low for parameters such as ESRH, SCP, and STP.

- The majority of reanalyses reasonably reproduced the kinematic environment with nearly uniform correlations for the various parameters. The lone exception was 20CR, which had significant negative biases. Regardless of reanalysis, performance was better for fixed-level shear parameters vs. ESRH.
- The net result of negative biases led to more overlapped distributions when the analysis was segregated by storm type. While no reanalysis can exactly reproduce the work of T03 and T07, many of the reanalyses can broadly distinguish between environments that are significantly tornadic versus nontornadic. Better segregation is found for fixed-layer parameters versus those that use the effective layer.

No effort is made to objectively or subjectively rank reanalyses. As the results show, performance varies across these datasets. Broadly, JRA55 and NARR do the best job representing the thermodynamic environment, and this performance positively impacts results for composite parameters and kinematic properties that involve the effective layer. Despite this result, analysis segregated by storm type demonstrates the reanalyses' ability to broadly distinguish between significantly tornadic/severe environments and those that are unlikely to produce supercells.

The lone exception to the above statement is 20CR, which performed poorly across all categories. It should be stressed that for this reanalysis, the ensemble mean field was used creating a somewhat unfair comparison. While this decision was made on the premise that no a-priori knowledge of an event should be needed, it is conceivable that a user could arbitrary pick ensemble members that better match available observations.

The presented work offers itself as an initial intercomparison of reanalyses against a well-documented dataset of severe convective cases. Outstanding questions include how

representative these results are for other regions of the world that are less constrained by observations.

Climatology and Trends in Severe Weather Parameters

The second goal of this thesis was to investigate a reanalysis-based climatology of severe weather parameters and determine whether trends exist over the 30-year period from 1986-2015. It is again important to realize that the climatology and trends in these parameters does not necessarily represent the climatology or trends in actual severe weather occurrence, but rather the *potential* for severe weather occurrence. Severe weather parameters were calculated at each grid point, at synoptic times for the full domain of NARR excluding points over water and points where 2m temperature was below freezing. Provided below is a bulleted list of findings:

- Overall spatial distribution of severe weather climatologies closely matched what was found in previous studies such as Brooks et al. (2003) and Gensini and Ashley (2011). However, in the case of days with MUCAPE exceeding 2000 J/kg, the magnitude of the values in this study was higher than that of either study. The reasons for this are likely due to the increased resolution of NARR (compared to R1 used in Brooks et al. 2003) as well as differences in CAPE calculation and the different time periods used for each study.
- Trends in CAPE over this period tended to be positive for spring months and negative for summer months. The strongest negative trends exist for the Southern and Central Plains regions for MLCAPE during the summer.

- Days with favorable values of CIN are shown to be decreasing over most of the Southern, Central, and Northern Plains regions (i.e. traditional "Tornado Alley" regions). This decrease is also statistically significant for the majority of these regions.
- Trends in composite parameters such as SCP followed a similar pattern to CAPE with positive trends over the Midwest during the spring, and negative trends over the Plains regions during the Summer. This positive trend seem to be at least partially due to an increase in low level shear due to the inclusion of the effective layer.

Overall, trends were mixed, and a closer inspection of specific locations show large amounts of inter-annual variability. The exact reasons for the trends depicted by this thesis are beyond the scope of this study but will need to be further investigated before any definitive conclusions can be made. Additionally, more advanced methods to determine trends could be employed rather than just using a simple linear regression as was used in this study. Additional future work will include the production of severe weather datasets for multiple additional reanalyses. Code written in the Compute Unified Device Architecture (CUDA) platform is in development to enable processing of these additional reanalyses at a much faster rate by utilizing graphical processing units (GPUs). This will enable the creation of much larger severe weather parameter datasets (ex. global reanalyses such as JRA55 or upcoming ERA5), as well as the ability to utilize all available times (ex. all 8 daily times for NARR).

REFERENCES

Allen, J. T., and Karoly, D. J. 2013: A climatology of Australian severe thunderstorm environments 1979-2011: inter-annual variability and ENSO influence. *International Journal of Climatology*, **34**(1), 81–97. doi:10.1002/joc.3667

Benjamin, S.G., D. Dévényi, S.S. Weygandt, K.J. Brundage, J.M. Brown, G.A. Grell, D. Kim, B.E. Schwartz, T.G. Smirnova, T.L. Smith, and G.S. Manikin, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, doi: 10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

Blamey, R. C., Middleton, C., Lennard, C., & Reason, C. J. C. 2016: A climatology of potential severe convective environments across South Africa. *Climate Dynamics*. doi:10.1007/s00382-016-3434-7

Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L. Thompson, and J. A. Hart, 2017: SHARPpy: An Open Source Sounding Analysis Toolkit for the Atmospheric Sciences. *Bull. Amer. Meteor. Soc*. doi:10.1175/BAMS-D-15-00309.1, in press.

