Evaluating the Ability of Remote Sensing Observations to Identify Significantly Severe and Potentially Tornadic Storms

Thea N. Sandmæl* and Cameron R. Homeyer

School of Meteorology, University of Oklahoma, Norman, Oklahoma.

Kristopher M. Bedka

NASA Langley Research Center, Hampton, VA 23666.

Jason M. Apke

Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO 80521.

John R. Mecikalski

Department of Atmospheric Sciences, University of Alabama in Huntsville, Huntsville, AL 35805.

Konstantin Khlopenkov

Science Systems and Applications, Inc., Hampton, VA 23666.

**Corresponding author address:* Thea N. Sandmæl, 120 David L Boren Blvd #2400, Norman, OK 73072.

E-mail: thea@ou.edu

Early Online Release: This preliminary version has been accepted for publication in *Journal of Applied Meteorology and Climatology*, may be fully cited, and has been assigned DOI 10.1175/JAMC-D-18-0241.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT

Remote sensing observations, especially those from ground-based radars, have been used extensively to discriminate between severe and non-severe storms. Recent upgrades to operational remote sensing networks in the United States have provided unprecedented spatial and temporal sampling to study such storms. These networks help forecasters subjectively identify storms capable of producing severe weather at the ground; however, uncertainties remain in how to objectively identify severe thunderstorms using the same data. Here, three large-area datasets (geostationary satellite, ground-based radar, and ground-based lightning detection) are used over 28 recent events in an attempt to objectively discriminate between severe and non-severe storms, with an additional focus on severe storms that produce tornadoes. Among these datasets, radar observations, specifically those at middle and upper levels (altitudes at and above 4 km), are shown to provide the greatest objective discrimination. Physical and kinematic storm characteristics from all analyzed datasets imply that significantly severe (≥2-in. hail and/or ≥65-kt straight-line winds) and tornadic storms have stronger upward motion and rotation than non-severe and less severe storms. In addition, these metrics are greatest in tornadic storms during the time in which tornadoes occur.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

1. Introduction

Severe and tornadic storms have been extensively studied using ground-based weather radar and satellite observations during the past four decades. A common goal of past research efforts has been enabling improvements in tornado prediction, which can save lives. Substantial efforts are almost always underway to improve tornado warnings, including ongoing projects like Warn-on-Forecast and PROBSEVERE (Stensrud et al. 2009; Cintineo et al. 2018). Despite previous efforts, the time from a warning being issued to a tornado occurring, commonly known as the warning lead time, has stayed the same from 1986-2011, averaging 18.5 min (Stensrud et al. 2013; Brooks and Correia Jr. 2018).

To distinguish tornadic storms from non-tornadic storms, forecasters and researchers have commonly utilized unique radar signatures at low levels (within a few kilometers of the Earth's surface) that often precede tornadogenesis, such as hook echoes, weak echo regions, inflow notches, bowing line segments, and rotation visible through radial velocity couplets, which were key to early improvements in tornado warnings (Fujita 1958; Browning and Donaldson 1963; Lemon and Doswell III 1979; Przybylinski 1995). More recently, tornado warning decision making has increasingly leveraged the development and strength of low-level rotation, visual reports from human spotters, and the presence of unique signatures in dual-polarization radar, such as the tornadic debris signature (Ryzhkov et al. 2005). For broader discrimination between severe and non-severe storms using radar observations, weak echo regions, mesocyclones, vertically integrated parameters based on radar reflectivity, and dual-polarization signatures have been used (Greene and Clark 1972; Lemon et al. 1977; Amburn and Wolf 1997; Kumjian and Ryzhkov 2008). In comparison, remote sensing observations of the upper levels of storms (especially those from satellite) have been increasingly used for severe storm detection due to recent improvements in spatiotempo-

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

ral sampling (e.g., Bedka et al. 2015; Gravelle et al. 2016). Satellite-observed cloud-top features associated with severe storms include rapid cloud-top cooling, anomalous cloud-top flow characteristics (strong divergence and couplets of high positive and negative vorticity), overshooting storm tops (OTs), and the "Enhanced-V" signature and other signatures related to above-anvil cirrus plumes (McCann 1983; Mecikalski and Bedka 2006; Cintineo et al. 2013; Bedka et al. 2015; Apke et al. 2016; Line et al. 2016; Homeyer et al. 2017). All of these features are hypothesized to be associated with strong upward motion within severe storms.

