Wilson, M. B., and M. S. Van Den Broeke, 2021: Using the Supercell Polarimetric Observation Research Kit (SPORK) to examine a large sample of pretornadic and nontornadic supercells. *Electronic J. Severe Storms Meteor.*, **17** (2), 1–38.

Electronic Journal of
SEVERE STORMS METEOROLOGY

Using the Supercell Polarimetric Observation Research Kit (SPORK) to Examine a Large Sample of Pretornadic and Nontornadic Supercells

MATTHEW B. WILSON AND MATTHEW S. VAN DEN BROEKE

University of Nebraska-Lincoln, Lincoln, Nebraska

(Submitted 20 September 2021; in final form 02 August 2022)

ABSTRACT

Previous work examining dual-polarization (dual-pol) radar signatures of supercells has shown that differential reflectivity (Z_{DR}) column area, hailfall area and specific differential phase (K_{DP})–Z_{DR} separation angles may differ between tornadic and nontornadic storms. However, these signatures often can be difficult to quantify quickly enough to enable their use in forecasting operations, and little work has been done examining how these characteristics vary with environmental parameters in large samples of observed storms. This paper introduces the Supercell Polarimetric Observation Research Kit (SPORK) as an update to our automated Z_{DR} arc-detection algorithm. This update adds the capability to identify Z_{DR} column and inferred hailfall signatures in supercells automatically, and quickly quantify their characteristics. Dual-pol metrics calculated by SPORK are compared to manually calculated dual-pol metrics from previous work. SPORK is run on a large sample of supercells to examine whether SPORK-calculated dual-pol metrics exhibit the same differences between tornadic and nontornadic supercells seen in manual analyses. Storm-mean dual-pol metrics obtained from SPORK also are used to evaluate how supercell dual-pol metrics vary with different environmental parameters. Results from SPORK support previous findings that tornadic supercells have larger Z_{DR} column areas, smaller hailfall areas, and larger K_{DP}–Z_{DR} separation angles than nontornadic storms. Additionally, Z_{DR} columns tend to be larger and deeper in more conditionally unstable environments. Hailfall areal extents are larger in environments with lower environmental 0ºC levels, higher LCLs and LFCs, and less SRH. Separation angles are larger in environments with larger low-level shear vectors, SRH and lower lifted condensation levels (LCLs) and levels of free convection (LFCs). However, none of these correlations exceed $r = 0.52$. Overall, our results indicate that SPORK can quantify supercell dual-pol signatures accurately enough to detect potentially useful differences between dual-pol signatures of pretornadic and nontornadic supercells, and provide a first look at how dual-pol signatures vary with environmental characteristics in a large sample of supercells.

1. Introduction

Supercell storms produce the most tornadoes of any convective mode, and the vast majority of strong tornadoes (Duda and Gallus 2010; Smith et al. 2012), yet most are nontornadic (Trapp et al. 2005). Thus, distinguishing between tornadic and nontornadic supercells is an important area of ongoing research. Supercells are more likely to produce tornadoes in environments with large low-level storm relative helicity (SRH), strong low-level shear, low lifted condensation levels (LCLs), and large values of low-level instability (Thompson et al. 2003; Hampshire et al. 2018; Coffer et al. 2019). However, tornadic and nontornadic supercells often exist near each other, in ostensibly the same environment (Klees et al. 2016), limiting the utility of these environmental parameters in determining which storms are likely to produce a tornado. Supercells can also induce perturbations to their

Corresponding author address: Matthew B. Wilson, University of Nebraska-Lincoln, E-mail: mwilson41@huskers.unl.edu

environments that may affect a given storm's likelihood of producing a tornado (Parker 2014; Wade et al. 2018; Coniglio and Parker 2020; Flournoy et al. 2020). Environmental boundaries or heterogeneities (that may not be resolved well by numerical models) also can play an important role in tornado formation (e.g., Maddox et al. 1980; Markowski et al. 1998; Atkins et al. 1999; Ziegler et al. 2010; Magee and Davenport 2020).

Furthermore, the internal dynamics of a supercell—for example, the buoyancy of the rear-flank outflow (Markowski et al. 2002; Weiss et al. 2015), the baroclinic generation of surface vertical vorticity and near-surface streamwise vorticity in the storm's cold pool (Klemp and Rotunno 1983, Rotunno and Klemp 1985, Markowski et al. 2008, Orf et al. (2017)), and cyclic mesocyclogenesis (Lemon and Doswell 1979; Adlerman et al. 1999; Dowell and Bluestein 2002)—can also influence whether a given storm is likely to produce a tornado in the near future. Since these factors are often not detectable using operationally available observations or model output, any source of data that might provide insight into these processes with a given supercell may be useful to forecasters.

One source of data that may help is dual-polarization (hereafter dual-pol) radar. Much recent work has focused on examining the unique dual-pol signatures found in supercell storms (Romine et al. 2008; Kumjian and Ryzhkov 2008, 2009; Van Den Broeke et al. 2008; Kuster et al. 2019; Homeyer et al. 2020; Van Den Broeke 2016, 2020, (hereafter VDB16, VDB20)). Several of these signatures have exhibited potentially useful differences between severe and nonsevere storms and between tornadic and nontornadic supercells, including differential reflectivity (Z_{DR}) columns (VDB20, Kuster et al. 2019, 2020), inferred hailfall area (VDB16, VDB20), and the separation between areas of enhanced Z_{DR} and specific differential phase (K_{DP}) (Crowe et al. 2012; Loeffler and Kumjian 2018; Loeffler et al. 2020; Homeyer et al. 2020). A brief description of each of these signatures and the potential connections between their characteristics and a given supercell's hazard production is given below.

Z_{DR} columns are areas of positive Z_{DR} that extend above the environmental 0ºC level within a storm, and represent large raindrops and/or small, wet ice particles lofted above the environmental 0ºC level by the storm's updraft (Brandes et al. 1995; Loney et al. 2002; Kumjian et al. 2014; Snyder et al. 2015). Z_{DR} column location, area, and depth can be used as a proxy for updraft location and strength, which may provide valuable information about a given storm's potential to produce severe weather. Kuster et al. (2019) used an automated Z_{DR} column depth algorithm developed by Snyder et al. (2015) to examine Z_{DR} column characteristics in a large sample of storms from central Oklahoma, and found that severe storms had wider and deeper Z_{DR} columns than nonsevere storms. A peak in Z_{DR} column area preceded a majority of severe hail and wind reports in their sample of storms; however, they did not find substantial differences between the characteristics of Z_{DR} columns associated with tornadic and nontornadic mesocyclones, possibly due to the small number of tornadoes (17) included in their sample.

In contrast, recent work (VDB20, VDB17, and French and Kingfield 2021) comparing larger samples ($n = 63$, $n = 35$, and $n = 198$, respectively) of tornadic and nontornadic supercells has found that Z_{DR} columns in tornadic supercells, just prior to tornadogenesis, tend to be larger than in nontornadic supercells. Supercells that produce stronger tornadoes also have larger Z_{DR} columns. VDB20 found that tornadic supercells had larger and less variable Z_{DR} column areas in the 30 min prior to tornadogenesis than nontornadic supercells. VDB17 also found that supercells producing EF3 or stronger tornadoes had larger, steadier Z_{DR} columns than those with only EF0 tornadoes. Similarly, French and Kingfield (2021) found Z_{DR} columns were larger in tornadic supercells on radar scans during or immediately preceding tornadogenesis compared to nontornadic supercells during or immediately preceding peak 0–1-km azimuthal shear. Furthermore, they also found that supercells that went on to produce EF3 or stronger tornadoes had larger Z_{DR} columns on radar scans at or immediately preceding tornadogenesis compared to storms with weaker tornadoes. As discussed by French and Kingfield (2021), these results match well with previous work showing that stronger tornadoes come from broader pretornadic mesocyclones. Physically, this may be due to the larger size of updrafts and associated downdrafts in storms with broader mesocyclones. Greater updraft size allows larger areas of near-surface vertical vorticity in the

outflow, through the feet-first tilting of baroclinic horizontal vorticity. That, in turn, converges and stretches into stronger low-level mesocyclones and tornadoes (Trapp et al. 2017, Sessa and Trapp 2020).

Areas within a storm where the dominant scatterer type is large hail can often be identified at S band by large values of reflectivity (>50–55 dBZ) paired with relatively low values of Z_{DR} (<0.5–1 dB, Ryzhkov et al. 2013, Snyder et al. 2017). Kumjian and Ryzhkov (2008) examined a small sample (n = 15) of tornadic and nontornadic supercells and found that the presence of a large-hail signature in the storm core was less consistent in tornadic supercells. This may be due to more frequent disruption of the midlevel updraft by rotationally induced downward-directed perturbation pressure gradient forces, from the stronger low-level rotation in tornadic storms. Results from a similar small sample of supercells (n = 25 across 12 environments) examined by VDB16 also noted greater variability in the area of polarimetrically inferred large hail (referred to here as hailfall area for brevity) in tornadic supercells than in nontornadic storms. In a larger sample of supercells, VDB20 found that the inferred hailfall area was smaller in pretornadic supercells than in nontornadic storms. Maximum storm-core reflectivity was also higher in nontornadic storms than in pretornadic storms, potentially due to the greater quantities of hail present in nontornadic supercells (VDB20). The larger amounts of hail present at low levels in nontornadic supercells could indicate that these storms have colder outflow temperatures, which may hinder tornadogenesis (Markowski et al. 2002; Weiss et al. 2015).

Size sorting of rain and melting hail by the storm-relative wind often leads to a pronounced area of enhanced Z_{DR} along the edge of the storm's forward flank, known as the Z_{DR} arc (e.g, Kumjian and Ryzhkov 2008, 2009; Dawson et al. 2014, 2015), as well as a pronounced horizontal separation between the Z_{DR} arc and an area of enhanced K_{DP} in the storm core known as the K_{DP} foot (Romine et al. 2008; Crowe et al. 2010, 2012; Loeffler and Kumjian 2018; Loeffler et al. 2020). While the maximum value of Z_{DR} in the Z_{DR} arc was initially thought to be correlated to the strength of low-level storm-relative helicity in a supercell's inflow (Kumjian and Ryzhkov 2009) and could be of use in identifying storms that are more likely to be tornadic, subsequent work revealed that Z_{DR} arc size and intensity tend to differ little between tornadic and nontornadic storms (VDB20).

However, the separation between the Z_{DR} arc and K_{DP} foot has shown more promise in distinguishing between tornadic and nontornadic storms. Loeffler and Kumjian (2018) found that a vector drawn from the centroid of the K_{DP} foot to the centroid of the Z_{DR} arc (referred to as the separation vector) in tornadic nonsupercell storms tended to be closer to perpendicular to storm motion for a given vector length when low-level SRH was higher. In a larger sample of supercells (n = 116), Loeffler et al. (2020) found that the separation vector was closer to perpendicular to storm motion in tornadic storms than in nontornadic storms. Similarly, Homeyer et al. (2020) found a separation between areas of enhanced K_{DP} and Z_{DR} that was more perpendicular to storm motion in storm-centered radar composites of tornadic supercells than storm-centered composites of nontornadic supercells in a large sample of supercells (n = 772, with 490 having dual-pol data).

Idealized modeling by Loeffler and Kumjian (2020) showed that separation-vector length and direction were well-correlated to the 0–3-km storm-relative wind speed and direction, and that a custom size-sorting parameter combining separation-vector length and orientation angle correlated well with low-level SRH. However, it is not known how well the correlations between separation-vector characteristics and environmental parameters from these idealized experiments will generalize to observed supercells, in a variety of thermodynamic and kinematic environments. Furthermore, Van Den Broeke (2021) also found that separation angles did not differ meaningfully between pretornadic supercells and nontornadic storms with strong low-level mesocyclones, in the 30 min before their peak low-level rotation. Healey and Van Den Broeke (2022) found that tornadic supercells did not have meaningfully different separation angles from nearby nontornadic supercells on the same day. However, direct comparisons between these two studies and the results of Homeyer et al. (2020) and Loeffler et al. (2020) are difficult, since they have smaller sample sizes (n = 63 and n = 46 tornadic-nontornadic pairs, respectively), and different selection criteria for nontornadic supercells, than Homeyer et al. (2020) and Loeffler and Kumjian (2020).

While all three of these signatures show promise in helping to differentiate between tornadic and nontornadic supercells, quantifying them quickly and accurately enough to be useful in analyzing a supercell in real time can be difficult. Two approaches can help alleviate this problem: choosing metrics for the signatures that are easier to calculate in real time, or developing automated methods to detect and analyze these signatures. The first approach was taken by Kuster et al. (2020), who evaluated the feasibility of replicating the results of Kuster et al. (2019) with Z_{DR} column depth metrics that can more easily be determined in real time by forecasters. They found that these Z_{DR} column depth metrics are still useful in differentiating between severe and nonsevere storms; however, they did not analyze Z_{DR} column area, since doing so manually would be too time-consuming in real-time warning operations.

Several recent papers have taken the second approach of designing automated or semi-automated algorithms to detect and quantify supercell dual-pol signatures. Snyder et al. (2015) designed an automated algorithm to calculate Z_{DR} column depth and display it as a 2-dimensional field for forecasters to examine. French and Kingfield (2021) introduced an automated algorithm to calculate Z_{DR} column area in supercells. However, the latter still required manual intervention in cases where a Z_{DR} column was composed of multiple areas of enhanced Z_{DR}. For low-level signatures of size sorting, Loeffler and Kumjian (2018) built a semi-automated algorithm to calculate the separation vector between areas of enhanced Z_{DR} and K_{DP} in nonsupercell storms, that was then applied to supercells by Loeffler et al. (2020).

Finally, Wilson and Van Den Broeke (2021, hereafter WV21) presented an automated algorithm for detecting Z_{DR} arcs and K_{DP}–Z_{DR} separation signatures in right-moving supercells, and quantifying their characteristics. For the convenience of researchers and potentially the eventual use of forecasters, it would be advantageous to have an algorithm that could automatically identify and quantify multiple supercell dual-pol signatures at once. Thus, this paper will build on the algorithm presented by WV21, with an improved Python algorithm called the Supercell Polarimetric Observation Research Kit (SPORK). SPORK is capable of quantifying the supercell dual-pol features that may be the most useful in differentiating between tornadic and nontornadic storms: Z_{DR} columns, inferred hailfall areas, and K_{DP}–Z_{DR} separation signatures. SPORK then will be used to answer the following questions:

1. Can SPORK successfully identify the differences between tornadic and nontornadic supercell dual-pol signatures noted by VDB20, Loeffler et al. (2020) and Homeyer et al. (2020)?
2. How do dual-pol signatures vary with the characteristics of a supercell's environment?

Question 1 extends research comparing manual analyses of supercell dual-pol signature differences between tornadic and nontornadic supercells (VDB20, Loeffler et al. 2020; Homeyer et al. 2020), to focus on automated analyses of these signatures in a larger sample of storms. Question 2 focuses on conducting a novel analysis of how dual-pol signatures vary in different environments. The relationship between Z_{DR} arc and K_{DP}–Z_{DR} separation signature characteristics and low-level wind field characteristics has been examined (e.g., Kumjian and Ryzhkov 2008, 2009; Dawson et al. 2014, 2015; Loeffler and Kumjian 2018, 2020); however, comparatively little is known about how these and other dual-pol signatures vary with other environmental parameters outside of the small sample of storms examined by VDB16. To answer these questions, Section 2 of this paper describes the design of SPORK and its validation against manually calculated dual-pol metrics from the dataset used by VDB20. Section 2 also describes methods used to select the 206 supercell cases, and proximity soundings used in the comparison between pretornadic and nontornadic storms and in the analysis of how the dual-pol metrics vary with different environmental parameters. Section 3 describes how algorithm-calculated dual-pol signatures differ between the pretornadic and nontornadic datasets, how the differences in algorithm-identified signatures compare to those found in previous work, and how they correlate to each other and various environmental parameters. Finally, section 4 summarizes our findings and discusses potential future research questions and algorithm applications based on these results.

2. Data and methods

a. Algorithm design

SPORK uses the Z_{DR} arc and K_{DP}–Z_{DR} separation algorithm presented by WV21 as a starting point and adds the capability to automatically identify Z_{DR} columns and inferred hailfall signatures. For this study, Z_{DR} columns are defined as areas of Z_{DR} >1 dB (instead of the 0.5-dB threshold used by VDB20) at 1 km above the environmental 0ºC level in a supercell's updraft region. Polarimetrically inferred hailfall is defined as regions with Z above 50 dBZ and Z_{DR} below 1 dB, within the storm's core (chosen to be similar to the high Z/Z_{DR} below 1-dB criteria used by VDB16 and VDB20). The 1-dB Z_{DR} threshold for automated column identification was chosen after initial testing, with a 0.5-dB threshold on a number of cases, showed that it tended to identify large, amorphous areas as columns that extended well outside of what would typically be subjectively identified as the Z_{DR} column core. A 1-dB threshold did a much better job of identifying features consistent with subjectively defined Z_{DR} columns.

Radar-data ingestion, Z_{DR} calibration, storm tracking, assignment of dual-pol signatures to storms, and K_{DP}–Z_{DR} separation signature calculations are all conducted as in WV21, with two key differences. Firstly, full volumetric radar data are used instead of just the lowest-level scan, since full radar volumes are needed to identify Z_{DR} columns. These data are interpolated to a grid with 500-m horizontal and 250-m vertical grid spacing, centered on the radar site using a Barnes interpolation scheme from the Python ARM Radar Toolkit (PyART, Helmus and Collis 2016). Secondly, storm identification and K_{DP}–Z_{DR} signature calculations use 1-km above radar level (ARL) constant altitude plan position indicators (CAPPIs), instead of the lowest radar tilt. To remove regions of high Z_{DR} associated with three-body scattering signatures (TBSSs, Zrnić 1987), Z_{DR} data used for Z_{DR} column detection are masked where Z is <20 dBZ or cross-correlation coefficient (CC) is <0.70.

As a final pre-processing step before the identification of Z_{DR} column and hailfall objects, a planar field of Z_{DR} column depth is calculated using a simplified Python implementation of the Z_{DR} column depth calculation described by Snyder et al. (2015). Starting from the grid level closest to the environmental 0ºC level obtained from a proximity Rapid Refresh model (RAP) sounding, the number of continuous grid points in the vertical above that level with Z_{DR} >1 dB are counted for each grid point in the horizontal. These counts are then multiplied by the 250-m vertical grid spacing to estimate Z_{DR} column depth at each point. One possible limitation of this approach is that the vertical-continuity requirement may lead to underestimated maximum column depth in highly tilted Z_{DR} columns of fast-moving supercells.

