AUTOMATICALLY DERIVED RADAR ATTRIBUTES OF TROPICAL CYCLONE TORNADOES FROM 2013-2020

A Thesis

by

JUSTIN RELTON SPOTTS

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

MASTER OF SCIENCE

Chair of Committee,	Christopher Nowotarski
Committee Members,	Courtney Schumacher
	Matthias Katzfuss
Head of Department,	Ramalingam Saravanan

May 2023

Major Subject: Atmospheric Sciences

ABSTRACT

Tropical cyclone tornadoes (TCTORs) pose a warning challenge and are prone to having a high number of false-alarm warnings. Previous studies have found weather-radar attributes (i.e., rotational velocity (V_{rot}) or azimuthal shear (AzShear)) show some discriminating power between tornadic and nontornadic, but warned events. Given the time-consuming nature of a manual analysis, these previous studies often use a limited sample size. This work builds upon previous studies by developing and performing an automated method for determining radar attributes of 334 tornadic and 721 nontornadic initial events within 29 tropical-cyclone environments from 2013 to 2020. Rotation below 10 kft above radar level, vertically integrated liquid, and echo tops are examined to determine if these attributes can discriminate between tornadic and nontornadic events using an automated method. Results show that rotation, particularly rotation at lower elevation angles, is generally stronger in tornadic than nontornadic events whereas vertically integrated liquid and echo tops do not show much discriminating power. Warning statistics at 10 minutes before the tornado start time or nontornadic maximum rotation at an elevation angle of 0.5° show a best AzShear discriminator of about ~$1 \times 10^{-2}\,\mathrm{s}^{-1}$ within 40 n mi. (74.1 km) of the nearest radar and ~$8 \times 10^{-3}\,\mathrm{s}^{-1}$ beyond 40 n mi. Warning only for potentially more impactful events yields a higher probability of detection at higher rotation thresholds and the maximum Critical Success Index is shifted to higher thresholds as well, although with lower values.

ACKNOWLEDGMENTS

There are many without whom this project would not be possible. First, I would like to thank my parents. I would not have gotten to this point without their support. Second, I would also like to thank my advisor, Dr. Nowotarski, for his mentorship and the opportunity to work on this project. Thank you to my committee members: Dr. Christopher Nowotarski, Dr. Courtney Schumacher, and Dr. Matthias Katzfuss for their support and suggestions for this project. To my fellow grad students for their support and listening ear, thank you. I would like to especially thank Leland MacDonald for her friendship, support, and listening to me ramble.

There are also a number of collaborators without whom this project would not be possible. I would like to thank Lance Wood and Dan Reilly (NWS HGX) for helping make this project possible. Scott Overpeck (NWS ABQ) his guidance and work for this project and my undergraduate work. Roger Edwards (SPC) for his mentorship, suggestions, and previous work in TCTORs including the TCTOR database for both this project and my undergraduate work as well. I would also like to thank Jared Klein and Kevin Gilmore (NWS LIX) for sharing their TCTOR experience with Hurricane Ida (2021).

Peyton Stevenson (UNL), Shane Lewis (TAMU), and Daniel Alverez deserve thanks for their work cataloging tornado warnings and QCing the TCTOR database. Dr. Don Conlee for his teaching and mentorship during my undergraduate years. Kiel Ortega (CIWRO) and Brandon Smith (CIWRO/NSSL) for their assistance with the WDSS-II algorithms and John Cintineo for answering my questions on Prob. Severe. Lastly, I would like to thank those that make the ADRAD's operation possible.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of advisor Dr. Christopher Nowotarski and Dr. Courtney Schumacher from the Department of Atmospheric Sciences and Dr. Matthias Katzfuss from the Department of Statistics.

The TCTOR database was compiled by Roger Edwards. Tornado warnings and warning classifications were cataloged by Peyton Stevenson and TCTOR reports were quality controlled by Peyton Stevenson and Shane Lewis. Some cataloging was also performed by Daniel Alvarez and Justin Spotts. The manual radar analysis for the tornadic and some nontornadic cases during Hurricane Harvey (2017) was performed by Scott Overpeck. Algorithms used to process radar and model data were included with WDSS-II (`wdssii.org`).

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

Funding Sources

This project was supported under funding NOAA/NWS CSTAR grant number NA19NWS4680007.

NOMENCLATURE

AGL	Above Ground Level
AzShear	Azimuthal Shear
BTRT	Best Track Real Time
BWD	Bulk Wind Difference
CSI	Critical Success Index
FAR	False Alarm Ratio
IS	Interest Score
MSL	Mean Sea Level
POD	Probability of Detection
RP	Reference Point
RT	Reference Time
Sfc	Surface
TCTOR	Tropical Cyclone Tornado
VIL	Vertically Integrated Liquid
WDSS-II	Warning Decision Support System - Integrated Information
V_r	Radial Velocity
V_{rot}	Rotational Velocity

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Tropical cyclones (TCs) produce many hazards including flooding, intense wind, storm surge, and tornadoes. Of these hazards, tropical cyclone tornadoes (TCTORs) tend to occur in the outer rain bands of the TCs (Edwards 2012). The typically shallow and weak nature of their parent convective cells, often supercells, contributes to an increased difficulty in the radar detection of TCTORs (McCaul and Weisman 1996). TCs with at least one tornado and ten tornado warnings yielded an average false alarm ratio (FAR) of 0.858 during the 2008 to 2013 period (Martinaitis 2017). More recently, Hurricane Harvey's (2017) FAR was approximately 0.84 (Nowotarski et al. 2021). This statistic becomes particularly important when both the TC hazards of flooding and tornadoes impact the same area resulting in overlapping warnings known as Tornado-Flash Flood warnings (TORFFs; Nielsen et al. 2015; Burrow et al. 2021).

As Burrow et al. (2021) and others have pointed out, the dilemma produced by a TORFF is that the safety messaging for tornadoes ("go to your basement, safe room, or an interior room"; NWS 2022b) conflicts with the safety messaging for flooding ("Get to higher ground"; NWS 2022a) resulting in the potential for confusion when warnings for both phenomena are issued simultaneously (Nielsen et al. 2015; Burrow et al. 2021). Burrow et al. (2021) identified 619 TORFFs for 32 TCs over the 2008 to 2018 period. Of those 619 TORFFs, 209 were found during Hurricane Harvey (2017) and 113 were found during Hurricane Gustav (2008). In addition to the confusion caused by messaging during TORFFs, increased issuance of false-alarm warnings could result in a decreased likelihood of people taking shelter (Trainor et al. 2015), although there is still some ambiguity in the literature around whether this finding is universally applicable (e.g., Lim et al. 2019).

Given the high FAR for TCTORs, the impact of TORFFs, and the potential impact of false-alarm warnings on the public, further research efforts to reduce FARs for TCTOR warnings is

warranted. Previous studies have performed manual analyses of radar data for several TCs to improve discrimination between TCTORs and potential false-alarm warnings (e.g., Spratt et al. 1997; McCaul et al. 2004; Schneider and Sharp 2007; Martinaitis 2017; Nowotarski et al. 2021). However, these studies use a relatively limited sample size, likely in part due to the time-consuming nature of a manual analysis. Therefore, this project creates a climatology of automatically derived radar attributes of TCTORs and false-alarm (null) events for U.S. landfalling TCs from 2013, when polarimetric capabilities were added to the WSR-88D radars, to 2020. Areas of interest within azimuthal shear (AzShear; Smith and Elmore 2004; Mahalik et al. 2019), 20-dBZ echo tops (Cifelli et al. 2007; Lakshmanan et al. 2013a; Edwards and Picca 2016), and vertically integrated liquid (VIL; Greene and Clark 1972) fields are determined. Attributes from these fields as well as spectrum width and the presence of a tornadic-debris signature (TDS; Ryzhkov et al. 2005; Edwards and Picca 2016) are recorded. After comparing an estimate of V_{rot} from AzShear and echo tops to the manual analysis of Nowotarski et al. (2021), these attributes and their temporal trends are used to determine if there are any differences between tornadic and nontornadic (but warned), TC rainband convection from 2013 to 2020.

1.2 Background

TCTORs are primarily found in the northeast (north-relative framework) or front-right (TC motion relative framework) quadrant of a TC with the distribution of TCTORs tending to shift clockwise about the TC with decreasing TC intensity (McCaul 1991; Schultz and Cecil 2009; Edwards 2012; Edwards and Mosier 2022). Furthermore, most TCTORs occur 100 to 500 km from the TC center (Edwards 2012). An analysis of the spatial distribution of TCTORs in Hurricane Harvey by Nowotarski et al. (2021) was consistent with this pattern (Fig. 1.1). A composite sounding analysis by McCaul (1991) shows increased low-level wind shear and larger convective available potential energy (CAPE) values in the front-right quadrant. In comparison to other quadrants of the TC, the front-right quadrant had the most favorable wind shear and third-most favorable CAPE. Their study also found that vertical wind shear was larger and CAPE was smaller within 300 km of the TC center, but the differences were not statistically significant. More recent work by Schenkel

et al. (2020) showed an enhanced secondary circulation, increased vertical wind shear, more favorable 0-1 and 0-3 km cell-relative helicity, and larger 0-3 km CAPE in the downshear-right quadrant of strongly-sheared TCs. Nowotarski et al. (2021) showed decreasing storm-relative helicity and increasing CAPE values with distance from Hurricane Harvey suggesting a favorable region for tornadoes within 100-300 km of the TC, consistent with previous studies.

The TCTOR-spawning convection within outer rainbands of TCs can take on a variety of convective modes. Edwards et al. (2012) conducted an analysis of 730 TCTOR events, finding that supercells were the most common storm mode (79% of events) with 34% of events consisting of discrete, right-moving supercells and 45% of events consisting of supercells embedded in a quasi-linear convective system (QLCS) or cluster (QLCS with an aspect ratio less than 3/1 representing disorganized convection). Moreover, 9% of cases were marginal supercells. Numerical simulations show that TC supercells tend to be shallower than their Great Plains counterparts, with weaker near-surface rotation despite comparable low-level updraft strengths (McCaul and Weisman 1996).

One tool that has been available to operational meteorologists and researchers for tornado detection is Doppler weather radar (Brown et al. 1978). From the Doppler radar, the rotational velocity (V_{rot}) of a circulation is calculated by halving the sum of the absolute values of the maximum local outbound and inbound radial velocities (V_r; Smith et al. 2015). Smith et al. (2015) revealed that as the 0.5° elevation angle (tilt) V_{rot} increased, the probability that a tornado exceeds a given Enhanced Fujita (EF) scale rating increases, conditional on the formation of a tornado. Furthermore, dividing the difference of the maximum and minimum V_r by the distance between the peaks yields the gradient of velocity, or shear (Smith and Elmore 2004). This peak-to-peak method of calculating shear can suffer from radar sampling limitations, however (Smith and Elmore 2004; Mahalik et al. 2019). A more robust, yet more complicated, way to calculate the shear is to first smooth the V_r field with a median filter, then fit the smoothed values to a linear model using local least squares regression (LLSD; Smith and Elmore 2004; Mahalik et al. 2019). This model is then used to solve for the derivatives of V_r across the kernel that yield the smallest residuals. The

component of the gradient in the azimuthal direction is referred to as azimuthal shear (AzShear).

In addition to reflectivity and velocity data, modern WSR-88D radars retrieve phase information from the horizontal and vertical components of the returned signal creating polarimetric variables (see Kumjian 2013). These variables include differential-reflectivity (Z_{DR}), which measures the logarithm ratio of the horizontal and vertical returned powers whereas the specific differential phase (K_{DP}) measures the half-range derivative of the difference in phase between the two polarizations. Previous studies have shown that spatial separation between these regions can be indicative of tornadic potential (Romine et al. 2008; Crowe et al. 2010; Loeffler et al. 2020). These regions, known as the Z_{DR} arc and K_{DP} foot, form as different sized hydrometeors are advected different distances by the storm-relative winds. Specifically, the orientation of the vector between the centroids of the two regions were shown to be closer to orthogonal to the storm's motion in tornadic cases when compared to nontornadic cases (Loeffler et al. 2020).

Given the differences between TC supercells and their mid-latitude counterparts, several studies have examined Doppler weather radar signatures of TCTORs and their parent mesocyclones via manual analysis over a limited number of events (e.g., Spratt et al. 1997; McCaul et al. 2004; Schneider and Sharp 2007; Martinaitis 2017; Nowotarski et al. 2021). These studies support concerns raised by McCaul and Weisman (1996) in that the shallow, weak nature of the typical TC supercell and its rotation may complicate radar detection of these features. Several studies found V_{rot} and shear thresholds that improve discrimination between tornadic and nontornadic cells (e.g., Schneider and Sharp 2007; Martinaitis 2017; Nowotarski et al. 2021).

Spratt et al. (1997) examined four TC supercells using level IV radar data during Tropical Storm Gordon (1994) and Hurricane Allison (1995). Their analysis found that the mesocyclones' circulation diameters decreased and rotational shear increased prior to tornadogenesis. The magnitude of both reflectivity and rotational signatures decreased with distance from the radar, but velocity spectrum width was shown to be useful for tracking potential tornadic cells, even at farther ranges. During these TCTOR events, the operational automated mesocyclone detection algorithm frequently failed to detect the mesocyclone, demonstrating the difficulty TC mesocyclones can

cause for automated detection.

McCaul et al. (2004) performed a time-height analysis of three TC supercells during Tropical Storm Beryl (1994). Specifically, they examined the rotational velocity, rotational shear, mesocyclone diameter, 30-dBZ echo tops, and angular momentum per unit mass of the three cells. Their findings show values of V_{rot} were lower in TC mini-supercells than their mid-latitude counterparts and tended to increase during tornadogenesis. However, if V_{rot} was the only variable used, the false alarm ratio (FAR) would have increased. Additionally, they found the diameters of TC mesocyclones were smaller than their Great Plains counterparts with diameters from 2-4 km (1-2 n mi.) and modal echo tops of 8-11 km (26-36 kft), occasionally reaching heights of 13-14 km (43-46 kft). The authors note they limited their sample size for radar data to three cells because of sampling limitations and the work necessary to perform an analysis on all thirteen cells in their study was too great.

Schneider and Sharp (2007) performed an analysis of 20 TC convective cells, 12 of which were tornadic, during TCs Frances, Gaston, Ivan, and Jeanne in 2004. Using data from the Raleigh-Durham, North Carolina WSR-88D radar (KRAX), they determined three radar signatures that could precede TCTORs: a rotational velocity of at least 20 knots (10.3 $\mathrm{m\,s^{-1}}$), a supercellular reflectivity signature such as an appendage or hook echo, or a velocity enhancement signature [VES; local enhancement of radial velocity between 7 and 12 kft (2.1 and 3.7 km) AGL] of at least 30 knots (15.4 $\mathrm{m\,s^{-1}}$).

Martinaitis (2017) analyzed 76 tornado-warned circulations in Tropical Storm Debby (2012) and 48 in Tropical Storm Andrea (2013), where they examined velocity-based attribute trends and K_{DP}/Z_{DR} separations. Their results indicate that the rate at which V_{rot} and shear increases over time was significantly larger in the two volume scans prior to the peak value or time of tornadogenesis, but only for events within 74.1 km (40 n mi.) of the nearest radar. Their subjective analysis of K_{DP} and Z_{DR} separation found that 70.6% of tornadic and 55.5% of nontornadic storms showed horizontal separation at the 0.5° elevation angle. While a significant number of nontornadic cases showed separation, the separation in the tornadic cases was more distinct. This

study suggested warning criteria using V_{rot} and other radar signatures. The V_{rot} portion of the criteria includes a 0.5° V_{rot} of at least 20 knots (10.3 $\mathrm{m\,s^{-1}}$) at ranges less than 74.1 km and a threshold of at least 15 knots (7.7 $\mathrm{m\,s^{-1}}$) at ranges greater than 74.1 km.

Nowotarski et al. (2021) examined warning skill, near-cell environments, and radar attributes between TCTOR producing convection (collectively named ALL TOR) for all of Hurricane Harvey (2017) and convection that resulted in a false-alarm tornado warning (NON TOR) for the Houston/Galveston county warning area (HGX; CWA) during Harvey. The TCTOR cells were further classified by the presence of a TDS into two groups: TOR TDS and TOR NO TDS. Differences in low-level V_{rot} were more pronounced between tornadic cells that produced a TDS than those that did not, more so than between tornadic and nontornadic cells. However, the TOR TDS mean V_{rot} was over 10 knots (5.1 $\mathrm{m\,s^{-1}}$) higher than NON TOR. At higher elevation angles, V_{rot} decreased for NON TOR, but remained larger for TCTORs. The authors note that a smaller sample size may be the cause of this finding, rather than a true indication of deeper rotation in tornadic cells. Moreover, K_{DP}/Z_{DR} separation signatures were present in more tornadic than nontornadic cases but did not discriminate between the two as well as V_{rot}. Finally, they found that the skill for tornado warnings was maximized when a V_{rot} threshold of 35 knots (18 $\mathrm{m\,s^{-1}}$) was used, larger than that recommended by Martinaitis (2017). However, using this threshold resulted in a decreased probability of detection (POD). Because of the time required for manual analysis, the authors only limited their analysis to 49 tornadic and 64 nontornadic cells.

Beyond low-level rotation, Doppler radars are capable of deriving reflectivity-based attributes such as echo tops (e.g., Lakshmanan et al. 2013a) and vertically integrated liquid (VIL; Greene and Clark 1972). Downward-pointing radar measurements have shown some correlation of 30 and 40-dBZ constant-reflectivity heights and the maximum updraft strength (Heymsfield et al. 2010) suggesting echo-top heights as a proxy for convective intensity. Moreover, Greene and Clark (1972) noted rapid increases in VIL tend to occur with storm intensification. These proxies for convective intensity have been studied in the context of TCTORs (e.g., Spratt et al. 1997; McCaul et al. 2004; Schneider and Sharp 2007; Martinaitis 2017; Nowotarski et al. 2021). In general, these

studies have not found any discrimination between tornadic and nontornadic events using VIL or echo tops, although these have not been examined over a larger sample size.

1.3 Objectives and Hypotheses

Given the limited sample sizes of prior TCTOR radar analyses, a larger sample is needed to build a more robust climatology (both in terms of the total number of cells and also a wider variety of parent TCs) of radar attributes to improve criteria for the warning decision process. The manual collection of radar attributes is time consuming and the TC supercells are potentially ambiguous in nature, leading to a subjective analysis. Therefore, this project creates an automatically derived climatology of radar attributes of TC tornadic and nontornadic convection from 2013 to 2020. Using this larger climatology, this study will test the following hypotheses:

- (H1) Tornadic cells will show stronger radar-detected rotation (AzShear and V_{rot}) and higher values of spectrum width compared to nontornadic cells at a given elevation angle.
- (H2) Tornadic cells with larger values of V_{rot} and AzShear correspond to higher tornado damage ratings and are more likely to display a TDS.
- (H3) Echo-top heights and vertically integrated liquid (VIL) will not be significantly different between tornadic and nontornadic cells.
- (H4) Temporal trends of radar-derived rotation, spectrum width, echo tops, and VIL will be different between tornadic and nontornadic cells.
- (H5) Using a higher V_{rot} threshold than that suggested by Martinaitis (2017) will improve TCTOR warning skill as suggested by the more limited analysis of Nowotarski et al. (2021).

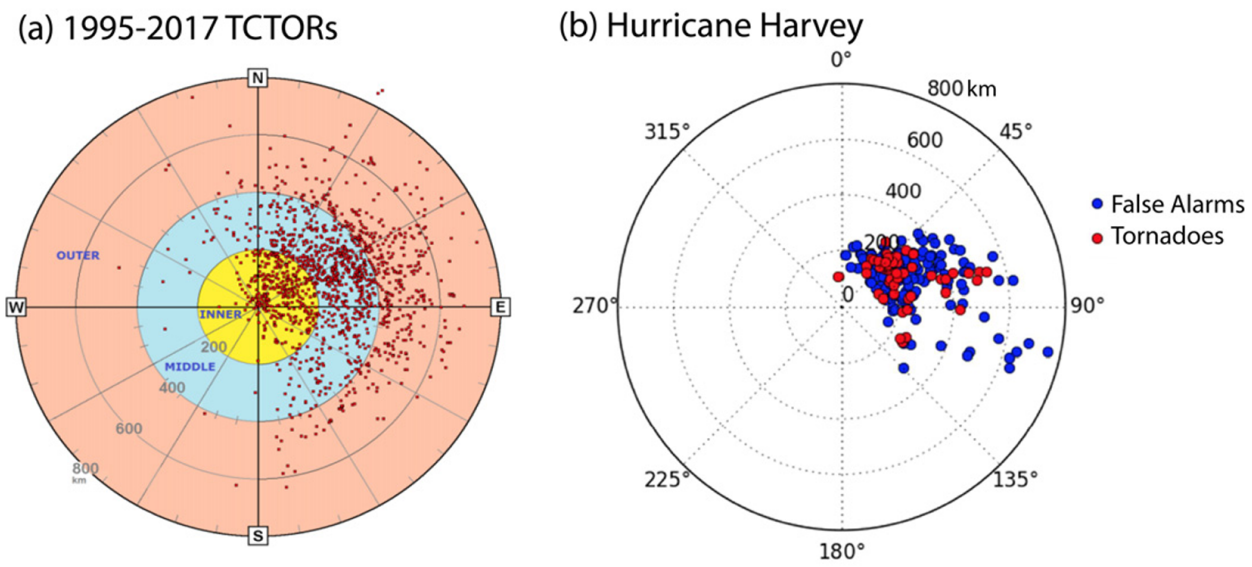

Figure 1.1: Polar plots of tornado position (red dots) with respect to true-north-relative azimuth and range (km) from center in (a) all U.S. landfalling tropical cyclones from 1995 to 2017 and (b) Hurricane Harvey (including false alarm cell position as blue dots). [**©American Meteorological Society. Used with permission.** Figure and caption reprinted with permission from the American Meteorological Society (Nowotarski et al. (2021) Figure 2).]

2. DATA AND METHODS

2.1 Overview and Case Selection

The goal of this project is to build a climatology of automatically, objectively determined radar attributes of TCTORs and nontornadic events. Events considered for this study are retrieved from landfalling TCs of at least tropical storm strength during the 2013-2020 period. Geographically, this corresponds to a region within the United States, east of the Rocky Mountains. Building such a climatology consists of five steps: 1) the creation of a database of individual times, dates, and locations of TCTORs and false-alarm warnings (hereafter referred to as cases), 2) the retrieval and processing of radar and model analysis data for each case, 3) automatically identifying the area of convection and rotation associated with the case, 4) calculation and recording of the desired attributes for each area of convection and rotation, and 5) compilation and analysis of the recorded attributes.

Tornadic cases are identified using the Edwards and Mosier (2022) TCTOR database. Each case is examined to ensure that a radar signature is present and determine whether a TCTOR is included in a tornado warning. Tornado warning statistics following the 2-by-2 contingency table model (Wilks 2019) are calculated where the probability of detection (POD) is the ratio of "hits" to the total number of TCTOR events. A "hit" is defined as the start point of a TCTOR existing within a warning polygon at or before the time of the warning. These include cases where a warning is issued for a different feature, but the TCTOR still occurred within the warning. A "miss" is defined as a TCTOR that is not warned or the start coordinate fell outside of the warning polygon at the TCTOR's start time. A false alarm is a tornado warning with no TCTOR reported and the false alarm ratio (FAR) is the number of false-alarm cases divided by the sum of the number of hits and false alarms. Cases showing an ambiguous or no radar signature at or near the TCTOR's starting coordinate at the start time are removed from the analysis, rather than attempting to correct the report (10 tornadic cases are removed from the analysis). An additional 10 cases are also removed

due to missing radar data. While clearly erroneous or ambiguous cases are removed, errors in the time and location of TCTOR reports could still be present (see Edwards and Mosier (2022) for more details on potential database errors). From these data, the time, date, start-point latitude and longitude, EF rating, and whether the tornado is a "hit" or "miss" are recorded for the tornadic cases.