Model forecasts and simulations have played a large role in understanding the processes and environments that lead to severe and tornadic storms (e.g., Thompson et al. 2003; Cintineo et al. 2014; Coffer et al. 2017). The probability of all severe weather (tornadoes, hail, and straight-line winds) is known to increase with increasing values of convective available potential energy (CAPE) and vertical wind shear (typically in a layer 0-6 km AGL). For tornadic storms, additional environmental variables such as the significant tornado parameter, helicity, or the supercell composite parameter, have shown skill in distinguishing regions with favorable conditions for tornadic storm formation and where the most intense tornadic storms are likely to form (e.g., Stensrud et al. 1997; Rasmussen and Blanchard; Thompson et al. 2003, 2012). High-resolution modeling studies demonstrate that low-level streamwise horizontal vorticity is a key ingredient in environments favorable for tornadogenesis, as tilting of this vorticity into the vertical dimension helps maintain a strong, steady, low-level mesocyclone (e.g., Coffer et al. 2017; Orf et al. 2017). In addition to the tornadogenesis process, simulations of tornadic supercells have further indicated that wider updrafts can lead to more intense tornadoes if it is assumed that the scale and intensity of the tornadic circulation is associated with the scale and intensity of the rotating updraft at higher altitudes (Trapp et al. 2017), but these model results have been demonstrated to be sensitive to the design of the model simulations (Coffer and Markowski 2018).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Forecasting the potential for severe and tornadic storms hours to days in advance has largely been accomplished using predicted or measured properties of the near-storm environment (e.g., Cintineo et al. 2013; Parker 2014). These include winds, temperature, moisture, and related variables such as CAPE and vertical wind shear. While both individual environmental variables and unique combinations of different variables have proven to be useful predictors of severe storms and tornadoes, their utility in the warning process is limited in part by the lack of observations available at scales necessary to resolve the near-storm variability in real time (Thompson et al. 2003; Parker 2014). In addition, the stochastic nature of internal storm dynamics results in considerable overlap in the parameter spaces occupied by tornadic and non-tornadic storms, particularly in the case of weak tornadoes. This overlap makes it challenging for a forecaster to determine which storms will and will not be tornadic within a given environment (Anderson-Frey et al. 2016).

Operational observing systems in the United States provide measurements of storms at high spatial and temporal resolution and for many years. The Next-Generation Weather Radar (NEXRAD) network provides three-dimensional observations of storms at approximately 5-min increments (Crum and Alberty 1993). Satellite imagery from the Geostationary Operational Environmental Satellite (GOES) constellation provides cloud-top visible and infrared (IR) wavelength measurements of storms at intervals of 15 min or less (Menzel and Purdom 1994). The GOES-16 Advanced Baseline Imager provides imagery with temporal resolutions of 30 seconds to 1 min over 1000 km $\times$1000 km regional domains, and every 5 min over much of North America (Schmit et al. 2017). Prior to GOES-16, GOES-14 was used in experimental mode to acquire 1-min resolution data, with a focus on severe-storm and high-impact weather analyses (Schmit et al. 2013).

This study seeks to evaluate the utility and limitations of remote sensing observations to objectively discriminate between severe and non-severe storms using a fusion of recent high-resolution radar, satellite, and lightning datasets. In addition, tornadic storms are evaluated separately from

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

the remaining population of severe storms (those producing severe hail and straight-line winds) given their unique impacts and societal relevance. The goal of this work is to determine the value of modern remote sensing observations for early objective discrimination between tornadic, severe and non-severe storms. Below, novel results are presented that reveal significant differences in inferred upward motion and rotation between a large sample of severe and non-severe storms. These metrics reach a maximum in tornadic storms during the time tornadoes occur. Based on these results, an objective data-based approach for tornadic storm identification and short-term prediction is developed for performance evaluation.