Table 1: Variables saved for each potential column object for use in the random-forest (RF) algorithm. Normalized storm centroid–object centroid distances are calculated by dividing the distance between the column-object centroid and the centroid of the associated storm object by the radius of a circle with the same area as the storm object. The forward-flank downdraft (FFD) vector is calculated as in WV21 as a vector perpendicular to the edge of the FFD Z gradient pointing from the core into the inflow region. Storm area is obtained from the tracking algorithm, and is the area within the Z contour used by the tracking algorithm to define the storm object.

Column-object RF Predictors
Column Area (km^2)
Normalized Storm Centroid–Column-object Centroid Distance
Mean Z_{DR} Value at 1 km above 0ºC (dB)
Max Z_{DR} Value at 1 km above 0ºC (dB)
Mean CC Value at 1 km above 0ºC
Mean K_{DP} Value at 1 km ARL (deg km^{-1})
Mean Z Value at 1 km above 0ºC (dBZ)
Mean Z Gradient Direction Relative to FFD Vector Direction at 1 km ARL (deg)
Mean Z Gradient Value at 1 km ARL (dBZ km^{-1})
Max Column Depth (km)
Mean Column Depth (km)
Mean K_{DP} Value at 1km above 0ºC (deg km^{-1})
Storm Area (km^2)
Normalized X-Component of Storm Centroid–Column-object Centroid Distance
Normalized Y-Component of Storm Centroid–Column-object Centroid Distance

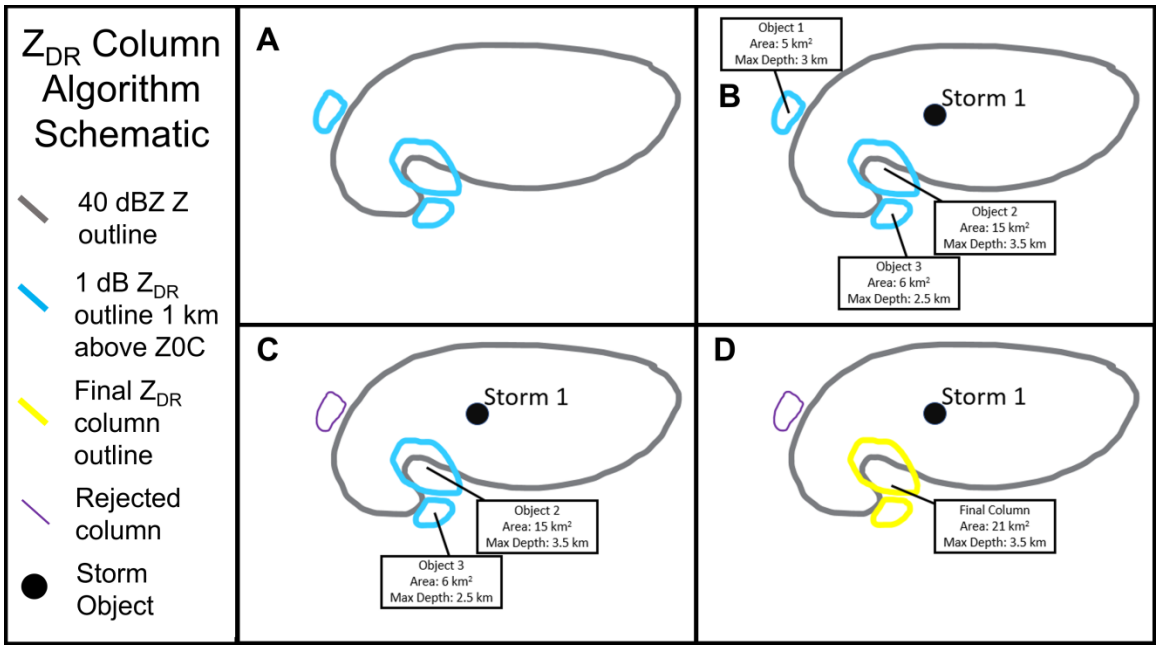


Figure 1: Schematic showing the process of identifying Z_{DR} column objects for each storm. In step 1 (panel a), a 1-dB Z_{DR} contour is drawn at 1 km above the environmental 0ºC level (Z0C). In step 2 (panel b), that contour is split into separate polygons, and each polygon is then assigned to the nearest storm object from the storm tracking algorithm before the metrics listed in Table 1 are calculated for it. In step 3 (panel c), the random forest algorithm is used to remove spurious column objects (denoted here with thin purple outlines). Finally, in step 4 (panel d), the remaining column objects associated with each storm object are combined into a single object, and column area, mean column depth, and maximum column depth are then calculated using all points within the multi-polygon final object.

Once the data pre-processing is completed, preliminary Z_{DR} column objects are identified by drawing a 1-dB Z_{DR} contour at 1 km above the environmental 0ºC level, splitting the contour into individual polygons, and assigning the polygons to the nearest storm object following the methodology of WV21 (Fig. 1a, b). At this stage, supercells often have multiple associated Z_{DR} column objects, not all of which are actually associated with that particular storm's updraft region. For example, these spurious objects may be associated with a nearby, nonsupercellular storm, or with a left-moving storm in the process of splitting from the storm of interest. To remove these spurious column objects, a random-forest algorithm was trained on a dataset of manually labeled actual and spurious Z_{DR} column objects, using predictors extracted from each object and listed in Table 1. Once trained on labeled objects from the 20 supercells used in WV21's training dataset, the resulting random forest achieved a probability of correct detection (POD) of 91.2 percent, on a dataset of 1129 actual and 956 spurious column objects from the 51-supercell test dataset in WV21, with a false alarm rate (FAR) of 13.2 percent. Plotting the location of all classified objects from the testing dataset along with the 40-dBZ outline of a supercell for reference shows that the algorithm correctly classifies almost all Z_{DR} column objects located near the supercell's hook echo and inflow notch regions as Z_{DR} columns (Fig. 2). Additionally, many of the missed detections are small areas of Z_{DR} located on the periphery of this region that might be easily missed by a human analyst as well, and many of the false detections classified as columns by the algorithm are large areas of Z_{DR} that are in a storm-relative location where one might normally expect to find a Z_{DR} column. Once trained, this random forest algorithm is used to remove spurious column objects. Once these are removed, any remaining column objects associated with a given storm are combined into a single column object, and Z_{DR} column area, mean depth, and maximum depth are then calculated for that object (Fig. 1c, d)

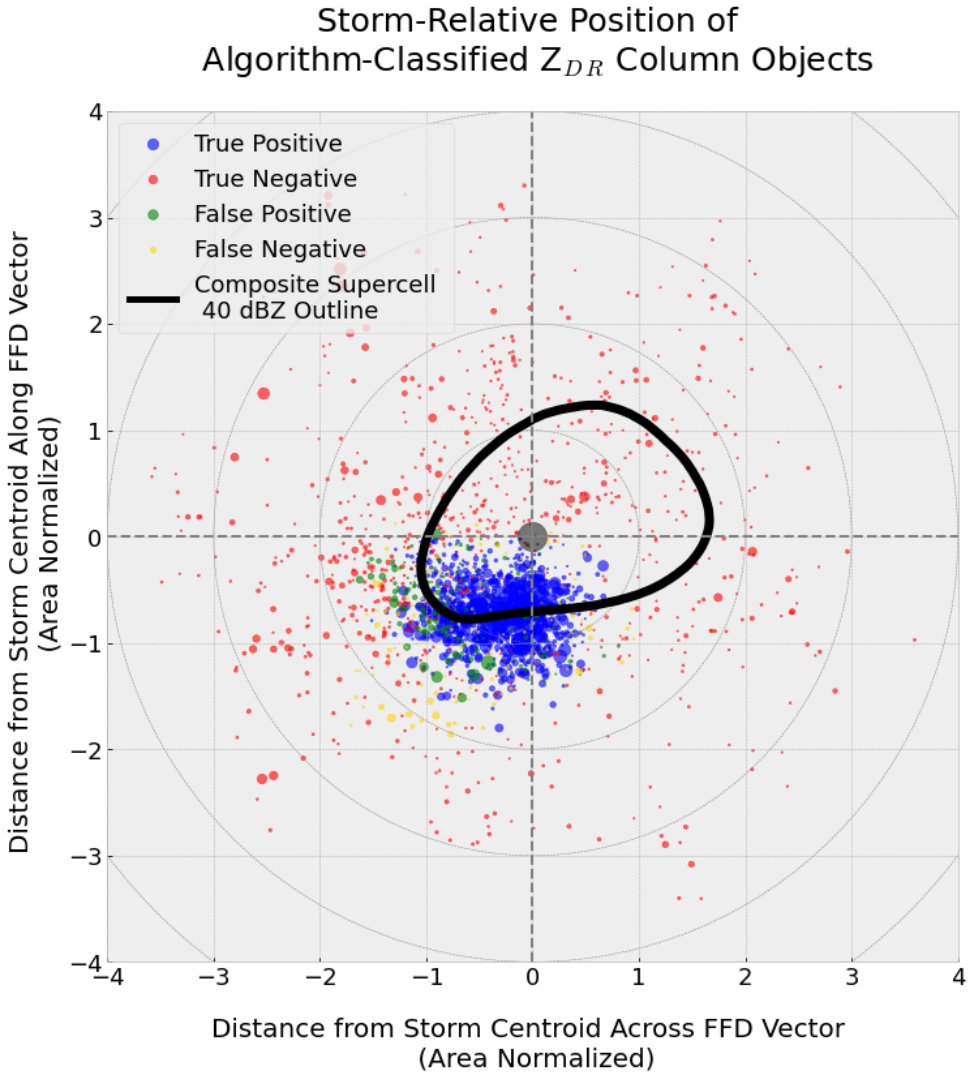


Figure 2: Storm-relative plot of preliminary Z_{DR} column objects classified by the random forest from the testing dataset. Centroids of Z_{DR} column objects correctly detected by the algorithm are plotted in blue, false detections are plotted in green, missed Z_{DR} columns are plotted in yellow, and correctly classified non-column objects are red. Dot sizes are proportional to the object areas. The thick black outline is a composite 40-dB*Z* outline from the 206 supercells examined in section 3b of this paper, and the FFD vector points to the top of the plot. The grey dot in the middle is the storm centroid. Range rings of storm-radius-normalized distance from storm centroid are plotted in increments of 1 and are unitless.

To identify dual-pol hailfall signatures at 1 km ARL, a contour is drawn on the 1 km ARL CAPPI around all areas where Z >50 dBZ is collocated with Z_{DR} <1 dB. This contour is broken into individual closed polygons to create hailfall objects. Each hailfall object is then assigned to the closest storm object within 15 km. This distance threshold is chosen to help ensure that hailfall objects from neighboring storms are not accidentally included. Manual inspection of initial test cases showed that false detections of hailfall objects tended to be much less common than for Z_{DR} column objects; thus, a random-forest algorithm was not needed to remove them. Once all hailfall objects are assigned to the closest storm, hailfall objects assigned to the same storm object are consolidated into a single hailfall object, and inferred hailfall area values are then calculated and saved for each storm object.

The current version of SPORK is available at https://github.com/mwilson14/SPORK-SPIN and can be run using example Python scripts available on that site. To run SPORK for a given case, a user must input the radar site, a start time, case duration, environmental 0ºC level, a storm-motion estimate, and information about the orientation of the forward-flank reflectivity gradient for the supercell of interest. Options are also available to include a Z_{DR} calibration and modify the thresholds used to define the Z_{DR} arc and K_{DP} foot. Once SPORK has been run for a given case, it saves all dual-pol metrics for each tracked storm to a Pandas dataframe (McKinney 2010), and a placefile that can be displayed in Gibson Ridge Analyst 2 (GR2) software. Detailed instructions on how to run SPORK and access saved SPORK output are provided on the GitHub page.

In addition, a real-time version of SPORK at https://github.com/mwilson14/SPORK-Realtime automatically extracts environmental 0ºC level information and anticipated FFD reflectivity gradient direction from RAP analysis fields and obtains an observed storm motion for each cell from its included tracking algorithm, requiring the user to only enter the radar site of interest. When running on a laptop computer, the real-time SPORK can produce placefiles for viewing in GR2 about a minute after a given volume scan is finished. Currently, the real-time SPORK displays only metrics for a given volume scan, instead of 30-min time averages as presented in the results of this paper; however, future algorithm updates may support real-time generation of these averages.

b. Algorithm validation

To verify that the dual-pol metrics objectively calculated by SPORK are reasonable compared to manual analyses of dual-pol signature characteristics, SPORK was run on the 64 supercells examined by VDB20, and the objectively calculated dual-pol metrics from SPORK were compared to VDB20's manually calculated metrics. Scatterplots comparing manual and algorithm-generated Z_{DR} column areas, Z_{DR} column maximum depths, and hailfall areas averaged over the analysis period for each storm are shown in Fig. 3, and similar plots for all available analysis times are shown in Fig. 4.

Overall, the algorithm performed well, with a Spearman's correlation of $r = 0.817$ between manual and automated storm-mean Z_{DR} column areas and $r = 0.620$ for individual scans. Excluding cases where a given signature was not detected by SPORK, the median SPORK Z_{DR} column area is 18 km^2 smaller than the median manual Z_{DR} column area, which was expected with the change in Z_{DR} threshold from 0.5 dB with the manual analyses to 1.0 dB with SPORK. SPORK also performed well in calculating hailfall areal extent, with a correlation of $r = 0.852$ between the manual and algorithm storm-mean calculations, and $r = 0.839$ for the individual scans. A low bias also exists with SPORK-calculated hailfall areas, with the median SPORK hailfall area being 13 km^2 smaller than its manually calculated counterpart. Algorithm performance was lower with Z_{DR} column depth values, with a correlation of $r = 0.551$ between the manual and algorithm storm-mean Z_{DR} column depths and $r = 0.404$ for individual scans. The median SPORK-calculated column depth was also 192 m higher than its manual counterpart.

Several explanations for the lower correlation between manual and algorithm-calculated Z_{DR} column depths emerged upon close examination of cases with particularly large (>1-km) differences between storm-mean manual and SPORK-calculated column depths. High Z_{DR} values aloft associated with TBSSs missed by SPORK's quality control were most often responsible for cases where SPORK's column depths were substantially higher than manual column depths, along with the extension of Z_{DR} >1 dB above what was present in the raw data by the interpolation involved in gridding the raw data for SPORK (as one of the reviewers of this paper noted is fairly common in interpolated Z_{DR} analyses). Cases where SPORK calculated substantially lower column depths than the manual analyses were most often narrow, tilted columns where SPORK's vertical continuity requirement prevented it from capturing the full column depth.

In some instances, Z_{DR} columns analyzed manually were not detected by SPORK. Upon close examination of cases where a column deeper than 3.5 km or larger than 30 km^2 was missed, several failure modes became evident. Firstly, deep but narrow columns occasionally were missed by SPORK, likely due to their small area. That made the random-forest algorithm

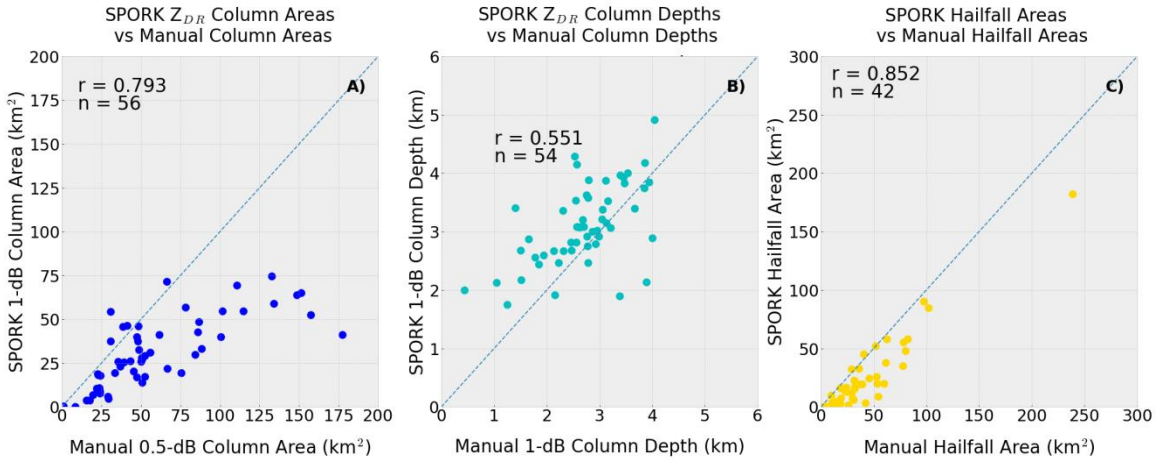


Figure 3: Scatterplots comparing storm-mean SPORK-calculated values of a) Z_{DR} column area, b) Z_{DR} column maximum depth and c) hailfall area, to manually calculated storm-mean values from VDB20 for each metric. Spearman's rank-order correlation coefficients are included on each plot, as well as the number of storms included in each comparison.

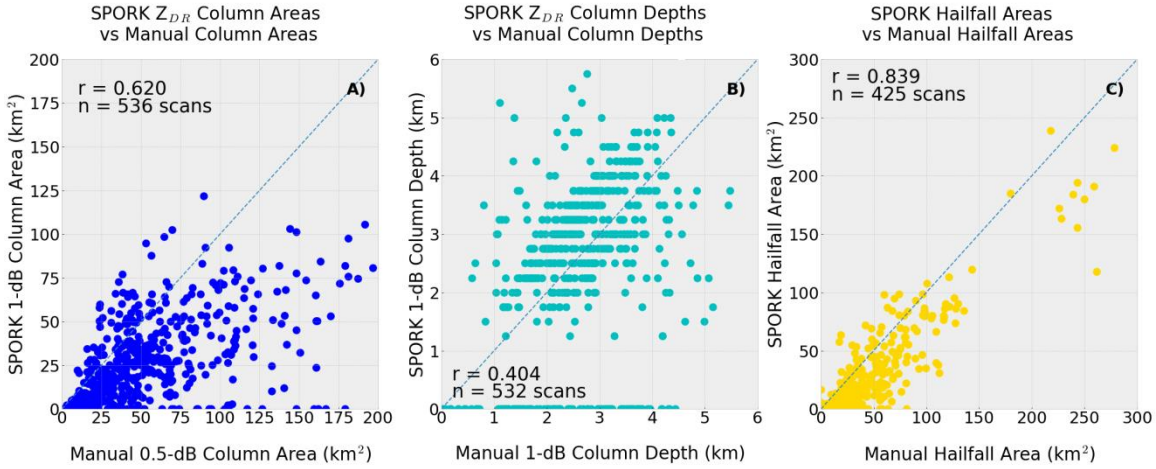


Figure 4: As in Fig. 3, but for all individual volume scans for each signature. On each panel, n indicates the number of volume scans included in each comparison.

less likely to classify them correctly as a column object. Secondly, columns were occasionally missed because they were assigned to a spurious storm object from the tracking algorithm that was closer to the column than the actual supercell. Finally, in a few cases, very large, elongated supercells resulted in missed column objects because the Z_{DR} column was so far away from the storm centroid that they were not properly assigned to that storm object.