Nontornadic cases are compiled by determining which county warning areas (CWA) intersect an 800-km (432-n mi.) buffer of the TC's best track (Landsea and Franklin 2013). The 800-km buffer is chosen following McCaul (1991). Radar data are then examined to determine if the warned convective cells are associated with the TC using the Iowa State Cow (ISC; `https://mesonet.agron.iastate.edu/cow/`) archive of tornado warnings. The locations and times of the feature points from the tornado warnings of each associated cell are recorded. The feature points are examined with radar data, and warnings associated with one cell are all assigned to that cell. The time, date, latitude, and longitude of the first feature point of each cell are used to define the nontornadic cases.

The methodology used here builds upon previous work (Spotts et al. 2020; Nowotarski et al. 2021; Spotts et al. 2022). These studies performed an analysis for the volume scan that is most representative of the tornado (tornadic) or the volume scan with maximum low-level rotation of the two volume scans leading up to and including the first warning's feature time (approximately the warning issuance time). Here, the analysis is performed relative to a reference point (RP) defined as the location of TCTOR start point from the Edwards and Mosier (2022) TCTOR database for tornadic cases. The nontornadic RP is the feature point of the cell's first warning.

2.2 RAP and Radar Data

Radar data from the nearest two available WSR-88D radars are downloaded using the Next Generation Radar (NEXRAD) Amazon Web Services (AWS) Python module (`https://github.com/aarande/nexradaws`) from 24 minutes before to 10 minutes after the TCTOR start time (for tornadic cases) or the time associated with the feature point (for nontornadic cases). This time is hereafter referred to as the reference time (RT). The nearest radar is used for velocity and dual-

polarization attributes relevant to the case. The second radar is incorporated for the calculation of composite (the maximum value of all vertical levels), integrated, and interpolated reflectivity-based attributes (e.g., composite reflectivity, VIL, 20-dBZ echo tops, etc.). Additionally, the use of two radars helps alleviate gaps in the data during cases where valid echo tops and reflectivity exist above the nearest radar's maximum tilt (Lakshmanan et al. 2007). To be consistent with the analysis of Nowotarski et al. (2021), supplementary adaptive intra-volume low-level scans (Chrisman 2014) are excluded from the analysis.

Vertical wind profiles are retrieved from the 13-km Rapid Refresh (RAP) model (Benjamin et al. 2016) analysis to aid in dealiasing velocity data (Losey et al. 2017) and cluster tracking. These data are processed using a set of Warning Decision Support System-Integrated Information (WDSS-II; Lakshmanan et al. 2007) algorithms. WDSS-II is a framework for processing data from multiple radars, other sensors, and model data and serves as the backbone for part of the Multi-Radar Multi-Sensor (MRMS) operational products (Smith et al. 2016). Data are indexed for the algorithms starting from the time of the nearest radar's first scan. Each of the key WDSS-II algorithms, the Best Track Real Time (BTRT; `https://github.com/arkweather/BestTrackRT`; Harrison 2018) algorithm, and the parameters used for each are described in Appendix A.

Once ingested into WDSS-II, non-meteorological echoes are removed using the w2qcnndp algorithm yielding "ReflectivityQC" (Lakshmanan et al. 2014). This reflectivity is merged onto a 0.6°x0.6°x20-km grid with 0.0025°x0.0025°x0.5-km spacing using the w2merger algorithm (Lakshmanan et al. 2006; Lakshmanan and Humphrey 2014). From this grid, VIL and composite reflectivity are calculated. Additionally, echo tops are calculated by vertically interpolating ReflectivityQC between elevation angles to the 20-dBZ level (Lakshmanan et al. 2013a). For echo tops, the larger values between the nearest two radars are then merged onto the same grid.

The location and magnitude of rotation associated with TCTORs and nontornadic cells are determined by using a local linear least squares (LLSD) technique to determine the gradient of velocity in the azimuthal direction ("AzShear"; Smith and Elmore 2004; Mahalik et al. 2019). This

technique is considered more robust than traditional methods such as rotational velocity (V_{rot}), ΔV, or peak-to-peak azimuthal shear where the difference in velocity between inbound and outbound gates are divided by the distance between them as more gates are used in the calculation, making it more resistant to noise (Smith and Elmore 2004; Mahalik et al. 2019). To reduce unnecessary AzShear recorded from non-meteorological echoes, the ReflectivityQC field is dilated and gates of AzShear with dilated ReflectivityQC values less than 20 dBZ are discarded.

To identify the areas of rotation and convection coherent with the desired case, the AzShear, VIL, and echo-top fields are clustered using an enhanced watershed algorithm (Lakshmanan et al. 2009) and tracked across time (Lakshmanan and Smith 2010, see the "Cell Tracking" section below). Specifically, clusters for AzShear are determined for the $0.5°$, $0.9°$, $1.3°$, $1.45°$, $1.8°$, and $2.4°$ elevation angles (an example of the different clusters is shown in Figure 2.1). Once clusters have been identified, attributes such as maximum AzShear, maximum spectrum width, mean and maximum VIL, and 90^{th}-percentile and maximum echo tops as well as radar coordinates such as range, azimuth, and beam height (Doviak and Zrnic 1993) are determined. Following Nowotarski et al. (2021), tornadic cases are further separated into those that do and do not have a TDS. To determine the presence of a TDS, the w2tds algorithm is used to find gates that meet TDS criteria. These criteria include having a correlation coefficient (ρ_{HV}) value less than 0.90, an AzShear value of at least $2.5 \times 10^{-3}\,\mathrm{s}^{-1}$, and a reflectivity value of at least 35 dBZ consistent with the 2016 Warning Decision Training Division guidance (WDTD 2016). The "count" attributes of gates within the "TDS_CC" output of w2tds are recorded for each AzShear cluster to aid in determining which cases may have a TDS. The addition of cases with a "detected" TDS yields four categories of cases: NON TOR (nontornadic), ALL TOR (tornadic), TOR TDS, and TOR NO TDS.

2.3 Cell Tracking

2.3.1 Determining the Initial Reference Cluster

In order to record attributes of the rotation and convection clusters, the clusters must be tracked and associated with the rotation of interest. Initially, the $0.5°$ AzShear clusters are tracked using the

w2segmotionll (Lakshmanan and Smith 2010) algorithm with the Cintineo et al. (2020) parameters with a tolerance distance of 7.5 km (4.0 n mi.) instead of 10 km (5.4 n mi.) for AzShear clusters. The archive Best Track Real Time program (archiveBTRT; Harrison 2018) is then used to connect a cluster's original track to a best-fit track. After clusters are initially tracked, the first desired cluster, or "reference cluster," must be identified. Clusters within 7.5 km of the RP and 5 minutes of the RT are determined using the w2hailtruth_size algorithm. The clusters from the volume scan closest to the reference time are then scored using a modified interest score (IS) which has been used to associate forecast and observed rainfall (Davis et al. 2006, 2009) and objects within the Warn-on-Forecast System (WOF) and observed objects (e.g., areas of low-level rotation; Skinner et al. 2016, 2018). The score for the reference cluster is calculated according to Equation 2.1 where d is the great-circle distance between the RP and the candidate cluster and $AzShear_{abs}$ is the absolute value of the minimum or maximum AzShear, whichever is greater. The coefficients of 0.9 and 1.8 (as well as the coefficients of other IS calculations) are determined heuristically. Using this equation, more intense clusters that are at similar distances from the reference point receive a higher score. The cluster with the highest score is chosen as the reference cluster and used as the starting point for further tracking. If no reference cluster is found for cyclonic or anticyclonic clusters, the case is not recorded in the analysis.

$$IS_{ref} = 0.90\frac{7.5km - d}{7.5km} + 1.80\frac{AzShear_{abs}}{0.005\text{s}^{-1}} \tag{2.1}$$

2.3.2 AzShear Cluster Tracking and Association

Once a reference cluster has been identified at the lowest elevation angle (0.5°), that cluster is then tracked across time, initially using the w2semgotionll tracking and BTRT post-processing. Starting from the reference time, the reference cluster is tracked backwards in time. At each previous time step, a point is determined 61% (determined heuristically) along a projection from the reference cluster to its expected location (point A in Figure 2.2). This vector is initially the Sfc-to-6-km mean wind, then becomes the average of all motion vectors accumulated up to that

point in the tracking. If the ID of the next cluster matches the current cluster's ID, a sector is defined from the candidate cluster within 50° of the motion vector and a radius of 7.5 km. If the projected point falls within this sector or within 3.675 km (2.0 n mi.) of the candidate cluster, that cluster is selected. If no matching ID is found or the cluster with the same ID does not meet this criteria, the remaining candidate clusters are scored per Equation 2.2.

$$IS_{shear} = 0.90\frac{7.5km - d}{7.5km} + 0.25\frac{180° - bearing_{diff}}{180°} + 1.80\frac{AzShear_{minabs}}{AzShear_{maxabs}} \tag{2.2}$$

Here, the first term is the same as in Equation 2.1. The second term gives a higher score to the clusters that are more aligned with the average motion vector where $bearing_{diff}$ is the difference between the bearing from the older (earlier time step) candidate cluster, to the projected point in degrees. If the distance between the cluster centroid and projected point is more than 3.675 km and $bearing_{diff}$ is greater than 50° or the distance between the cluster centroid and projected point is greater than 7.5 km, the candidate cluster is disqualified. Finally, $AzShear_{minabs}$ and $AzShear_{maxabs}$ are the minimum and maximum $Azshear_{abs}$ between the two clusters. The candidate cluster with the highest score greater than 0.2 (following Skinner et al. (2016, 2018)) that meets the distance and bearing criteria is selected. This process repeats, stepping backwards in time, until an association has been made at all possible time steps or no association can be made. Once the tracking process is completed backwards in time, the same tracking is performed forwards in time, beginning from the initial reference cluster. The difference is that the bearing and distance criteria are now evaluated from the projected point, rather than the candidate cluster. This tracking is performed on both cyclonic and anticyclonic clusters, with the group with more intense rotation at more time steps being chosen for analysis. Figure 2.3 shows an example of the 0.5° cluster being tracked.

After the 0.5° AzShear clusters are tracked and chosen, clusters at higher elevation angles are then associated with their corresponding 0.5° clusters and added to an "(Az)Shear group." The process for vertical AzShear cluster association is similar to the tracking at 0.5°. For each elevation angle, a point is projected 61% along the 0.5° motion vector using the time between the 0.5° and

higher-level elevation angle. Each cluster at the higher elevation angle is scored using the interest score defined in Equation 2.2, with a search radius of 5 km (2.7 n mi.). The cluster with the highest score of at least 0.2 within 3.5 minutes after the 0.5° cluster is added to the shear group.

2.3.3 VIL and Echo-Top Clusters

To determine the location and intensity of the storm's convection associated with the previously described AzShear clusters, VIL and echo-top clusters are created. An interest-score methodology is also adopted for determining which cluster is associated with the shear group. For VIL clusters, a search radius of 15 km (8.1 n mi.) is used and an IS score is calculated using Equation 2.3. The first term of this equation is the same as in previous IS calculations except the maximum distance is 15 km. The second term is the same as in Equation 2.2 except rather than the averaged motion vector, the direction of the Sfc-6-km MSL bulk wind difference in degrees (BWD) is used to yield the bearing difference ($BWD_{6kmdiff}$). The cluster closest in time within five minutes of the 0.5° AzShear cluster and with the highest score of at least 0.2 is selected and added to a "storm group." The storm group consists of a shear group and the associated VIL and echo-top clusters. If a VIL cluster is found, any potential echo-top clusters are scored to determine which cluster to add to the storm group using a similar process to VIL, but using Equation 2.4. This process is illustrated in Figure 2.4.

$$IS_{VIL} = 1.15\frac{15km - d}{15km} + 0.75\frac{180° - BWD_{6kmdiff}}{180°} \tag{2.3}$$

$$IS_{ET} = 1.2\frac{20km - d}{20km} + 1.16\frac{180° - BWD_{6kmdiff}}{180°} \tag{2.4}$$

2.4 Rotational Velocity Estimation

Nowotarski et al. (2021) calculated V_{rot} following Smith et al. (2015) where the rotational velocity is half the difference between the peak inbound and outbound velocities (or half the sum of the absolute values of the peak inbound and outbound velocities) within 5 n mi. (9.26 km) and

45° of each other. To the author's knowledge, there is not a program that directly calculates V_{rot} per the Smith et al. (2015) criteria. Early attempts to automate the calculation proved ineffective. Therefore, V_{rot} is estimated from AzShear.

Mahalik et al. (2019) tested the calculation of AzShear by comparing AzShear calculated for a combined Rankine vortex (Wood and Brown 1997) against the vortex's true half-vorticity value. Using this relationship, V_{rot} is calculated by multiplying both sides of the equation by the radius to yield Equation 2.5.

$$V_{rot} = \frac{\Delta V_r}{2} = AzShear * r \tag{2.5}$$

To determine a general radius (r), the manually determined V_{rot} of cases from Nowotarski et al. (2021) are compared against the new automated method for estimating V_{rot} for the same cases. The V_{rot} for each case is calculated using Equation 2.5 and compared to the manually calculated values. The radius is varied from 500 m (0.3 n mi.) to 2500 m (1.3 n mi.) in 50-m (0.03 n mi.) increments. The radius that provides the best-fit slope closest to 1 is used.

2.5 Final Analysis and Hypothesis Testing

Each cluster is placed in a five-minute bin based on the cluster's radar sweep or grid time, similar to Davis and Parker (2014), but offset by 2.5 minutes from their analysis. For example, the zero-minute bin covers 2.5 minutes before to 2.5 minutes after the time of TCTOR onset or warning feature time. The data retrieved using the previously described methodology are sorted and aggregated by time bin. Data retrieved from AzShear clusters are further sorted by elevation angle and EF rating. Trends between time bins for each elevation angle, echo tops, and VIL are also calculated. If multiple clusters from the same case fall within the same tilt and time bin, the cluster with the more favorable value (e.g., more intense AzShear, higher echo tops, or stronger trend) is chosen. While both cyclonic and anticyclonic AzShear clusters are considered, a separate subset of the data is created with only cyclonic AzShear clusters. For NON TOR cases, the time difference between the time of maximum 0.5° AzShear and the reference time is subtracted from

each elevation angle, VIL, and echo-top cluster time such that the new reference time corresponds to the maximum low-level rotation. Examining NON TOR cases in this way is consistent with Martinaitis (2017). In a few cases, this time shifting creates time bins near $\pm$30 minutes where ALL TOR cases are not present. Therefore, only the -20 to 10-minute bins are examined. Finally, for each projected point, the number of clusters within the tracking search radius, not including the directional tracking criteria, of the point is recorded. This number is hereafter referred to as the number of "potential clusters" for the potential number of clusters that could be tracked. Clusters are only eligible to be included in the analysis if they have a maximum of three potential clusters. This criteria is employed to reduce the number of potentially incorrect tracks.

For TDS cases, objective TDS criteria are determined using the Harvey (2017) manual analysis. To determine the optimal number of TDS pixels to trigger TDS detection, the Critical Success Index (CSI; Wilks 2019) is calculated at increasing thresholds from the time bin 20 minutes before the RT. A pixel count of 25 yields the best CSI of 0.41. Because a TDS requires a TCTOR to have already developed and be lofting debris, it may not be reasonable to look for a TDS before the beginning of the TCTOR. Therefore for this analysis, the search for a minimum pixel count of 25 begins at the zero-minute time bin. This criteria yields a CSI of 0.31 from the manual analysis cases with 13 tornadic and 4 nontornadic cases triggering a TDS detection. One possible reason for this decrease is that the manual analysis reference time may be slightly different than the reference time from the automated analysis. Moreover, for determining TDS cases used to calculate the proportion of TDS cases at each EF rating, a beam-height criteria of 10 kft (3 km) and maximum of 3 potential clusters for the AzShear object that meet the other criteria is imposed. To be considered a NO TDS case, at least one 0.5° tracked AzShear cluster must meet the beam-height and potential-cluster criteria.

From this data, the Mann-Whitney-U test is applied with a null hypothesis that the AzShear and V_{rot} distributions for the ALL TOR and NON TOR cases (or for H2, differences between tornado EF ratings) are equal with an alternative hypothesis that the ALL TOR cases (or more intense tornadoes) will have stronger rotation. Additionally, the proportion of tornadoes that have a TDS

signature at higher EF ratings are tested for larger proportions of cases with a TDS than lower EF ratings using a Z-test for proportions. Differences, or lack thereof, of echo-top heights and VIL values between ALL TOR and NON TOR cells (H3) are tested in a similar way to H1 but using a two-sided test on echo-top heights and VIL instead of velocity-based attributes. The temporal trends between ALL TOR and NON TOR cells for the velocity-based attributes and echo tops (H4) are tested by taking the change of the attribute between each pair of time bins and performing the Mann-Whitney-U test for differences between categories and the Wilcoxon Signed-Rank to test for differences from 0. The trends are calculated in a manner similar to the analysis of temporal trends performed by Martinaitis (2017) who tested differences of V_{rot}, rotational diameter, and shear between volume scans prior to tornadogenesis (for ALL TOR cases) and time of maximum V_{rot} (for NON TOR cases). The effect of increasing the V_{rot} threshold on warning performance is tested by calculating the True Skill Score (TSS; Wilks 2019) with increasing V_{rot} values, until a maximum TSS is reached.

Additionally, the probability of detection (POD), false alarm ratio (FAR), and Critical Success Index (CSI; Wilks 2019) are calculated. Automated V_{rot} thresholds are varied from 5 to 60 knots. These warning metrics are also calculated for hypothetical warning-decision schemes where only TDS TOR or EF1 and higher-rated TCTORs (EF1+) are warned. TSS values are calculated following the Martinaitis (2017) V_{rot} criteria where thresholds for cases beyond 40 n mi. (74.1 km) are reduced by 5 knots (2.6 $\mathrm{m\,s^{-1}}$; AzShear = ~$2 \times 10^{-3}\,\mathrm{s^{-1}}$). Warning statistics for the thresholds that yield the largest TSS values are also compared to statistics from the Martinaitis (2017) V_{rot} criteria and the observed statistics. Observed statistics are calculated on a case-by-case rather than warning-by-warning basis, with the 10-kft beam height and at least a known number of potential clusters criteria applied. The observed statistics for EF1+ events are also calculated.

Nowotarski et al. (2021) performed a similar comparison using their data and found that by increasing the V_{rot} threshold, a local maximum in CSI was obtained at the expense of a lower POD. By performing the comparison using only TDS cases, POD was better preserved at larger thresholds. While the methodology used herein does not explicitly allow for the same comparison,

cases are selected using similar criteria. This criteria includes using the volume with the maximum 0.5° AzShear from the -10 to 0-minute bin for NON TOR cases, and the zero-minute bin for the ALL TOR cases. The warning statistics are also calculated on these cases for all three warning-decision schemes. Finally, for a discussion on some of the limitations and sources of error for the methodology discussed in this chapter, see Appendix B.

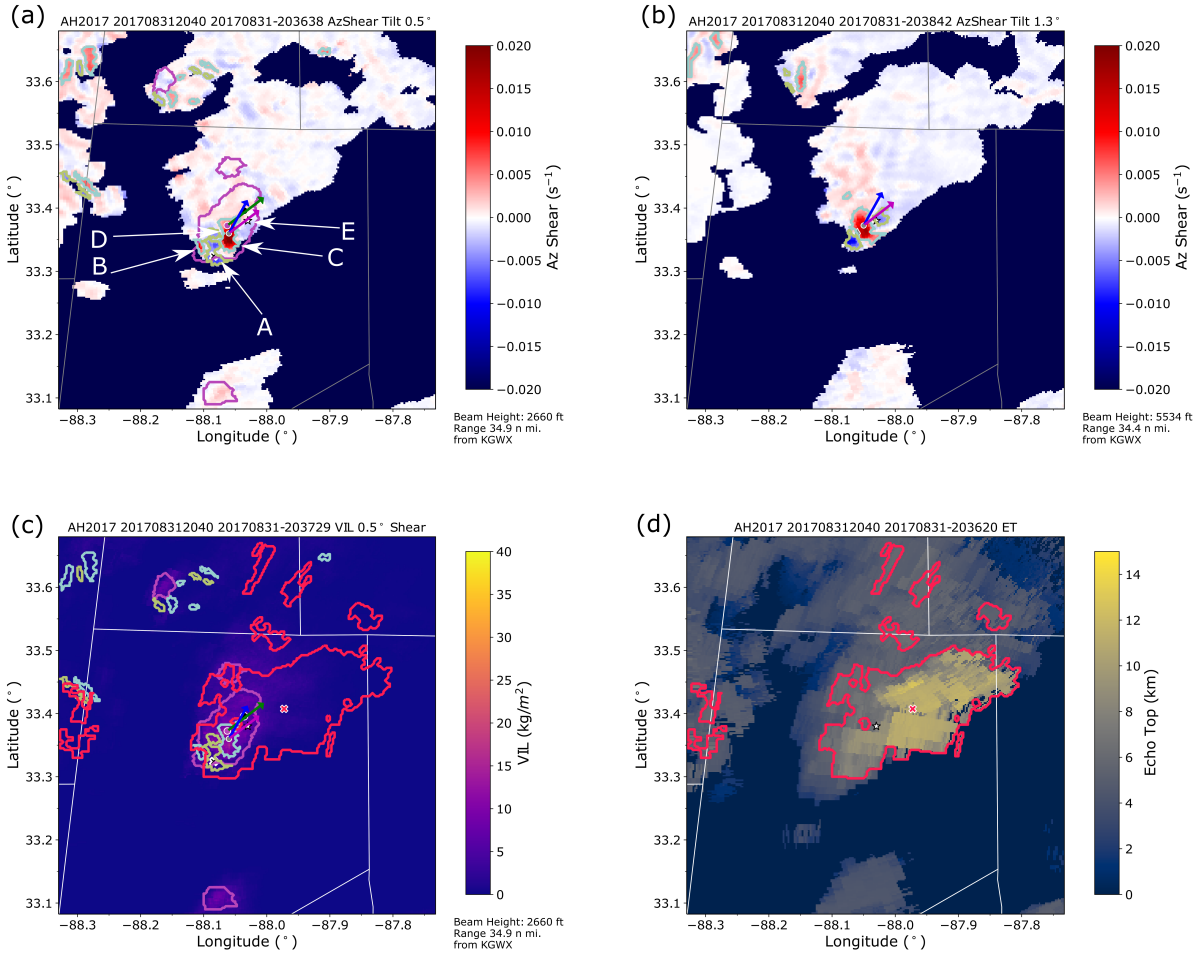

Figure 2.1: **(a)** 0.5° AzShear for an ALL TOR case from Hurricane Harvey on 31 August, 2017 at 2037Z with negative (anticyclonic) clusters contoured in green (Point A), VIL clusters contoured in purple (Point B) positive (cyclonic) clusters contoured in cyan (Point C), the tracked AzShear cluster indicated by the grey dot and origins of the blue (accumulated average motion) and purple (Sfc-to-6-km MSL shear vector) arrows (Point D), and the reference point is indicated by the grey star (Point E). The approximate beam height and range from the nearest radar of the tracked AzShear cluster and the radar's ID are shown in the lower-right corner. The forward and backward projected points are shown by the black circle and X respectively. The purple circle represents the centroid of the tracked VIL cluster and the green arrow is the Sfc-to-6-km MSL shear vector at the time closest to the shear clusters. **(b)** Same as in **(a)** except at the 1.3° elevation angle at 2039Z. The centroid of the 0.5° AzShear cluster is shown by the grey diamond. **(c)** Same as in **(a)** but for VIL at 2037Z. The nearest echo-top cluster in time is shown in red. The red X represents the centroid of the tracked cluster. **(d)** Same as in **(c)**, but with 20-dBZ echo tops at 2036Z.