2. Data and Methods

a. Cases

This study examines 27 single-day severe weather events in the United States that occurred during 2011-2016. These cases comprise more than 7000 storms defined using NEXRAD data, 273 of which produced tornadoes (Table 1). Severe weather days were chosen to capture a wide range of environmental conditions, severe weather frequencies, and tornado intensity. Nine of the 27 days were chosen due to the availability of GOES-14 super rapid scan data (1-min intervals), which is necessary to calculate satellite-based cloud-top divergence (Apke et al. 2016, 2018). The days when GOES-14 data were available are in bold in Table 1. Additional case studies were added to represent a variety of severe weather events from widespread tornado outbreaks in late spring to wintertime mesoscale convective systems. Radar-derived storm tracks (see Section 2f) from all 27 cases are shown in Fig. 1. Most storms analyzed in this study are clustered in the central U.S., but some events extend into the eastern U.S. and the Mississippi Valley.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

b. Radar Data

NEXRAD Level II data (i.e., volumes in range, azimuth and elevation relative to the location of a radar) were retrieved from the National Centers for Environmental Information (NCEI) (NOAA/NWS/ROC 1991). The NEXRAD network in the contiguous United States consists of more than 140 WSR-88D S-band (10-11 cm wavelength) radars that observe precipitation particles. All NEXRAD observations used in this study were obtained at a range resolution of 250 m, an azimuthal resolution of 0.5 degrees for the lowest 3-4 elevations and 1.0 degree otherwise, and typically at 14 elevations per volume. Each Level II volume includes (at a minimum) the radar reflectivity at horizontal polarization Z_H that is related to the size and/or density of cloud and precipitation particles in a radar volume and is in units of dBZ, and the radial velocity V_R, a measure of the motion of cloud and precipitation particles toward and away from the radar location, in units of m s^{-1}. Depending on the characteristics of the operational scanning strategy, the expected uncertainty in NEXRAD observations is up to 1 dB for Z_H and up to 1 m s^{-1} for V_R. These uncertainties can lead to even greater uncertainties in many of the derived variables outlined below, but such errors are typically smaller than observed differences between storm types (e.g., see documented errors in observables and derived variables in OFCM 2005, 2006).

The radar data are processed using the four-dimensional space-time merging methods described in Homeyer et al. (2017) and references therein, which resulted in volumes of the radar variables at 2-km horizontal resolution, 1-km vertical resolution, and 5-min temporal resolution over the entire extent of each analysis domain (see also information available at `http://gridrad.org`). Merging of V_R from multiple radar volumes onto a common grid is challenging, largely due to the fact that V_R is a measure of the motion of scatterers toward and away from the radar, such that any given measurement has a unique geometry and thus can vary significantly in magnitude

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

and sign compared to a measurement made at the same location from a different radar. In order to overcome this challenge, derivatives of V_R must be merged instead. For this study, the radial derivative of V_R (radial divergence) and the azimuthal derivative (azimuthal shear) are merged into multi-radar volumes, both of which are computed using centered differencing. These yield the approximate half-components of the divergence and rotation, which will be referred to as simply divergence and rotation in the remainder of the paper. Given the expected uncertainties in V_R, the resulting uncertainties in divergence and rotation estimates should be less than 0.004 s^{-1}, with uncertainties in derived rotation decreasing by more than an order of magnitude out to the farthest ranges observed by a radar (due to increasing azimuthal length scales; see also the discussion at the end of Section 3). This estimate is based on calculations using fixed range resolution, varying azimuthal resolution and assuming maximum error in winds: ± 1 m s^{-1} at each bound of the derivative, such that the maximum ΔV_R error expected is 2 m s^{-1}. For the azimuthal derivative, the distance is $2\Delta\theta$ for the derivative. For 0.5° azimuthal sampling, $\Delta\theta$ increases $\sim$875 m per 100 km range. For 1° azimuthal sampling (most elevations), $\Delta\theta$ increases $\sim$1750 m per 100 km (i.e., twice that of 0.5° resolution). To estimate the expected uncertainty in the azimuthal derivative, it is simply (2 m s^{-1})/($2\Delta\theta$). For ranges beyond 30 km, the uncertainty for the azimuthal derivative is much less than 0.004 s^{-1} in all cases. For the radial (i.e., range) derivative, the uncertainty is (2 m s^{-1})/(500 m) everywhere (i.e., 0.004 s^{-1}). While it is not possible to evaluate the uncertainties of these and other derived variables in greater detail due to a lack of finer-resolution auxiliary data sets, we expect the errors in rotation and divergence in our multi-radar merged data to be reduced further by following several quality-control steps outlined below.