Sample sizes for each of these metrics differ in the verification dataset, since not all metrics were calculated for each storm at every analysis time by VDB20. In addition, verification of SPORK's Z_{DR} arc and K_{DP}–Z_{DR} separation metrics is not performed here, since those metrics calculated using the same methodology have already been extensively compared to manually calculated values by WV21, and because manual K_{DP}–Z_{DR} separation metrics were not calculated for the supercells included in VDB20's analysis.

c. Supercell case selection

A dataset of 206 supercells was compiled to evaluate whether SPORK can capture the differences in dual-pol signatures between tornadic and nontornadic supercells observed by VDB20, Loeffler et al. (2020), Homeyer et al. (2020), and French and Kingfield (2021), and to provide a sample of storms to evaluate

Table 2: Spearman's rank-order correlations between environmental parameters obtained from proximity RAP soundings and storm-mean dual-pol metrics from SPORK. Correlations with a single star have p <0.05, and correlations with two stars have p <0.01. SRH is calculated with both Bunkers et al. (2000) and observed storm motions, the latter marked with an (obs) in the variable name. Correlations ≥0.30 are bolded. "Shear" represents vector magnitudes of wind differences.

Variable	Z_{DR} Column Area	Z_{DR} Column Max Depth	Z_{DR} Column Mean Depth	Hail Area	Separation Angle
MLCAPE	**0.35****	**0.52****	**0.52****	–0.09	–0.04
MUCAPE	**0.32****	**0.50****	**0.49****	–0.10	–0.06
0–6-km Shear	0.06	–0.12	–0.19**	0.00	0.11*
0–3-km Shear	0.11	–0.04	–0.07	–0.12	0.22**
0–1-km Shear	0.18*	–0.04	–0.02	**–0.34****	**0.45****
Effective Shear	0.18*	0.00	–0.06	–0.05	0.08
Sfc–3-km SRH	0.17*	0.01	–0.04	–0.20**	**0.31****
Sfc–1-km SRH	0.24**	0.04	0.01	**–0.35****	**0.41****
Effective SRH	0.27**	0.15*	0.09	–0.28**	**0.33****
Sfc–3-km SRH (obs.)	0.20**	0.00	–0.03	–0.26**	**0.33****
Sfc–1-km SRH (obs.)	0.19*	0.04	–0.05	–0.12	0.22**
Effective SRH (obs.)	**0.30****	0.19**	0.10	–0.22**	0.25**
LCL Height	–0.13	0.13	–0.02	0.27**	**–0.32****
LFC Height	–0.22**	–0.13	–0.25**	0.28**	**–0.31****
EL Height	**0.31****	**0.41****	**0.44****	–0.27**	–0.01
0ºC Height	**0.37****	**0.35****	**0.41****	**–0.41****	–0.01
1–3-km RH	0.07	–0.09	0.02	–0.17*	0.13
3–6-km RH	–0.11	–0.18*	–0.07	–0.08	0.12
6–9-km RH	–0.14*	–0.24**	–0.14*	–0.12	0.19**
CIN	0.16*	0.10	0.21**	–0.21**	0.17*
SCP	**0.42****	**0.44****	**0.37****	**–0.30****	0.24**
STP	**0.42****	**0.35****	**0.34****	**–0.44****	**0.35****
0–3-km CAPE	0.17*	0.14	0.25**	–0.20**	0.29**
0–1-km SR Wind	0.09	0.04	–0.15*	0.20**	–0.26**
0–2-km SR Wind	0.06	0.05	–0.12	0.26**	**–0.35****
0–3-km SR Wind	0.05	0.03	–0.11	0.25**	**–0.37****
2–4-km SR Wind	0.10	0.06	0.03	0.13	–0.10
0–6-km SR Wind	0.12	0.06	–0.05	0.19**	**–0.31****
6–9-km SR Wind	–0.04	–0.07	–0.17*	0.17*	–0.07

relationships between different dual-pol metrics and near-storm environmental parameters. This dataset was split evenly between 103 tornadic and 103 nontornadic supercells during 2012–2020, to ensure that each case had dual-pol data. To create this dataset, an initial collection of supercells was gathered from the VDB16, VDB17, VDB20, and WV21 databases. Additional cases were gathered by searching for linear segments of storm reports on the Storm Prediction Center archived maps of preliminary storm reports from the years in question (https://www.spc.noaa.gov/climo/reports/), and examining convective morphology using the radar composites on the UCAR Radar Archive (http://www2.mmm.ucar.edu/imagearchive/).

Cases that appeared to contain one or more supercells were examined more closely in GR2 using NEXRAD Level II data downloaded from Amazon Web Services (available at https://aws.amazon.com/public-datasets/nexrad/). To be included in the dataset, a storm had to: 1) consistently display typical supercell reflectivity structures such as hook or pendant echoes or a bounded weak echo region (BWER), and dual-polarization supercell signatures such as Z_{DR} arcs, and 2) maintain a midlevel mesocyclone for ≥30 min. Each storm was also required to have at least four full volume scans where most of the forward-flank region was sampled at ≤1 km ARL, to ensure adequate sampling of the shallow size sorting in the Z_{DR} arc. All 206 of these

supercells are used in the analysis comparing dual-pol signatures in different environments presented in section 3b.

The analysis presented in section 3a comparing pretornadic and nontornadic supercells uses a slightly smaller dataset. Following the methodology of VDB20, the pretornadic dataset included only storms where the 30-min period prior to the storm's first tornado could be analyzed to remove possible effects from cyclic tornadogenesis. At least three radar scans where the low-level mesocyclone and most of the forward flank were sampled at ≤1 km ARL, in the 30 min leading up to tornadogenesis, also were required, to enable analysis of pretornadic dual-pol metrics. Tornadogenesis times were taken from NCEI's Storm Event Database (https://www.ncdc.noaa.gov/stormevents/). This reduced the sample size of the pretornadic dataset to 95 storms from the original 103. The EF-scale ratings of all tornadoes produced over a supercell's life cycle also were recorded.

To be included in the nontornadic supercell dataset, a supercell had to not be associated with any tornado reports during its entire life cycle as determined by matching NCEI storm reports to the supercell's location in GR2. For consistency, nontornadic storms were analyzed over a 30-min period prior to peak low-level rotation [defined as the peak value of the Normalized Rotation (nrot) parameter in GR2 associated with a valid low-level circulation on the 0.5° tilt], and also were required to have at least three radar scans with most of the forward flank sampled at ≤1 km ARL, leading up to and including the time of peak low-level rotation. Storms for which a peak nrot value could not be defined—either due to the absence of a low-level circulation or data quality issues such as vertical sidelobe contamination—were discarded, resulting in a nontornadic dataset of 77 storms.

Environmental data also were gathered for each supercell, using proximity soundings from the National Center for Environmental Information (NCEI) archive of Rapid Refresh (RAP) model analyses. RAP soundings were obtained for each supercell case from the RAP archive THREDDS server maintained by NCEI (https://www.ncei.noaa.gov/thredds/catalog/rap130anl/). Each sounding was required to be within 80 km of the supercell, and had to be located on the same side of any mesoscale boundaries as the storm in question, and away from the outflow of other storms. Such positioning was considered representative of the supercell's inflow airmass. For cases that span multiple hours and have multiple possible proximity soundings, the sounding closest to the middle of the analysis period was used for that storm, when possible. Multiple thermodynamic and kinematic variables and derived parameters were calculated for all soundings using the Sounding/Hodograph Analysis and Research Program in Python (SHARPpy, Blumberg et al. 2017) and Meteorological Python (MetPy, May et al. 2017) packages. A list of these parameters is available in Table 2.

3. Results and discussion

a. Pretornadic-nontornadic supercell comparison

To compare the values of dual-pol metrics calculated by SPORK in pretornadic and nontornadic supercells, SPORK was run for all 77 nontornadic and 95 pretornadic supercell cases from the filtered dataset described above. For the tornadic supercells, mean values over all scans within the 30-min pretornadic window for each storm are calculated for Z_{DR} column area, mean depth and maximum depth, hailfall areal extent, and K_{DP}–Z_{DR} separation distance and angle. For nontornadic supercells, average values were calculated over the 30 min prior to peak low-level rotation (specifically, the peak nrot value associated with a valid low-level couplet in GR2.) In the comparisons in the following section, all values discussed are mean values for the 30-min analysis period for each storm. The statistical significance of any resulting differences was assessed using a Wilcoxon-Mann-Whitney test following the methods of VDB16, with significance defined as $p < 0.05$.

Time steps where a given signature was not detected by SPORK presented a problem for this analysis. This is because the absence of a signature could be interpreted differently for different signatures, with a missing Z_{DR} arc reasonably assumed to represent a Z_{DR} arc area of 0 km^2, while a missing K_{DP}–Z_{DR} separation vector does not correspond to a 0° separation

Table 3: Mean values for SPORK metrics for pretornadic and nontornadic supercells. Wilcoxon-Mann-Whitney (WMW) p-values are also listed for comparisons of the pretornadic and nontornadic datasets.

SPORK Metric	Pretor	Nontor	Pretor-Nontor *p*-value
Column Area (km^2)	38.87	27.02	1.77×10^{-3}
Column Mean Depth (km)	1.69	1.57	0.01
Column Max Depth (km)	3.27	3.08	0.14
Hailfall Area (km^2)	17.42	28.32	0.02
Hailfall area (w/zeros, km^2)	0.00	6.17	3.87×10^{-4}
Separation Angle (°)	69.09	38.42	3.97×10^{-7}

angle. Thus, for all signatures besides the K_{DP}–Z_{DR} separation, the comparisons of the SPORK metrics were performed twice: once with all metrics for a given signature set to 0 at times when that signature was not detected, and once with those times not included in the analysis for that signature. For all metrics besides hailfall area, these two analyses did not differ substantially. Thus, both analyses are presented for the hailfall areal extent analysis, and comparisons for all other signatures exclude radar scans when that signature was not present.

To examine whether storms producing significant tornadoes have meaningfully different dual-pol signatures than those producing less-damaging (i.e., "weaker", EF0–EF1) tornadoes, this analysis was repeated with the tornadic supercell category split into two subsets. A "weak tornado" subset consisted of storms where the tornado produced immediately after the analyzed pretornadic period was rated EF0–EF1. A significantly tornadic subset contained storms with the tornado rated EF2 or higher. Finally, analyses also were run using 15-min and 10-min analysis periods, as well as all individual scans from these analysis periods, instead of temporal averages. Since the results did not differ substantially between these analyses, only those using 30-min storm means are presented below, beginning with the Z_{DR} column metrics. Summary statistics from the pretornadic-nontornadic comparisons are listed in Table 3.

Z_{DR} column area and depth can be thought of as a proxy for the size and strength of a supercell's updraft, and recent work (Coffer and Parker 2017; Sessa and Trapp 2020) has found that tornadic supercells (and in particular those producing strong tornadoes) tend to have larger, steadier mesocyclones. As such, we expect pretornadic supercells to have larger and deeper Z_{DR} columns than nontornadic supercells. Furthermore, we expect significantly tornadic supercells to have larger and deeper Z_{DR} columns than weakly tornadic supercells, assuming that the former have larger, stronger updrafts. Our analysis supports this expectation, because pretornadic supercells had notably larger Z_{DR} column areas than did nontornadic supercells. In our dataset, tornadic supercells have a mean Z_{DR} column area of 39 km^2, while nontornadic supercells have mean Z_{DR} column area of 27 km^2 (Table 3). Although substantial overlap does exist between the pretornadic and nontornadic Z_{DR} column area distributions (Fig. 5), the difference *between pretornadic and nontornadic* mean Z_{DR} column areas is statistically significant (Table 3). Three-quarters of the nontornadic supercells in our analysis have Z_{DR} column areas <38 km^2, while half of the tornadic supercells have Z_{DR} column areas >36 km^2. That suggests that this metric may be useful to warning forecasters in determining which supercells may be most likely to become tornadic. Splitting the tornadic storms into significant and weak tornadoes (Fig. 5b) did not reveal statistically significant differences between them. These results are similar to those found by VDB20 and French and Kingfield (2021); however, the nontornadic subset examined herein has notably larger Z_{DR} column areas than their nontornadic subset, leading to greater overlap in the pretornadic and nontornadic Z_{DR} column areas. This could be due to: 1) the larger nontornadic sample used in this analysis [n = 73 compared to their n = 30], 2) the possibility that the nontornadic storms chosen for this paper were systematically stronger than those examined by French and Kingfield (2021), or 3) the possibility that SPORK may systematically calculate larger areas in weak Z_{DR} columns than their algorithm. Regardless, their guidance for using Z_{DR} column area in the warning decision process (namely, that a supercell with a Z_{DR} column area >40 km^2 has a greater chance of producing a tornado soon

compared to a supercell with a Z_{DR} column area <40 km^2) works with our dataset as well. The vast majority (71 %) of column areas >40 km^2 in our dataset occurred with pretornadic supercells. However, almost half (47%) of storms with column areas <40 km^2 in our dataset produced tornadoes, so a column area <40 km^2 does not necessarily mean a storm will be nontornadic.

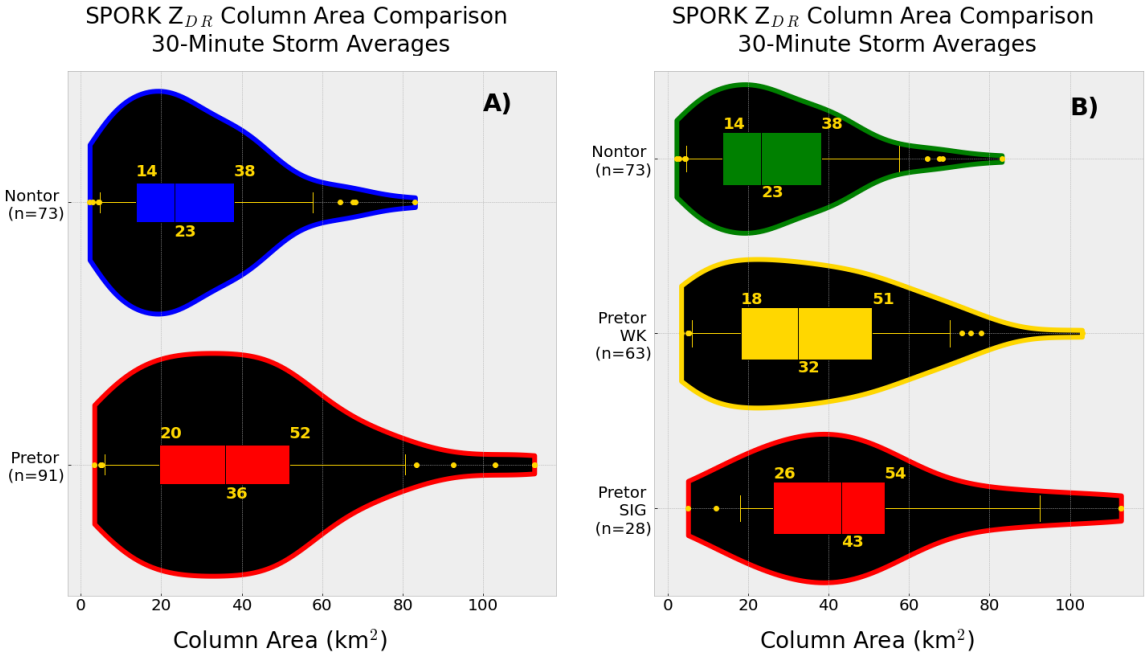


Figure 5: Violin plots comparing a) pretornadic and nontornadic Z_{DR} column areas obtained from SPORK and b) nontornadic column areas to pretornadic areas from storms that produce weak or significant tornadoes. Gold text indicates the 25th, 50th, and 75th percentile values for each sample. Whiskers extend to the 5th and 95th percentiles, and outliers are plotted as gold circles.

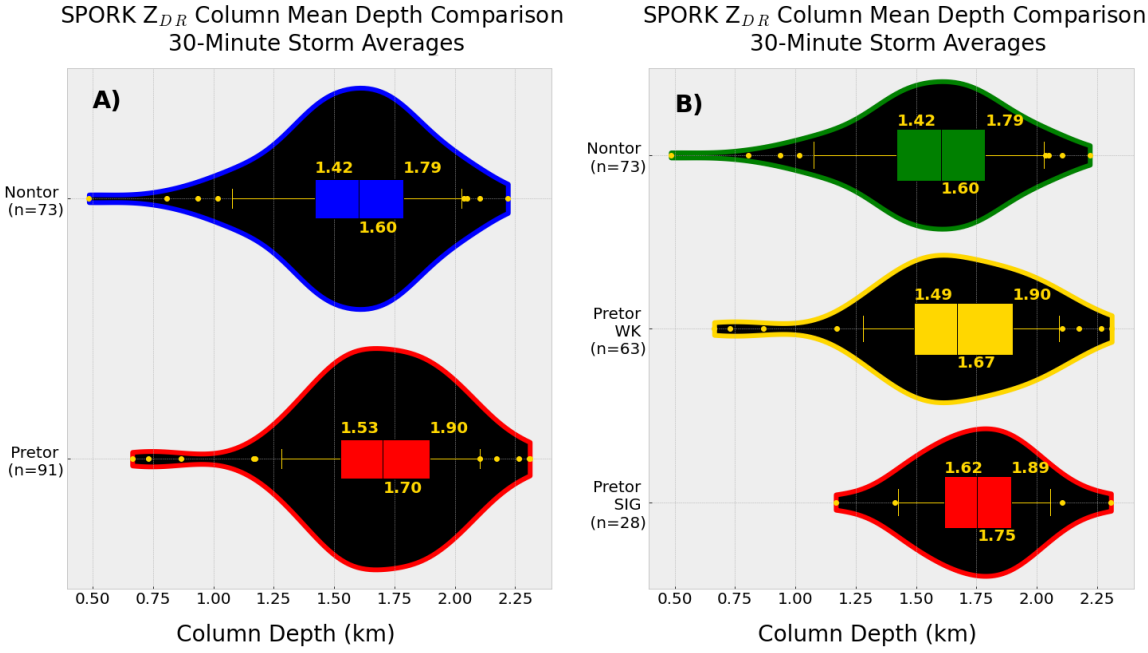


Figure 6: As in Fig. 5, but for Z_{DR} column mean depth.