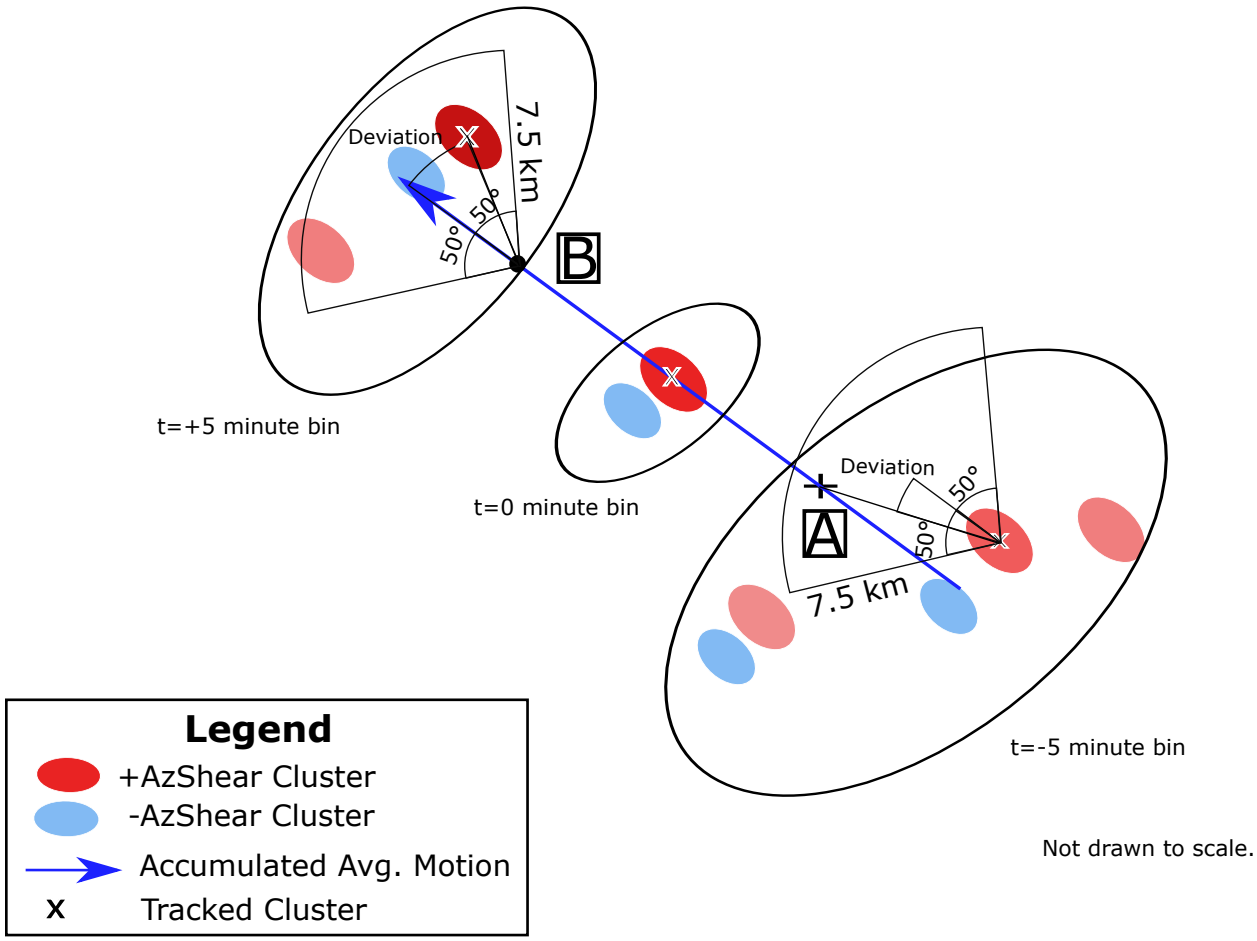

Figure 2.2: An idealized depiction of the process of tracking a cyclonic 0.5° cluster (tracked cluster is marked with an X) after utilizing the MULTISTAGE and archiveBTRT methods. Positive (cyclonic) AzShear clusters are shown in red while the negative (anticyclonic) clusters are shown in blue. The shading indicates the relative intensity of the AzShear with darker colors indicating more intense rotation. Each black ellipse indicates a grouping of clusters for a given volume scan. The blue arrow represents the averaged accumulated motion over the tracking period. The cross and circle along this vector are the points projected for interest score calculations for the backward (Point A) and forward (Point B) directions respectively.

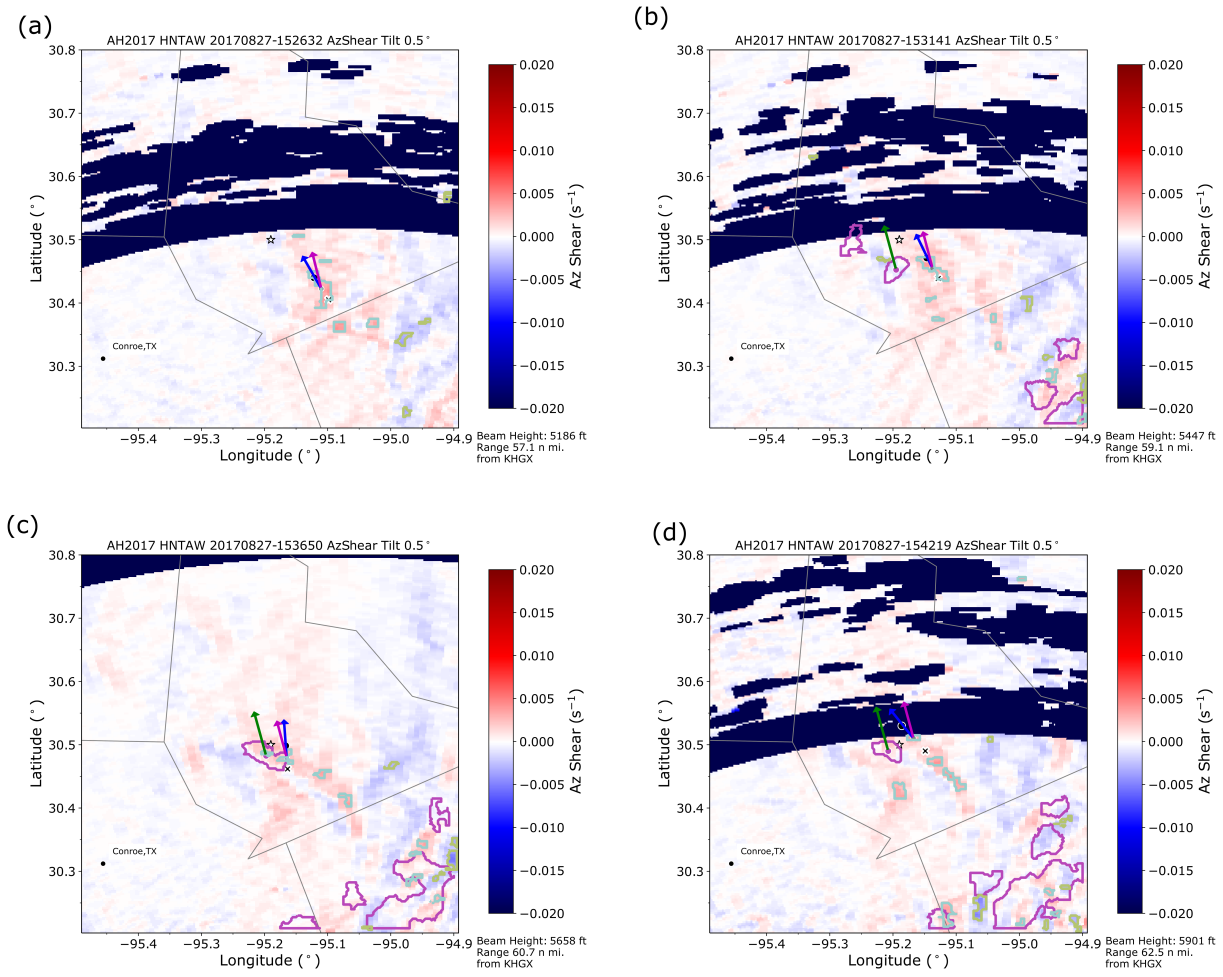

Figure 2.3: Same as in Figure 2.1 **(a)** except for a NON TOR case during Hurricane Harvey on 27 August, 2017 **(a)** 1527Z, **(b)** 1532Z, **(c)** 1537Z, and **(d)** 1542Z. A region of range-folded values or the "purple haze" can be seen in **(a)**, **(b)**, and **(d)** as a band of missing values just north of the reference point.

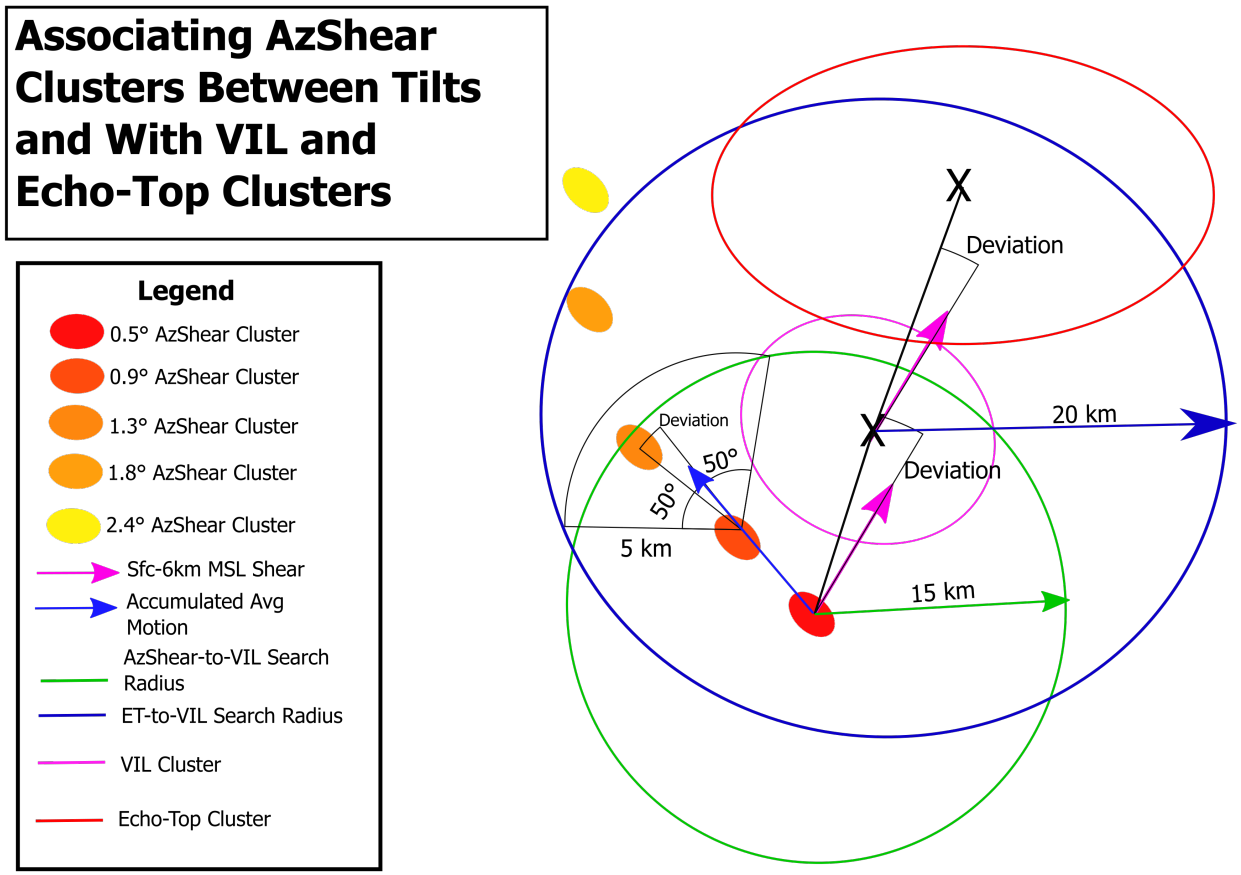

Figure 2.4: An idealized depiction (not drawn to scale) of the association of a 0.5° AzShear cluster with VIL, echo-top, and higher elevation-angle clusters. The red through yellow ellipses represent different clusters at different elevation angles. The VIL cluster is shown in purple and the echo-top cluster is shown in red. The green and blue circles represent the search radii for the VIL and echo-top clusters respectively. The blue arrow from the 0.5° cluster is the accumulated average motion of the 0.5° elevation-angle cluster and the purple arrows represents the Sfc-to-6-km MSL shear vector. Centroids of the VIL and echo-top clusters are represented by a black X.

3. RESULTS

3.1 Comparison to Manual Analysis of Harvey Cases

To test the reliability of Equation 2.5, the automated analysis is compared to V_{rot} values obtained by Nowotarski et al. (2021). Of the 113 cases processed, 111 are manually examined at the zero-minute time bin to determine if the initial association was performed correctly. Two cases did not have a valid reference cluster or necessary model data. AzShear and echo-top clusters are examined and those that do not appear to show a correct association (generally less than 10%) are flagged and removed from the comparison.

Figure 3.1 shows scatter plots of V_{rot} estimated from the automated method using Equation 2.5 compared to the manually calculated values for each elevation angle and all elevation angles (all tilts) combined. The lower-level and all-tilt comparisons show well-correlated values with Pearson correlation coefficients ranging from 0.781 to 0.904 (Fig. 3.1a-c,f). The all-tilts comparison has a correlation coefficient of 0.802. However, the samples are not independent as multiple elevation angles can be attributed to the same case. Upper-level elevation angles show slightly lower correlation between manual and automated V_{rot} (0.634 and 0.760 for the 1.8° and 2.4° elevation angles respectively; Fig. 3.1d,e). Similar to Spotts et al. (2022), the resulting root-mean-square error (RMSE) values with respect to the best-fit line are larger than ideal. The 1.3/1.45° elevation angle has the lowest RMSE of 4.6 knots (2.4 $\mathrm{m\,s^{-1}}$) whereas the 1.8° elevation angle has the highest RMSE of 11.1 knots (5.7 $\mathrm{m\,s^{-1}}$). The all-tilts comparison has an RMSE of 6.6 knots (3.4 $\mathrm{m\,s^{-1}}$). Per Figure 3.2a, the errors compared to the manual analysis for all tilts range from a 26.7 knot (13.7 $\mathrm{m\,s^{-1}}$) underestimation to a 19.7 knot (10.1 $\mathrm{m\,s^{-1}}$) overestimation. The interquartile range, however, is 3.3 knots (1.7 $\mathrm{m\,s^{-1}}$) underestimated to 4.4 knots (2.3 $\mathrm{m\,s^{-1}}$) overestimated with median and mean values of 0.40 (0.21 $\mathrm{m\,s^{-1}}$) and 0.62 knots (0.27 $\mathrm{m\,s^{-1}}$) respectively. Despite the relatively small mean and median errors, the range of errors and RMSE values indicate that a precise value cannot be derived from these relationships. However, the derived values may be

useful for a general qualitative analysis. With a best-fit slope radius of 1300 m (0.7 n mi.) yielding a slope closest to 1 for the lowest two elevation angles and all tilts, this radius is used in subsequent analysis for the estimation of V_{rot}.

Figure 3.3 shows histograms of the radii used in Equation 2.5 that minimizes the error between the manual and automated analysis using the same 50-m increments. Overall, Figure 3.3f shows the most common radius with the least amount of error is near 1250 m (0.7 n mi.). Looking at the distribution by elevation angle, the 0.5°, 0.9°, 1.3/1.45°, and 2.4° show peaks near 1050 m (0.6 n mi.), 1250 m, 1150 m (0.6 n mi.), and 1250 m respectively. These values are near or lower than the overall radius that yields the best-fit slope closest to one. At 1.8° the most common radius is near 1750 m (0.9 n mi.), near its best-fit radius of 1700 m (0.9 n mi.). While these radii minimize the error between the values from the automated and manual analysis, the individual vortices' radii may differ from this number. In addition to the limitations discussed in Appendix B, it is not known if the most common best radius near 1250 m or the best-fit radius of 1300 m is dependent on the kernel size (2500 m or a half-length of 1250 m).

The maximum and 90^{th}-percentile echo tops are also compared to the manual analysis. Figure 3.4 shows histograms of the manually retrieved value subtracted from the value retrieved from the cluster (automated). The clusters' maximum echo tops tend to overestimate the manually retrieved values with a median error of an 8.3-kft (2.5-km) overestimation with a maximum value of 30.7 kft (9.4 km; Fig. 3.2b). The 90^{th}-percentile values are more prone to underestimation with errors reaching values near -23.2 kft (7.0 km); however, the errors are more closely centered near 0 ft with a median and mean error of 1.0 kft (0.3 km) and 1.1 kft (0.3 km) respectively. In subsequent analysis, both the maximum and 90^{th}-percentile echo-top values are considered. Despite some larger errors, the automated analysis generally compares well to the manual analysis with automated V_{rot} typically within 5 knots of the manual-analysis values.

3.2 Rotation and Spectrum Width

A total of 370 ALL TOR (280 HIT and 90 MISS) along with 782 NON TOR events for a total of 1,152 cases are initially provided to the program as shown in Figure 3.5. After quality control

during the automated analysis, a total of 1,055 cases (334 ALL TOR and 721 NON TOR) remain. Sixty-eight ALL TOR cases display a TDS using the criteria defined in Section 2.2. Fifty-eight met the criteria to be included in the testing of Hypothesis 2. A total of 44 cases (26 NON TOR and 18 ALL TOR) track anticyclonic rotation. Examination of only cyclonic cases compared to all cases reveal a larger distinction in low-level rotation between ALL TOR and NON TOR distributions. Therefore, the remaining analysis only focuses on cases where cyclonic rotation is tracked, with the exception of the TDS proportion tests.

Figure 3.6 shows AzShear at different elevation angles and time bins. As the automatic V_{rot} is simply scaled AzShear, both variables have the same relative differences in time and between NON TOR and ALL TOR events. Therefore, most of the subsequent analyses focuses on only AzShear to avoid redundancy. Comparison of the distributions of NON TOR and ALL TOR rotation reveals that the NON TOR distribution is significantly less than the ALL TOR distribution at a 0.05 level of significance at most time bins using the Mann-Whitney-U test for the lowest elevation angle (Fig. 3.6a). As the elevation angle increases, the sample size, separation between distributions, and number of significantly different bins decrease. At 2.4°, only the -5-minute bin shows the NON TOR AzShear significantly less than the ALL TOR AzShear with considerable overlap in their distributions (Fig. 3.6c). In the lower-levels, ALL TOR cases appear to reach larger values of AzShear earlier, indicating ALL TOR storms may have a persistence of stronger rotation across time compared to NON TOR cases (Fig. 3.6a). For example, at the -10-minute bin, the median value of ALL TOR AzShear is approximately $1 \times 10^{-2}\,\mathrm{s}^{-1}$ with a median NON TOR value of approximately $7.2 \times 10^{-3}\,\mathrm{s}^{-1}$. The median ALL TOR AzShear remains above $1 \times 10^{-2}\,\mathrm{s}^{-1}$ through the zero-minute bin while the NON TOR AzShear only exceeds this value at the time of maximum rotation with a median of approximately $1.1 \times 10^{-2}\,\mathrm{s}^{-1}$. Moreover, stronger rotation at higher tilts may indicate an area worth watching, but not fully discriminate between ALL TOR and NON TOR cases as overlap in the distributions tends to increase.

At the 0.5° elevation angle, NON TOR spectrum width is less than ALL TOR spectrum width at most time bins (Fig. 3.6d). Furthermore, the lower-tilt ALL TOR spectrum width tends to

increase slightly just prior to the TCTOR start time. Despite the statistically significant differences between the two distributions, significant overlap is present. As the elevation angle increases the number of significant time bins tends to decrease, especially at higher tilts. However, at 1.3°, the NON TOR distribution is significantly less than the ALL TOR distribution at most time bins (not shown). Separation between the two distributions is less at earlier time bins and becomes more distinct than at the 0.5° elevation angle starting at the zero-minute bin, but overlap between the two distributions still exists.

These results support H1 suggesting that ALL TOR convection has larger values of AzShear than NON TOR convection at most time bins, particularly in the low-levels 10-15 minutes prior to the TCTOR start or maximum rotation. The larger values of low-level rotation at earlier time bins suggest that when AzShear reaches a particular threshold and continues to increase it may be indicative of an impending TCTOR more than a single volume scan with enhanced rotation. However, this finding may not be universally applicable as some TCTORs spin up quickly. Additionally, some cases have relatively low values of AzShear. While rotation can be weak, it is possible that some AzShear clusters could suffer from tracking errors and lead to an underestimation of rotation. Finally, while the distributions of spectrum width values may be significantly different at times, this result does not strongly support H1 because there is generally considerable overlap in these distributions and a lack of operationally relevant differences in median values.

3.3 TDS Signatures, Damage Rating, and Rotation Magnitudes

Figure 3.7 shows AzShear for each EF rating at different elevation angles for the zero and five-minute bins. At these times, the EF0 rating AzShear tends to be significantly less than the EF1 and EF2 rating AzShear values at most elevation angles including 2.4° at the zero-minute bin. However, the EF1 rating is generally not less than the EF2 rating. Within each EF-rating group, a Mann-Whitney-U test for the NO TDS AzShear less than the TDS AzShear is performed at a 0.05 significance level. At times, the EF0 and occasionally higher-rated TOR NO TDS AzShear values are significantly less than the TOR TDS values. While not always significant, ALL TOR AzShear tends to be larger upon visual inspection of the distributions at higher ratings, although

this varies depending on elevation angle and time bin. These results are more commonly found at the 0.5° elevation angle starting from -10 minutes (not shown) and at higher elevation angles from 0 minutes. While these findings suggest the possibility that rotation could be stronger for higher-rated TCTORs, a lack of samples at higher ratings and bin-to-bin variability prevent these results from definitively supporting H2. Moreover, caution should be used when interpreting significance in Figure 3.7 when sample sizes are small.

H2 also postulates that higher-rated TCTORs are more likely to produce a TDS than lower-rated TCTORs. Table 3.1 shows the number of cases that meet TDS and other criteria, the total of TDS and NO TDS cases, and the proportion of the total cases that meet the criteria for EF ratings 0 through 2. A one-tailed Z-test for the proportion of lower-rating TCTORs less than the proportion of higher-rated TCTORs that meet the TDS and other criteria is performed between each pair of EF ratings (excluding EF3 and unknown ratings). Table 3.2 shows the results of the Z-test for the three combinations of EF ratings. The proportion of EF0 cases with a detected TDS (~11.9%) is significantly less than the proportion of EF1 cases (~24.8%) with a p-value of 0.002 and EF2 cases (~33.3%) with a p-value of 0.006. The AzShear of EF1 cases, however, are not significantly less than EF2 cases with a p-value of ~0.221. More EF2 samples may be necessary to determine if this finding is the result of the small EF2 sample size. Moreover, given the minimum-pixel count criterion, the number of TDS cases is likely underestimated. With respect to EF0 cases, these findings support H2. While the proportion of EF1 TDS cases is smaller than EF2 cases, the difference is not significant and cannot be used to support H2.

Table 3.1: Number and proportion of cases that meet QC'd TDS criteria for EF0 through EF2 rated tornadoes including both cyclonic and anticyclonic cases.