Brooks, H. E., A. R. Anderson, K. Riemann, I. Ebbers, and H. Flachs, 2007: Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. *Atmos. Res*., **83**, 294–305, doi: 10.1016/j.atmosres.2005.08.005.

Brooks, H.E., Doswell III, C.A., 2001: Some aspects of the international climatology of tornadoes by damage classification. Atmos. Res. 56, 191 – 201. https://doi.org/10.1175/WAF866.1

Brooks, H. E., Lee, J. W., Craven, J. P., 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.,* **67-68**, 73-94, doi: 10.1016/S0169-8095(03)00045-0

Burgess, D. W., and coauthors 2002: Radar Observations of the 3 May 1999 Oklahoma City Tornado. *Wea. Forecasting*, **17**, 456–471, doi: 10.1175/152 0434(2002)017<0456:ROOTMO>2.0.CO;2.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Q.J.R. Meteorol. Soc*., **137**, 1–28, doi:10.1002/qj.776

Coniglio, M. C., 2012: Verification of RUC 0–1-h forecasts and SPC mesoscale analyses using VORTEX2 soundings. *Wea. Forecasting*, **27**, 667–683

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. Q. J. R. Meteorol. Soc. 120: 1367–1388.

Courtier P, Andersson E, Heckley W, Pailleux J, Vasiljevic D, Hamrud M, Hollingsworth A, Rabier F, Fisher M. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. Q. J. R. Meteorol. Soc. 124: 1783–1807.

Davies, J. M., 1993: Small tornadic supercells in the central plains. Preprints, 17th Conf. on Severe Local Storms, St. Louis, MO, *Amer. Meteor. Soc.*, 305-309.

Davies, J. M., 2004: Estimations of CIN and LFC associated with tornadic and nontornadic supercells. *Wea. Forecasting*, **19**, 714-726.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q*.J.R. Meteorol. Soc.,* **137**, 553–597, doi:10.1002/qj.828

Doswell III, C.A., Brooks, H.E., Kay, M.P., 2005: Climatological Estimates of Daily Local Nontornadic Severe Thunderstorm Probability for the United States. Wea. Forecasting, 20, 577-595,

ECMWF, 2016: "New EFI parameters for forecasting severe convection" Newsletter 144, Summer 2015, 9pg.

Edwards, Roger and Thompson, Richard L. NWS/Storm Prediction Center, Norman, Oklahoma , 2009: Comments on "The North Dakota Tornadic Supercells of 18 July 2004: Issues Concerning High LCL Heights and Evapotranspiration". *Wea. Forecasting*, **24**, 1149

Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1

Gensini, V. A., and W. S. Ashley, 2011: Climatology of potentially severe convective environments from North American regional reanalysis. *Electronic J. Severe Storms Meteor* **6** (8), 1–40.

Hoogewind, K.A., Baldwin, M.E., Trapp, R.J., 2017: The Impact of Climate Change on Hazardous Convective Weather in the United States: Insight from High-Resolution Dynamical Downscaling. J. Climate, 30, 10081-10100, https://doi.org/10.1175/JCLI-D-16-0885.1

Hong, SY. and Kanamitsu, M., 2014: Dynamical downscaling: Fundamental issues from an NWP point of view and recommendations, *Asia-Pacific J Atmos* **50**, doi: 10.1007/s13143-014-0029-2

Johns, R. H., J. M. Davies, and P. M. Leftwich, 1993: Some wind and instability parameters associated with strong and violent tornadoes. Part II: Variations in the combinations of wind and instability parameters. The Tornado: Its Structure, Dynamics, Hazards, and Prediction*, Geophys. Monogr*., **79**, Amer. Geophys. Union, 583–590.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kellenbenz, David J. and Grafenauer, Thomas J. NOAA/National Weather Service/Weather Forecast Office, Eastern North Dakota, Grand Forks, North Dakota, Davies, Jonathan M., Wichita, Kansas , 2007: The North Dakota Tornadic Supercells of 18 July 2004: Issues Concerning High LCL Heights and Evapotranspiration. *Wea. Forecasting*, **22**, 1200–1213, doi: 10.1175/2007WAF2006109.1.

Kennedy, A., X. Dong, B. Xi, S. Xie, Y. Zhang, and J. Chen, 2011: A Comparison of MERRA and NARR Reanalyses with the DOE ARM SGP Data. *J. Climate*, **24**, 4541–4557, doi: 10.1175/2011JCLI3978.1.