First, since V_R is prone to large errors in magnitude and sign due to aliasing (i.e., winds that exceed the maximum detectable V_R at a given operating frequency – the Nyquist velocity – and become “folded”), the winds must be de-aliased prior to computing the derivatives (Doviak and

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Zrnić 1993). De-aliasing is performed using the Python ARM Radar Toolkit (Py-ART; Helmus and Collis 2016). For use in this merging procedure, a Py-ART routine is invoked that does not require a reference atmospheric wind profile and is more computationally efficient than alternative approaches – dealias_region_based, which accomplishes de-aliasing by modeling the problem as a dynamic network reduction.

Following de-aliasing, random fluctuations of V_R in each azimuthal sweep (a 360-degree scan made at a single elevation) are further suppressed by applying a 3×3 median filter and by using a 5-gate running-mean range filter prior to computing the radial and azimuthal derivatives (in that order). The derivatives (divergence and rotation) are then calculated using the quality controlled V_R and merged into the large-area, multi-radar dataset following the procedure in Homeyer et al. (2017). In order to avoid potential artifacts within weak or non-meteorological radar echo, V_R derivatives are only analyzed within $Z_H \geq 30$ dBZ in this study. Similar techniques describe known uncertainties that occur with V_R derivatives in range and azimuthal distance (Smith and Elmore 2004), which can be as large as ±20% relative to a known (or prescribed) value. The divergence maximum above an altitude of 8 km (upper-level; example in Fig. 2A) and the convergence maximum – or divergence minimum – below 3 km (lower-level; Fig. 2B), as well as their column maximum values, are calculated for each storm at each time step. Maximum cyclonic rotation is also calculated for the lower- and upper-level altitudes (Figs. 2D & E), as well as for the mid-levels (4-7 km). Due to the nature of radar sampling, the low-level variables will be limited by the distance to the radar, and thus will have much fewer data points than the mid- and upper-level variables.

Echo-top altitudes are computed for this study using multiple Z_H thresholds, with the majority of analysis conducted using 40-dBZ echo-top altitudes (Fig. 2G). The echo-top altitudes are com-

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

puted at every horizontal grid point by finding the highest altitude where Z_H exceeds the specified threshold, provided that Z_H is also greater than the threshold in the next two lowest altitude layers.

Velocity spectrum width, or the standard deviation of V_R estimates within a radar volume, is also extracted from the radar data where $Z_H \geq 30$ dBZ (Fig. 2H). Spectrum width is influenced by several factors, including substantial contributions from horizontal shear in V_R at low-levels and turbulence at any level (Doviak and Zrnić 1993). The turbulence component has been linked to updraft strength within convection and is often a major contributor to spectrum width observations at altitudes in the middle and upper troposphere (Feist et al. 2019). The column maximum spectrum width at each time step of each storm is calculated for analysis in this study.

c. Satellite Data

GOES imagery was retrieved from University of Wisconsin-Madison Space Science and Engineering Center (`http://www.ssec.wisc.edu/`) and NOAA (1994). GOES is primarily a constellation of two operational satellites that continuously monitor the weather over the United States: GOES-West stationed at 135°W and GOES-East at 75°W nadir longitudes. For the time period analyzed in this study, GOES-15 was operational in the West position and GOES-13 was operational in the East position. GOES-13 and -15 provide visible and IR imagery at 5- to 15-min intervals. A spare GOES satellite (GOES-14), positioned at 105°W, has been used for experimental super rapid scan observations in preparation for GOES-R (1-min frequency; SRSOR) during various periods since late summer 2012 (Schmit et al. 2013). For nine severe weather days (bolded in Table 1), 1-min imagery from GOES-14 is used for analysis. For the remaining severe weather days, imagery from GOES-13 is used. The GOES-13 and -14 Imager 0.65 μm visible wavelength channel has a horizontal resolution of $\sim$1 km at nadir, while the 10.7 μm IR channel has a

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

horizontal resolution of $\sim$4 km at nadir and an absolute accuracy of $\leq$1 K (Menzel and Purdom 1994).