EF Rating	**# of TDS Cases**	**# of Total Cases**	**TDS Proportion**
EF0	24	201	11.9%
EF1	27	109	24.8%
EF2	6	18	33.3%

Table 3.2: Results of a one-tailed Z-test for the proportion of cases that triggered a TDS detection from the cases in Table 3.1. Five cases with an unknown EF rating and one case rated an EF3 are not included.

Smaller EF Rating	Proportion	Larger EF Rating	Proportion	Z	p-Value
0	11.9%	1	24.8%	-2.909	0.002
0	11.9%	2	33.3%	-2.529	0.006
1	24.8%	2	33.3%	-0.767	0.221

3.4 Echo Tops and VIL

Echo tops and VIL do not show much separation between ALL TOR and NON TOR cases. Figure 3.8 shows the maximum and 90^{th}-percentile echo tops and maximum and average VIL for their respective clusters across time for cases with cyclonic rotation. For most bins, the NON TOR distributions are not significantly different than the ALL TOR distributions. The exceptions are the time bins around and including the zero-minute bin for echo tops (Fig. 3.8a,b) and the -15-minute bin for VIL (Fig. 3.8c,d). ALL TOR maximum and 90^{th}-percentile echo tops initially appear to be generally similar to NON TOR cases. Starting at the -5-minute bin, ALL TOR distributions begin to shift upwards compared to the NON TOR distributions, creating a statistically significant difference.

VIL distributions are more similar between ALL TOR and NON TOR cases. At the -15-minute bin, the ALL TOR VIL distributions increase to be significantly above the NON TOR VIL distributions. This difference, however, does not continue. Generally, these results support H3 that ALL TOR and NON TOR VIL are not significantly different. While echo-top distributions between the two categories are significantly different for a few bins, considerable overlap still exists. At the 5-minute bin, maximum echo tops have a difference of means of ~3.4 kft (1.0 km) similar to the difference found by Nowotarski et al. (2021). This small difference along with the overlap between distributions suggests that these variables have a lack of discriminating power.

However, this does not preclude areas of high echo tops or VIL from highlighting potential areas of interest to a forecaster.

3.5 Temporal Trends of Radar Attributes

In addition to values at individual time bins, the changes in these attributes between time bins for each case are examined. Trends are examined only if both volume scans have valid data and meet beam-height and potential-cluster criteria for a particular case. Figure 3.9 shows changes in AzShear between time bins at the 0.5°, 1.3°, and 2.4° elevation angles and spectrum width at 0.9° for cyclonic cases. Both 0.5° NON TOR and ALL TOR AzShear tend to increase leading up to the time of maximum NON TOR rotation or the TCTOR start, without a significant difference between the two distributions using a two-tailed Mann-Whitney-U test (Fig. 3.9a). At the zero-minute bin, the ALL TOR AzShear values tend to stop increasing, then begin decreasing while the NON TOR cases experience a sharp increase followed by a sharp decrease. This inflection is expected as the NON TOR zero-minute bin represents the time of maximum rotation. ALL TOR cases also experience this inflection when sorted by their time of maximum rotation (not shown). The inflection results in a significant difference between the distributions from the -5 to 0 and 0 to 5-minute bins. These patterns progressively diminish at higher elevation angles (Fig. 3.9b-c). In Figure 3.9b, a similar pattern can be observed with generally increasing values leading up to the zero-minute bin, decreasing afterwards. Only the changes of AzShear from the -5 to zero-minute bin show any significant differences between the ALL TOR and NON TOR distributions. At higher elevation angles, fewer trends are significantly different from zero and none are significantly different between ALL TOR and NON TOR distributions. Meanwhile, some cases show outliers in AzShear trends. These outliers, particularly the more extreme values, may be caused by errors in tracking.

With respect to AzShear trends, when comparing NON TOR maximum rotation and the TC-TOR initial rotation, H4 is supported near the zero-minute bin, but not overall as most bins do not show a significant difference between ALL TOR and NON TOR distributions. Spectrum width shows few trends significantly different from zero and only one where the distributions for the two

categories are different at the 0.9° (Fig. 3.9d) and 1.3° elevation angles (not shown). Therefore, the trends found in spectrum width do not support H4.

Trends in echo tops show few time bins different from zero and none that show differences between trends in ALL TOR and NON TOR distributions (Fig. 3.10a,b). Trends in VIL attributes are generally not different from zero with the exception of the -5 to 0 and 0 to 5-minute bins for maximum VIL and the 5 to 10-minute bin for average VIL. The only bin where the ALL TOR and NON TOR distributions are significantly different and at least one distribution is significantly different from zero is the -5 to 0-minute bin where the ALL TOR distribution median is less than zero and significant.

From the -15-minute to the zero-minute bin, NON TOR maximum and average VIL increases by an amount significantly different from zero while the ALL TOR distributions stay closer to zero (not shown). From the -15 to the 5-minute bin, ALL TOR average VIL decreases by an amount significantly less than zero and different from the NON TOR average VIL. For maximum VIL from -15 minutes, the ALL TOR distributions are significantly different from the positive NON TOR distributions, but not different from zero to the -5 and zero-minute bins. From the -15 to 5-minute bin, the ALL TOR cases are significantly less than zero and different from NON TOR cases which are not different from zero. For these cases, trends in NON TOR VIL tend to be positive to neutral while ALL TOR VIL trends tend to be neutral to negative. Maximum cluster echo tops show a significant increase in the ALL TOR distribution from the -15 to 0-minute bin, however there are only seven samples indicating the test should be interpreted with caution. Both VIL and echo-top trends on a case-by-case basis tend to not support H4 and are generally more consistent with the findings of Schneider and Sharp (2007) that there is no discernible pattern in echo tops for ALL TOR cases. One exception may be the significant increase found in Figure 3.10a from the -5 to 0-minute bin and the upward shift in the Figure 3.8a ALL TOR distributions which suggest the possibility of increasing echo tops with the TCTOR onset, although this is not as visible in the case-by-case trends. Another exception may be the maximum-VIL trends from the -15-minute bin, although the sample size is more limited.

3.6 V_{rot} Thresholds and Warning Skill

Nowotarski et al. (2021) selected NON TOR case volume scans that showed the strongest 0.5° rotation two volumes before and including each cell's first warning. ALL TOR cases were selected as the volume most representative of the TCTOR. They then calculated warning statistics for the TDS warning-decision scheme to examine the impacts of only warning for hypothetically more impactful events. Warning statistics are calculated on the automated-climatology dataset for each warning-decision scheme using similar volumes by selecting the strongest 0.5° rotation within the -10 to 0-minute bins before the tornado warning for NON TOR cases and the zero-minute bin for ALL TOR and TOR TDS cases. The best threshold for NON TOR versus ALL TOR is 37 knots (19.0 $\mathrm{m\,s^{-1}}$) with a TSS of ~0.06. Using this threshold shows lower FAR than than the Martinaitis (2017) V_{rot} criteria and observed warning skill. Meanwhile the Martinaitis (2017) V_{rot} yields the highest POD and CSI of the three schemes and a slightly lower FAR than the observed warning skill (Fig. 3.11a,b). As the threshold increases, POD begins to decrease without a significant reduction in FAR while CSI steadily decreases. These results differ from Nowotarski et al. (2021) in that their analysis of only Hurricane Harvey shows that POD begins decreasing at lower V_{rot} thresholds, FAR experiences a larger decrease, and CSI experiences a local maximum at 35 knots (18.0 $\mathrm{m\,s^{-1}}$).

Similar to Nowotarski et al. (2021), calculating the statistics as if only warning for TDS cases (Fig. 3.11c,d) leads to a larger POD at larger V_{rot} thresholds, but yields a large FAR and low CSI (likely due to the small number of TDS cases and consequently, potential for many false alarms). When comparing statistics for warnings targeted at EF1+ tornadoes, the best threshold is 36 knots (18.5 $\mathrm{m\,s^{-1}}$) with a TSS of ~0.19. This threshold yields a higher POD and FAR than the ALL TOR versus NON TOR cases and the highest CSI compared to the Martinaitis (2017) V_{rot} criteria and observed statistics. POD for these tornadoes also does not decrease at as low of a V_{rot} threshold as the POD for ALL TOR versus NON TOR. These results support H5 for the TDS and EF1+ cases with only a slightly larger CSI than the Martinaitis (2017) V_{rot} criteria at larger thresholds. The ALL TOR versus NON TOR comparison does not support H5 as the CSI is lower and the POD is

significantly lower than that found using the Martinaitis (2017) V_{rot} criteria.

Another method of determining which V_{rot} values may be best discriminators is to examine each individual time bin. Examination of Figure 3.6a suggests that the 0.5° -10 and -5-minute bins may have potential in discriminating between ALL TOR and NON TOR cases before the zero-minute bin. Indeed, the -10-minute bin yields the highest TSS with a score of 0.337 while maintaining a sufficient number of samples. Figure 3.12 shows the TSS and automated V_{rot} with range for the 0.5° elevation angle and -10-minute bin. The best threshold is 26 knots (13.4 $\mathrm{m\,s^{-1}}$) or an AzShear of ~$1 \times 10^{-2}\,\mathrm{s^{-1}}$. Examination of Figure 3.12b shows most ALL TOR cases, particularly TDS cases, are above this and the Martinaitis (2017) threshold.

Figure 3.13 shows the warning statistics for the same three warning-decision schemes as Figure 3.11, but at the -10-minute bin. For all three schemes, the best threshold is 26 knots (AzShear = ~$1 \times 10^{-2}\,\mathrm{s^{-1}}$) with TSS values ranging from 0.302 to 0.376. For the ALL TOR versus NON TOR scheme (Fig. 3.13a,b), the maximum CSI of ~0.47 is observed at the lowest threshold values. However, a secondary maximum of ~0.46 is observed at 24 knots (12.3 $\mathrm{m\,s^{-1}}$). In all three warning-decision schemes, CSI for the best threshold of 26 knots is larger than the Martinaitis (2017) V_{rot} criteria and, except for the TDS scheme, the observed CSI. Warning only for TDS (Fig. 3.13c,d) or EF1+ (Fig. 3.13e,f) cases also results in increased POD at higher thresholds, albeit with a higher FAR. All three schemes show in a local maximum of CSI at a higher threshold than the Martinaitis (2017) threshold. For the TDS and EF1+ schemes, the maximum occurs at 33 knots (17.0 $\mathrm{m\,s^{-1}}$) and 28 knots (14.4 $\mathrm{m\,s^{-1}}$) respectively. Both the 1.3° and 2.4° elevation angles show a best threshold of 19 knots (9.8 $\mathrm{m\,s^{-1}}$; AzShear = ~$8 \times 10^{-3}\,\mathrm{s^{-1}}$) and 21 knots (10.8 $\mathrm{m\,s^{-1}}$; AzShear = ~$8 \times 10^{-3}\,\mathrm{s^{-1}}$) at the -10-minute bin with CSI values of ~0.38 and ~0.31 respectively (not shown). Therefore, lower values of rotation at higher elevation angles should be examined as well.

When comparing cases 10 minutes before the time of maximum NON TOR rotation or TC-TOR start time, results from all three schemes support H5 as CSI increases at higher thresholds. However, this conclusion is nuanced as the marginal increase in CSI and reduced POD without a

notable reduction in FAR observed in Figure 3.11 and lower FAR and POD in Figure 3.13 suggest that these higher thresholds may not be desirable. The cost and benefits of these hypothetical warning-decision schemes are beyond the scope of this work. Finally, some cases show relatively weak rotation. While this is expected, it is possible that some of these values are due to tracking errors and are underestimated. Assuming that an ALL TOR case is more likely to be underestimated than a NON TOR case, the warning skills presented here may be pessimistic. Moreover, the range criteria chosen by Martinaitis (2017) was chosen somewhat arbitrarily and may not fit for all TCs. Determining the ideal set of range bins is beyond the scope of this work, however.

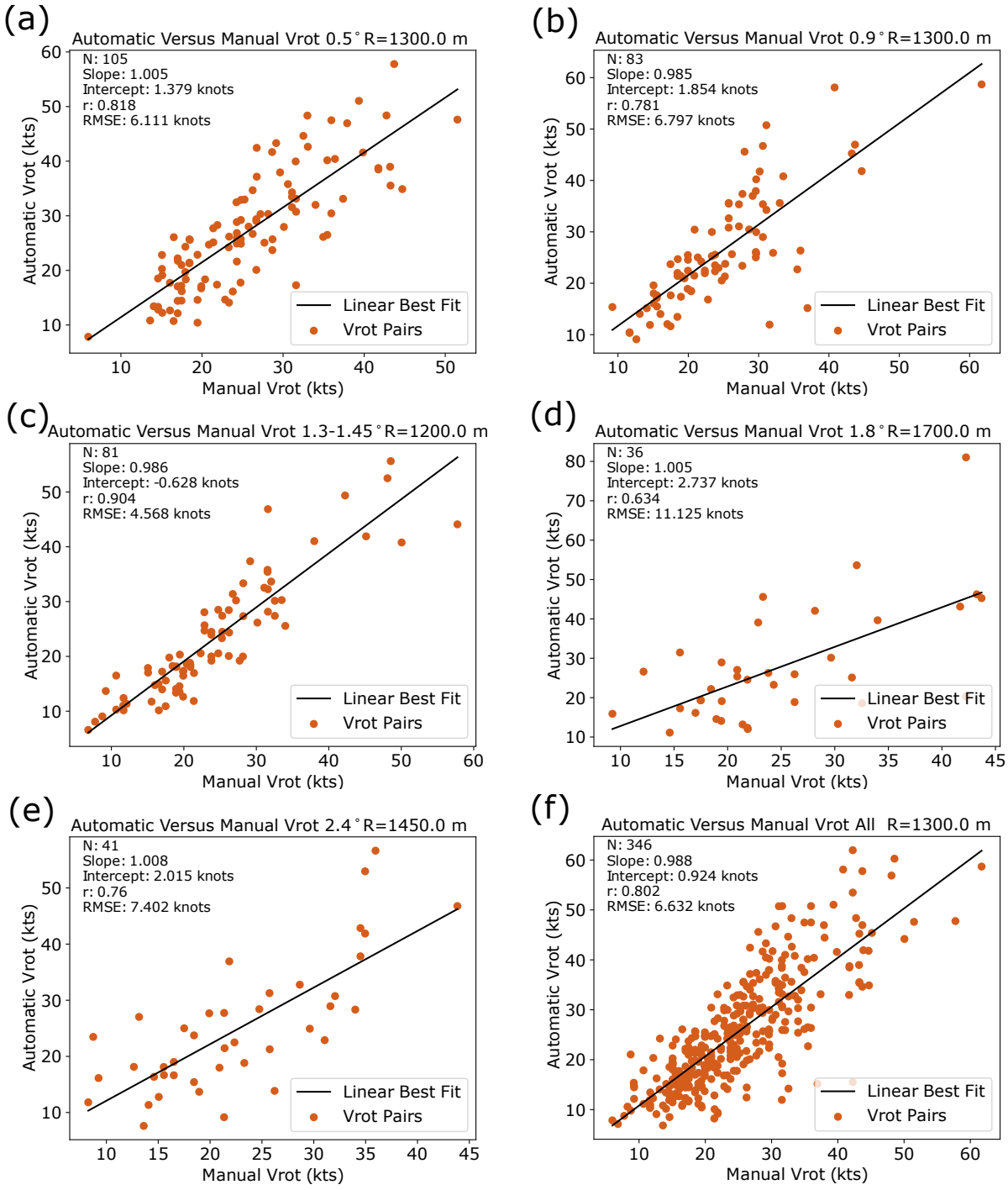

Figure 3.1: Automatically derived V_{rot} values using AzShear from the Nowotarski et al. (2021) cases compared to the manually derived values **(a)** the 0.5°, **(b)** the 0.9°, **(c)** 1.3/1.45°, **(d)** 1.8°, **(e)** 2.4° tilts and **(f)** samples at all tilts. The radius (R) used for each tilt is the one that yielded a best-fit slope closest to 1.0 when examined in 50-m increments. The number of samples (N), slope and intercept of the linear best fit line, Pearson correlation coefficient (r), and root-mean-square-error (RMSE) with respect to the best-fit line are shown in the upper-left corner of each figure.

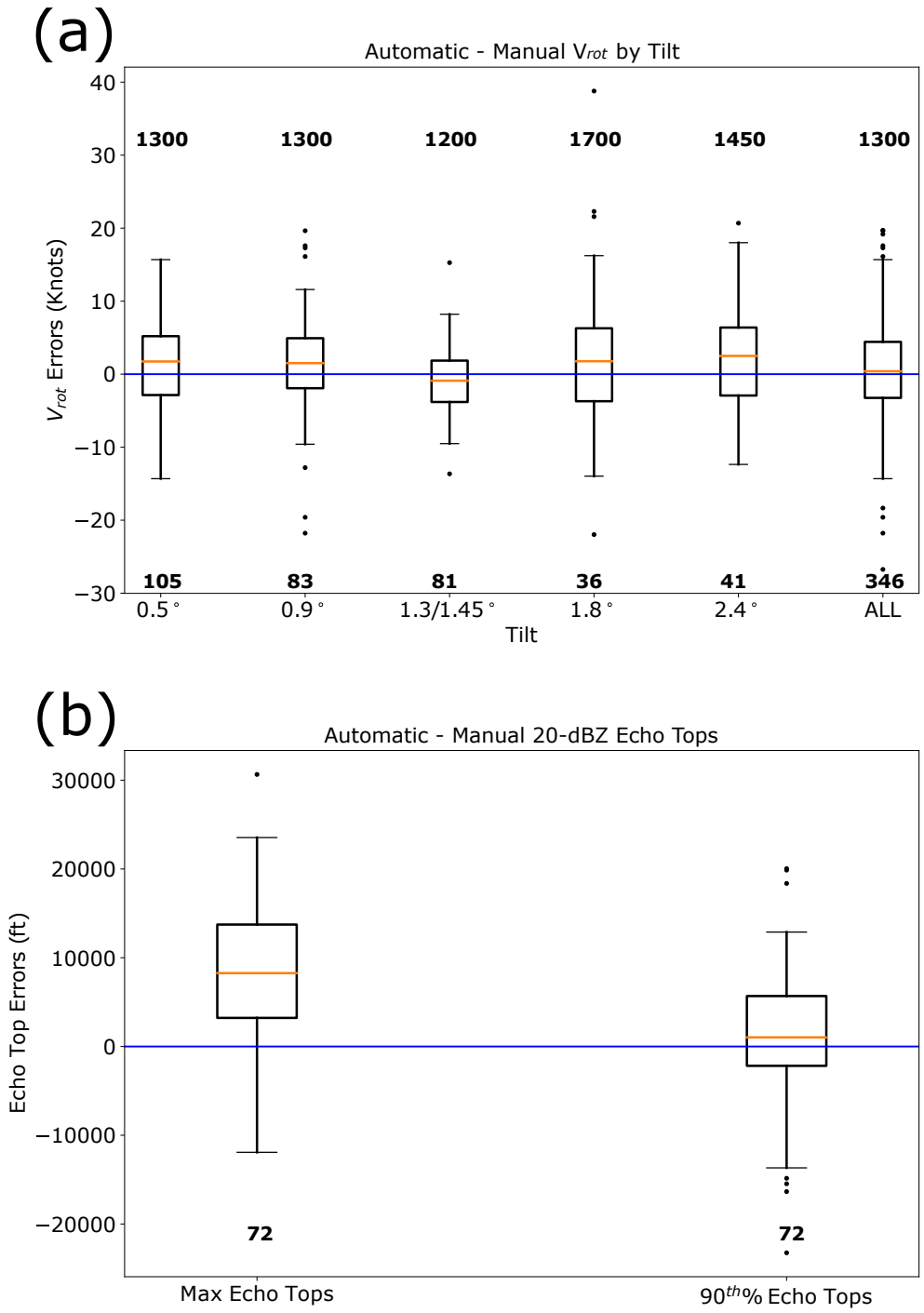

Figure 3.2: Box-and-whisker plots showing the errors for **(a)** automatic (automated) V_{rot} - manual V_{rot} for each tilt and all tilts combined. The sample size is shown in bold underneath each box and the radius used to calculate V_{rot} (yields best fit closest to 1) is shown above each box in meters. **(b)** The maximum and 90^{th}-percentile echo tops minus the manual analysis. The sample size is also shown under each box.

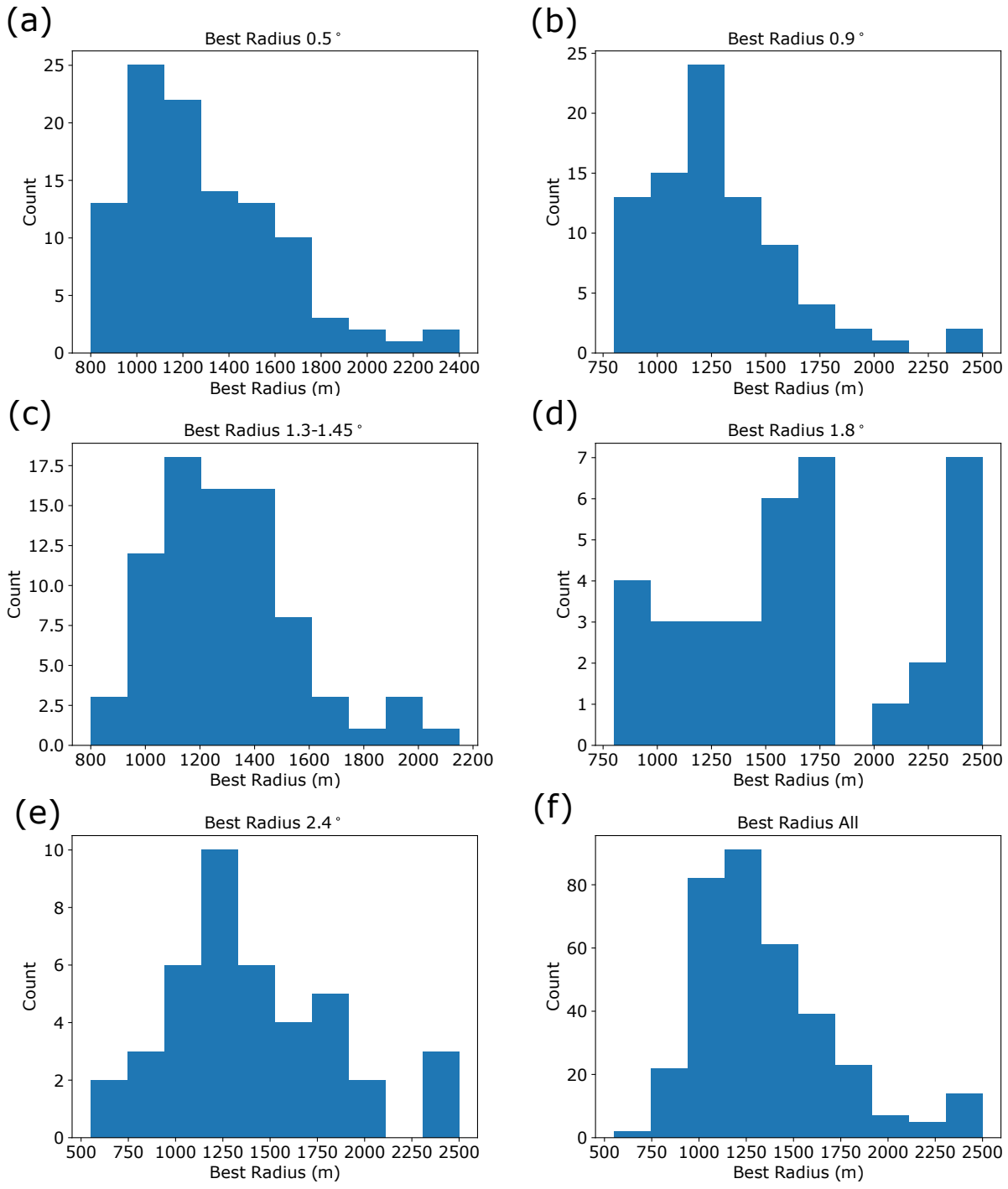

Figure 3.3: Histograms showing the radius that resulted in the smallest error between the automatically and manually derived V_{rot} values for **(a)** the 0.5°, **(b)** 0.9°, **(c)** 1.3/1.45°, **(d)** 1.8°, **(e)** 2.4°, and **(f)** all tilts.