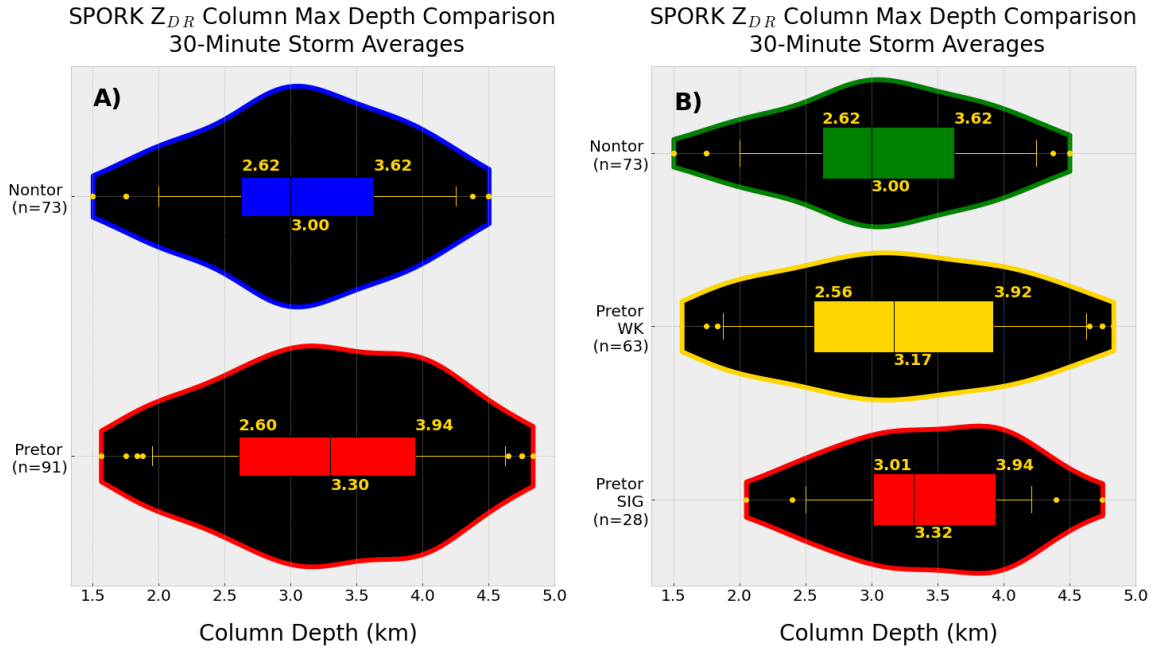


Figure 7: As in Fig. 5, but for Z_{DR} column maximum depth.

Z_{DR} column-depth metrics were not found to be usefully different between pretornadic and nontornadic supercells, as shown in Table 3 and Figs. 6 and 7. The mean value of maximum Z_{DR} column depth was 3.27 km for pretornadic supercells and 3.08 km for nontornadic supercells, and the mean value of mean Z_{DR} column depth was 1.69 km for tornadic storms and 1.57 km for nontornadic storms (Table 3), and the distributions of both metrics for pretornadic and nontornadic supercells almost entirely overlap (Figs. 6 and 7). Although the difference between pretornadic and nontornadic Z_{DR} column mean depths was significant ($p = 0.01$), the two distributions have substantial overlap. The magnitude of the difference (120 m), compared to the typical vertical beam spacing with which supercell updrafts are usually sampled by WSR-88Ds and the 250-m vertical resolution of the objective analyses used by SPORK, is quite small. Thus, this difference is likely not operationally useful. These results concur with Kuster et al. (2019), who found no significant differences between Z_{DR} column depths of the tornadic and nontornadic mesocyclones. Our results also concur with VDB20, who also found that Z_{DR} column depth did not exhibit a statistically significant difference between pretornadic and nontornadic supercells. Finally, as with Z_{DR} column area, significantly tornadic supercells did not display a notable, statistically significant increase in column depth compared to weakly tornadic storms (Figs. 6b, 7b).

Work by VDB20 found that hailfall areal extent was larger in nontornadic supercells than in pretornadic supercells. SPORK displayed similar differences in hailfall areal extent between pretornadic and nontornadic supercells to those found by VDB20. However, these differences were most apparent when timestamps of an undetected hailfall signature were considered to have a hailfall area of 0 km^2, instead of being excluded from the analysis (Table 3, Figs. 8 and 9). This is because a substantial fraction of radar scans in pretornadic supercells did not display a detectable hailfall signature, with 54 % of pretornadic storms displaying no SPORK-detected hailfall signature compared to 28 % of nontornadic storms. When these scans are excluded, pretornadic supercells have a mean hailfall area of 17 km^2 compared to 28 km^2 in nontornadic supercells, and the pretornadic and nontornadic hailfall area distributions share a fair amount of overlap (Fig. 8). When these scans are included, the difference between the pretornadic and nontornadic hailfall area distributions becomes even more apparent (Fig. 9).

SPORK-detected hailfall signatures are even more uncommon in pretornadic storms that go on to produce significant tornadoes than in

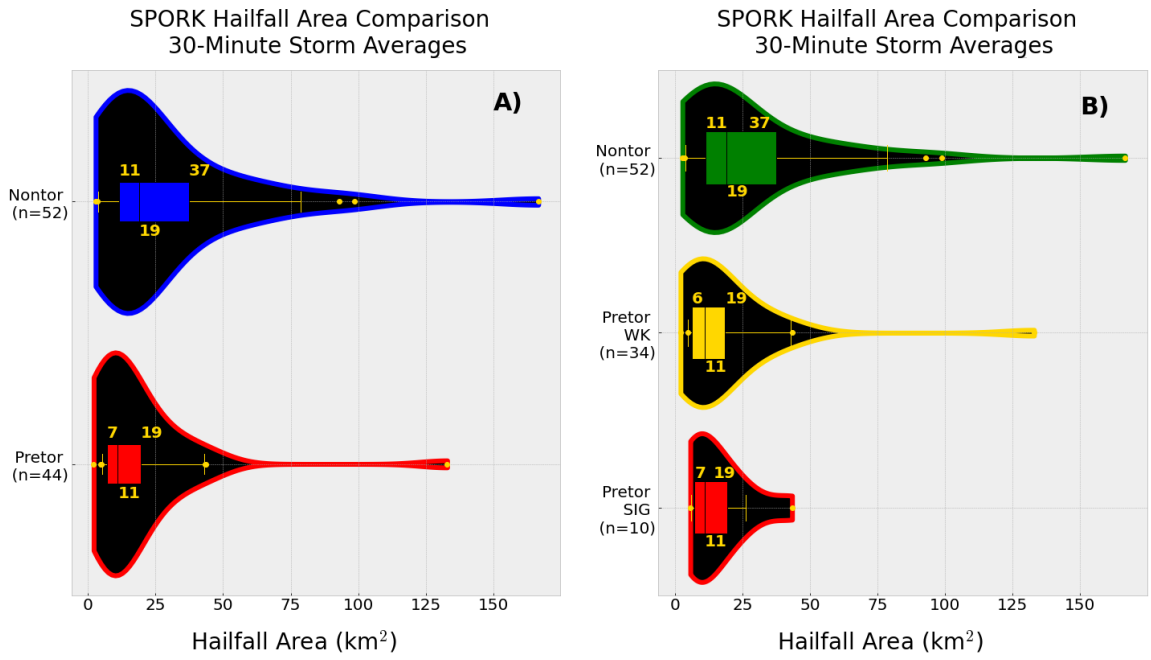


Figure 8: As in Fig. 5, but for hailfall extent excluding scans with no detected hailfall.

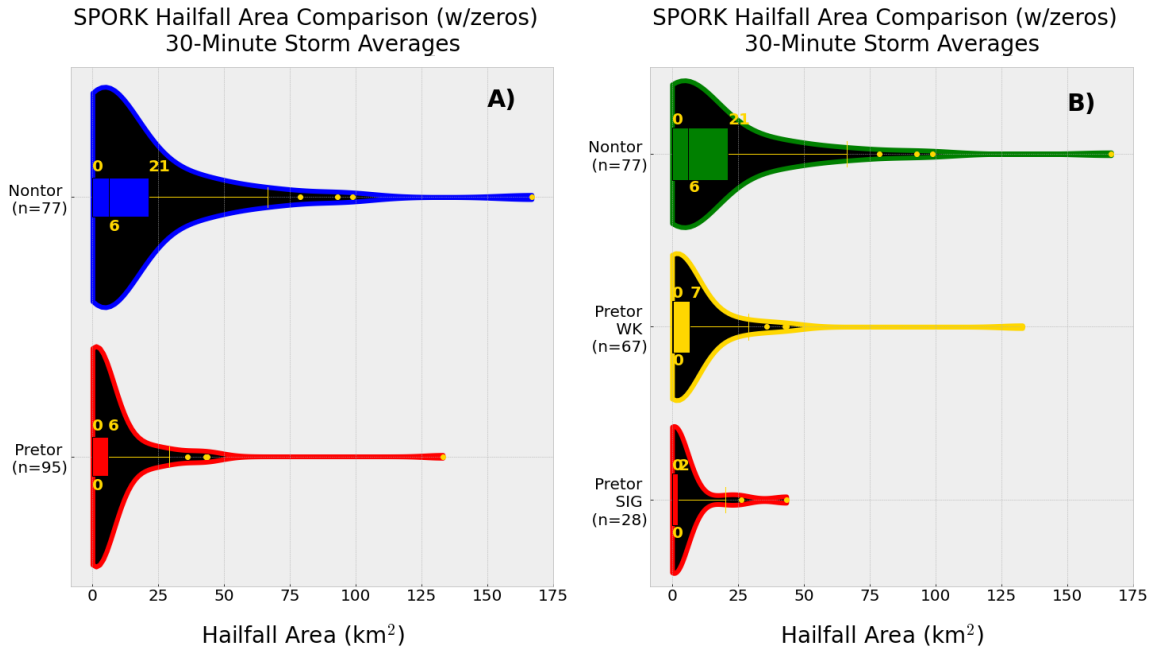


Figure 9: As in Fig. 5, but for hailfall extent calculated including hailfall areal extents of 0 km^2.

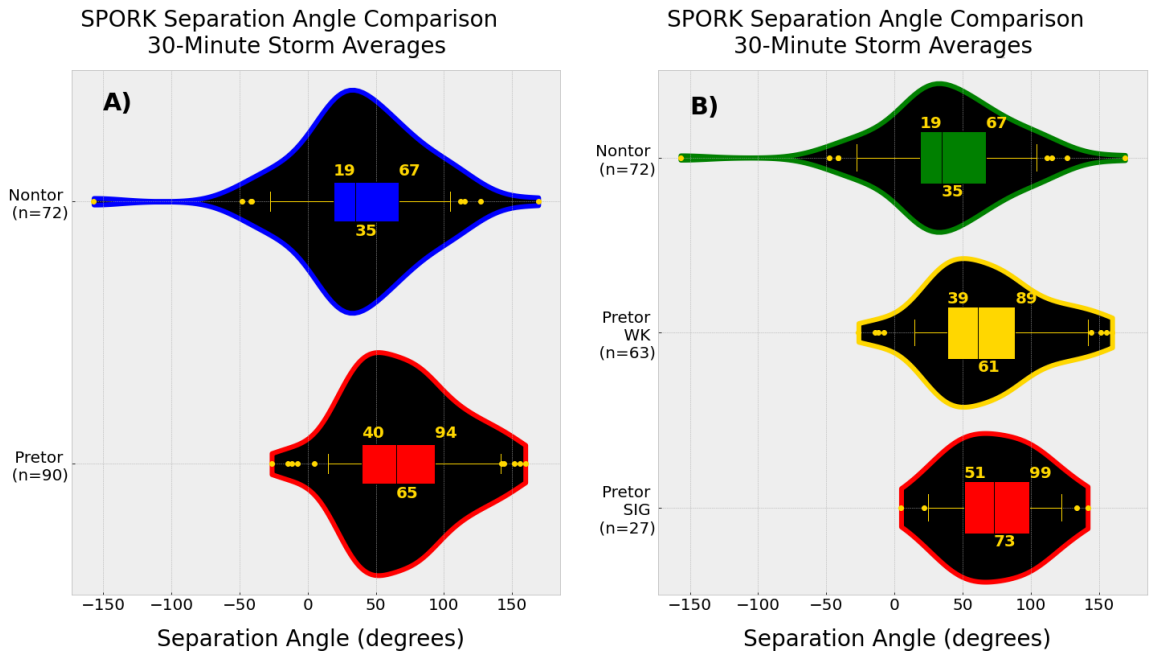


Figure 10: As in Fig. 5, but for separation angle.

pretornadic storms overall, with 64 % of such storms having no detectable hailfall signature compared to 49 % of storms producing weak tornadoes. Nontornadic supercells may be more likely to have bigger hailfall areas because the same relatively dry low-level environments with high LCLs that are favorable for delayed melting of hail (Rasmussen and Pruppacher 1982) also support excessively cold, dense outflow. Such outflow could impede the stretching of near-surface vorticity to tornado strength by the low-level updraft (Markowski et al. 2002). Furthermore, the tendency for nontornadic storms to have polarimetrically inferred hailfall signatures noted here is similar to Kumjian and Ryzhkov (2008), who found that 48 % of volume scans in their 9-storm tornadic sample had no dual-pol hail signature, compared to 15 % of the volume scans in their 6-storm nontornadic sample.

Loeffler et al. (2020) and Loeffler and Kumjian (2018) found that K_{DP}–Z_{DR} separation angles tend to be larger and closer to perpendicular to storm motion in tornadic than nontornadic storms. Separation angles closer to perpendicular hypothetically are associated with tornadic storms, because they may indicate larger values of storm-relative helicity in a given storm's inflow (Loeffler and Kumjian 2018). Separation angles also may be smaller in nontornadic storms because the K_{DP} foot is located farther to the rear of the precipitation shield and closer to the updraft, which allows colder outflow air to advect under the low-level mesocyclone and impede tornadogenesis (Loeffler et al. 2020). The automated separation angles calculated by SPORK show similar, statistically significant differences between pretornadic and nontornadic supercells as found by Loeffler et al. (2020); pretornadic storms have a mean separation angle of 69°, compared to 38° for nontornadic supercells (Table 3, Fig. 10). Moreover, three-quarters of pretornadic supercells had separation angles >40°, while over half of nontornadic supercells had separation angles smaller than 40° (Fig. 10). As with Z_{DR} column-area and depth metrics, pretornadic storms that produced significant tornadoes did not have meaningfully larger separation angles than weakly tornadic storms (Fig. 10b)

These results differ from those of Van Den Broeke (2021), who found that separation angles in pretornadic supercells and in nontornadic supercells with a strong low-level mesocyclone were fairly similar. However, the average separation angles found by Van Den Broeke (2021) were large for both pretornadic and tornadogenesis failure cases (78.4° for tornadic storms and 79.4° for the nontornadic cases), which might be expected since all of the storms examined developed a strong low-level mesocyclone. Limiting the nontornadic storms

in this dataset to only the 24 storms that meet the nrot >1 criterion from Van Den Broeke (2021) still results in a notable difference between separation angles in this sample (36°/40° mean/median) and the pretornadic mean separation angle of 69°. Thus, more work is likely needed comparing separation angles in a large sample of nontornadic supercells that develop strong low-level rotation to tornadic supercells. This could determine if the separation angle is used best as simply a diagnostic indicator of a storm-scale environment favorable for strong low-level mesocyclones, or if the signature can provide more information about a given storm's likelihood of producing a tornado, given that it produces a strong low-level mesocyclone.

b. Correlations between dual-pol metrics and environmental variables

To examine how each of the dual-pol signatures explored in this study is affected by environmental parameters, Spearman's rank-order correlations were calculated for each of the dual-pol metrics, and for all of the environmental variables calculated from each storm's RAP proximity sounding. For a consistent analysis across all cases, and to ensure the dual-pol metrics calculated represented the environment shown by the proximity sounding for each storm, all dual-pol metrics in this analysis were calculated as storm means for all available scans within the hour centered on the sounding time. To be included in this analysis, a storm metric was required to have at least 3 valid SPORK values within the hour-long analysis window. The results of these calculations are displayed in Table 2, and some of the more notable correlations and possible physical explanations are noted in the following discussion.

Z_{DR} column area had moderate positive correlations with MLCAPE (r = 0.35) and MUCAPE (r = 0.32), as well as the environmental 0ºC level (r = 0.37). Weaker but still statistically significant positive correlations existed with effective bulk shear vector magnitudes (r = 0.18), SRH calculated using several combinations of depths and storm motions (r = 0.17 to 0.30), and EL height (r = 0.31). A weak negative correlation was also found with LFC height (r = –0.22). The moderate correlations between Z_{DR} column area and CAPE agree with modeling work showing that supercell updrafts are larger in environments with more CAPE (Lin and Kumjian 2022). Furthermore, the supercell composite parameter [SCP, Thompson et al. (2003)] and significant tornado parameter [STP, Thompson et al. (2003, 2012)] both had larger correlations with Z_{DR} column area than any individual parameter (r = 0.42 for both), likely reflecting the mutually reinforcing influence of both CAPE and SRH in the calculations of these metrics.

Although Z_{DR} column area was found to have a weak positive correlation with SRH, this was likely not due to stronger storm-relative flow leading to larger updrafts, as in Peters et al. (2020), and thus larger Z_{DR} columns. This is because Z_{DR} column area and storm-relative flow had no statistically significant correlation over any layer in this analysis. The contrast with the findings of Peters et al. (2020) raises the possibility that Z_{DR} column areas may not correlate well with updraft width, or that the updraft speed contour that might be approximated by the Z_{DR} column is different from that used in their analysis. Future work should compare a large sample of Z_{DR} columns to other available updraft size proxies (such as overshooting top areas) to test Z_{DR} column-area correlations with updraft areas, especially in cases where large amounts of hail may mask the Z_{DR} column signature partially.

Z_{DR} column mean and maximum depths were both most strongly correlated with MLCAPE (r = 0.52/0.52, mean depth/maximum depth, respectively) and MUCAPE (r = 0.49/0.50). Both also displayed moderate positive correlations with the environmental 0ºC level (r = 0.41/0.35) and equilibrium level (EL) heights (r = 0.44/0.41). Outside of a weak positive correlation between maximum column depth and effective SRH (r = 0.15/0.19 using Bunkers et al. (2000)/observed storm motions) and a weak negative correlation between surface–6-km bulk shear vector magnitudes and mean column depth (r = –0.19), column-depth metrics generally displayed little or no correlation with shear and SRH metrics. As with column area, both column-depth metrics displayed modest positive correlation with SCP (r = 0.37/0.44) and STP (r = 0.34/0.35). However, these correlations were both smaller than the correlations between column depth and CAPE, possibly due to the lack of a correlation between column depth and SRH, as was present for column area. Overall, these results indicate that Z_{DR} column depth in supercells is influenced

strongest by environmental thermodynamics, including CAPE, the environmental 0ºC level, and EL height.