Kobayashi, S, and Coauthors, 2015: The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II,* **93**, 5-48, doi: 10.2151/jmsj.2015-001

Markowski, Paul M., Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania , 2002: Mobile Mesonet Observations on 3 May 1999. *Wea. Forecasting*, **17**, 430–444, doi: 10.1175/1520-0434(2002)017<0430:MMOOM>2.0.CO;2.

Marsh, P. T., H. E. Brooks, and D. J. Karoly, 2007: Assessment of the severe weather environment in North America simulated by a global climate model. *Atmos. Sci. Letters*, **8**, DOI: 10.1002/asl.159.

Marsh, P. T., Brooks, H. E., & Karoly, D. J. (2009). Preliminary investigation into the severe thunderstorm environment of Europe simulated by the Community Climate System Model 3. Atmospheric Research, **93**(1-3), 607–618. doi:10.1016/j.atmosres.2008.09.014

Mesinger, F., and Coauthors, 2006: North Blamey, R. C., Middleton, C., Lennard, C., & Reason, C. J. C. 2016: A climatology of potential severe convective environments across South Africa. *Climate Dynamics*. doi:10.1007/s00382-016-3434-7

Moorthi, S., and M. Suarez, 1992: Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev*, **120**, 978–1002.

Rasmussen, E. N., and Blanchard, D. O., 1998: A Baseline Climatology of Sounding-Derived Supercell and Tornado Forecast Parameters. *Wea. Forecasting*, **13**(4), 1148–1164. doi:10.1175/1520-0434(1998)013<1148:abcosd>2.0.co;2

Roebber, Paul J., Atmospheric Science Group, Department of Mathematical Sciences, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin Schultz, David M. and Romero, Romualdo, NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma , 2002: Synoptic Regulation of the 3 May 1999 Tornado Outbreak. *Wea. Forecasting*, **17**, 399–429, doi: 10.1175/1520-0434(2002)017<0399:SROTMT>2.0.CO;2.

Romero, R., M. Gayà, and C. A. Doswell III, 2007: European climatology of severe convective storm environmental parameters: A test for significant tornado events. *Atmos. Res*., **83**, 389–404, doi:10.1016/j.atmosres.2005.06.011.

Saha, S, and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi: 10.1175/2010BAMS3001.1.

Sherburn, K., M. Parker, J. King, and G. Lackmann, 2016: Composite Environments of Severe and Nonsevere High-Shear, Low-CAPE Convective Events. *Wea. Forecasting*, **31**, 1899–1927, doi: 10.1175/WAF-D-16-0086.1.

Stensrud, David J., NOAA/National Severe Storms Laboratory, Norman, Oklahoma Weiss, Steven J., NOAA/NWS/NCEP/Storm Prediction Center, Norman, Oklahoma , 2002: Mesoscale Model Ensemble Forecasts of the 3 May 1999 Tornado Outbreak. *Wea. Forecasting*, **17**, 526–543, doi: 10.1175/1520-0434(2002)017<0526:MMEFOT>2.0.CO;2.

Tang, Y., Winkler, J., Zhong, S., Bian, X., Doubler, D., Yu, L., and Walters, C., 2017: Future changes in the climatology of the Great Plains low-level jet derived from fine resolution multi-model simulations. *Scientific Reports,* **7**, doi: 10.1038/s41598-017-05135-0

Thompson, R. L. and Edwards, R., Storm Prediction Center, Norman, Oklahoma , 2000: An Overview of Environmental Conditions and Forecast Implications of the 3 May 1999 Tornado Outbreak. *Wea. Forecasting*, **15**, 682–699, doi: 10.1175/1520-0434(2000)015<0682:AOOECA>2.0.CO;2.

Thompson, R. L., Edwards, R., Hart, J. A., Elmore, K. L., and Markowski, P., 2003: Close Proximity Soundings within Supercell Environments Obtained from the Rapid Update Cycle., *Wea. Forecasting*, **18**(6), 1243–1261. doi:10.1175/1520-0434(2003)018<1243:cpswse>2.0.co;2

Thompson, R. L., C. M. Mead, and Edwards R., 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102-115.

Trapp, R. J., Diffenbaugh, N. S., Brooks, H. E., Baldwin, M. E., Robinson, E. D., & Pal, J. S. (2007). Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proceedings of the National Academy of Sciences,* **104**(50), 19719–19723. doi:10.1073/pnas.070549410

Veersé F, Thépaut J-N. 1998. Multiple-truncation incremental approach for four-dimensional variational data assimilation. Q. J. R. Meteorol. Soc. 124: 1889–1908.

Verbout, S.M., Brooks, H.E., Leslie, L.M., Schultz, D.M., 2006: Evolution of the U.S. Tornado Database: 1954–2003. Wea. Forecasting, 22, 86-93, https://doi.org/10.1175/WAF910.1