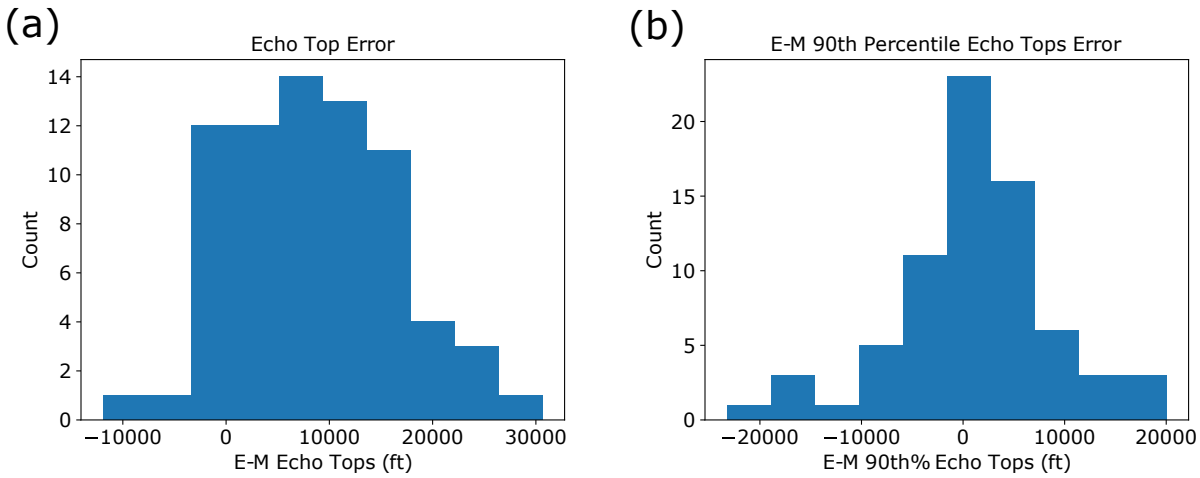

Figure 3.4: Histograms showing the echo-top errors for manually derived values (M) subtracted from automatically derived values (E) for **(a)** the cluster maximum echo-top and **(b)** the cluster 90^{th}-percentile echo top.

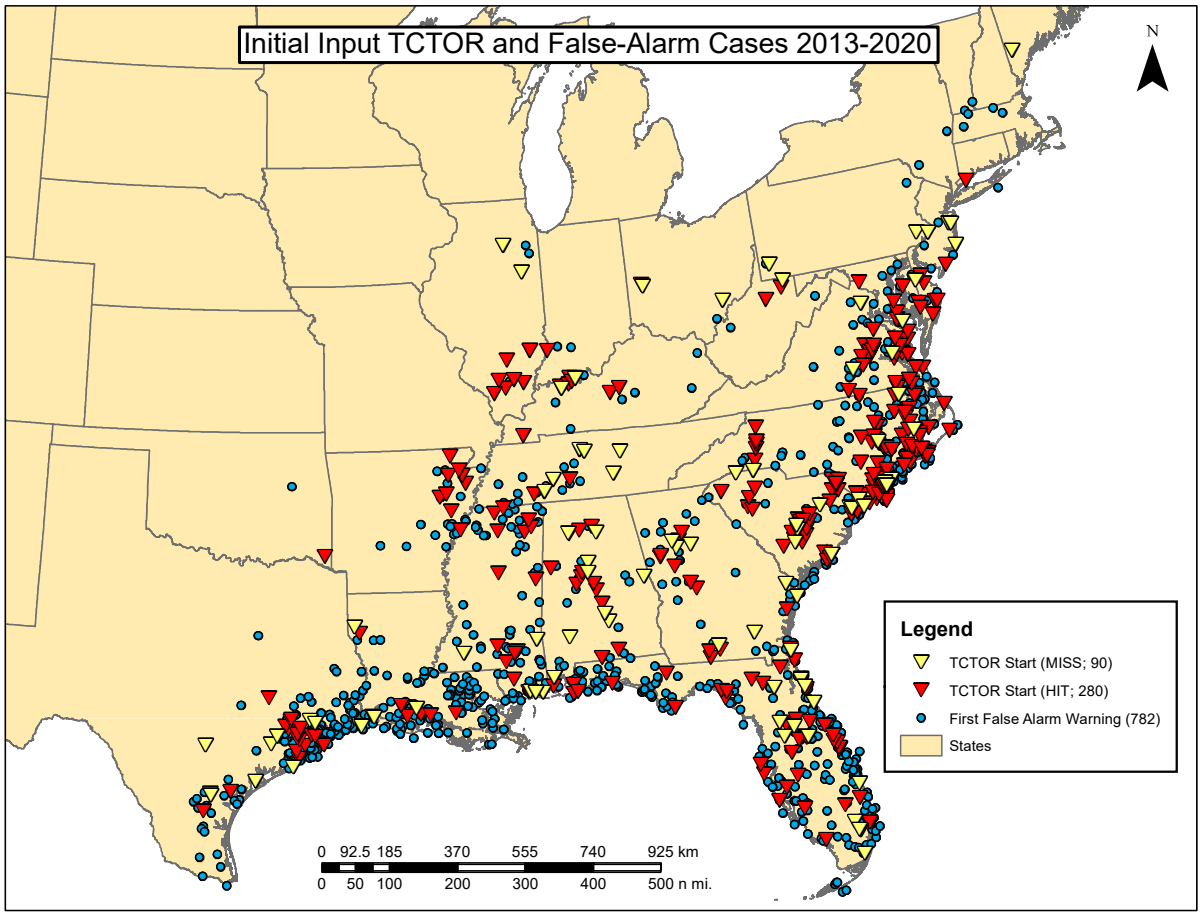

Figure 3.5: A map showing the initial ALL TOR and NON TOR cases from 2013-2020. The triangles represent TCTOR cases where red indicates the TCTOR is within a warning at the TCTOR start time and yellow is not within a warning. The blue dots represent the warning feature points of the first warning of each NON TOR cell.

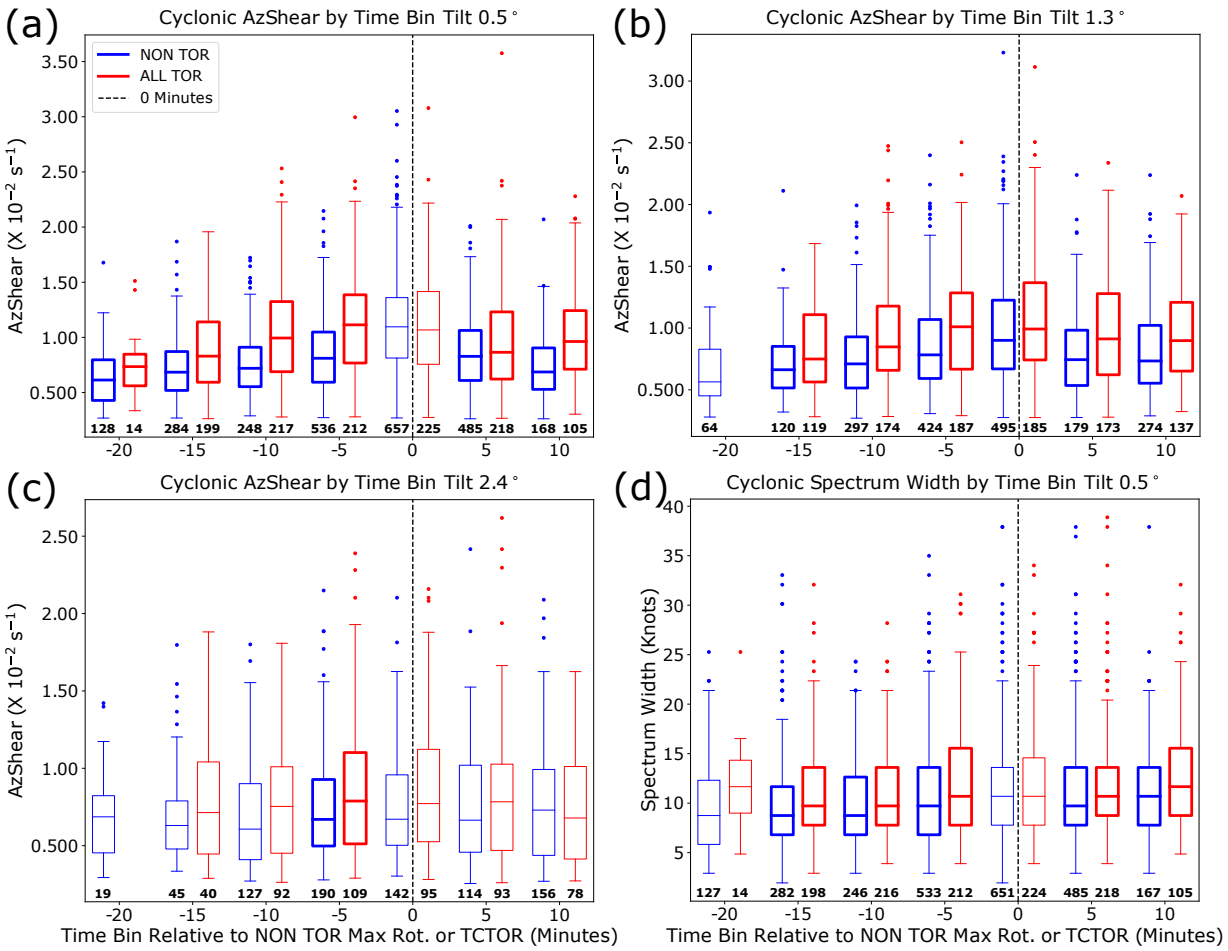

Figure 3.6: Box-and-whisker plots showing **(a)** AzShear cluster maximum AzShear at 0.5° from only cases having cyclonic rotation. Bold boxes indicate where the NON TOR distribution is less than the ALL TOR distribution at a 0.05 significance level using a Mann-Whitney-U test. The sample sizes of each distribution are shown in bold along the bottom. **(b)** Same as in **(a)**, but at the 1.3° elevation angle. **(c)** Same as in **(a)**, but at 2.4°. **(d)** Same as in **(a)**, but with spectrum width.

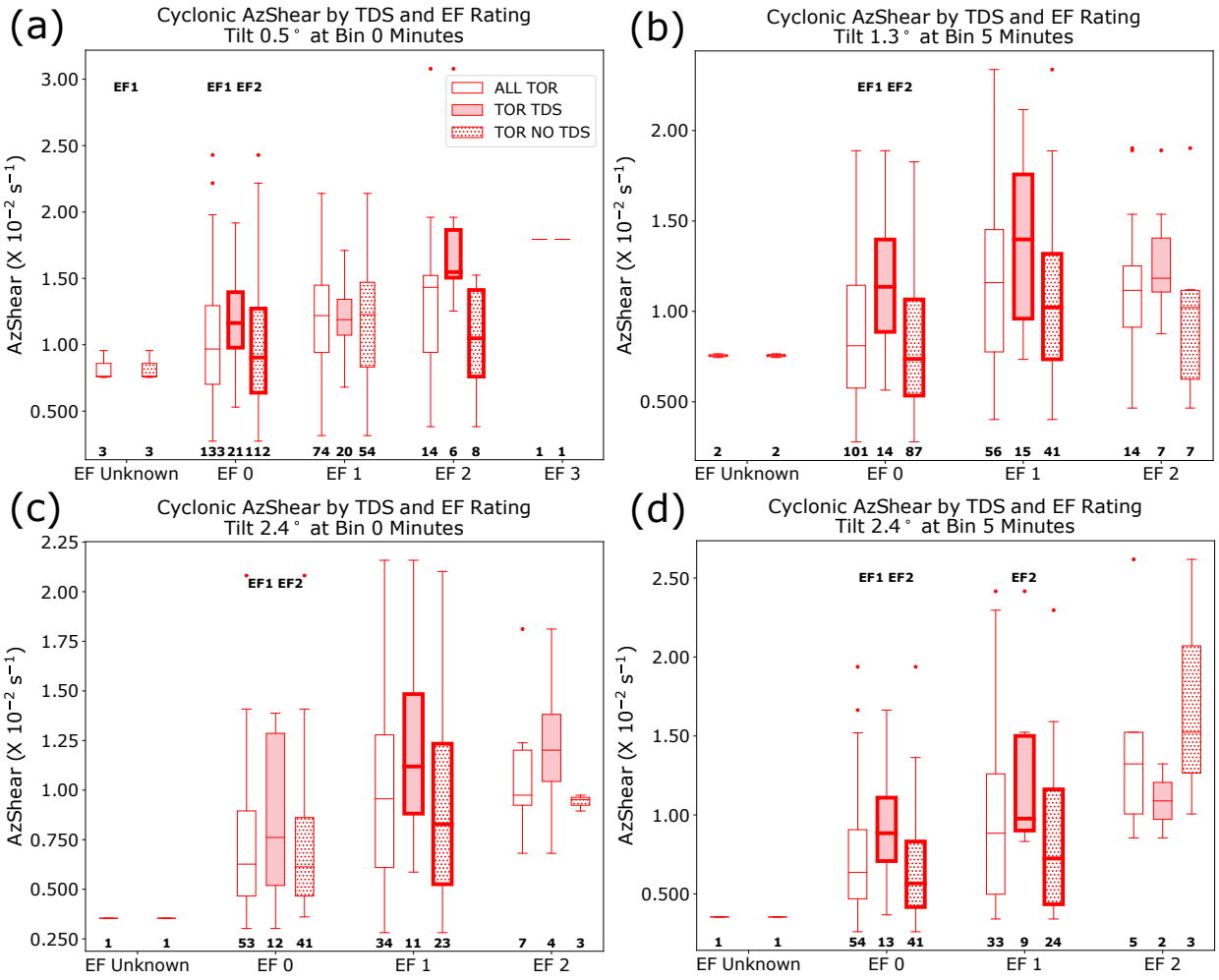

Figure 3.7: Box-and-whisker plots of AzShear cluster maximum AzShear for only cases having cyclonic rotation by EF rating for **(a)** 0.5° at bin zero minutes, **(b)** 1.3° at five minutes after, **(c)** 2.4° at zero minutes, and **(d)** five minutes after. Bold TOR TDS and TOR NO TDS boxes indicate that the TOR NO TDS boxes are significantly less than the TDS boxes at the 0.05 significance level using a Mann-Whitney-U test. The EF ratings above a box group indicate which ratings that group's ALL TOR distribution is significantly less than using the Mann-Whitney-U test with a 0.05 significance level.

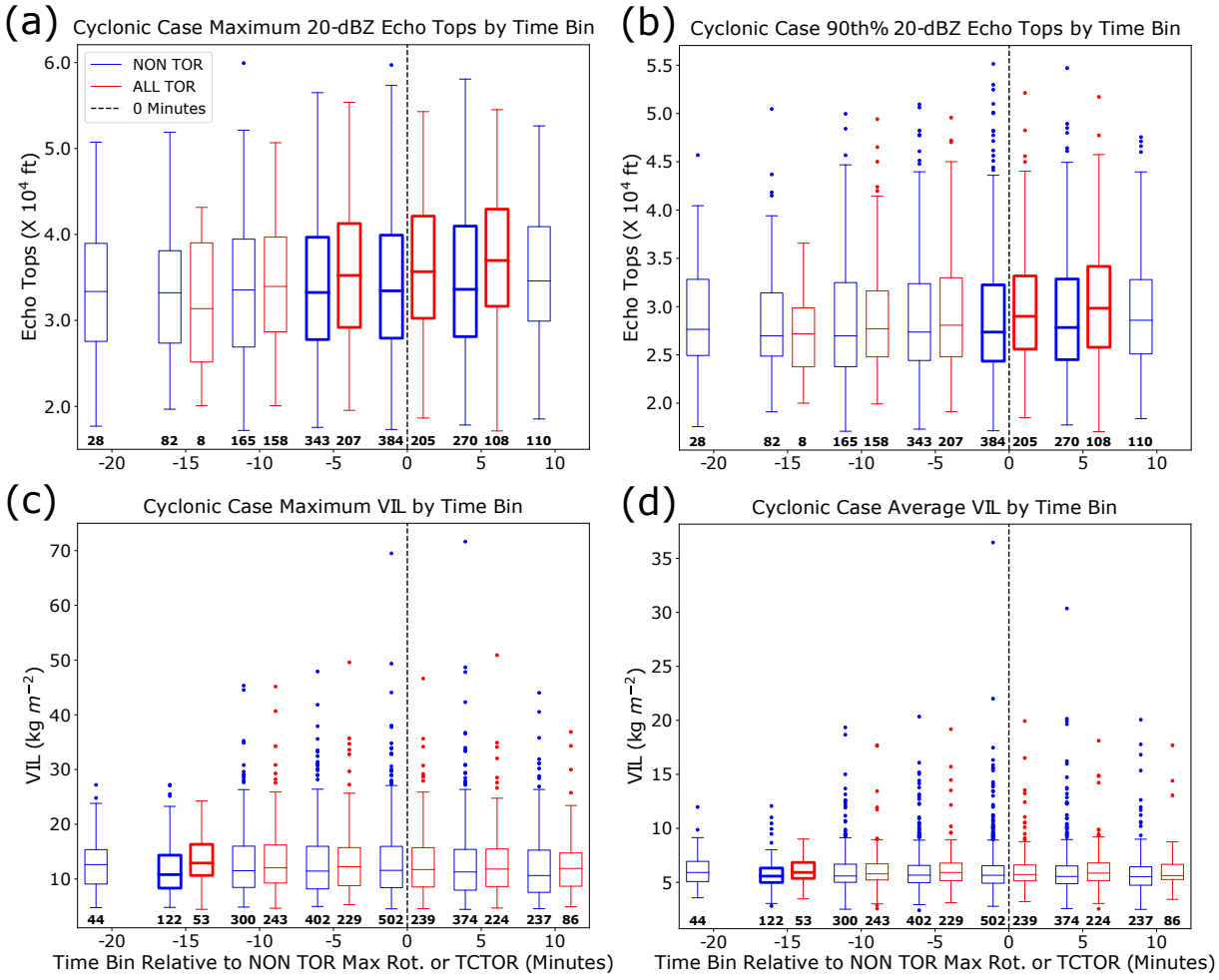

Figure 3.8: As in Figure 3.6 except with **(a)** 20-dBZ echo-top cluster maximum and **(b)** 90^{th}-percentile 20-dBZ echo tops and **(c)** maximum and **(d)** average VIL-cluster VIL. Bold boxes indicate where the ALL TOR and NON TOR distributions are significantly different using a two-tailed Mann-Whitney-U test with a 0.05 significance level.

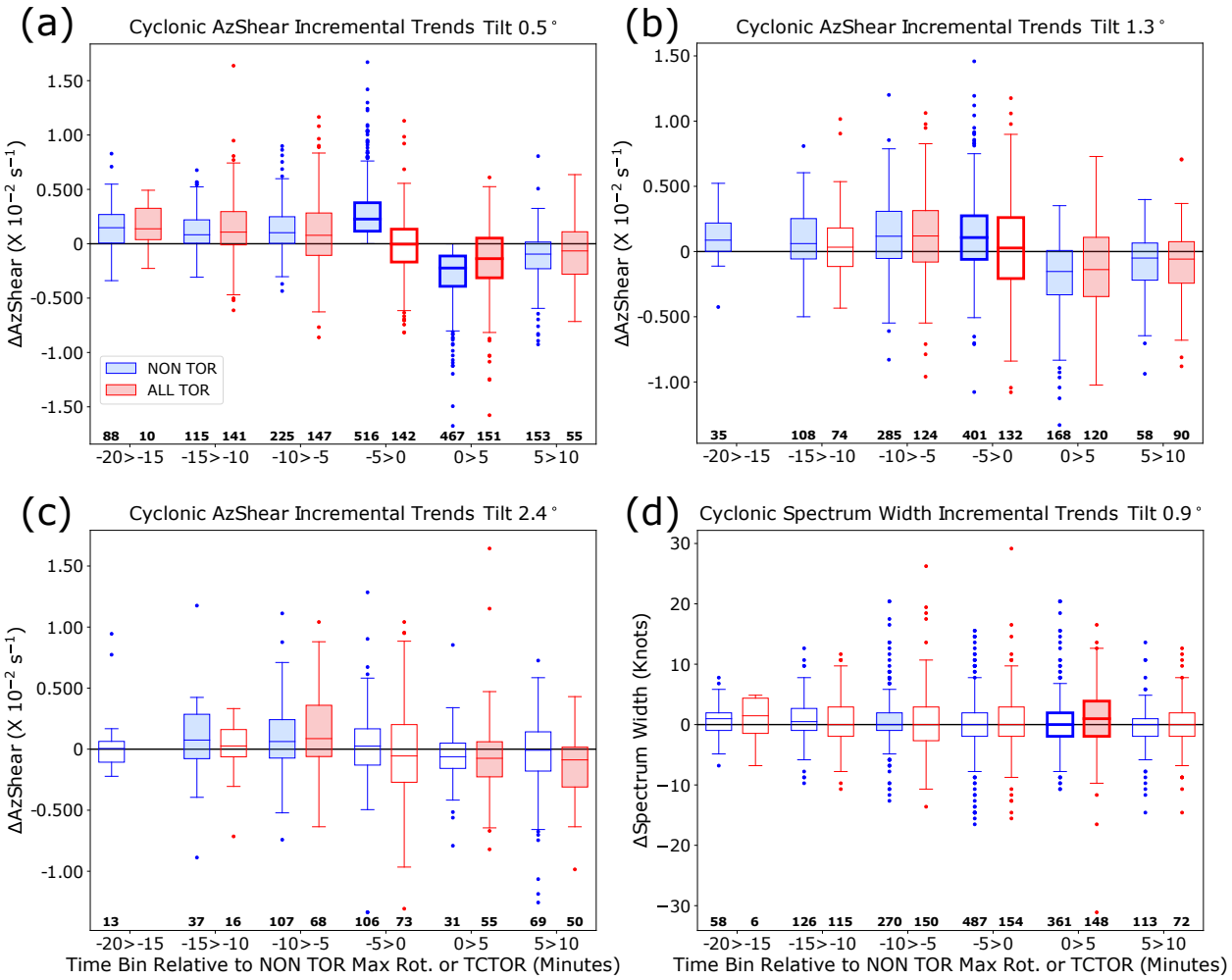

Figure 3.9: Trends between sequential time bins for **(a)** 0.5°, **(b)** 1.3°, and **(c)** 2.4° cyclonic-cluster maximum AzShear and **(d)** cyclonic-case 0.9° spectrum width. The x-axis shows the two bins between which the change is calculated. Bold boxes show where the ALL TOR and NON TOR distributions are significantly different using a two-sided Mann-Whitney-U test. Shaded boxes indicate that the particular distribution is significantly different from zero using a Wilcoxon Signed Rank test. Both tests are performed at the 0.05 significance level.