The individual variable with the strongest influence on hailfall area was the environmental 0ºC level, with a moderate negative correlation (r = –0.41). Weak to moderate negative correlations also were noted between hailfall area and SRH calculated over several different layer and storm motion combinations (r = –0.20 to –0.35), as well as with CIN (r = –0.21), SCP (r = –0.30), STP (r = –0.44), 0–3-km CAPE (r = –0.20), and EL height (r = –0.27). Weak to moderate positive correlations, meanwhile, were noted for several layers of low-level storm-relative flow below 3 km (r = 0.20 to 0.26) as well as LFC (r = 0.28) and LCL heights (r = 0.27). Taken together, these correlations indicate a moderate tendency for supercells with large hailfall areas to occur in environments less favorable for tornadogenesis, with higher LFC and LCL heights, weaker low-level shear vectors and SRH, and less 0–3-km CAPE. Indeed, examining a scatterplot of hailfall area and STP shows that hailfall areas of 25 km^2 or larger are very uncommon in environments of STP ≥2 (Fig. 11a). These results also concur with modeling work by Kumjian et al. (2021) and Lin and Kumjian (2022), where less curved hodographs with lower SRH were more favorable for the production of large amounts of hail, and that hail production did not scale monotonically with the amount of environmental SBCAPE.

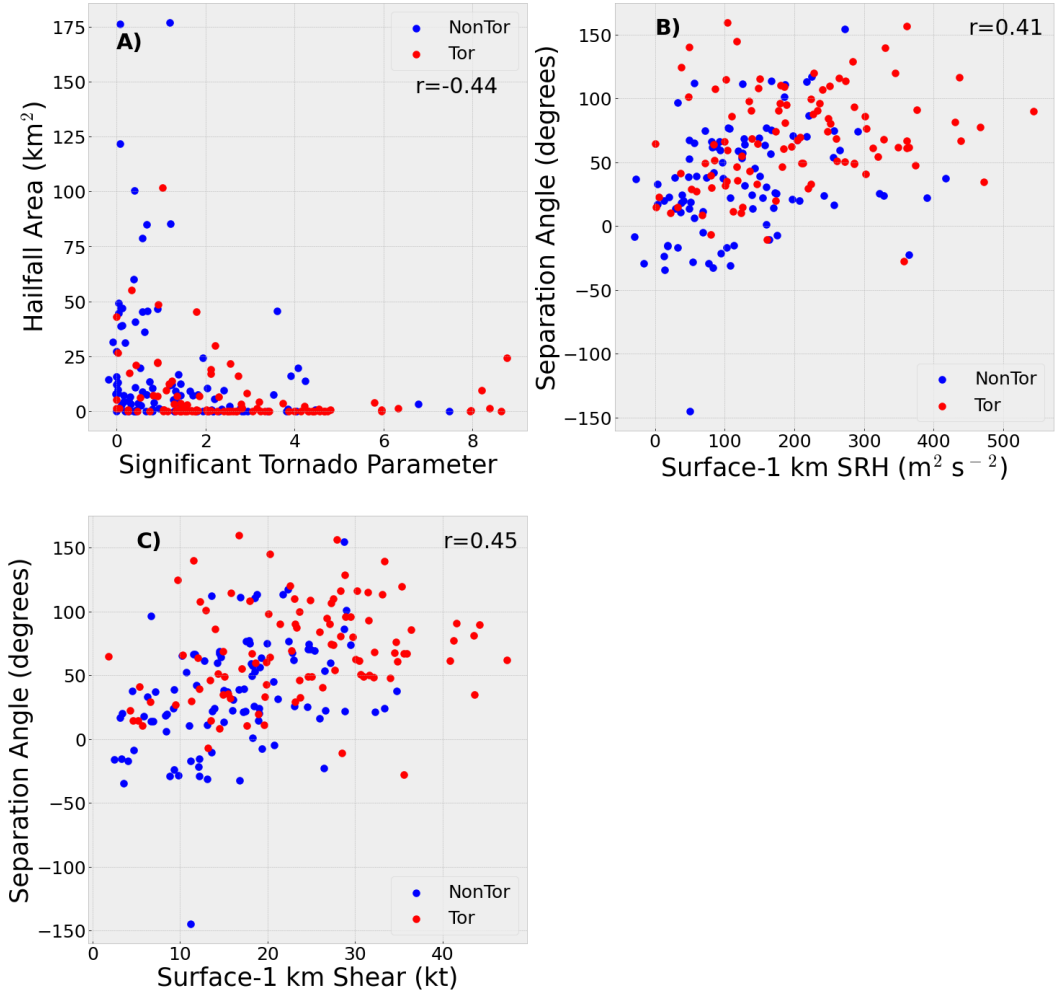


Figure 11: Scatterplots of a) hailfall area and the significant tornado parameter, b) K_{DP}–Z_{DR} separation angle and surface–1-km SRH and c) K_{DP}–Z_{DR} separation angle and surface–1-km shear vector magnitude.

Separation angle displayed a moderate positive correlation with surface–1-km shear vector magnitude (r = 0.45) and a weak correlation with surface–3-km shear vector magnitude (r = 0.22), as well as weak to moderate correlations with all SRH metrics examined (r = 0.22 to 0.41) and with 0–3-km CAPE (r = 0.29). Scatterplots of separation angle and surface–1-km SRH (Fig. 11b) and surface–1-km shear vector magnitude (Fig. 11c) show that separation angles smaller than 50° tend to be uncommon with surface–1-km SRH >300 m^2 s^{-2} or surface–1-km shear vectors are >30 kt (15 m s^{-1}), and that separation angles smaller than 0° tend to be uncommon with surface–1-km SRH >200 m^2 s^{-2} or surface–1-km shear vectors >20 kt (10 m s^{-1}). Weak to moderate negative correlations also were observed between separation angle and LFC (r = –0.31) and LCL heights (r = –0.32), as well as with all storm-relative wind metrics for which the layer started at the surface (r = –0.26 to –0.37). Likely due to the influence of the low-level shear and SRH terms, separation angle also showed weak to moderate positive correlations with SCP (r = 0.24) and STP (r = 0.35).

The moderate positive correlations between low-level SRH and separation angle are generally consistent with the results of Loeffler and Kumjian (2018, 2020) and Loeffler et al. (2020), which indicated that the K_{DP}–Z_{DR} separation angle should be more perpendicular to storm motion in higher-SRH environments. Furthermore, the correlations between more favorable 0–3-km CAPE, LCL height, and LFC height values for tornadic supercells and larger separation angles makes sense in light of: 1) Loeffler and Kumjian (2020)'s showing that separation angles tend to be larger in tornadic storms, and 2) findings presented earlier in this paper that SPORK-calculated separation angles tend to be larger in pretornadic supercells in the 30 min prior to tornadogenesis than in nontornadic supercells. While Loeffler and Kumjian (2020) convincingly demonstrate the physical link between changes in size-sorting behavior and larger separation angles in higher-SRH environments, theoretical explanations of the observed correlations between 0–3-km CAPE, LCL height, LFC height, and the separation angle have yet to be explored. Thus, the correlations between these variables and the separation angle may simply be a function of these variables and the separation angle both being larger in tornadic supercell environments, rather than a physical connection between larger values of these parameters and larger separation angles.

4. Summary and conclusions

In this study, we have built on the work of WV21 to introduce the Supercell Polarimetric Observation Research Kit (SPORK), an automated Python algorithm capable of identifying and quantifying supercell dual-pol signature metrics, including Z_{DR} column area and depth, hailfall area, and the K_{DP}–Z_{DR} separation angle. We evaluated the dual-pol signature metrics produced by SPORK by comparing them to manual calculations from VDB20. Then we ran the algorithm on a sample of 206 supercells to determine whether automatically calculated dual-pol metrics from SPORK displayed the same differences between tornadic and nontornadic storms found in VDB20 and Loeffler et al. (2020). We also examined how each dual-pol metric was related to various environmental parameters derived from RAP proximity soundings for each storm. Our main findings are summarized below:

- SPORK's calculations for Z_{DR} column area and hailfall area matched manual calculations from VDB20 well (r = 0.620 to 0.852); however, correlations between manually calculated Z_{DR} column depth metrics and values from SPORK were less strong (r = 0.404 to 0.551).

- The differences between the dual-pol signatures of tornadic and nontornadic supercells found by VDB20, Loeffler et al. (2020), Homeyer et al. (2020) and French and Kingfield (2021) also showed up in SPORK's automated analysis of a larger sample of storms. Pretornadic supercells had larger Z_{DR} column areas, smaller hailfall areas, and larger separation angles than nontornadic storms. Some of these differences, particularly those involving column area and separation angle, may be large enough to be operationally useful.

- Storm-mean dual-pol metrics from SPORK mainly displayed moderate to weak correlations with various environmental parameters. Notably, Z_{DR} columns were larger and deeper in more unstable environments (r = 0.32 to 0.53), and slightly larger in environments with higher SRH (r = 0.17 to 0.30). Hailfall area was larger

in environments with lower environmental 0ºC levels (r = –0.41), lower SRH (r = –0.20 to –0.35), and higher LCLs and LFCs (r = 0.27 to 0.28). Separation angles were larger in environments with stronger low-level shear vectors (r = 0.45) and SRH (r = 0.22 to 0.41), as well as lower LCLs and LFCs (r = –0.31 to –0.32). Accordingly, Z_{DR} columns and separation angles were larger in environments with higher STP (r = 0.42 and 0.35), while hailfall area was smaller (r = –0.44). These tendencies may be useful to forecasters, since they may indicate that a storm with a larger Z_{DR} column, smaller hailfall area, and larger separation angle is in an environment more favorable for tornadogenesis than vice versa. These results could be especially useful in cases with multiple proximal supercells, as differences in dual-pol signatures may indicate that one storm is experiencing a more favorable storm-scale environment for a tornado in the near term than its neighbors, and that may not be otherwise resolved in operationally available data.

The work presented here represents a first step toward expanding the WV21 algorithm into automatically quantifying multiple dual-pol metrics in supercells. SPORK shows promise in quickly identifying and quantifying supercell dual-pol signatures with similar results to VDB20, and SPORK-identified dual-pol metrics display similar differences between tornadic and nontornadic supercells as found using smaller samples and less-automated analyses by VDB20 and Loeffler et al. (2020). Still, the results presented here come with some caveats that will be addressed in future work. Namely, since SPORK is designed for right-moving supercells, it may miss or misclassify dual-pol signatures in nonsupercell storms or left-moving supercells. Additionally, the potential covariances between some of the environmental parameters examined in section 3b complicate the drawing of connections between individual environmental parameters and dual-pol signature characteristics.

Future studies using SPORK should examine much larger supercell samples for fuller relationships between supercell dual-pol signatures and storm environments. Larger samples of more intense tornadoes also would provide a better understanding of how dual-polarization signatures vary with tornado rating than can be gleaned from the relatively small sample here. They should also compare dual-pol signatures of large samples of nearby tornadic and nontornadic storms to see if any useful dual-pol differences exist between them on the same day. Furthermore, the use of dual-pol metrics from SPORK, or similar algorithms as predictors in machine-learning algorithms to produce probabilistic short-term forecasts of convective hazards (such as ProbSevere (Cintineo et al. 2020)), should be explored.

ACKNOWLEDGMENTS

This research was supported by NOAA (grant numbers NA18OAR4590307 and NA19OAR4590340). The authors would also like to thank Nicholas Humrich for helpful work in consolidating and streamlining the code for SPORK, as well as Charles Kuster, Joey Picca, Dr. Jeffrey Snyder, and Dr. Matthew Kumjian for their insightful reviews that greatly improved this manuscript.

REFERENCES

Adlerman, E. J., K. K. Droegemeier, and R. Davies-Jones, 1999: A numerical simulation of cyclic mesocyclogenesis. *J. Atmos. Sci.*, **56**, 2045–2069, https://doi.org/10.1175/1520-0469(1999)056<2045:ANSOCM>2.0.CO;2.

Atkins, N. T., M. L. Weisman, and L. J. Wicker, 1999: The influence of preexisting boundaries on supercell evolution. *Mon. Wea. Rev.*, **127**, 2910–2927, https://doi.org/10.1175/15200493(1999)127<2910:TIOPBO>2.0.CO;2.

Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L Thompson, and J. A. Hart 2017: SHARPpy: An open-source sounding analysis toolkit for the atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **98**, 1625–1636, https://doi.org/10.1175/BAMS-D-15-00309.1.

Brandes, E. A., J. Vivekanandan, J. D. Tuttle, and C. J. Kessinger, 1995: A study of thunderstorm microphysics with multiparameter radar and aircraft observations. *Mon. Wea. Rev.*, **123**, 3129–3143, https://doi.org/10.1175/1520-0493(1995)123<3129:asotmw>2.0.co;2.

Bunkers, M. J., B. A. Klimowski, J. W. Zeitler, R. L. Thompson, and M. L. Weisman, 2000: Predicting supercell motion using a new hodograph technique. *Wea. Forecasting*, **15**, 61–79, https://doi.org/10.1175/1520-0434(2000)015<0061:PSMUAN>2.0.CO;2.

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.

Coffer, B. E., and M. D. Parker, 2017: Simulated supercells in nontornadic and tornadic VORTEX2 environments. *Mon. Wea. Rev.*, **145**, 149–180, https://doi.org/10.1175/MWR-D-16-0226.1.

——, M. D. Parker, R. L. Thompson, B. T. Smith, and R. E. Jewell, 2019: Using near-ground storm relative helicity in supercell tornado forecasting. *Wea. Forecasting*, **34**, 1417–1435, https://doi.org/10.1175/WAF-D-19-0115.1.

Coniglio, M. C., and M. D. Parker, 2020: Insights into supercells and their environments from three decades of targeted radiosonde observations. *Mon. Wea. Rev.*, **148**, 4893–4915, https://doi.org/10.1175/MWR-D-20-0105.1.

Crowe, C. C., W. A. Petersen, L. D. Carey, and D. J. Cecil, 2010: A dual-polarization investigation of tornado-warned cells associated with Hurricane Rita (2005). *Electronic J. Oper. Meteor.*, **4**, 1–25.

——, C. Schultz, M. Kumjian, L. Carey, and W. Petersen, 2012: Use of dual-polarization signatures in diagnosing tornadic potential. *Electronic. J. Oper. Meteor.*, **13**, 57–78.

Dawson, D. T., E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-level Z_{DR} signatures in supercell forward flanks: The role of size sorting and melting of hail. *J. Atmos. Sci.*, **71**, 276–299, https://doi.org/10.1175/JAS-D-13-0118.1.

——, ——, and M. R. Kumjian, 2015: Does wind shear cause hydrometeor size sorting? *J. Atmos. Sci.*, **72**, 340–348, https://doi.org/10.1175/JAS-D-14-0084.1.

Dowell, D. C., and H. B. Bluestein, 2002: The 8 June 1995 McLean, Texas, storm, Part I: observations of cyclic tornadogenesis. *Mon. Wea. Rev.*, **130**, 2626–2648, https://doi.org/10.1175/1520-0493(2002)130<2626:TJMTSP>2.0.CO;2.

Duda, J., and W. Gallus, 2010: Spring and summer Midwestern severe weather reports in supercells compared to other morphologies. *Wea. Forecasting*, **25**, 190–206, doi: https://doi.org/10.1175/2009WAF2222338.1

Flournoy, M. D., M. C. Coniglio, E. N. Rasmussen, J. C. Furtado, and B. E. Coffer, 2020: Modes of storm-scale variability and tornado potential in VORTEX2 near- and far-field tornadic environments. *Mon. Wea. Rev.*, **148**, 4185–4207, https://doi.org/10.1175/MWR-D-20-0147.1.

French, M. M. and D. M. Kingfield, 2021: Tornado formation and intensity prediction using polarimetric radar estimates of updraft area. *Wea. Forecasting*, **36**, 2211–2231, https://doi.org/10.1175/WAF-D-21-0087.1

Hampshire, N. L., R. M. Mosier, T. Ryan, and D. Cavanaugh, 2018: Relationship of low-level instability and tornado damage rating based on observed soundings. *J. Oper. Meteor.*, **6**, 1–12, https://doi.org/10.15191/nwajom.2018.0601.

Healey, D. J. and M. S. Van Den Broeke, 2022: Comparing polarimetric signatures of proximate tornadic and nontornadic supercells in similar environments. Proc., *38th Conf. on Environmental Information Processing Technologies*, Houston, TX (remote), Amer. Meteor. Soc., J4B.5, [Available online at https://ams.confex.com/ams/102ANNUAL/meetingapp.cgi/Paper/396210.]

Helmus, J. J., and S. M. Collis, 2016: The Python ARM radar toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *J. Open Res. Software*, **4**, https://doi.org/10.5334/jors.119.

Homeyer, C. R., T. N. Sandmæl, C. K. Potvin, and A. M. Murphy, 2020: Distinguishing characteristics of tornadic and nontornadic supercell storms from composite mean analyses of radar observations. *Mon. Wea. Rev.*, **148**, 5015–5040, https://doi.org/https://doi.org/10.1175/MWR-D-20-0136.1.

Klees, A. M., Y. P. Richardson, P. M. Markowski, C. Weiss, J. M. Wurman, and K. Kosiba, 2016: Comparison of the tornadic and nontornadic supercells intercepted by VORTEX2 on 10 June 2010. *Mon. Wea. Rev.*, **144**, 3201–3231, https://doi.org/10.1175/MWR-D-15-0345.1.

Klemp, J. B., and R Rotunno, 1983: A study of the tornadic region within a supercell thunderstorm, *J. Atmos. Sci.*, **40**, 359–377, https://doi.org/10.1175/1520-0469(1983)040<0359:ASOTTR>2.0.CO;2

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47**, 1940–1961, https://doi.org/10.1175/2007JAMC1874.1.

——., and ——, 2009: Storm-relative helicity revealed from polarimetric radar measurements. *J. Atmos. Sci.*, **66**, 667–685, https://doi.org/10.1175/2008JAS2815.1.

——, A. P. Khain, N. Benmoshe, E. Ilotoviz, A. V. Ryzhkov, and V. T. Phillips, 2014: The anatomy and physics of Z_{DR} columns: Investigating a polarimetric radar signature with a spectral bin microphysical model. *J. Appl. Meteor. Climatol.*, **53**, 1820–1843, https://doi.org/10.1175/JAMC-D-13-0354.1.

Kuster, C. M., T. J. Schuur, T. T. Lindley, and J. C. Snyder, 2020: Using Z_{DR} columns in forecaster conceptual models and warning decision making. *Wea. Forecasting*, **35**, 2507–2522, https://doi.org/10.1175/WAF-D-20-0083.1.