Convective updrafts that penetrate through a thunderstorm anvil, known as overshooting tops or OTs, produce texture in GOES visible-channel imagery due to turbulent flow and shadowing induced by the updraft penetration. An algorithm to detect and quantify this texture has recently been developed that produces a "visible texture rating" product (Bedka and Khlopenkov 2016). Anvil clouds are identified using a two-step process and then a search is performed within the anvils to identify texture associated with penetrative updrafts. The first step in anvil detection is based on thresholding of GOES visible reflectance based upon an empirical model used to define how bright an anvil should be at a given time of day and day of year. Spatial and statistical analysis of the pixels that meet the day/time-dependent threshold is performed to eliminate singular pixels and preserve those within a broad area (greater than or equal to approximately 10 km^2) of near-uniform reflectance characteristic of anvil clouds. Fourier-transform analysis of visible reflectance within small (32 pixel) windows is then performed, yielding a power spectrum for varying wavelengths in a 32$\times$32-pixel domain. Typical OT signatures and concentric gravity waves that often surround OTs produce the strongest signal in a ring-like pattern with a wavelength of $\sim$4-8 km. Pattern recognition is applied to the power spectrum to identify ring patterns within this wavelength range. The results of the pattern recognition analysis define the unitless visible texture rating (Fig. 2C); the most coherent ring patterns are assigned a high rating.

Another method for convective updraft identification by GOES satellite involves objective identification of vigorous anvil outflow in $\leq$1-min scanning rate information. This is achieved here using the Super Rapid Scan Anvil Level flow system (SRSAL; Apke et al. 2016, 2018, and references therein). SRSAL objectively identifies deep convection cloud-top flows with mesoscale

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

atmospheric motion vectors (mAMVs; Bedka and Mecikalski 2005), which are point-source wind estimates based on pattern recognition in a sequence of GOES visible images.

SRSAL contains a cloud-top horizontal divergence (CTD; Fig. 2F) product output to a $0.02^{\circ}\times 0.02^{\circ}$ longitude-latitude grid. When associating SRSAL CTD with individual storms, only data points with final smoothing parameter (α from Hayden and Purser 1995, and Apke et al. 2018) values less than 0.5 are considered for analysis, as points with higher values are not densely sampled by mAMVs. In order to mitigate sampling errors in storms obscured by cirrus at higher altitudes, the data points for CTD, as well as visible texture rating, are also filtered by using only those points with a maximum visible texture rating greater than 7, which is indicative of a convective OT and gravity waves generated by the OT (Bedka and Khlopenkov 2016). Note that SRSAL, like visible texture rating, is a visible-only product as it requires the Visible channel to operate. The maximum CTD is calculated at each time step for each storm.

In order to extract satellite data along the path of the radar-based storm tracks, corrections for parallax error (owing to the viewing geometry of the satellite) are required. Parallax error increases as the cloud-top altitude and distance from satellite nadir increases (Vicente et al. 2002). Methods typically used to correct for parallax involve converting IR cloud-top temperature to cloud-top altitude using a reference tropospheric temperature profile. However, these methods are prone to large errors for deep convective anvils because high-altitude clouds may either be: i) thermally adjusted to stratospheric temperatures that are warmer than the upper troposphere, or ii) be optically thin and thus mostly transparent in IR. In this study, the merged radar observations are used to correct for parallax error. In particular, the $Z_H = 5$ dBZ echo-top altitude is used as a proxy for cloud-top height to estimate parallax. These estimates are used to correct the coordinates of the satellite imagery in order to extract values coincident with the storm tracks.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

d. Lightning Data

The Earth Networks Total Lightning Network (ENTLN) detects lightning using pulses in vertical electric field measurements from parts of the 1 Hz to 12 MHz frequency range from over 700 sites across the contiguous United States (Liu and Heckman 2011). Individual pulses are located in space and time by statistically solving over-determined electrical signal time-of-arrival equations using measurements from at least 5 stations. Sources close together in space and time are grouped into flashes, which are binned into $0.08^{\circ}\times0.08^{\circ}$ longitude-latitude ($\sim$64 km^2) flash density grids for analysis, designed to emulate the spatial resolution of data to be provided by the Geostationary Lightning Mapper instrument (Goodman et al. 2013). Lightning activity is correlated with intensification of updrafts (Schultz et al. 2017). When upward motion in the mixed-phase (liquid and ice) region of a cloud increases, hydrometeor collision charging mechanisms typically become more efficient and thus, lightning flashes become more frequent (Deierling and Petersen 2008). ENTLN data were available for eight of the nine GOES-14 severe weather days. The maximum of the total lightning flash density is extracted along each storm track for analysis in this study, which consists of both cloud-to-ground and intracloud flash density (Fig. 2I).