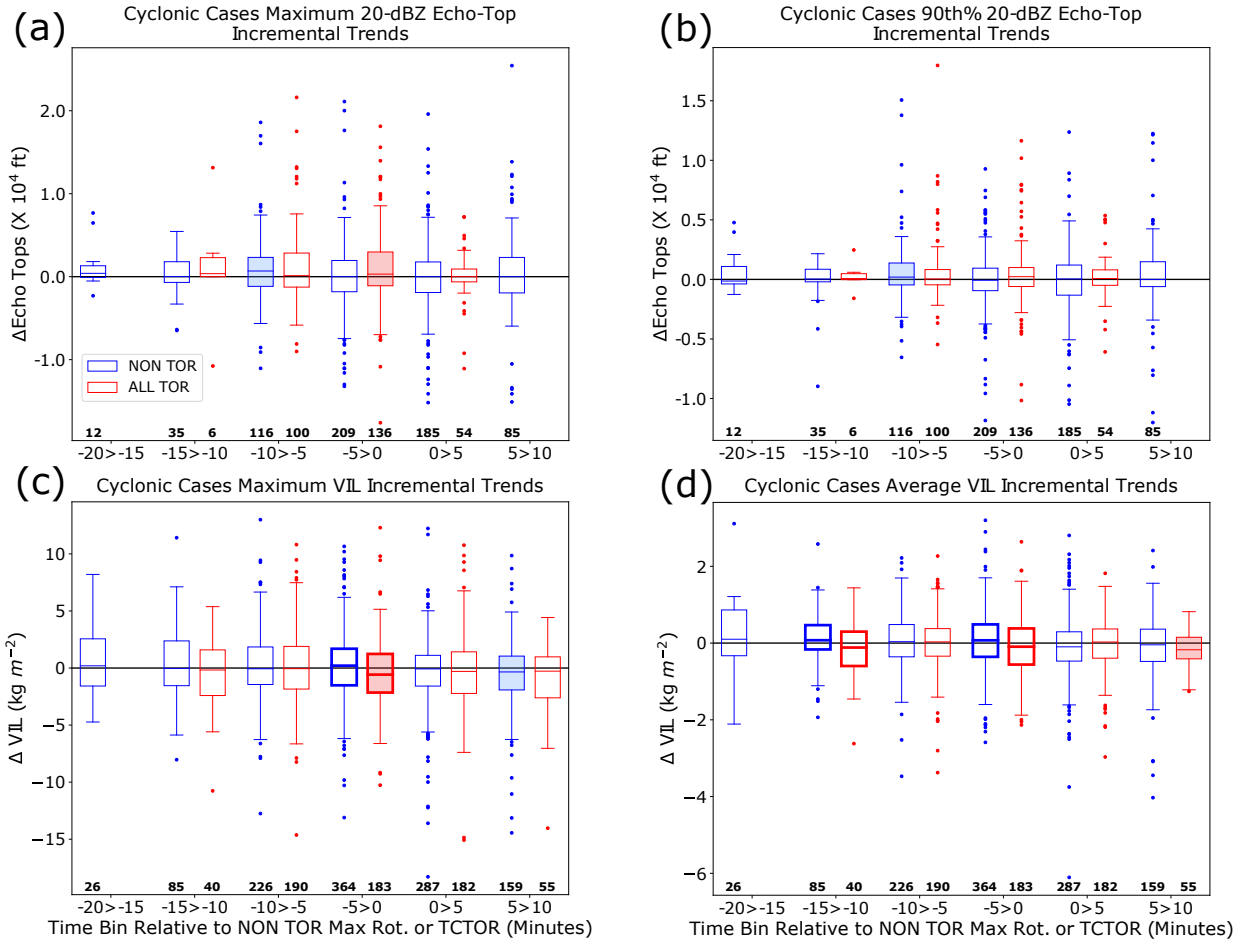

Figure 3.10: As in Figure 3.9 except for **(a)** 20-dBZ echo-top cluster maximum and **(b)** 90^{th}-percentile 20-dBZ echo tops and **(c)** maximum and **(d)** average VIL-cluster VIL.

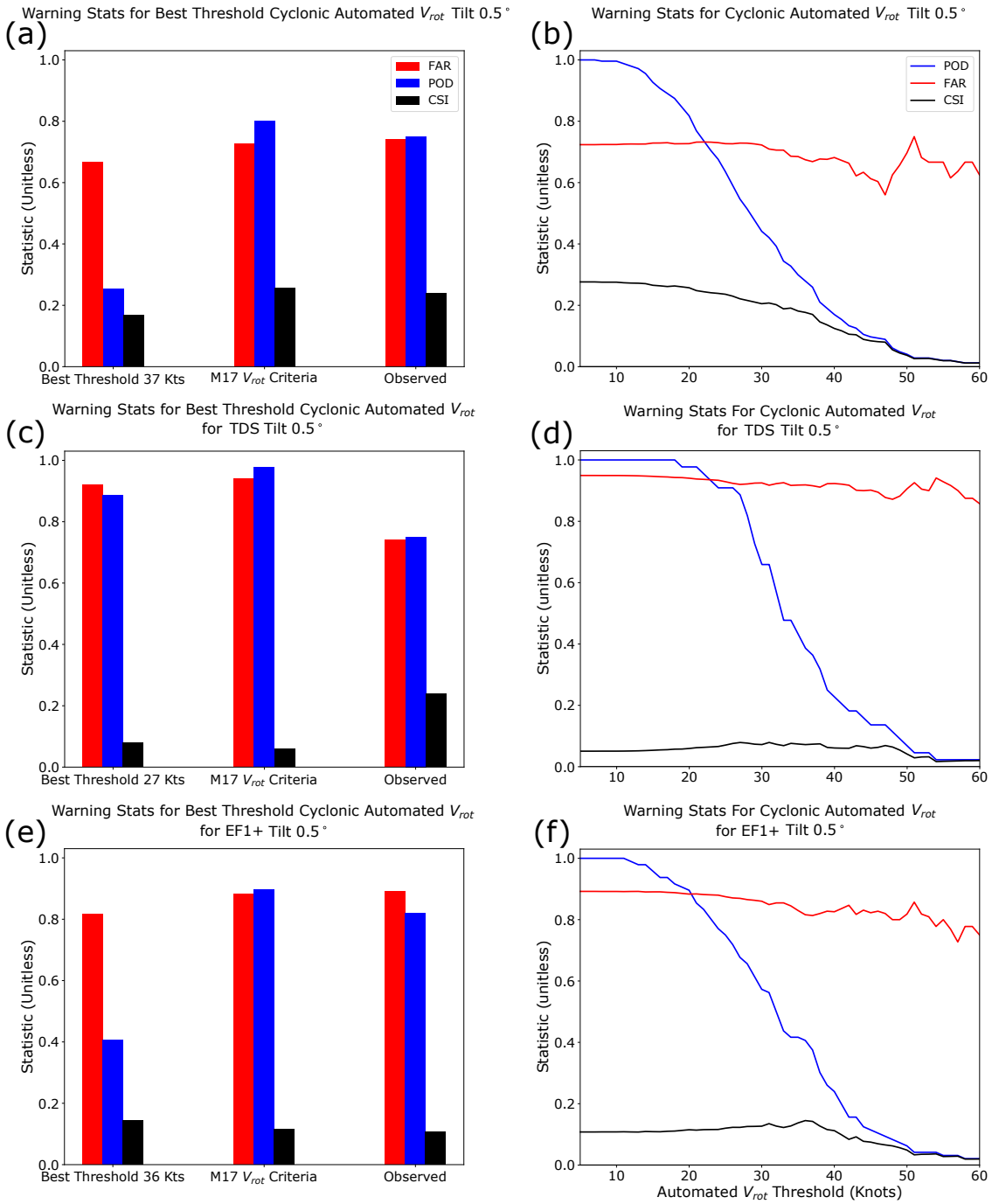

Figure 3.11: Warning statistics for the NON TOR volume scan with maximum rotation in the -10-0-minute bins before the tornado warning and ALL TOR zero-minute bin. **(a)** Warning statistics for the best threshold, M17 V_{rot} criteria, and observed statistics. **(b)** Warning statistics by automatic-V_{rot} threshold. **(c)** and **(d)** as in **(a)** and **(b)** except calculated for a hypothetical warning scheme designed to warn only TDS cases. **(e)** and **(f)** as in **(a)** and **(b)** except calculated for a warning scheme designed to warn only EF1+ cases. Tornadoes without a known rating are excluded from EF1+ and TDS statistics.

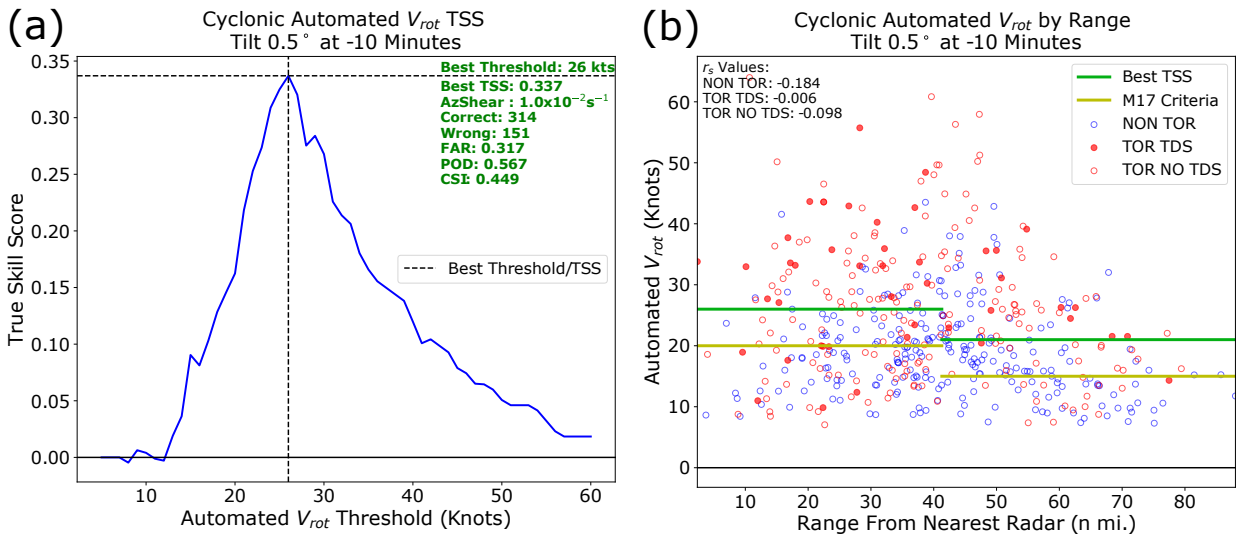

Figure 3.12: **(a)** True Skill Score (TSS) as a function of an automatic V_{rot} threshold for ALL TOR against NON TOR cases at the -10-minute bin. The threshold with the highest TSS is highlighted and warning statistics at that threshold are shown in the green text. Correct indicates the number of ALL TOR cases above and NON TOR cases below the threshold. Wrong indicates the number of ALL TOR below and NON TOR above. **(b)** Automatic V_{rot} as a function of range from the nearest radar at the -10-minute bin. The Martinaitis (2017) criteria are shown by the yellow lines and the best-TSS threshold is shown by the green line. Spearman Rank correlation coefficients for each category are shown in the upper-left (Wilks 2019).

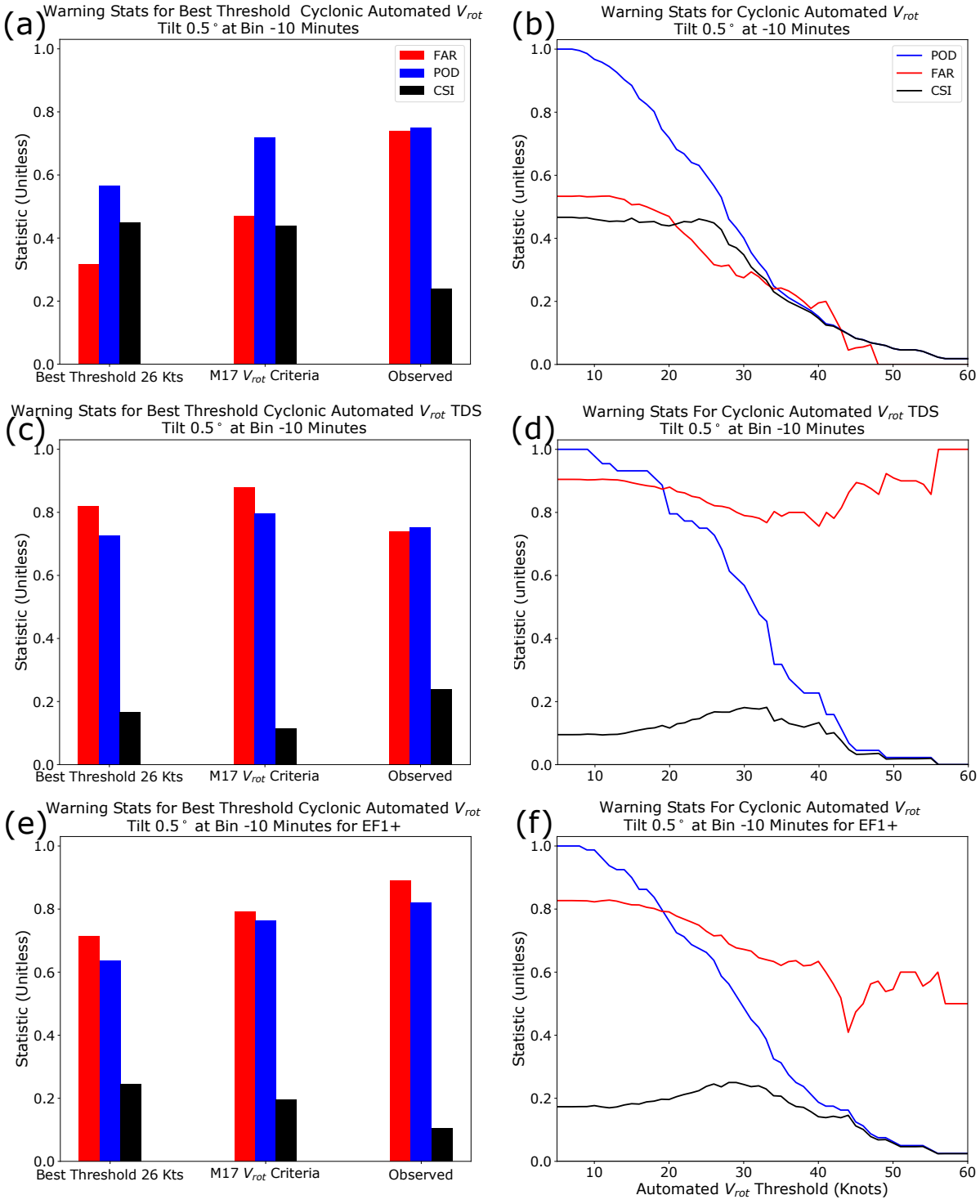

Figure 3.13: As in Figure 3.11 but at the -10-minute bin.

4. SUMMARY AND CONCLUSIONS

Previous studies have examined radar attributes of tornadic and nontornadic cells in landfalling TC rainbands, however they have often been limited given sample size constraints of a manual analysis. Thus, this project sought to develop an automated method for calculating radar attributes and apply it to a larger dataset of tornadic and nontornadic cells. A total of 334 tornadic and 721 nontornadic, but tornado-warned, storms from 29 tropical cyclones from 2013 to 2020 are processed using the Warning Decision Support System-Integrated Information algorithms. Areas of enhanced rotation below 10 kft (3.0 km) above radar level, vertically integrated liquid, and 20-dBZ echo tops are tracked before and after each warning or tornado start time and placed into 5-minute bins relative to their respective times. The times for nontornadic cases are adjusted such that the zero-minute bin corresponds to the time of maximum low-level rotation. Radar attributes and their trends from each of these areas are recorded and used in subsequent analysis. Tornadic cases are then classified as TOR TDS or TOR NO TDS depending on if their low-level rotation contains a minimum number of pixels that meet TDS criteria.

A comparison of the automated method to the manual analysis shows that the automated V_{rot} values tend to be within about 5 knots (2.6 $\mathrm{m\,s^{-1}}$) of their manual counterparts, with a maximum error of about 26.7 knots (13.7 $\mathrm{m\,s^{-1}}$). Echo tops from the automated analysis tend to overestimate the manual analysis, while the 90^{th}-percentile values are closer to the manual analysis. Generally, the automated and manual methods compare well. The following conclusions are made from the results of the 2013-2020 automated analysis.

- At low elevation angles, ALL TOR cases tend to have stronger rotation than NON TOR cases at most times. Moreover, ALL TOR cases tend to reach larger values of rotation sooner than NON TOR cases. Examination of previous volume scans for continuity of enhanced rotation values, especially with AzShear values of 8×10^{-3} to $1 \times 10^{-2}\,\mathrm{s^{-1}}$, may indicate cells that are more likely to be tornadic. The pattern of enhanced rotation for ALL TOR

cases is present at higher elevation angles with less pronounced differences and statistically significant bins. Spectrum-width values are slightly higher in ALL TOR cases than NON TOR cases at 0.5°, although significant overlap in the distributions exist and would not make a robust discriminator. These results support H1, although limited to lower-level elevation angles and rotation more so than spectrum width.

- Rotation is also found to be stronger in higher-rated tornadoes, particularly starting at the time of the TCTOR. Moreover, TOR TDS cases tend to have stronger rotation than TOR NO TDS cases for a given EF rating at times. While these findings may indicate that stronger rotation is associated with potentially more impactful tornadoes, a larger sample size of higher-rated TCTORs and robust TDS detection is needed to solidify this finding within the TCTOR context and fully support H2.

- Echo tops are slightly larger in ALL TOR than NON TOR cases near the zero-minute bin. While these bins technically do not support H3, significant overlap still exists between the distributions such that there is not much operational relevance to the differences in the distributions. Maximum and average VIL do not show any significant differences between ALL TOR and NON TOR cases with the exception of one time bin, supporting H3.

- Trends in AzShear at low levels show mostly increasing rotation for both ALL TOR and NON TOR cases until the zero-minute bin where NON TOR cases experience an expected sharp increase followed by a sharp decrease in rotation whereas ALL TOR cases appear to stop increasing. In general, there do not appear to be significant differences between trends in rotational magnitude except near the zero-minute bin. Trends in spectrum width are also not supportive of H4 with most time bins showing no significant difference from zero or between ALL TOR and NON TOR cases. Similar to spectrum width, echo tops and VIL show only very few differences between ALL TOR and NON TOR distributions where at least one distribution is significantly different from zero. Therefore, H4 is generally not supported with the exception of lower elevation-angle AzShear near the zero-minute

bin, some maximum-VIL trends from the -15-minute bin (although with a more limited sample size), and -5 to 0-minute bin where the ALL TOR maximum VIL experiences a slight decrease whereas the NON TOR VIL does not experience a significant change.

- Calculations of the false alarm ratio (FAR), probability of detection (POD), and Critical Success Index (CSI) show that, when only potentially more impactful events (tornadoes with a TDS or EF1 or higher rating) are warned, the warning statistics only marginally improve compared to the Martinaitis (2017) V_{rot} criteria. These results are valid for when cases are selected for NON TOR maximum rotation within ten minutes of and including the tornado warning or the TCTOR start time. However, warning only for more impactful tornadoes increases POD. Focusing on the period 10 minutes before the tornado or NON TOR maximum rotation yields higher CSI values than the Martinaitis (2017) V_{rot} criteria, but with lower FAR and POD. This time bin is likely of more relevance to the warning-decision process as a decision to issue a warning at this time would provide a lead time of about 10 minutes. Maximum CSI values and the best threshold based on the True Skill Score occur with values greater than that of Martinaitis (2017) indicating threshold values of 26 knots (13.4 $\mathrm{m\,s^{-1}}$; AzShear = ~$1 \times 10^{-2}\,\mathrm{s^{-1}}$) for cases within 40 n mi. (74.1 km) of the nearest radar or 21 knots (10.8 $\mathrm{m\,s^{-1}}$; AzShear = ~$8 \times 10^{-3}\,\mathrm{s^{-1}}$) for ranges beyond 40 n mi. (74.1 km) may yield the best overall results, but the resulting reduction in POD and potential dataset noise should be considered before utilizing these thresholds. These results are supportive of H5 in that the best discriminating threshold is larger than that of the Martinaitis (2017) V_{rot} criteria.

This study shows that TCTORs tend to have stronger rotation over a longer period of time, increasing over time, and reaching a maximum just before the TCTOR start time whereas NON TOR and ALL TOR (not shown) cases tend to experience a sudden increase and decrease around the time of maximum rotation. The best automatic V_{rot} threshold is 26 knots (AzShear = ~$1 \times 10^{-2}\,\mathrm{s^{-1}}$), 10 minutes before the TCTOR or maximum NON TOR rotation. At this threshold CSI increases, but FAR and POD decrease. Warning only for cases that meet the TDS criteria or are

at least an EF1 or higher-rated tornado would increase the POD. Echo tops, VIL, and trends in the examined attributes do not show much separation between ALL TOR and NON TOR cases with the exception of 0.5° rotation near the zero-minute bin. The recommendation to consider the temporal continuity of the cells is consistent with Spratt et al. (1997) who suggested cells with persistent shear, including weak shear, should be suspect. McCaul et al. (2004) also suggested that long-lived cells in a favorable environment should be monitored. The best-threshold (in terms of the TSS) rotation values found at the -10-minute bin are less than those found by Nowotarski et al. (2021). One possible reason for this difference is the lowered rotation strength at earlier times resulting in overall lower thresholds whereas the Nowotarski et al. (2021) methodology compared a volume representative of the tornado time (likely stronger rotation) to a maximum NON TOR rotation.

Despite a large sample size, climatologies of TCTOR radar attributes could still benefit from future work. First, events from future years should be added to the dataset. The Edwards and Mosier (2022) TCTOR database is updated annually allowing for additional events to be processed and added to the climatology each year. The automation of the process of cataloging warnings and quality controlling TCTOR events could be explored, although doing so may introduce its own set of limitations.

This analysis is meant to replicate the analysis performed by Nowotarski et al. (2021), but in an automated fashion. The V_{rot} component of that analysis is estimated from AzShear which in itself is not a perfectly accurate estimate. Additionally, several attributes such as bounded weak echo regions, hook echoes, VES, and K_{DP}/Z_{DR} separation signatures are not included in this study. Automated methods of extracting K_{DP}/Z_{DR} and other radar attributes have been developed (e.g., Wilson and Broeke 2021) and could be incorporated. However, it is not clear how a fixed set of algorithm parameters would perform globally for all cases in a TCTOR context. Finally, future work on this project could benefit from the exploration of different, potentially more robust tracking methodologies (e.g., Multiple Hypotheses Tracking; Reid 1979; Lakshmanan et al. 2013b; Miller et al. 2013).

REFERENCES

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Brown, R. A., L. R. Lemon, and D. W. Burgess, 1978: Tornado detection by pulsed doppler radar. *Mon. Wea. Rev.*, **106**, 29–38, https://doi.org/10.1175/1520-0493(1978)106<0029:TDBPDR>2.0.CO;2.

Burrow, D., K. Ellis, and L. Tran, 2021: Simultaneous and collocated tornado and flash flood warnings associated with tropical cyclones in the contiguous United States. *International Journal of Climatology*, **41**, 4253–4264, https://doi.org/10.1002/joc.7071.

Chrisman, J., 2014: Multiple elevation scan option for SAILS (MESO-SAILS). NOAA, 27 pp, https://www.roc.noaa.gov/wsr88d/PublicDocs/NewTechnology/MESO-SAILS_Description_Briefing_Jan_2014.pdf.

Cifelli, R., S. W. Nesbitt, S. A. Rutledge, W. A. Petersen, and S. Yuter, 2007: Radar characteristics of precipitation features in the EPIC and TEPPS regions of the East Pacific. *Mon. Wea. Rev.*, **135**, 1576–1595, https://doi.org/10.1175/MWR3340.1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, L. Cronce, and J. Brunner, 2020: NOAA ProbSevere v2.0-ProbHail, ProbWind, and ProbTor. *Wea. Forecasting*, **35**, 1523–1543, https://doi.org/10.1175/WAF-D-19-0242.1.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous united states. *Wea. Forecasting*, **27**, 1235–1248, https://doi.org/10.1175/WAF-D-11-00151.1.