Kuster, C. M., J. C. Snyder, T. J. Schuur, T. T. Lindley, P. L. Heinselman, J. C. Furtado, J.W. Brogden, and R. Toomey, 2019: Rapid-update radar observations of Z_{DR} column depth and its use in the warning decision process. *Wea. Forecasting*, **34**, 1173–1188, https://doi.org/10.1175/WAF-D-19-0024.1.

Lemon, L. R. and C. A. Doswell III., 1979: Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Wea. Rev.*, **107**, 1184–1197, https://doi.org/10.1175/1520-0493(1979)107<1184:STEAMS>2.0.CO;2.

Lin, Y., and M. R. Kumjian, 2022: Influences of CAPE on hail production in simulated supercell storms, *J. Atmos. Sci.*, **79,** 179–204, https://doi.org/10.1175/JAS-D-21-0054.1

Loeffler, S. D., and M. R. Kumjian, 2018: Quantifying the separation of enhanced Z_{DR} and K_{DP} regions in nonsupercell tornadic storms. *Wea. Forecasting*, **33**, 1143–1157, https://doi.org/10.1175/WAF-D-18-0011.1.

——, and ——, 2020: Idealized model simulations to determine impacts of storm-relative winds on differential reflectivity and specific differential phase fields. *J. Geophys. Res. Atmos*, **125**, e2020JD033 870, https://doi.org/10.1029/2020JD033870.

——, ——, M. Jurewicz, and M. M. French, 2020: Differentiating between tornadic and nontornadic supercells using polarimetric radar signatures of hydrometeor size sorting. *Geophys. Res. Lett.*, **47**, https://doi.org/10.1029/2020GL088242.

Loney, M. L., D. S. Zrnić, J. M. Straka, and A. V. Ryzhkov, 2002: Enhanced polarimetric radar signatures above the melting level in a supercell storm, *J. Appl. Meteor. Climatol.*, **41**, 1179–1194, https://doi.org/10.1175/1520-0450(2002)041<1179:EPRSAT>2.0.CO;2

Maddox, R. A., L. R. Hoxit, and C. F. Chappell, 1980: A study of tornadic thunderstorm interactions with thermal boundaries. *Mon. Wea. Rev.*, **103**, 322–336, https://doi.org/10.1175/15200493(1980)108<0322:ASOTTI>2.0.CO;2.

Magee, K. M., and C. E. Davenport, 2020: An observational analysis quantifying the distance of supercell-boundary interactions in the Great Plains. *J. Oper. Meteor.*, **8**, 15–38, https://doi.org/10.15191/nwajom.2020.0802.

Markowski, P. M., E. N. Rasmussen, and J. M. Straka, 1998: The occurrence of tornadoes in supercells interacting with boundaries during VORTEX-95. *Wea. Forecasting*, **13**, 852–859, https://doi.org/10.1175/1520-0434(1998)013<0852:TOOTIS>2.0.CO;2.

——, J. M. Straka, and E. N. Rasmussen, 2002: Direct surface thermodynamic observations within the rear-flank downdrafts of nontornadic and tornadic supercells. *Mon. Wea. Rev.*, **130**, 1692–1721, https://doi.org/10.1175/1520-0493(2002)130<1692:DSTOWT>2.0.CO;2.

——, Y. Richardson, E. Rasmussen, J. Straka, R. Davies-Jones, and R. J. Trapp, 2008: Vortex lines within low-level mesocyclones obtained from pseudo-dual-doppler radar observations. *Mon. Wea. Rev*., **136**, 3513–3535, https://doi.org/10.1175/2008MWR2315.1 .

May, R. M., S. C. Arms, P. T. Marsh, E. Bruning, and J. R. Leeman, cited 2017: MetPy: A Python package for meteorological data. Unidata. [Available online at https://github.com/Unidata/MetPy.] https://doi.org/10.5065/D6WW7G29.

McKinney, W., 2010: Data structures for statistical computing in Python. Proc., *9th Python in Science Conf.*, S. van derWalt and J. Millman, Eds., 56–61, https://doi.org/10.25080/Majora-92bf1922-00a.

Orf, L., R. Wilhelmson, B. Lee, C. Finley, and A. Houston, 2017: Evolution of a long-track violent tornado within a simulated supercell. *Bull. Amer. Meteor. Soc.*, **98**, 45–68, https://doi.org/10.1175/BAMS-D-15-00073.1.

Parker, M. D., 2014: Composite VORTEX2 supercell environments from near-storm soundings. *Mon. Wea. Rev.*, **142**, 508–529, https://doi.org/10.1175/MWR-D-13-00167.1.

Peters, J. M., C. J. Nowotarski, J. P. Mulholland, and R. L. Thompson, 2020: The influences of effective inflow layer streamwise vorticity and storm-relative flow on supercell updraft properties, *J. Atmos. Sci.*, **77**, 3033–3057, https://doi.org/10.1175/JAS-D-19-0355.1

Rasmussen, R., and H. R. Pruppacher, 1982: A wind tunnel and theoretical study of the melting behavior of atmospheric ice particles. I: A wind tunnel study of frozen drops of radius <500 μm. *J. Atmos. Sci*, **39**, 152–158, https://doi.org/10.1175/1520-0469(1982)039<0152:AWTATS>2.0.CO;2

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.

Rotunno, R and J.B. Klemp, 1985: On the rotation and propagation of simulated supercell thunderstorms. *J. Atmos. Sci*., **42**, 271–292, https://doi.org/10.1175/1520-0469(1985)042<0271:OTRAPO>2.0.CO;2

Ryzhkov, A.V., M.R. Kumjian, S.R. Ganson, and A.P. Khain, 2013: Polarimetric radar characteristics of melting hail. Part I: theoretical simulations using spectral microphysics modeling. *J. Appl. Meteor. Climatol.*, **52**, 2849–2870, https://doi.org/10.1175/JAMC-D-13-073.1

Sessa, M. F., and R. J. Trapp, 2020: Observed relationship between tornado intensity and pretornadic mesocyclone characteristics. *Wea. Forecasting*, **35**, 1243–1261, https://doi.org/10.1175/WAF-D-19-0099.1.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Snyder, J. C., A. V. Ryzhkov, M. R. Kumjian, A. P. Khain, and J. Picca, 2015: A Z_{DR} column detection algorithm to examine convective storm updrafts. *Wea. Forecasting*, **30**, 1819–1845, https://doi.org/10.1175/WAF-D-15-0068.1.

——, H. B. Bluestein, D. T. Dawson, and Y. Jung, 2017: Simulations of polarimetric, x-band radar signatures in supercells. Part I: description of experiment and simulated ρ_{hv} rings. *J. Appl. Meteor. Climatol.*, **56**, 1977–1999, https://doi.org/10.1175/JAMC-D-16-0138.1

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

——, B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous united states. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, https://doi.org/10.1175/WAF-D-11-00116.1.

Trapp, R. J., G. J. Stumpf, and K. L. Manross, 2005: A reassessment of the percentage of tornadic mesocyclones. *Wea. Forecasting*, **20**, 680–687, https://doi.org/10.1175/WAF864.1.

——., G. R., Marion, and S. W. Nesbitt, 2017: The regulation of tornado intensity by updraft width, *J. Atmos. Sci*., **74,** 4199–4211, https://doi.org/10.1175/JAS-D-16-0331.1

Van Den Broeke, M. S., 2016: Polarimetric variability of classic supercell storms as a function of environment. *J. Appl. Meteor. Climatol.*, **55**, 1907–1925, https://doi.org/10.1175/JAMC-D-15-0346.1.

Van Den Broeke, M. S., 2017: Polarimetric radar metrics related to tornado life cycles and intensity in supercell storms. *Mon. Wea. Rev.*, **145**, 3671–3686, https://doi.org/10.1175/MWR-D-16-0453.1.

——, 2020: A preliminary polarimetric radar comparison of pre-tornadic and nontornadic supercell storms. *Mon. Wea. Rev.*, **148**, 1567–2584, https://doi.org/https://doi.org/10.1175/MWR-D-19-0296.1.

——, 2021: Polarimetric radar characteristics of tornadogenesis failure in supercell thunderstorms. *Atmosphere*, **12**, 581, https://doi.org/10.3390/atmos12050581.

——, J. M. Straka, and E. N. Rasmussen, 2008: Polarimetric radar observations at low levels during tornado life cycles in a small sample of classic southern Plains supercells. *J. Appl. Meteor. Climatol.*, **47**, 1232–1247, https://doi.org/10.1175/2007JAMC1714.1.

Wade, A. R., M. C. Coniglio, and C. L. Ziegler, 2018: Comparison of near- and far-field supercell inflow environments using radiosonde observations. *Mon. Wea. Rev.*, **146**, 2403–2415, https://doi.org/10.1175/MWR-D-17-0276.1.

Weiss, C. C., D. C. Dowell, J. L. Schroeder, P. S. Skinner, A. E. Reinhart, P. M. Markowski, and Y. P. Richardson, 2015: A comparison of near-surface buoyancy and baroclinity across three VORTEX2 supercell intercepts. *Mon. Wea. Rev.*, **143**, 2736–2753, https://doi.org/10.1175/MWR-D-14-00307.1.

Wilson, M. B., and M. S. Van Den Broeke, 2021: An automated Python algorithm to quantify Z_{DR} arc and K_{DP}–Z_{DR} separation signatures in supercells. *J. Atmos. Oceanic Technol.*, **38**, 371–386, https://doi.org/10.1175/JTECH-D-20-0056.1.

Ziegler, C. L., E. R. Mansell, J. M. Straka, D. R. MacGorman, and D.W.Burgess, 2010: The impact of spatial variations of low-level stability on the life cycle of a simulated supercell storm. *Mon. Wea. Rev.*, **138**,1738–1766, https://doi.org/10.1175/2009MWR301.1.

Zrnić, D. S., 1987: Three-body scattering produces precipitation signature of special diagnostic value. *Radio Sci*., **22**, 76–86, https://doi.org/10.1029/RS022i001p00076.

REVIEWER COMMENTS

[Authors' responses in *blue italics.*]

REVIEWER A (Charles M. Kuster):

Initial Review:

Recommendation: Accept with major revisions.

General Comments: The authors present information about an automated method to quantify several important supercell dual-pol signatures. The method is validated using past manual analyses and then used on a relatively large sample of supercells to examine differences in dual-pol signatures between tornadic and nontornadic supercells. Relationships between environmental conditions and dual-pol signatures are also examined. Overall, the paper is interesting, generally well written, and results are important and relevant for the research and operational community. Figures and tables are also generally very well done. I have several substantive (major) comments regarding consistent wording, terminology, and additional data analysis as well as many technical (minor) comments, so I recommend the paper be accepted after major revisions. Nice work!

First, thanks for your insightful review of our paper! We've made almost all of the changes you suggested. We'd also like to note that we discovered and fixed a bug in our code which was making SPORK's area calculations too large across the board. After fixing this bug and retraining the random-forest algorithm to use to new area calculations, our results comparing pretornadic and nontornadic supercells and examining correlations between dual-pol metrics and environmental variables did not materially change (except for the Z_{DR} column and hailfall areas getting smaller across the board). However, the verification of SPORK metrics against the manual calculations from VDB20 did change somewhat, with a notable low bias appearing in the Z_{DR} column area calculations (somewhat expected given the differing Z_{DR} thresholds between VDB2020 and this work), and a smaller low bias appearing in SPORK's hail area calculations. Correlations between SPORK-calculated and manually calculated metrics showed only small changes. We've added a mention of these biases in the algorithm verification section.

Substantive comments: I understand why the 30-min window prior to tornadogenesis was selected for tornadic supercells, but I am concerned about making direct comparisons between tornadic and nontornadic storms when the analysis windows are not the same. For the nontornadic supercells, results could be biased downwards (e.g., lower Z_{DR} column depth) since the full mesocyclone lifecycle analysis window likely includes mesocyclone development and dissipation and maybe even storm dissipation. Z_{DR} columns are more likely to be smaller and weaker during those times. Tornadic supercell results are less likely to have this bias since just a 30-min window prior to tornadogenesis is looked at, and that very well may not include mesocyclone development and dissipation. For a better comparison (i.e., more "apples to apples"), perhaps choosing a 30-min window prior to low-level mesocyclone maximum intensity (as measured by rotational velocity) would be better for the nontornadic supercells, similar to Van Den Broeke (2021). Or comparisons could be made when looking at the entire lifecycle of the tornadic supercell's mesocyclone. Do the results change if you apply either of these time windows? Was Z_{DR} column depth more likely to be lower near the beginning and end of the nontornadic supercell mesocyclone's lifetime?

These are all good points! We've reworked the pretornadic/nontornadic analysis to use a 30-min window prior to peak low-level rotation for each nontornadic storm (except for storms where no clear low-level rotation exists, or where the apparent rotation is an artifact of vertical sidelobe contamination, for which a 30-min window prior to the RAP sounding used in the environmental analysis was used). This reduced the sample size for the nontornadic dataset somewhat, but did not otherwise change the results much except for Z_{DR} column depth, which is no longer notably different between pretornadic and nontornadic supercells using the new analysis window. The methods and results section have been rewritten with these updates. Applying the new window doesn't greatly change the results, except for Z_{DR} column depth, which is now nearly the same for pretornadic and nontornadic storms. This likely indicates that column depth were lower for the nontornadic supercells during the beginning or end of their life cycles.

Similarly, this difference in time windows could make it more difficult to compare real-time algorithm output to the results of the paper. A forecaster won't be able to know magnitudes of the dual-pol metrics for the full life cycle of a mesocyclone until it has dissipated. Therefore, real-time values could not be compared to the results for nontornadic storms in a meaningful way. If a 30-min window were used for both types of supercells, at least the starting point and associated limitations would be the same.

Addressed in the response to the comment above by going with a 30-min window for both subsets.

This text [original p. 13] does not appear to match with the text in section 3a. Did you mean, "As opposed to the previous section…", or was the analysis in 3a actually carried out over the full analysis period for both tornadic and nontornadic supercells?

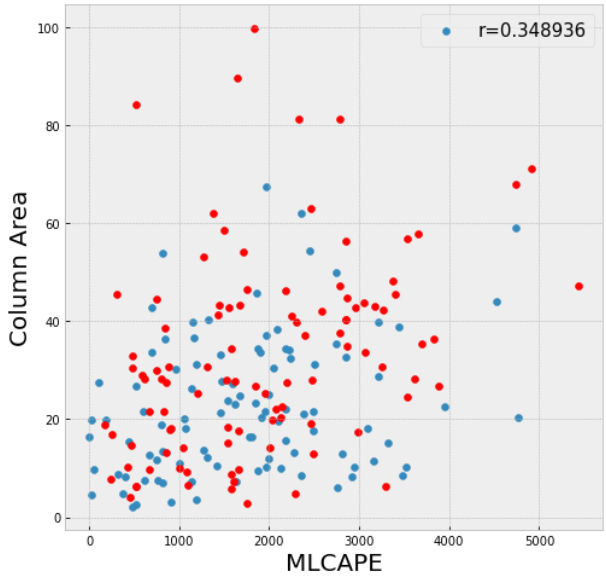
I meant "as opposed to" (the wording in that sentence was not the best, my mistake). This section has now been updated to use a consistent 1-h window centered on the RAP sounding used for each storm for consistency, and to limit the impact of changing environments on storms with long analysis periods in the previous iteration of the analysis comparing dual-pol metrics to environmental parameters.

How might the real-time algorithm be used by forecasters in the context of the results in this paper? Does the real-time algorithm output statistics for the past 30 min of data for all storms or is it based on each individual volume scan? If it is the latter, it could be difficult for forecasters to compare the real-time output with the results of this paper. If the real-time output is not intended to be used with the results of this paper, that should be explicitly stated somewhere, likely on p. 6. If the algorithm is intended to be used with the results of this paper and it is based on individual scans, would using all of the dual-pol metric values from individual scans (same idea as Fig. 3 from algorithm validation) in the analysis make more sense here rather than 30-min or storm-averaged values? Do the results change at all if you use values from individual volume scans rather than averaged values? If the hope is for the paper to be used operationally, analysis methods and the associated results should be as close as possible to what a forecaster could use in real-time operations.

The real-time algorithm currently displays results from each individual volume scan, since that was the simplest way to code it up and get it to work in real time. We've experimented with running the comparisons using individual scans as well as 10-, 15-, and 30-min pretornadic/pre-peak-nrot averaging windows, and the comparisons don't change appreciably. Future iterations of the algorithm could be configured to use a 30-min (or other duration) time window, but for the proof of concept presented here we thought it would be good to keep it simple, especially since I'd really want to get some operational forecasters' opinions on whether single scans or a given averaging window might be best, before settling on a final choice. For now, we've added a note that the current real-time algorithm displays only the metrics for a given volume scan, and that future algorithm updates may allow generation of time-average stats in real time.

It might be interesting to see how magnitudes of various environmental parameters compare to magnitudes of various dual-pol signatures. For example, a scatter plot of K_{DP}– Z_{DR} separation angle to 0–1-km shear and/or Z_{DR} column depth to MLCAPE could be added. Knowing a rough range of the values that could be expected for the environment and radar could help complete the conceptual model applied during forecast/warning operations. For example, what are typical CAPE values when Z_{DR} column area is less than and greater than 50 km^2 (mentioned in 3a)?

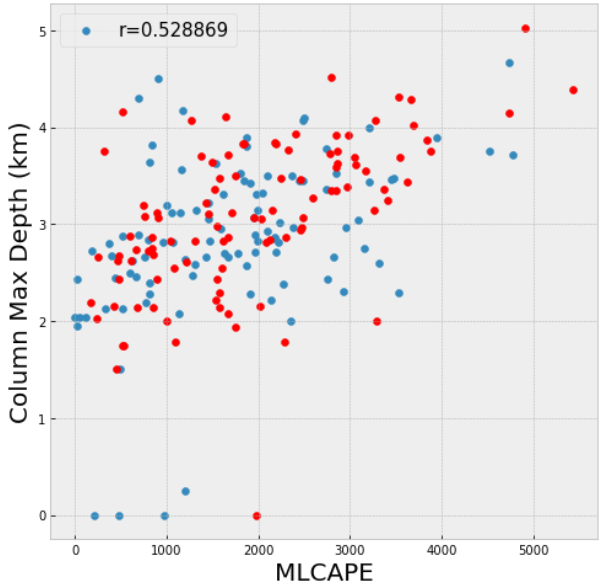
We considered adding such figures, but we weren't sure how much value they would add to the manuscript and we were concerned that they would end up making it too long. Here's a few examples of them:



(The red dots are tornadic storms, blue nontornadic). To answer your question about column area and MLCAPE, there doesn't appear to be a single threshold value above which column areas of a certain size become much more common; however, columns larger than 40 km2 are very uncommon with <1000 J kg^{-1} of CAPE, and very small columns (<20 km^2) appear quite uncommon above 3000 J kg^{-1} of CAPE.