Cressman, G. P., 1959: An operational objective analysis system. *Mon, Wea. Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Crowe, C. C., L. D. C. Walter A. Petersen, and D. J. Cecil, 2010: A dual-polarization investigation of tornado-warned cells associated with Hurricane Rita (2005). *Electron. J. Severe Storms*

Meteor., **2010-EJ4**, http://nwafiles.nwas.org/ej/pdf/2010-EJ4.pdf.

Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.

Davis, C. A., B. G. Brown, R. Bullock, and J. Halley-Gotway, 2009: The method for object-based diagnostic evaluation (mode) applied to numerical forecasts from the 2005 NSSL/SPC spring program. *Wea. Forecasting*, **24**, 1252–1267, https://doi.org/10.1175/2009WAF2222241.1.

Davis, J. M., and M. D. Parker, 2014: Radar climatology of tornadic and nontornadic vortices in high-shear, low-cape environments in the Mid-Atlantic and Southeastern United States. *Wea. Forecasting*, **29**, 828–853, https://doi.org/10.1175/WAF-D-13-00127.1.

Doviak, R. J., and D. S. Zrnic, 1993: *Doppler Radar and Weather Observations*. 2nd ed., Academic Press, 562 pp.

Edwards, R., 2012: Tropical cyclone tornadoes: A review of knowledge in research and prediction. *Electron. J. Severe Storms Meteor.*, **7**, https://ejssm.org/archives/wp-content/uploads/2021/09/vol7-6.pdf.

Edwards, R., A. R. Dean, R. L. Thompson, and B. T. Smith, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part III: Tropical cyclone tornadoes. *Wea. Forecasting*, **27**, 1507–1519, https://doi.org/10.1175/WAF-D-11-00117.1.

Edwards, R., and R. M. Mosier, 2022: Over a quarter century of TCTOR: Tropical cyclone tornadoes in the WSR-88D era. *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 171, https://ams.confex.com/ams/30SLS/mediafile/Manuscript/Paper407018/27yr-sls.pdf.

Edwards, R., and J. C. Picca, 2016: Tornadic debris signatures in tropical cyclones. *28th Conf. on Severe Local Storms*, Portland, OR, Amer. Meteor. Soc., 162, https://ams.confex.com/ams/28SLS/webprogram/Paper300633.html.

Greene, D. R., and R. A. Clark, 1972: Vertically integrated liquid watera new analysis tool. *Mon. Wea. Rev.*, **100**, 548–552, https://doi.org/10.1175/1520-0493(1972)100<0548:VILWNA>2.3.CO;2.

Harrison, D., 2018: Correcting, improving, and verifying automated guidance in a new warning paradigm. M.S. thesis, School of Meteorology, University of Oklahoma, 109 pp.

Heymsfield, G. M., L. Tian, A. J. Heymsfield, L. Li, and S. Guimond, 2010: Characteristics of deep tropical and subtropical convection from nadir-viewing high-altitude airborne doppler radar. *J. Atmos. Sci*, **67**, 285–308, https://doi.org/10.1175/2009JAS3132.1.

Kumjian, M. R., 2013: Principles and applications of dual-polarization weather radar. Part I: Description of the polarimetric radar variables. *J. Operational Meteor.*, **1**, 226–242, http://dx.doi.org/10.15191/nwajom.2013.0119.

Lakshmanan, V., B. Herzog, and D. Kingfield, 2015: A method for extracting postevent storm tracks. *J. Appl. Meteor. Climatol.*, **54**, 451–462, https://doi.org/10.1175/JAMC-D-14-0132.1.

Lakshmanan, V., K. Hondl, C. K. Potvin, and D. Preignitz, 2013a: An improved method for estimating radar echo-top height. *Wea. Forecasting*, **28**, 481–488, https://doi.org/10.1175/WAF-D-12-00084.1.

Lakshmanan, V., K. Hondl, and R. Rabin, 2009: An efficient, general-purpose technique for identifying storm cells in geospatial images. *J. Atmos. Oceanic Technol.*, **26**, 523–537, https://doi.org/10.1175/2008JTECHA1153.1.

Lakshmanan, V., and T. W. Humphrey, 2014: A MapReduce technique to mosaic continental-scale weather radar data in real-time. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7**, 721–732, https://doi.org/10.1109/JSTARS.2013.2282040.

Lakshmanan, V., C. Karstens, J. Krause, and L. Tang, 2014: Quality control of weather radar data using polarimetric variables. *J. Atmos. Oceanic Technol.*, **31**, 1234–1249, https://doi.org/10.1175/JTECH-D-13-00073.1.

Lakshmanan, V., M. Miller, and T. Smith, 2013b: Quality control of accumulated fields by applying spatial and temporal constraints. *J. Atmos. Oceanic Technol.*, **30**, 745–758, https://doi.org/10.1175/JTECH-D-12-00128.1.

Lakshmanan, V., and T. Smith, 2009: Data mining storm attributes from spatial grids. *J. Atmos. Oceanic Technol.*, **26**, 2353–2365, https://doi.org/10.1175/2009JTECHA1257.1.

Lakshmanan, V., and T. Smith, 2010: An objective method of evaluating and devising storm-tracking algorithms. *Wea. Forecasting*, **25**, 701–709, https://doi.org/10.1175/2009WAF2222330.1.

Lakshmanan, V., T. Smith, K. Hondl, G. J. Stumpf, and A. Witt, 2006: A real-time, three-dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity, and derived products. *Wea. Forecasting*, **21**, 802–823, https://doi.org/10.1175/WAF942.1.

Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The warning decision support system integrated information. *Wea. Forecasting*, **22**, 596–612, https://doi.org/10.1175/WAF1009.1.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, https://doi.org/10.1175/MWR-D-12-00254.1.

Lim, J. R., B. F. Liu, and M. Egnoto, 2019: Cry wolf effect? Evaluating the impact of false alarms on public responses to tornado alerts in the southeastern United States. *Wea. Climate Soc.*, **11**, 549–563, https://doi.org/10.1175/WCAS-D-18-0080.1.

Loeffler, S. D., M. R. Kumjian, M. Jurewicz, and M. M. French, 2020: Differentiating between tornadic and nontornadic supercells using polarimetric radar signatures of hydrometeor size sorting. *Geophysical Research Letters*, **47**, e2020GL088 242, https://doi.org/10.1029/2020GL088242.

Losey, A., W. D. Zittel, and Z. Jing, 2017: Use of mid-level model data and vad winds to improve WSR-88D velocity dealiasing. *AMS 38th Conf. on Radar Met.*, Chicago, IL, 167, https://ams.confex.com/ams/38RADAR/webprogram/Manuscript/Paper319960/38thAMS_Radar_Conf_ExtAbs_ALosey.pdf.

Mahalik, M. C., B. R. Smith, K. L. Elmore, D. M. Kingfield, K. L. Ortega, and T. M. Smith, 2019: Estimates of gradients in radar moments using a linear least squares derivative technique. *Wea. Forecasting*, **34**, 415–434, https://doi.org/10.1175/WAF-D-18-0095.1.

Martinaitis, S. M., 2017: Radar observations of tornado-warned convection associated with tropical cyclones over Florida. *Wea. Forecasting*, **32**, 165–186, https://doi.org/10.1175/

WAF-D-16-0105.1.

McCaul, E. W., 1991: Buoyancy and shear characteristics of hurricane-tornado environments. *Mon. Wea. Rev.*, **119**, 1954–1978, https://doi.org/10.1175/1520-0493(1991)119<1954:BASCOH>2.0.CO;2.

McCaul, E. W., D. E. Buechler, S. J. Goodman, and M. Cammarata, 2004: Doppler radar and lightning network observations of a severe outbreak of tropical cyclone tornadoes. *Mon. Wea. Rev.*, **132**, 1747–1763, https://doi.org/10.1175/1520-0493(2004)132<1747:DRALNO>2.0.CO;2.

McCaul, E. W., and M. L. Weisman, 1996: Simulations of shallow supercell storms in landfalling hurricane environments. *Mon. Wea. Rev.*, **124**, 408–429, https://doi.org/10.1175/1520-0493(1996)124<0408:SOSSSI>2.0.CO;2.

Miller, M. L., V. Lakshmanan, and T. M. Smith, 2013: An automated method for depicting mesocyclone paths and intensities. *Wea. Forecasting*, **28**, 570–585, https://doi.org/10.1175/WAF-D-12-00065.1.

Newman, J. F., V. Lakshmanan, P. L. Heinselman, M. B. Richman, and T. M. Smith, 2013: Range-correcting azimuthal shear in doppler radar data. *Wea. Forecasting*, **28**, 194–211, https://doi.org/10.1175/WAF-D-11-00154.1.

Nielsen, E. R., G. R. Herman, R. C. Tournay, J. M. Peters, and R. S. Schumacher, 2015: Double impact: When both tornadoes and flash floods threaten the same place at the same time. *Wea. Forecasting*, **30**, 1673–1693, https://doi.org/10.1175/WAF-D-15-0084.1.

Nowotarski, C. J., J. Spotts, R. Edwards, S. Overpeck, and G. R. Woodall, 2021: Tornadoes in Hurricane Harvey. *Wea. Forecasting*, **36**, 1589–1609, https://doi.org/10.1175/WAF-D-20-0196.1.

NWS, 2022a: Accessed 7 April 2022. During a flood. https://www.weather.gov/safety/flood-during.

NWS, 2022b: Accessed 7 April 2022. What to do during a tornado. https://www.weather.gov/safety/tornado-during.

Reid, D., 1979: An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, **24**, 843–854, https://doi.org/10.1109/TAC.1979.1102177.

Romine, G. S., D. W. Burgess, and R. B. Wilhelmson, 2008: A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City area tornadic supercell. *Mon. Wea. Rev.*, **136**, 2849–2870, https://doi.org/10.1175/2008MWR2330.1.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, and D. S. Zrnic, 2005: Polarimetric tornado detection. *J. Appl. Meteor*, **44**, 557–570, https://doi.org/10.1175/JAM2235.1.

Schenkel, B. A., R. Edwards, and M. Coniglio, 2020: A climatological analysis of ambient deep-tropospheric vertical wind shear impacts upon tornadoes in tropical cyclones. *Wea. Forecasting*, **35**, 2033–2059, https://doi.org/10.1175/WAF-D-19-0220.1.

Schneider, D., and S. Sharp, 2007: Radar signatures of tropical cyclone tornadoes in Central North Carolina. *Wea. Forecasting*, **22**, 278–286, https://doi.org/10.1175/WAF992.1.

Schultz, L. A., and D. J. Cecil, 2009: Tropical cyclone tornadoes, 1950-2007. *Mon. Wea. Rev.*, **137**, 3471–3484, https://doi.org/10.1175/2009MWR2896.1.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.*, **63**, 1379–1389, https://doi.org/10.1080/01621459.1968.10480934.

Skinner, P. S., L. J. Wicker, D. M. Wheatley, and K. H. Knopfmeier, 2016: Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Wea. Forecasting*, **31**, 713–735, https://doi.org/10.1175/WAF-D-15-0129.1.

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype warn-on-forecast system. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Smith, B. T., R. L. Thompson, A. R. Dean, and P. T. Marsh, 2015: Diagnosing the conditional probability of tornado damage rating using environmental and radar attributes. *Wea. Forecasting*, **30**, 914–932, https://doi.org/10.1175/WAF-D-14-00122.1.

Smith, T. M., and K. L. Elmore, 2004: The use of radial velocity derivative to diagnose rotation and divergence. *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., P5.6, https://ams.confex.com/ams/pdfpapers/81827.pdf.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Spotts, J., C. J. Nowotarski, S. Overpeck, and R. Edwards, 2022: Automatically derived radar attributes of tropical cyclone supercells. *AMS 31st Conf. on Weather and Forecasting/27th Conf. on Numerical Weather Prediction*, Virtual, Amer. Meteor. Soc., 668, https://ams.confex.com/ams/102ANNUAL/mediafile/Manuscript/Paper389618/AMS2022_extended_abstract_cor.pdf.

Spotts, J. R., C. J. Nowotarski, S. Overpeck, B. Fillipiack, and R. Edwards, 2020: Analysis of tornadic and nontornadic convective cell environments during Hurricane Harvey. *Tropical Meteorology and Tropical Cyclones Symp.*, Boston, MA, Amer. Meteor. Soc., 857, https://ams.confex.com/ams/2020Annual/webprogram/Paper363902.html.

Spratt, S. M., D. W. Sharp, P. Welsh, A. Sandrik, F. Alsheimer, and C. Paxton, 1997: A WSR-88D assessment of tropical cyclone outer rainband tornadoes. *Wea. Forecasting*, **12**, 479–501, https://doi.org/10.1175/1520-0434(1997)012<0479:AWAOTC>2.0.CO;2.

Theil, H., 1950: A rank-invariant method of linear and polynomial regression analysis (parts 1-2). *Ned. Akad. Wetensch. Proc. Ser. A*, **53**, 1397–1412, https://ir.cwi.nl/pub/8270/8270D.pdf.

Trainor, J. E., D. Nagele, B. Philips, and B. Scott, 2015: Tornadoes, social science, and the false alarm effect. *Wea. Climate Soc.*, **7**, 333–352, https://doi.org/10.1175/WCAS-D-14-00052.1.

WDTD, 2016: Tornado warning guidance 2016: Quick reference guide. 12 pp, www.wdtb.noaa.gov/courses/woc/documentation/severe/twg16-reference-sheets.pdf.

Wilks, D. S., 2019: *Statistical methods in the atmospheric sciences*. Elsevier, 818 pp.

Wilson, J. W., T. M. Weckwerth, J. Vivekanandan, R. M. Wakimoto, and R. W. Russell, 1994: Boundary layer clear-air radar echoes: Origin of echoes and accuracy of derived winds. *J. Atmos. Oceanic Technol.*, **11**, 1184–1206, https://doi.org/10.1175/1520-0426(1994)011<1184:BLCARE>2.0.CO;2.

Wilson, M. B., and M. S. V. D. Broeke, 2021: An automated python algorithm to quantify ZDR arc and KDP-ZDR separation signatures in supercells. *J. Atmos. Oceanic Technol.*, **38**, 371–386,

https://doi.org/10.1175/JTECH-D-20-0056.1.

Wood, V. T., and R. A. Brown, 1997: Effects of radar sampling on single-Doppler velocity signatures of mesocyclones and tornadoes. *Wea. Forecasting*, **12**, 928–938, https://doi.org/10.1175/1520-0434(1997)012<0928:EORSOS>2.0.CO;2.

APPENDIX A

KEY ALGORITHM DESCRIPTIONS

Nineteen WDSS-II algorithms and archiveBTRT.py are used in the process of retrieving the attributes of individual clusters. A list of these programs along with a short description is given in Table A.1. A flowchart for how data are passed between each program is shown in Figure A.1. The following subsections describe the methodology employed by key programs along with a brief discussion of key parameters used by the programs that deviated from their defaults.

A.1 w2qcnndp

Non-meteorological scatterers can cause reflectivity and radial-velocity to appear in a volume scan (Wilson et al. 1994). The WDSS-II algorithm w2qcnndp (Lakshmanan et al. 2014) is used to remove most non-meteorological echoes while retaining most echoes of interest (i.e., precipitation). To censor non-meteorological echoes, w2qcnndp employs four neural networks, for range gates where velocity data are unavailable, for when reflectivity is < 10 dBZ, for reflectivity between 10 and 20 dBZ, and for reflectivity greater than 20 dBZ. Before a neural network is employed, gates are pre-classified as meteorological and non-meteorological by whether they are beyond the range where moments other than reflectivity are available, as well as their correlation coefficient, differential reflectivity, or reflectivity. Gates on the edges of echoes are assigned a probability of precipitation of 0.5. Reflectivity, velocity, spectrum width, differential phase, and differential reflectivity are the six radar moments used by the neural networks. Additional inputs to the neural networks include: maximum reflectivity in the column containing the gate, the maximum height of an echo greater than -14 dBZ, the reflectivity at 3 km from elevation angles above 1°, the difference in reflectivity between the lowest elevation angle and the next elevation angle greater than 1°, the local variance of reflectivity, differential reflectivity, and correlation coefficient, and the output of the simple classifier.

The simple classifier checks each range gate and classifies it as meteorological if the reflectivity

is at least 3 dBZ, the correlation coefficient is at least 0.9, or the absolute value of differential reflectivity is less than 2.3 dB. At least two of these three conditions must be met. After the neural network has been applied, each elevation-angle sweep undergoes smoothing using a 2.5x2.5-km median filter with a minimum size of 5x5 gates. Gates within 150 km of the radar or 2 km in height undergo bimodal clustering. Areas are placed into two categories if they are separated enough by reflectivity and range. Within these areas, regions are then grown and the reflectivity-weighted average probability is calculated for each region. If the resulting averaged probability is greater than 0.5, the region is retained. For this study, the surface temperature at the radar site is not provided via model data and instead utilizes w2qcnndp's climatology temperature for the radar site. The surface temperature at the radar site is used to determine if the warm-weather reflectivity threshold of 0 dBZ or cool-weather threshold of -14 dBZ is used. Additionally, a digital elevation model (DEM) is not provided for beam-blockage calculations. However, the effects of beam blockage should be limited in much of the eastern two-thirds of the United States (Cintineo et al. 2012). Finally, the various QC modules ("sunstrobe", "electronicinterference", etc.) are also enabled as per the default settings.

A.2 w2circ

To derive the strength of the rotation in event cells, azimuthal shear ("AzShear") is calculated for the elevation angles of interest using the w2circ algorithm. A Linear Least Squares Derivative (LLSD; Smith and Elmore 2004; Mahalik et al. 2019) technique is used to calculate the azimuthal gradient of the radial velocity within a local kernel. Before the LLSD technique is applied, the velocity data are dealiased using the dealiasVel algorithm. First, the dealiased velocity data are smoothed with a 3x3 median filter assigning the center gate the median value of all nine gates if ⅝ of the gates within the kernel contain valid data. A first-order planar equation is defined in terms of the azimuthal and radial components of velocity as well as a constant. The residuals are then defined as the weighted sum of the square of the differences between each velocity value and the plane at that location and rewritten in terms of range and azimuth. The derivative of the residuals with respect to the constant, range, and azimuthal derivatives are set equal to zero

and rearranged creating a system of equations. This system of equations is then solved for the azimuthal component (AzShear) and other terms. The LLSD technique is considered more robust than more traditional methods due to the fact it uses a larger number of gates and is less susceptible to noise and sampling limitations of vortices (Smith and Elmore 2004). The area over which this calculation takes place is a 2500 m (1.4 n mi.; azimuthal) by 750 m (0.4 n mi.; radial) kernel. Near the radar, the maximum number of radials is 51 for super-resolution radar data and a 2500 m kernel. As range increases, the width of each bin increases and the number of radials within the kernel decreases. At farther ranges, the kernel expands to maintain a minimum of three gates in the azimuth and range directions.

The ReflectivityQC field for each tilt is then dilated using a 5x5 median filter B. Smith (2022, personal communication). AzShear values where the dilated ReflectivityQC field is less than 20 dBZ are then discarded. This thresholding procedure is applied to the AzShear fields to remove AzShear resulting from non-meteorological gates and reduce the potential number of undesired circulations tracked. The field is dilated to capture valid AzShear that could exist in a weak reflectivity area such as a hook echo and the threshold of 20 dBZ is chosen to match the operational value used by MRMS (Smith et al. 2016). Finally, removal of spikes of AzShear, typically caused by improperly dealiased velocity (Mahalik et al. 2019), is enabled.

A.3 w2merger and w2cropconv

The w2merger algorithm (Lakshmanan et al. 2006; Lakshmanan and Humphrey 2014) receives data from multiple radars and merges the data onto a three-dimensional grid. From this grid, various products such as composite reflectivity, vertically integrated liquid (VIL), vertically integrated ice (VII), reflectivity along an isothermal surface, etc. can be produced. The range gates received by a radar for each individual elevation angle are assigned an "intelligent agent" to each three-dimensional grid point, or voxel, the gate influences. Each intelligent agent collaborates to determine the value of a nearby voxel. In the range dimension, a moving average of gates is determined (the number of gates used is dependent on the voxel dimensions) and the nearest neighbor value is used for that voxel. In the elevation dimension, the value assigned is the result of a weighted

combination of values in the vertical where weight is dependent on the ratio of the difference in elevation angle between the range gate and voxel to the beamwidth or difference between radar elevation angles, whichever is larger. For agents from multiple radars, agents are given a weight based on the exponential distance from their radar. Finally, the merged data in various tiles across the domain are further merged using a map-reduce technique (Lakshmanan and Humphrey 2014).

The w2merger algorithm creates composite reflectivity, VIL, and merged echo tops from the nearest two available radars for the analysis. The composite reflectivity and VIL are calculated on a 0.6°x0.6° grid around each case with a resolution of 0.0025°x0.0025°. The vertical dimensions are 0 to 20 km with a vertical resolution of 0.5 km. Operationally, the horizontal resolution is 0.01°x0.01° for reflectivity-based products and 0.005°x0.005° for velocity-based products (e.g., AzShear; Smith et al. 2016). A finer resolution is used in this analysis to better capture data from individual range gates, thereby more closely matching the ideal horizontal resolution of 0.25 km as suggested by Lakshmanan and Humphrey (2014). The operational vertical resolution is staggered with altitude with the low levels spaced at 0.25 km, mid-levels at 0.5 km, and upper levels at 1 km (Smith et al. 2016). This staggered resolution was unable to be incorporated while determining the parameters of the different algorithms for this study. Temporally, the current state of the constantly-updating grid is written every 5 minutes allowing for full volumes to be incorporated and match the time bin sizes of Davis and Parker (2014). ReflectivityQC data from multiple radars are merged using the exponential-distance weight, whereas the echo tops are merged using the maximum value between the two radars.

A set of cursory sensitivity tests are performed for VIL on three cases varying the vertical and horizontal resolution. The VIL at three points are compared using a vertical resolution of 0.25 km, 0.5 km, and 1 km and a fourth comparison is performed using a horizontal resolution of 0.01° and a vertical resolution of 0.25 km. The average standard deviation between the three vertical resolutions is 0.29 $\mathrm{kg\,m^{-2}}$ with a maximum standard deviation of 0.75 $\mathrm{kg\,m^{-2}}$. The average difference between the 0.25-km vertical resolution cases the different horizontal resolution is 1.88 $\mathrm{kg\,m^{-2}}$ with a maximum difference of 7.1 $\mathrm{kg\,m^{-2}}$ indicating that the change of horizontal resolution may

have a greater effect than the change in vertical resolution.

All other variables besides VIL and echo-tops (i.e., AzShear, ReflectivityQC, spectrum width, etc.) are gridded for each of the analyzed elevation angles using the w2cropconv program at the same 0.0025°x0.0025° resolution. This program is employed using a gate width of 0.25 km and a Cressman interpolation (Cressman 1959) within a neighborhood size of four. An examination of the polar data and resulting grid values shows that values from the polar data tend to be preserved in the grids with the potential for a small difference in spatial location. Using w2cropconv over w2merger is preferred as these grids are created from a single elevation angle requiring no vertical interpolation.

A.4 w2segmotionll

The w2segmotionll algorithm identifies, tracks, and records the attributes of clusters (Lakshmanan et al. 2009; Lakshmanan and Smith 2009, 2010). The creation of clusters is accomplished using an enhanced watershed technique where individual regions are grown from a local maximum in the data field incrementally until they reach a minimum size. The local maximum must meet some minimum value in order to begin to form a cluster. Additionally, only pixels above a separate minimum value are eligible to be included in the cluster (Lakshmanan et al. 2009). For this analysis, these minimum and maximum values for AzShear clusters are 0.0024 s^{-1} and 0.0025 s^{-1} respectively. The VIL clusters' thresholds range from 4 $\mathrm{kg\,m}^{-2}$ to 5.5 $\mathrm{kg\,m}^{-2}$ and the echo-top clusters' thresholds range from 5 km (16.4 kft) to 6 km (19.7 kft).