[Editor's Note: 0–1-km shear scatterplot not shown, since it became part of Fig. 11 in the final paper.]

Similarly, separation angles >50° are uncommon with <10 kt of surface–1-km shear, and separation angles <50° are uncommon when surface–1-km shear is >30 kt.



Finally, column depths >3 km are fairly uncommon with <1000 J kg^{-1} of CAPE, and depths <3 km are fairly uncommon with >2000 J kg^{-1}.

We ended up adding a 3-panel figure with some of these correlations as Fig. 11.

This additional figure/analysis would be especially important if the authors believe that the analysis shows that dual-pol signature metrics can provide information about a storm's near-storm environment that might not be captured by observations or models due to inadequate resolution. This idea is suggested in the last bullet of the conclusions. However, statements [elsewhere] make it harder to determine how the dual-pol signatures might be used with environmental information. Do the environmental conditions drive forecaster expectations for what dual-pol signatures to anticipate (e.g., the environment is not favorable for tornadic supercells so I should expect to see larger hailfall areas), or do the dual-pol signatures provide more precise information about the storm-scale near-storm environment that might be missed by the spatial resolution of observations and models (e.g., there are two storms close to each other in what appears to be a relatively homogeneous environment but one has a large hailfall area and the other does not so perhaps in reality one is in a more favorable environment than the other)? I think the latter option is more likely and is what the authors are suggesting as well. If so, I recommend making this point clearer on p. 15 since this result would be operationally useful. If there is a lot of uncertainty here, then just leaving this hypothesis as part of a bullet in the conclusion section would be appropriate.

We were thinking a bit of both. The knowledge from this paper might drive forecaster expectations of what kind of dual-pol signatures to expect in a given environment (i.e., expect small Z_{DR} columns in an environment with low CAPE and weak SRH), and thus a storm which defies those expectations and has a strong Z_{DR} column in such an environment might warrant additional scrutiny. It could also be useful in cases of storm in close proximity with different dual-pol signatures. We'll clarify this in the text.

Figuring out the best terminology with the calculated statistics can be tricky, especially when there are median values of averages of maximum values. There is a fine line between being explicit enough to be clear and so explicit that there are a lot of terms for the reader to wade through (e.g., median average maximum mean value), but below are a few examples where it was not overly clear what was being presented. I don't ultimately have a great answer here other than being as clear and explicit as possible. Perhaps terminology could be introduced and explained at the beginning of 3a or somewhere in section 2 and then more simplified terms (i.e., median ZDR column depth, maximum ZDR column depth, etc.) could be used after that introduction.

We have addressed this comment (and its sub-comments below) by adding a note saying that the pretornadic-nontornadic analysis is conducted using 30-min time means of each signature, and then simplified the wording of these statistics in the locations below. For consistency, we also have moved away from using median values and used means in all cases, since the take-away points don't change with either statistic.

This text [original p. 3] appears to differ/contrast with text near the end of section 3a and in the second bullet of the conclusions regarding the K_{DP}–Z_{DR} separation angle. Why is Van Den Broeke (2021) not discussed in the 3a paragraph, especially since the results of SPORK simulations suggest something different? Clarification is needed here or in section 3a.

A discussion of VDB21 has been added to section 3a, as well as some examination of a possible reason for the differing results. Initially, we thought that the different criteria for the nontornadic sample might be the reason for the difference, since VDB21 required the tornadogenesis failure cases to still have a strong low-level meso and the sample in this paper did not. However, applying the VDB21 criteria to this paper's nontornadic dataset still yielded a notable difference between the pretornadic and nontornadic separation angles, and we're not quite sure why. This could be an interesting direction for future research repeating VDB21's work with a much larger sample size, or using a different criterion (for example, a false-alarm tornado warning) to identify tornadogenesis failure cases. A note of this has been added to the text.

I agree with the possible explanations for why there were lower correlations between automated and manual measurements of Z_{DR} column depth. I do think there is one more important one that should be added, though. I have noticed a difference when measuring Z_{DR} column depth when using interpolated and non-interpolated (i.e., just the raw data from each elevation angle) data. Z_{DR} column depth is typically a little lower when using non-interpolated data. In this case, the gridded data would be interpolated and it appears as if the manual data would not have been interpolated. Figures 2 and 3 show that the SPORK

values (interpolated) were typically just a little higher than the manual values, and I think the reason stated above might explain a lot of this. With interpolation, I have seen methods frequently fill 1+ dB above the top of a given elevation angle, which causes a higher measured Z_{DR} column depth. This is ok because I think Z_{DR} column depth is underestimated slightly by radar, especially as range increases and gaps between successive elevation angles get larger.

Thanks for pointing this out! This has been added to the list of reasons why the SPORK column depths may be different from the manual analyses.

[Minor comments omitted...]

Second review:

Recommendation: Accept with minor revisions.

General comment: The authors have done a very good job addressing my comments and the article is now even stronger and clearer. I now only have some minor (technical) comments. The article is important to the research and operational community in that it presents information about an algorithm that quantifies multiple potentially useful dual-pol signatures and I recommend it be accepted pending minor revisions. I commend the authors on a job well done!

Thanks again for your helpful comments! We've addressed them below, and we've also fixed a small error we found when making these edits—the r value on Fig. 3a should be 0.817 instead of 0.793, and this has been fixed on the figure and where it appears in the text. [Editor's note: repeated mention of this fix is omitted from subsequent review responses, since it is documented here.]

REVIEWER B (Joseph C. Picca):

Initial Review:

Reviewer recommendation: Accept with major revisions.

Summary: This paper introduces a novel toolkit that automates calculation of a number of relevant dual-pol signatures in supercells. It then uses these calculations from a number of tornadic and nontornadic cases for a statistical analysis of any potential connections among these signatures, supercellular tornadogenesis, and the near-storm environment. While the motivation and overall effort of the paper is sound and could be a worthwhile contribution to the literature, there are questionable and/or unclear analysis methods, and numerous inconsistencies and errors, in the paper. My primary concern is the difference in analysis windows between pretornadic and nontornadic supercells, which I think introduces a non-negligible bias in the data. Given these data are a centerpiece, I think major revisions are necessary.

First, thanks for your helpful review! We've made almost all of the changes you suggested.

[Editor's note: omitted identical description of fixed SPORK bug as in Reviewer A reply.]

Substantive Comments: Section 1—I think there could be a touch more discussion on why the separation angle may not be a great predictor of tornadogenesis, given it's a diagnostic tool of one variable; furthermore, it's well known that this variable (low-level SRH) isn't a direct predictor of tornadogenesis because there are further processes to take us from low/mid-level meso to tornado. I'm not saying anything you write here is incorrect, but rather, I think you could make clearer that it's understood that, from a physical standpoint, it is likely that the connection between separation angle and tornado potential is unclear, rather than just citing statistical analyses questioning the connection.

This is a good point, especially in context with the results from VDB21. We've updated this section to suggest that the separation angle is best used as a diagnostic tool that a storm is in a favorable storm-scale

environment for the production of strong low-level mesocyclones, and that tornadogenesis failure may still occur even if the storm does develop a strong low-level meso.

Section 2a—While I don't necessarily disagree with some of the thresholds chosen, I think a little more justification could be given. For example, what's the purpose behind changing the ZDR column threshold from 0.5 dB to 1 dB? Similarly, is the 1-dB hailfall threshold based in another study? Obviously the likelihood of hail is much higher at values less than 1 dB (given high *Z*), but if this is subjectively chosen in this paper, it should be stated. Additionally, the Z_{DR} filter (*Z* <20 dBZ, CC <0.7) seems fine for improving the analysis, but the justification isn't entirely accurate. You state "to remove regions of high Z_{DR} associated with NBF… ", but NBF does not necessarily bias Z_{DR} high or low. Rather, it leads to noisier Z_{DR} due to a decrease in data quality associated with the CC reduction. Similarly, Z_{DR} with a TBSS can transition from anomalously high (directly behind the hail core) to anomalously low at longer ranges so a TBSS is not purely associated with high Z_{DR}. Therefore I think the phrasing should be changed to something more like "To remove regions of questionable Z_{DR} associated with…" I would also consider breaking up the paragraph, perhaps at "As a final pre-processing step…" to improve the organization. Finally, what are the concerns about tilted updrafts when only looking in the vertical for continuous grid points? You reference them later on, but it's probably worth a mention here.

The 1-dB Z_{DR} threshold was chosen, instead of the 0.5-dB threshold used by VDB20 for Z_{DR} column identification, after initial testing showed that using a 0.5-dB threshold with the automated algorithm would often identify amorphous columns that extended much beyond what would subjectively be considered the Z_{DR} column core, while the 1-dB threshold was much better in identifying Z_{DR} column features that were more consistent with subjectively-identified Z_{DR} columns. This has been added to the text. The ≤1-dB Z_{DR} and ≥50-dBZ Z thresholds for large hail were chosen to be broadly consistent with the thresholds used by VDB16 and VDB20, and text to that effect has been added to the manuscript.

That's a good point that the biases from TBSSs/NBF aren't always positive—that has been clarified in the text using your suggested phrasing. A new paragraph now starts with "As a final pre-processing step." A caveat mentioning possible limitations of the vertical continuity requirement in fast-moving or highly tilted supercells has also been added.

Section 2a, second paragraph—I think this paragraph could hugely benefit from a figure showing the steps of drawing the Z_{DR} contour, splitting into objects, and then assigning to storms. Additionally, how are the remaining column objects combined into a single object and how does this impact the calculation of area, mean depth, and maximum depth? Is there any limit on how separated the individual objects can be before being combined into a single object? I think there could be more clarity regarding the overall process here.

We've added a figure depicting that process—thanks for the suggestion. The remaining objects are combined into one object by considering all points within the boundaries of the remaining object polygons to be part of one object. So, the area of the combined object is the sum of the areas of its component objects, the mean depth is the mean depth of all points within the component objects, and the maximum depth is the maximum depth of any of the component objects.

Section 2b—VDB20 uses different thresholds than SPORK (e.g., 0.5 dB vs. 1 dB for Z_{DR} columns). Is this an issue for validating SPORK's Z_{DR} column calculations against VDB20? Additionally, in Fig 2a, while it appears the 0.5-dB manual calculation generally has larger column areas than the algorithm does, I see a few instances where the algorithm recorded a much larger area, despite a more restrictive Z_{DR} threshold (1 dB vs. 0.5 dB). What's going on here that a tighter threshold is leading to larger column areas? In Fig 3a, there are of course even more instances of this phenomenon where the more restrictive threshold results in a considerably larger area. Furthermore, SPORK appears to calculate larger depths than the manual analysis, on average (Fig 2b), yet two of your three explanations in the text (vertical continuity and Z/CC threshold filtering) should cause SPORK to generally bias lower than the manual analysis. So those reasons don't explain the data in Fig 2b. Either there were issues with the prior manual analysis, or there are other contributing factors in the algorithm design…? Lastly, the r values presented in the text don't match the r values in Fig 3.

We initially considered the same 0.5-dB threshold as VDB20 for ZDR column area; however, the features identified with a 0.5-dB contour often extended well beyond what would be subjectively considered the Z_{DR} column, as mentioned in the response to the comment on page 4/paragraph 1. Once the area calculation issue mentioned in the text before the reviewer responses was corrected, the algorithm's 1-dB column areas did tend to be biased smaller than the manual 0.5-dB calculations, however, the correlation between the two datasets is still fairly good. With that issue fixed, fewer algorithm 1-dB areas are now present that are larger than the 0.5-dB manual areas. In the 5 cases where the average SPORK column area was larger than the manual area, two had elongated SPORK-identified Z_{DR} column objects which extended rearward along the storm's RFD, and the flanking line portion of these columns may have been excluded in the manual analyses. The remaining three had SPORK-identified columns composed of multiple Z_{DR} column objects. Subjective analysis of these objects shows them to be reasonable, but they may have been excluded in the original subjective analysis.

The reasons for possible errors in Z_{DR} column depth have been split into reasons for instances where the SPORK column depth is greater than the manual analyses, and reasons for the (less common) instances where the SPORK columns are shallower than the manual columns. One of the other reviewers pointed out that it's fairly common for the 1-dB level in interpolated Z_{DR} fields to be higher up than in the raw data, and we think this is likely the main reason why the SPORK Z_{DR} columns tend to be deeper than the manual analyses. R values have been updated to be correct. Thanks for catching that—those were from a previous version of the analysis and apparently didn't get updated in the text.

Section 3a—If I have this straight, the comparison involves 30 min prior to first tornadogenesis for pretornadic vs the entire period of supercellular characteristic (with radar coverage) for nontornadic. This seems like a sizable flaw in the analysis, given we're focused on a specific period with one group, but the entire period with the other group. For example, would this not introduce a bias in that we're looking at a period of time more likely for enhanced polarimetric signatures in the pretornadic cases, but smoothing everything out for the nontornadic cases? I think a much more robust analysis would involve choosing the time of max V_{rot} (or some other objective marker) in the nontornadic cases and then comparing the preceding 30 min with the preceding 30 min in pretornadic cases. I think either this analysis needs to be corrected or a strong explanation must be provided for why the current method is acceptable.

This is an excellent suggestion—we've reworked the analysis to use a 30-min window prior to peak low-level rotation (measured as the highest value of nrot in GR2 associated with a valid low-level circulation) for the nontornadic storms, except for storms where there was no clear low-level circulation or apparent circulations were the result of vertical sidelobe contamination, in which the analysis period was set to a 30-min window prior to the environmental sounding used for that case. The results are mostly the same, except for column depth, which is not smaller in nontornadic storms using the new analysis window.

Section 3b—You state, "As in the previous section, polarimetric signature metrics for the tornadic storms were calculated over the full analysis period for each storm instead of the 30 minutes prior to tornadogenesis." However, according to the description in 3a, the prior section compared the prior 30 minutes (pretornadic) with the full analysis period (nontornadic), for which I stated my concerns above, but here you suggest that might not be the case? Whichever way it is, there are numerous consistency issues that need to be addressed.

You're right—this wording could definitely have been clearer. In the submitted draft, the analysis comparing polarimetric metrics to environmental variables starting on p. 13 used dual-pol data from the full analysis period for each storm, whereas the pretornadic/nontornadic comparison used a 30-min pretornadic window for the pretornadic storms and the full analysis period for the nontornadic storms. In the revised draft, the environmental analysis has been reworked to consistently use a 1-h period centered on the RAP sounding used for each storm, and the text has been updated to make the analysis periods used for the pretornadic/nontornadic comparison and the environmental analysis much clearer.

When we see the correlations between separation angle and 0–3-km CAPE, LCL, LFC, etc., are we not just seeing that tornadic storms are more likely in environments with high 0–3-km CAPE and low LCL/LFC and that larger separation angles are also more likely with tornadic storms? You somewhat allude to this in

the following discussion and then again in the conclusion, but I don't see what these specific correlation values add besides what is already well established in literature (tornadoes are more often associated with higher CAPE, lower LCLs) given that there likely isn't much (or at best it's uncertain) physical connection between the separation angle and 0–3-km CAPE, LCLs, LFCs. If there is, it needs to be described/supported better here. Otherwise, is it worth the discussion/inclusion?

These correlations still exist to some degree when the tornadic and nontornadic datasets are analyzed separately. Thus, it's not entirely clear that these correlations are just due to separation angles being larger in tornadic storms, which are more likely in environments with large low-level CAPE and low LCLs/LFCs. We think the correlations are still worth noting, although we have strengthened the caveat that we don't quite know what the physical mechanism connecting them would be.

[Minor comments omitted...]

Second review:

Recommendation: Accept with minor revisions.

Summary: I think the authors have performed excellent work in revising the analysis and manuscript to address my concerns and the concerns of the other reviewers. Their work is an important step in further developing more nuanced dual-pol algorithms to assist warning forecasters, and it's encouraging to see their analysis build upon previous work. Therefore, I only have some minor comments on a few items to clean up, but I do not think it's necessary (for me at least) to see the paper again before acceptance. Nice work!

Thanks again for your comments--they've definitely helped build this into a better manuscript! We've made all the minor changes you requested except where noted.

[Minor comments omitted...]

REVIEWER C (Jeffrey C. Snyder):

Initial Review:

Reviewer recommendation: Accept with minor revisions.

Summary: This paper introduces a Python-based algorithm/software package that provides a set of dual-polarization radar signatures/quantities thought to be relevant to assessing supercell structure/evolution. The authors describe the algorithm, compare several of the algorithm's outputs for sets of tornadic and nontornadic supercells, and examine correlations between algorithm outputs and environmental characteristics. The content is appropriate for EJSSM. The paper is well written, with little in the way of meaningful grammar or spelling errors. The figures are clear, though the font size of the text in Figs. 1–3 is a bit small (≈8 pt). Past work is cited at an appropriate depth and breadth. This tool has the potential to substantially aid the analysis of polarimetric signatures in future studies (pardon the split infinitive), and this paper is worthy of publication pending minor revisions.

First, thanks for your insightful comments! We've made almost all of the changes you suggested.

[Editor's note: omitted identical description of fixed SPORK bug as in Reviewer A reply.]

Substantive comments: The authors include all data for a nontornadic supercell but only the 30 min preceding tornadogenesis for a tornadic/pretornadic supercell. How long are the nontornadic supercells typically tracked? The entire lifecycle? Until they passed beyond some range from the radar? Was there thought put towards using the same 30-min threshold for the nontornadic supercells as was used for the

tornadic/pretornadic supercells? Admittedly, it's hard to know which 30-min chunk of time to use, but one option would be to consider the time of strongest rotation (perhaps using something like AzShear), similar to what was done in Tuftedal et al. (2021).

Tuftedal, K. S., M. M. French, D. M. Kingfield, and J. C. Snyder, 2021: Observed bulk hook echo drop size distribution evolution in supercell tornadogenesis and tornadogenesis failure. *Mon. Wea. Rev.*, **149**, 2539–2557, https://doi.org/10.1175/MWR-D-20-0353.1.

These are all good questions/points! We've reworked the pretornadic–nontornadic comparison to use the 30 min prior to peak low-level rotation for the nontornadic storms as in VDB21 (as measured by the peak nrot value in GR2 for each storm), except for storms in which the apparent low-level circulation was likely a result of vertical sidelobe contamination or for which no clear low-level circulation was evident. Cases in which no clear circulation could be identified were discarded from the nontornadic dataset in the comparison. This didn't notably change the results for Z_{DR} column area, hailfall area, or separation angle, but did mostly eliminate the pretornadic/nontornadic differences in Z_{DR} column depth.

I'm not sure if there's been a sufficient amount of research to show that the SVC and LFVVS "play an important role in whether a given storm is likely to produce a tornado". The work that I'm aware of is generally from a small number of numerical simulations and a very limited number of observed cases. This isn't to say that the SVC or LFVVS isn't important—I'm just not sure we know enough right now to know whether that's true. If the authors feel otherwise, please add additional references as supporting evidence.