Before the watershed algorithm is applied to each data field in this analysis, the data field undergoes image preprocessing. Table A.2 summarizes the pre-clustering filters applied to each field. For AzShear clusters, the 96^{th} to 100^{th}-percentile values are eligible to be clustered. For anticyclonic AzShear, the field is multiplied by negative one to ensure clustering is not performed on missing data values. A 95^{th}-percentile threshold followed by a 7x7 75^{th}-percentile filter is applied to the VIL field to only cluster the stronger storms and prevent over clustering, while slightly expanding the size of the existing clusters. Echo-top clusters are filtered using a 72^{nd}-percentile filter then smoothed with a 3x3 median filter as the echo-top field can be quite noisy.

Combining clusters to meet the minimum size criteria is disabled (Lakshmanan and Smith 2009) following Cintineo et al. (2020). Finally, the minimum sizes to create clusters are 16, 160, and 225 pixels for AzShear, VIL, and echo-top clusters respectively.

The above parameters are chosen somewhat heuristically by adjusting the parameter values and observing the response in the performance of cluster tracking. Tuning the parameters for one case may degrade performance in another. Lowering the minimum threshold for AzShear allows for the possibility of tracking weaker rotation, but increases the likelihood of a false-association. Likewise, an increase in the minimum threshold of echo tops reduces the area that gets clustered, but may prevent the lowest-topped convection from being recorded.

Once the clusters are defined, they are initially tracked via the w2segmotionll MULTISTAGE tracking method with a distance tolerance of two times the size of the cluster or 7.5 km (4.0 n mi.; AzShear) or 10 km (5.4 n mi.; VIL and echo tops) to be considered overlapping. These parameters are chosen based on the ProbSevereV2 parameters (Cintineo et al. 2020), reducing the AzShear distance tolerance to match the AzShear cluster search radius. No coast frames are included in the tracking. The MULTISTAGE method forecasts the position of a cluster and matches that cluster with another based on the degree of overlap between the actual and forecast location, age, track length, and size of the cluster, and a cost function based on the cluster size, distance, and intensity between the projected and candidate clusters (Lakshmanan and Smith 2010). After the w2segmotionll association, the cluster with the same ID as the previously tracked frame using this method is checked first before determining what other clusters meet the tracking criteria. The attributes of each cluster are recorded (Lakshmanan and Smith 2009) and used in the final analysis.

A.5 archiveBTRT

Changes in cluster identification numbers from the splitting, merging, and changing of the identification of clusters is a known issue in storm tracking (e.g., Lakshmanan and Smith 2010; Harrison 2018). Thus, the archive Best Track: Real Time (archiveBTRT) algorithm (Harrison 2018) is utilized to reduce changes in cluster ID numbers and improve tracking. In BTRT, each cluster is compared to the track of previous clusters. Clusters with the same ID number are matched.

Those in the current frame are compared with groups of clusters with the same ID from previous frames. If the time between the cluster in the previous and current frames is less than eight minutes (changed from the default of three minutes), the track's trajectory is calculated using a Theil-Sen fit (Theil 1950; Sen 1968; Lakshmanan et al. 2015). Object centroids that fall within 8 km (AzS-hear; changed from the default of 10 km) of the previous object projected along the trajectory are associated, with closer centroids taking priority.

After an initial association, any remaining objects are processed via split and merge algorithms. The split algorithms performs the same operation, but adds an additional 5 km (2.7 n mi.) to the projected object's buffer. A similar method is employed for merging except the 5 km buffer is applied to the current cluster instead of the extrapolated cluster.

Table A.1: List and short descriptions of WDSS-II algorithms used.

Algorithm	**Description**
ldm2netcdf	Ingests Archive L2 NEXRAD data into WDSS-II format.
g2w2	Ingests RAP analysis data into WDSS-II format.
w2csv2table	Converts the reference point to a table to be used in finding the first reference cluster.
createCache	Performs precomputations for w2merger.
w2makeindex.py	Indexes existing data to be used by the next algorithm.
nse	Calculates various near-storm environmental parameters including Sfc-6 km mean wind and wind profiles for dealiasVel.
w2difference	Calculates the difference between two fields such as the u and v wind components at the Sfc and 6 km MSL.
w2qcnndp	Classifies and removes non-meteorological reflectivity echoes creating ReflectivityQC.
dealiasVel	De-aliases Velocity data using methods from RDA/RPG Build 19

Table A.1 Continued	
	and wind profiles from **nse**.
w2circ	Calculates AzShear via the LLSD method where dilated ReflectivityQC is greater than 20 dBZ.
w2echotop	Calculates the 20-dBZ echo tops from the nearest two available radars.
w2tds	Determines which range gates meet the TDS criteria and writes a new radial set of the ρ_{HV} of those gates.
w2threshold	Retains range gates from the differential reflectivity field where the ρ_{HV} from the output of **w2tds** is at least 0 (not used here).
w2simulator	Simulates a stream of data from multiple sources for **w2merger**.
w2merger	Continuously updates a three-dimensional grid of ReflectivityQC and echo tops and writes out the maximum value in the vertical (or maximum value between the two radars for echo tops).
w2cropconv	Creates two dimensional grids from the radar data for individual sweeps.
w2segmotionll	Performs enhanced watershed clustering on filtered VIL, echo tops, and AzShear and provides the initial tracking for each cluster.
w2hailtruth_size	Identifies 0.5° AzShear clusters that are within 7.5 km and 5 minutes of the warning point or TCTOR start.
w2table2csv	Converts the tables with cluster information from XML to CSV format.

Table A.2: Filtering applied to cluster fields before the watershed algorithm is applied.

Field	**Filters**
VIL	95^{th}-percentile threshold; 7x7 75^{th}-percentile filter
AzShear	96^{th}-percentile filter
Echo Tops	72^{nd}-percentile filter; 3x3 median filter

Table A.3: List of key parameters for key WDSS-II algorithms. Parameters listed do not include generic default options (i.e., **-i** for input index). Option flags are bolded in the parameters column. Phrases such as "upper left" and "lower right" are used to describe the coordinates and altitude, for w2merger, of their respective corners.

Program	**Parameters**
w2qcnndp	Defaults
dealiasVel	Defaults
w2difference	**-X** U(V)Wind3D:06.00 **-Y** SfcU(V)Wind:modelanalysis
w2circ	**-v** "Velocity" **-b** 0 **-az -ka** "2500:750" **-t** "20.0" **-z** "ReflectivityQC" **-D -S -m -sr**
w2echotop	**-I** "ReflectivityQC" **-R** "0.5x250x230" **-T** "20"
w2tds	**-a** AzShear_Storm **-c** RhoHV **-z** Reflectivity **-A** 0.0025 **-C** 0.9 **-Z** 35
w2threshold	**-d** Zdr **-t** TDS_CC **-R** 0 **-T** 10 **-v** 0 **-n** _Threshold **-k** none **-w**
w2simulator	**-r** 200 **-b** Beginning of case hour minus 1 minute **-e** End of case plus 5 minutes **-C** 1000 **-E**
w2merger	**-C** 1 (echo-tops) 7 (VIL/composite reflectivity) **-t** "upper left 20" **-b** "bottom right 0.0" **-s** "0.0025 0.0025 0.5"

Table A.3 Continued	
w2merger (cont.)	**-a** "VIL Composite" (ReflectivityQC) "Composite" (echo tops) **-I** "InterbpDbZEchoTop_20" or "ReflectivityQC" **-e** 300:1:1 **-V** **-S** "5 2500" (ReflectivityQC only)
w2cropconv	**-I** Product:tilt **-g** 0.25 **-t** "northwest corner" **-b** "southeast corner" **-s** "0.0025 0.0025" **-S** Radar ID
w2semgotionll (VIL)	**-T** "VIL" **-d** "400 550 1 -1.0 0.4" **-E** 0 **-O** 0 **-F** 0 **-p** 160,800,1200:0:0,0,0 **-m** MULTISTAGE:2:10:0:30 **-k** threshold:95:100:percent,scaling:100:0,percent:75:3:0:3 **-f** "MergedReflectivityQCComposite VIL UWind3D VWind3D UWindMean0-6km VWindMean0-6km udiff vdiff"
w2segmotionll (AzShear)	**-T** "AzShear_Storm_tilt:tilt" **-d** "2400 2500 1 -1.0 0.4" **-E** 0 **-O** 0 **-F** 0 **-p** 16,80,120:0:0,0,0 **-m** MULTISTAGE:2:7.5:0:30 **-k** (scaling:-1:0 for -AzShear), threshold:96:100:percent,scaling:1000000:0 **-f** "AzShear_Storm_tilt TDS_CC_tilt RhoHV_tilt Zdr_Threshold_tilt ReflectivityQC_Threshold_tilt SpectrumWidth_tilt udiff vdiff UWindMean0-6km VWindMean0-6km"
w2segmotionll (Echo Tops)	**-T** "MergedInterpDbZEchoTop_20Composite" **-d** "500 600 1 -1.0 0.4" **-E** 0 **-O** 0 **-F** 0 **-p** 225,1125,1688:0:0,0,0 **-m** MULTISTAGE:2:10:0:30 **-k** threshold:72:100:percent,scaling:100:0,percent:50:1:0:1 **-f** "MergedInterpDbZEchoTop_20Composite"
w2hailtruth_size	**-s** 2 **-d** 7.5 **-w** "-5 5" **-p** "-90 90" **-A** 480 **-C** ClusterBT **-T** Reports **-c** List of columns to retain

Table A.3 Continued	
w2hailtruth_size (cont.)	**-a** "RowName"
archiveBTRT.py	**-it** xml **-ot** xml **-bt** 8 **-bd** 8 (AzShear Only)

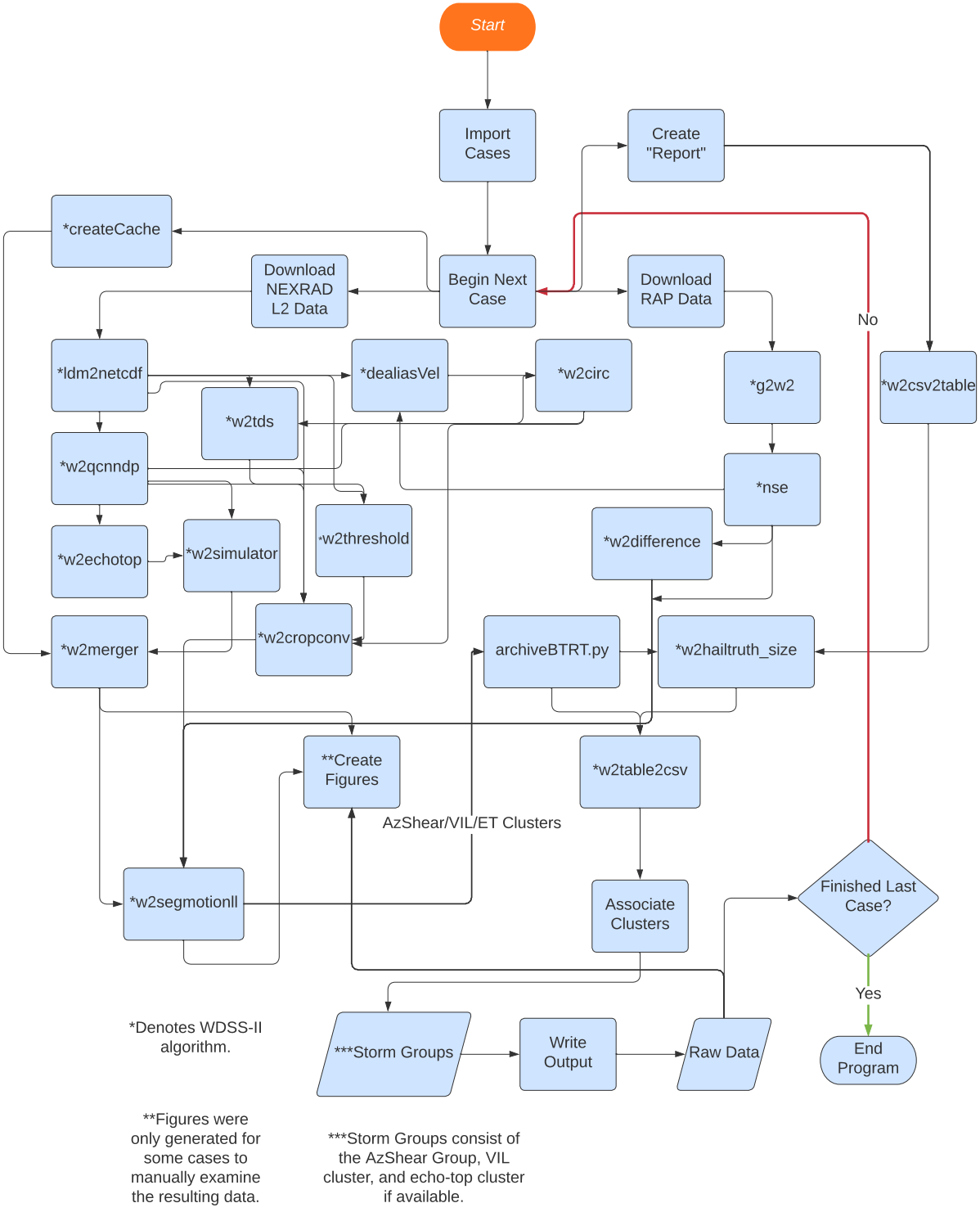

Figure A.1: A diagram showing the general flow of data through the program.

APPENDIX B

LIMITATIONS AND SOURCES OF ERROR

The methodology has several limitations and sources of error that should be considered when interpreting the results. Some limitations can occur with the calculation of AzShear and V_{rot} and with the association and tracking of clusters. Range plays a significant role in the calculation of AzShear (Newman et al. 2013). Beyond the effects of beam broadening with range on Doppler vortex measurements (Wood and Brown 1997) and increases in beam height with range from the radar, the kernel over which AzShear is calculated increases with range in order to maintain the minimum number of range gates necessary to perform the calculation (Mahalik et al. 2019). As pointed out by Davis and Parker (2014) and Newman et al. (2013), the smoothing in the azimuthal direction due to beam broadening with range or coarser azimuthal sampling at higher tilts can result in failure to fully resolve the mesocyclone or tornado and underestimate the true value of AzShear.

Furthermore, there are steps within the calculation that limit the comparison of the V_{rot} derived from AzShear to the V_{rot} derived from the manual analysis. V_{rot} is calculated as half the difference between the maximum or minimum or half the sum of the absolute values of the peak inbound and outbound velocities (Smith et al. 2015). Prior to calculating AzShear, a median filter is applied to the velocity field, which can smooth out peaks in the radial-velocity field (Newman et al. 2013; Mahalik et al. 2019). As a result, the velocity values of the gates used in the Smith et al. (2015) V_{rot} calculation may be reduced. Moreover, the Smith et al. (2015) criteria utilizes peak-velocity values within 5 n mi. and 45° of each other. Therefore, gates used to calculate V_{rot} may not be included in the kernel and the kernel required to include them may result in a significant underestimation of AzShear (Newman et al. 2013; Mahalik et al. 2019).

The parameters and process used for clustering has its own set of limitations. For AzShear clusters, the minimum value of AzShear for a cluster is $2.5 \times 10^{-3}\,\mathrm{s}^{-1}$ or the 96th-percentile

of AzShear within the domain, whichever is higher. While such a low value allows for weaker rotation to be detected and tracked, it can also increase the number of potential false associations by allowing the clustering of other weak, extraneous AzShear clusters. VIL clusters can also suffer from the diversity of cases. In some cases, the minimum value for clustering is just enough to create a VIL cluster, whereas in other cases, it is too large and allows for the merging of clusters. Finally, for echo tops the minimum value to begin growing a cluster is 6 km (19.7 kft) or the 72nd-percentile within the domain, whichever is higher. This can eliminate shallower convection from the echo-top clusters and may create a high bias in the data. Similar to AzShear clusters, VIL and echo-top clusters can merge together and cover a larger area or more convective cells than desired, despite the higher thresholds and percentile filter. In some cases, the values recorded from these clusters still correspond to the desired convection.

During the determination of the initial cluster there are two primary sources of error. First, despite using the 20-dBZ ReflectivityQC threshold, areas of AzShear generated from noisy, non-hydrometeorological echoes can still be included as potential reference clusters and overwhelm the intensity term of the initial interest score. Moreover, while the TCTOR database was examined and some entries removed, errors in the time and location of the reference point for tornadic and nontornadic cases may exist. For NON TOR cases, the warning feature point is assumed to be placed on the circulation, which may not always be the case. Depending on the case, this could potentially affect the selection of the initial reference cluster. A single tornadic cell can also be responsible for multiple TCTOR events. In these cases, each report is treated as a separate event. Given the large number of events, it is not clear in what proportion of events, if any, this occurs.

The tracking process is also subject to errors and limitations. Clusters are tracked or associated using their centroids. While a threshold filter is utilized to highlight only the more intense values, clusters can merge and split with one another resulting in changes in centroid location that may not be representative of the rotation's actual location. Although, the maximum or minimum AzShear may still be correct, depending on the case. Additionally, a change in the velocity field can result in AzShear clusters forming near the expected location of the next cluster resulting in a false

association. These events can result in a degradation of the motion vector's accuracy. Using the averaged motion vector removes some of the spontaneity in the motion vector compared to calculating the vector on a frame-by-frame basis. Nonetheless, sub-optimal tracking can result.

Along with the aforementioned, noisy clusters, other errors can exist in the AzShear field. During their analysis, Nowotarski et al. (2021) avoided including peak velocity values within the storm's inflow region. There is no mechanism implemented to differentiate between the inflow region from the rest of the storm beyond the 20-dBZ ReflectivityQC threshold for AzShear. Conversely, in cases where the updraft region or hook echo of the storm has very light reflectivity values, a small number of AzShear pixels for the rotation are created. If this number is below the minimum required for clustering (16), then a cluster is not created and a different cluster may be selected as the reference cluster or tracked. An example of this is shown in Figure B.1. Another potential point of error includes the clusters that are in or move into the region of velocity range-folding or the "purple haze." Like the effective inflow region, there is no technique known to the author or implemented to account for these cases.

The scoring component of the tracking is also limited. During the scoring of potential clusters for 0.5° tracking, the intensity term favors clusters that have similar intensities. This may help to reduce a false association with clusters of notably different intensities, however, this could also theoretically reduce the likelihood of making a correct association if the strength of the rotation changes too quickly. When clusters merge, the centroid of the merged cluster may not be where the scoring terms predict. For example, with a highly merged (covers a large area) echo-top cluster, the centroid may not be in or near the direction of the Sfc-to-6-km MSL BWD vector, reducing the effectiveness of the score and increasing the potential for a false association. Furthermore, the default range for echo-top calculations is set to 230 km. Though generally not a problem, at cases near this range, a discontinuity may be possible between the merged radars. For beam-height calculations, the nominal elevation is used versus the recorded elevation (e.g., 0.5° versus 0.482°) which may result in slightly different values than what would be found using software such as GR2Analyst (`http://grlevelx.com/gr2analyst/`). For TDS detection, the minimum

number of pixels required to meet TDS criteria to determine a TOR TDS case is 25. This may create a bias towards larger AzShear clusters as smaller clusters may not meet the minimum number of TDS pixels. Furthermore, non-TDS reductions in ρ_{HV} such as ground clutter or non-uniform beam filling may trigger a TDS detection. Finally, aircraft reconnaissance found evidence of TC-TORs that were not verified by a ground survey in the aftermath of Hurricane Ida (2021; J. Klein 2022, personal communication). This suggests that some NON TOR cases may be tornadic thereby impacting the warning statistics.

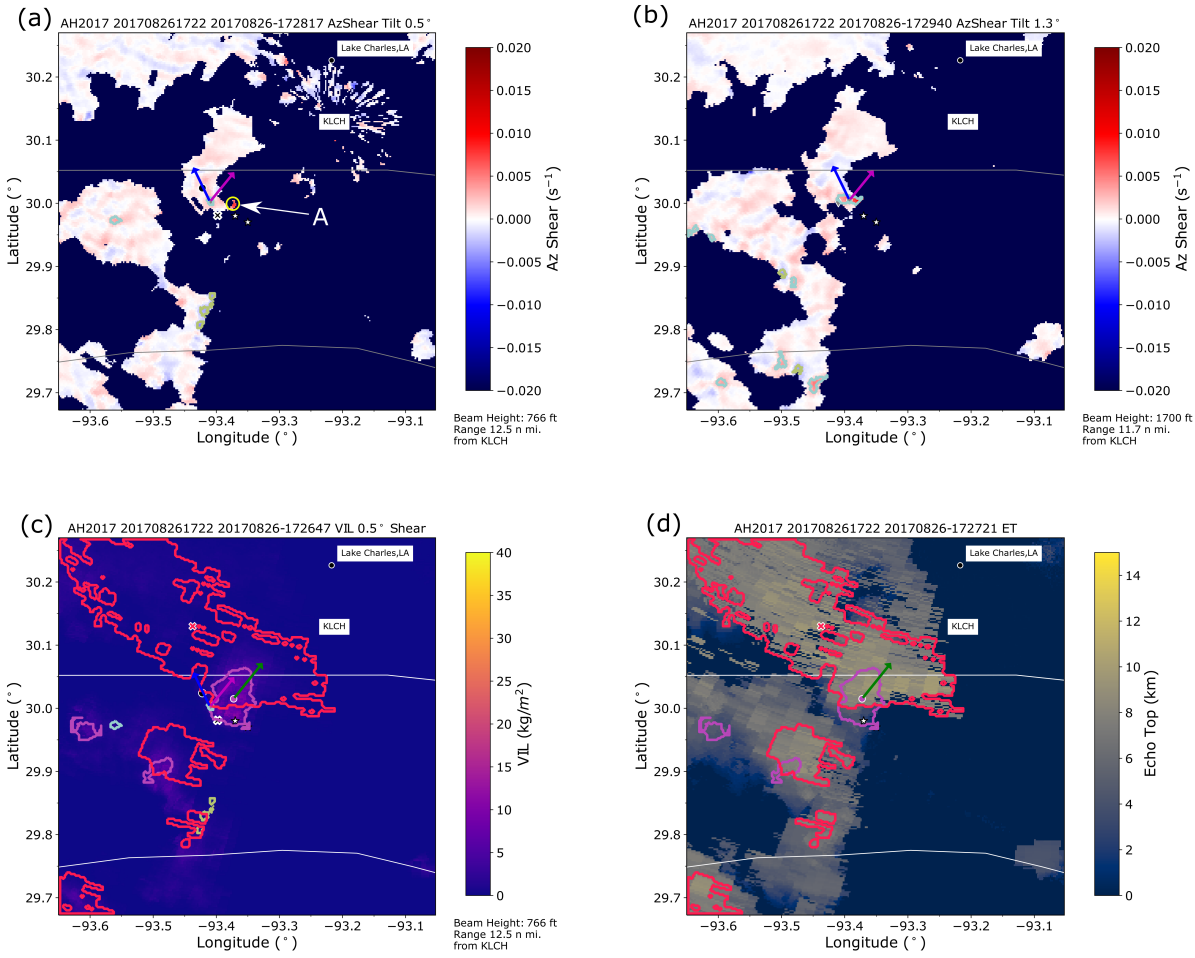

Figure B.1: Same as in Figure 2.1 except for a Hurricane Harvey ALL TOR case on 26 August 2017 at 1728z. This supercell had a low-reflectivity hook-echo resulting in an area of AzShear that did not meet the required size (denoted by the yellow circle at Point A in **(a)**).