Good point—this section was intended to provide examples of internal supercell features which could be important in tornado production, but we ended up lumping features which have a fair bit of evidence about their importance (RFD buoyancy, cyclic mesocyclogenesis) with features with comparatively less evidence (the SVC and LFVVS). We've replaced the mention of the SVC/LFVVS features with a mention of baroclinic generation of surface vertical and near-surface streamwise vorticity in the supercell cold pool and updated the references accordingly.

The hailfall signature in SPORK is established using Z and Z_{DR} thresholds. To a first order, this is probably fine, but I'm wondering if the authors considered using results from the Hydrometeor Classification Algorithm (HCA)/Hail Size Discrimination Algorithm (HSDA) instead, since the HCA/HSDA is likely to provide a more robust estimate of the location of hail given the added discriminatory power of CC/ρ_{hv} and K_{DP}. I recognize that this would mean either running the HCA on the base data or fetching the Level 3 data that have the HCA output in them, but it is worthwhile to look at how the simple Z–Z_{DR} thresholds compare to the HCA-determined hailfall areas, at least for a few cases. Regardless, it is probably relevant to add, at the end of the quoted sentence above, something about how this generally holds true at S band by adding, for example, the following: "… relatively low values of Z_{DR} (<0.5–1 dB), at least at S band (e.g., Ryzhkov et al. 2013; Snyder et al. 2017)." At X band, hail may be associated with much lower Z.

The S-band clarification has been added to the text. We also spent a while trying to get the Level-III HCA data into Python to compare hailfall areas calculated using the HCA's hail classification to our threshold-based methodology, but the first source we looked at didn't have the HCA files, and we had trouble figuring out which level III files from the second source contained the HCA output. As coding up the full Park et. al. (2009) HCA would be rather time-consuming, we ended up skipping this comparison.

The algorithm uses the 1-km ARL CAPPI instead of the lowest radar tilt (1st paragraph in section 2a). I may have missed this earlier, but what happens if that lowest radar tilt (usually 0.5°) is centered above 1 km ARL? This is essentially getting at "what's the maximum range of storms being examined", unless the objective analysis scheme is extrapolating the lowest-available-level down to 1 km ARL.

The objective analysis scheme does extrapolate data from the lowest-available scan down to 1 km, provided that scan is still within the vertical radius of influence that the Barnes analysis scheme is using.

Tables 1 and 2 in VDB20 indicate that there were 32 and 31 supercells examined, respectively (i.e., 63 total

supercells). Are all of these different than the 50 supercells used to train the algorithm noted in the previous section?

These are all different from the 20 supercells used in the training dataset; however, a number of storms are used in both the VDB20 dataset and the 51-storm testing dataset. Since the overlapping storms are not used to train the algorithm but only to evaluate it, this shouldn't lead to an unrealistic estimate of algorithm performance. The uses of the 20-storm and 51-storm datasets have been clarified in the text.

There are a lot of "may" speculative instances later in [section 2b] when differences between the SPORK and the manual results are discussed. Can you look at some of the "misses" to figure out why the gridded Z_{DR} columns differ so much from the subjectively analyzed Z_{DR} columns?

To look into this further, we conducted a detailed examination of all times from the verification dataset in which a column over 3.5 km deep or 30 km^2 was missed. This analysis revealed several main reasons that columns were missed. The most common was that a very narrow but deep manually-identified column was missed by the algorithm in a storm with lots of inferred hailfall. These columns were misclassified as spurious by the random forest due to their small area (or just weren't picked up as possible objects to begin with), and the gridding process also seemed to make them more difficult to identify by smoothing their higher Z_{DR} values together with the low Z_{DR} values in surrounding lofted hail/graupel. In a few other instances, part of the supercell was erroneously identified as a separate storm by the tracking algorithm, and the Z_{DR} column was erroneously assigned to that spurious storm object. Finally, in the rarest failure mode, large columns in very large, elongated storms were misclassified as spurious due to their great distance relative to the algorithm-detected storm center, even with that distance normalized by storm area in the random forest. Making small adjustments up or down in the Z thresholds used for the tracking algorithm can often eliminate the spurious storms which lead to the second failure case. A discussion of these failure cases has also been added to the manuscript.

[Minor comments omitted...]

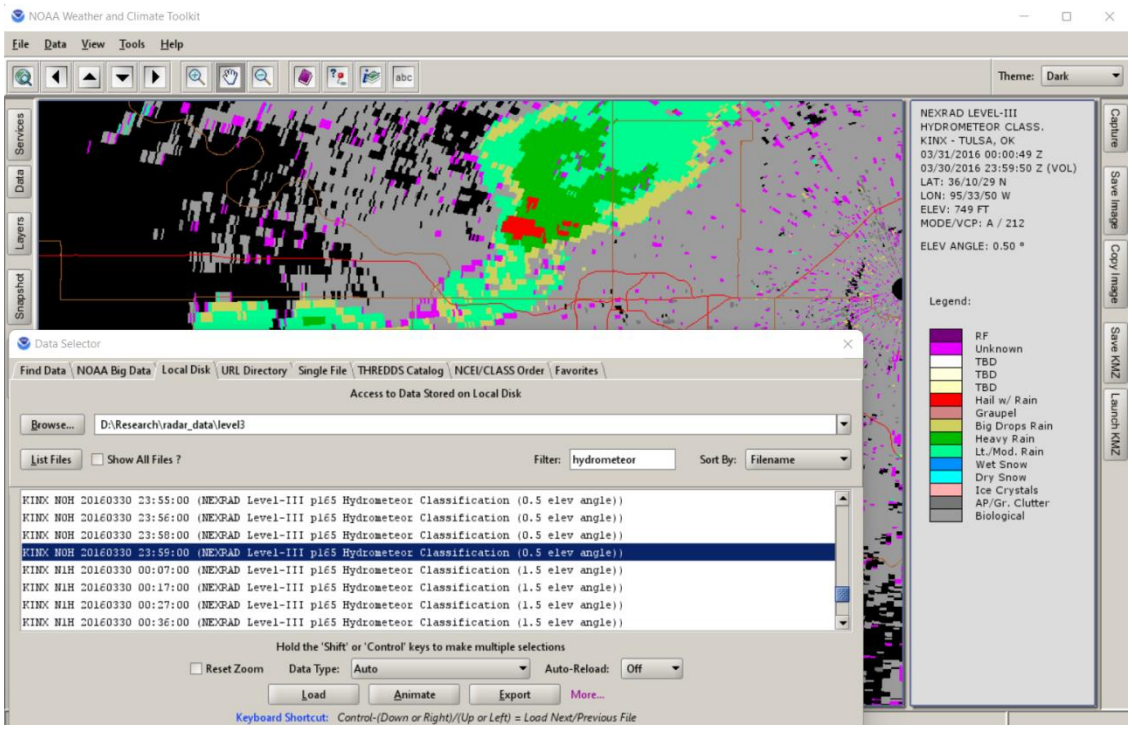
Second review:

Recommendation: Accept with minor revisions.

Summary: The paper nicely summarizes a new software tool that automates the collection of a handful of thought-to-be-relevant polarimetric characteristics of convective storms (and, more specifically, supercells). The authors' revisions to the paper are appreciated. My minor suggestions for changes to verbiage and punctuation are included in a separate document with tracked changes. Below are the very limited number of comments that do not pertain to such minor suggestions. This manuscript has improved, especially in terms illustrations and overall motivation. I have no further comments and recommend acceptance.

Thanks again for your helpful comments! We've made the minor changes you suggested to the document.

In my first review, I mentioned the possibility of leveraging the HCA results to assess hailfall extent rather than relying solely on the Z-ZDR pairs for hail identification. In response to this comment, the authors noted that they could not find the HCA data. For completeness, I'll point out that Level 3 data, in which the HCA output is contained, are available quite easily from the online archive accessible through the Weather & Climate Toolkit software package. You can download a tar file in the "NOAA Big Data" tab of the "Data Selector" (change source to "NEXRAD Level III") and then access the files in the "Local Disk" tab when files have been untarred. I've included an example screenshot below.



Thanks for showing us this—it's certainly easier to load the HCA files in WCT than trying to load them directly into Python. Subjectively comparing the HCA output to the SPORK hailfall areas for a few cases showed that the region SPORK identified as hailfall was generally smaller than the small hail (or hail w/ rain in older HCA data) region identified by the HCA, but larger than the "large hail" or "giant hail" regions identified in more recent HCA output.

Regarding the sentence "Several possible explanations exist for the lower correlation between the manual and algorithm-calculated Z_{DR} column depth values." Are there some cases in which the differences are particularly high? If so, can you look into those cases in a slightly more detailed manner to determine why the differences are so large? It's fine to provide reasonable speculations for why the differences may exist, but you have both the automated and manually identified data, so you should be able to look at the raw data to figure out why the differences exist. You probably don't want to do this for all cases, but it may be worth looking into if there are some particularly "egregious" cases with major differences.

That's a good point. An examination of all cases where the storm-mean value of maximum column depth was more than 1 km different from the manual value, showed that cases when SPORK's column depth was substantially larger most often were the result of high Z_{DR} from TBSSs just above a column which weren't removed by SPORK's QC, or from the vertical extension of >1 dB Z_{DR} by the gridding process already mentioned in this section. Cases where SPORK had a substantial low bias in column depth were mostly the result of tilted columns where SPORK's vertical continuity requirement caused it to underestimate the column depth. This section has been rewritten in light of these results.

What's the definition of "persistent" [midlevel mesocyclone] here? Please consider quantifying. On a similar note, consider quantifying what "the vast majority" means as it pertains to "column areas exceeding 40 km2 in our dataset" (even if it's simply a parenthetical statement that includes the specific proportion for which that holds).

We didn't set a specific time criterion for a persistent midlevel mesocyclone, but the storms did need to display one for the whole analysis period so a lower bound would be ~30 min for this dataset, and we've added this to the text. We've also added a parenthetical with the portion of column areas in our dataset exceeding 40 km^2 which were tornadic (71%).

[Minor comments omitted...]

REVIEWER D (Matthew R. Kumjian):

Initial Review:

Reviewer recommendation: Accept with major revisions.

Summary: The authors describe updates to the SPORK software package originally outlined in WV21. Namely, the new toolkit now calculates ZDR column metrics, in addition to those associated with Z_{DR}–K_{DP} signature separation at low levels. The new algorithm is then applied to ~200 supercells, roughly evenly split between tornadic and nontornadic storms. Various metrics quantifying differences in dual-pol radar signatures between tornadic and nontornadic supercells are presented, and compared to previous studies that used different methods and/or smaller sample sizes.

This is an interesting study that sheds some new light on the problem of dual-polarization radar analyses of supercell storms, particularly in trying to discern between those that produce tornadoes and those that do not. Some of the findings are quite striking, and, although consistent with previous works, do add to the confidence in the potential operational utility of the signatures. I have a large number of comments and suggestions for improving the manuscript, but don't see any show stoppers. Thus, I recommend Major Revisions, and would like to see the revised manuscript before recommending acceptance for publication.

First, thanks for your helpful review! We've made almost all of the changes you suggested, and we definitely think they've helped make this a better manuscript.

[Editor's note: omitted identical description of fixed SPORK bug as in Reviewer A reply.]

Substantive/General Comments:

1. Weaknesses in describing dynamics/processes. In general, I found some of the connections of signatures or behaviors of signatures to supercell processes to be weak or somewhat imprecisely worded. For example, it is unclear to me how results about the midlevel Z_{DR} column features are consistent with low-level mesocyclone steadiness, given the much different processes at play. Can the author please elaborate on how the dynamics of these two are linked?

Our initial thought was that the results of VDB17 and VDB20, which showed less variable Z_{DR} columns in tornadic storms compared to nontornadic storms and strongly tornadic storms compared to weakly tornadic storms, hint at steadier midlevel mesocyclones in these storms which are better able to intensify and maintain strong, steady low-level mesocyclones via tilting and stretching of storm-generated and environmental low-level vorticity. In storms where the Z_{DR} column is less steady, this may indicate disruptions in the midlevel updraft which may weaken its ability to support the low-level mesocyclone below via the processes mentioned above. However, to simplify this section and focus more on the updraft width arguments presented by Trapp et al. 2017 and Sessa and Trapp 2020, the mention of low-level mesocyclone steadiness has been removed.

"Nontornadic supercells may be…": I have a few problems with this sentence. First, this sentence has a few "mays" and comes across as a bit speculative. Does a storm's forward flank outflow often "make its way beneath the meso cyclone [sic]"? How does cold outflow impede stretching? (I understand what the authors are implying, but the wording is rather imprecise.) Does the low-level mesocyclone do the stretching? This sentence should be qualified that it is speculation, and should be cleaned up to accurately reflect the dynamics.

This section has been cleaned up to describe the dynamics more accurately.

Further, there is no physical justification presented by the authors that large areal extents of large hail would imply greater negative buoyancy generation than large areal extents of heavy rain. I would guess that it may depend on the low-level environment: drier environments would inhibit hail melting but

promote rain evaporation, and moister environments would do the opposite. Some explanation of the processes linking these thoughts should be included here or where the hypothesis is first introduced.

We've added text discussing the possibility that drier low-level environments with high LCLs that inhibit melting and allow large areas of hail to reach the surface are also favorable for cold RFDs that may hinder tornadogenesis to this section.

The discussion of Z_{DR} column width (and by proxy, updraft width) is interesting, but I'm left without a concrete explanation of the correlations. Instead of vague "expectations," connecting the environmental factors to the processes that lead to wider updrafts (and thus Z_{DR} columns) would be stronger here. Peters et al. studies have advanced our understanding of the controls of updraft width, and Lin and Kumjian (JAS, in press) also discuss how CAPE may influence updraft width. The inconsistency with Peters et al. deserves some further discussion, too. Why the inconsistency? Is it because Z_{DR} column width is not a robust proxy for updraft width?

We've reworked this section to mention the consistency with the results from Lin and Kumjian (2022) and elaborate on the inconsistency with Peters et al. (2020). We've also wondered how well Z_{DR} columns really represent updraft areas in supercells (especially in storms with large amounts of hail where the column may partially be masked out by frozen scatterers), so we've added a mention of this uncertainty to the paper as well. A study comparing overshooting top areas and Z_{DR} column areas in a large sample of storms would be a really interesting way to address this.

2. **Discussion of previous results with little to no statistical significance.** At times in the manuscript, especially when reviewing prior studies, the authors make claims about relationships between quantities or metrics, only to subsequently undercut those claims by saying there is no statistical significance to the relationship. [Two locations cited.] It's totally fine for papers to report null results, but I believe there is a difference in implication between saying "There were no statistically significant differences between Quantity A and Quantity B," and "Quantity A was larger than Quantity B, but the difference was not statistically significant." The former seems preferable, and perhaps more honest, in my opinion.

This is a good point. We ended up removing both of these mentions, as they were not critical to the background here.

3. **Data Presentation and Methods**. I have a few questions/concerns about the methods, and presentation of the data/results.

Figure 3: there are quite a few points where the manual detection identified large areas, and the algorithm detected zeroes. Were these cases inspected to see what the root cause was? Some discussion is warranted (i.e., a manually identified 4-km deep Z_{DR} column, which is extraordinary—not detected by the algorithm is somewhat concerning).

A detailed examination of all times from the verification dataset in which a column over 3.5 km deep or 30 km^2 was missed showed several root causes. The most common was that a very narrow but deep manually identified column was missed by the algorithm in a storm with plentiful inferred hailfall. These columns may have been misclassified as spurious by the random forest due to their small area, and the gridding process may also make them more difficult to identify by smoothing their higher Z_{DR} values together with the low Z_{DR} values in surrounding lofted hail/graupel. In a few other instances, part of the supercell was identified as a separate storm, and the Z_{DR} column was erroneously assigned to that spurious storm object. Finally, in the rarest failure mode, large columns in very large, elongated storms were misclassified as spurious due to their great distance relative to the algorithm-detected storm center, even with that distance normalized by storm area in the random forest. Making small adjustments up or down in the Z thresholds used for the tracking algorithm can often eliminate the spurious storms which lead to the second failure case. A discussion of these failure cases will also be added to the manuscript.

A general question about averaging over 30-min period: supercell storms can undergo significant variations in intensity, updraft width, etc., over a 30-min period. Is there any physical justification (aside from consistency with previous studies by the authors) for averaging over this period?

We mostly chose this averaging interval for consistency with previous work, but we've also run the tornadic-nontornadic comparisons with 15- and 10-min averaging intervals and the results didn't notably change.

4. **Conclusions.** I think some revision can help make the conclusion stronger and more informative for readers, particularly those who may skip to the end. For the bullets: try to be a bit more concrete, in case readers skip to this section. "Fairly well" and "less strong" are vague enough to not be informative. When describing the relationships to environmental factors in the final bullet, instead of lots of qualitative discussion of smaller/larger or higher/lower, you should explicitly include the correlations. A simple (r = xx) would suffice, probably. This is important because the correlations are, after all, rather low. Without that context, someone skimming the paper may take away a stronger message than the data suggest given the current wording.

The correlations have been added to this section of the conclusions, as well as to the environmental comparisons section of the results. Some of the qualitative wording has also been removed.

[Minor comments omitted...]

Second review:

Recommendation: Accept with minor revisions.

General comment: The authors have done considerable work in improving the manuscript, and fixing a notable error in the area calculations (which allows for improved comparisons with previous studies). The authors also re-did the analyses including a more fair comparison between time periods analyzed for the tornadic and nontornadic classes. Sensitivities to the thresholds chosen are discussed, and the language about weak correlations and statistical significance has been appropriately tempered. All of these amount to a manuscript that is in much better shape, and I have only a few relatively minor comments. Kudos to the authors for their hard work in improving the study!

Thanks again for all of your insightful comments! We've addressed the scientific comments below and made all the changes suggested in the technical/stylistic comments except where noted.

[Substantive] scientific comments: I still think this sentence needs further discussion: "..indicating that this signature may have limited utility in differentiating between tornadic supercells and nontornadic storms which still produce strong low-level rotation." The results quoted here are from a much smaller sample size than the Loeffler et al. (2020) data and the Homeyer et al. (2020) data, as well as your current results. Further, there is an implication that the previously studied non-tornadic supercells didn't have strong low-level rotation. That's clearly not explicitly stated in either if you read the case selection criteria for those two studies.

Those are fair points—we've reworked this section to note that comparisons between these studies are difficult due to their differing sample sizes and selection criteria for nontornadic supercells. We've also added a mention of new results from Healey and Van Den Broeke (2022)'s AMS presentation showing little difference between the separation angles of tornadic and nontornadic supercells in close proximity (albeit with a similarly small sample size to Van Den Broeke (2021)).

[Minor comments omitted...]