e. Tornado Warnings

Tornado warnings from the National Weather Service are used here to provide context on which storms produced physical indications of possible tornadogenesis and were publicly recognized by warning meteorologists. NWS warnings were obtained from the online archive maintained by Iowa State University (Iowa Environmental Mesonet 2017). The warnings are provided as shapefiles, with each warning consisting of a start (issuance) and end (expiration) time and coordinates of a polygon outlining the warned area. A warning was linked to all storm tracks that passed through the warning polygon during the time the warning was valid.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

f. Storm Tracking

Analysis of all datasets on an individual storm basis in this study is facilitated through objective radar-based storm tracking. In particular, individual storm tracks are computed for each severe weather day using an echo-top algorithm described in Homeyer et al. (2017). Local maxima in maps of Gaussian-smoothed echo-top altitude are identified in each 5-min radar observation and linked together in time if they lie within close proximity to each other ($\leq$12.5 km). For this study, tracking is accomplished through time linking of Z_H = 40 dBZ echo-top maxima, filtered by the convective echo classification output by the Storm Labeling in 3 Dimensions (SL3D) algorithm (Starzec et al. 2017). Tracked echo-top maxima are required to exceed an altitude of 4 km and be linked across 3 or more 5-min radar analyses. Radar reflectivity images of the objectively tracked storms were reviewed to manually identify and merge discontinuous tracks that correspond to the same storm. The quality-controlled storm tracks are then used to extract maximum or minimum (in the case of convergence and GOES IR brightness temperature) values from each dataset within a 10-km radius of the storm location at 1-min intervals, with observations made at coarser resolution than 1-min interpolated linearly in space and time to the storm track location. Such interpolation is only performed for data with time coverage gaps less than or equal to 5 min. Severe Weather Data Inventory (SWDI) tornado reports from NCEI are also added to the dataset and linked to the nearest storm within 3 km of the tornado path (National Centers for Environmental Information 2017).

g. Data Analysis

The tornadic storms are analyzed by extracting 1-min data points within a 5-min window centered on 30 and 15 min before the first tornado, 15 and 30 min after the last tornado, and during the entire time period of any tornado. This allows assessment of the potential for discrimination be-

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

tween tornadic and non-tornadic storms from each variable and for providing positive lead times. Time periods prior to only the first tornado in each storm are evaluated (rather than those prior to all individual tornadoes) to best isolate unique evolutionary characteristics of tornadic storms before they produce a tornado. Otherwise, time periods between successive tornadoes within a single storm may bias the perceived evolution in storm-based analyses and corresponding observational indicators of tornado potential. Similarly, time periods following the last tornado are analyzed to reveal the capacity for each variable to capture a decreasing tornado threat. The tornadic storms are compared to the most intense 30-min period of all tracked non-tornadic storms (i.e., any storm with a persistent 40-dBZ echo top exceeding 4 km) and of non-tornadic storms linked with severe hail or wind reports. The most intense 30-min period is defined as the $\pm$15-min window centered on the storm-maximum (or minimum) value observed for each separate variable. Therefore, the time periods considered to be the most intense for the non-tornadic storms could differ between variables. The non-tornadic storms are separated into categories containing non-severe, severe [those containing $\geq$1 in. (2.54 cm) diameter hail and $\geq$50 kt (25.7 m s^{-1}) wind speeds at ground level], and significant severe storms [those containing $\geq$2 in. (5.08 cm) diameter hail and $\geq$65 kt (33.4 m s^{-1}) wind speeds at ground level]. Significant severe non-tornadic storms were not included in the severe non-tornadic storm category and neither severe storm category was included in the non-severe non-tornadic category. While many variables were analyzed during the course of this study, the analysis presented here focuses on variables that provided the greatest discriminatory ability from each data source. Table 2 provides a concise list of all variables analyzed and included in the remainder of the paper.

The updraft strength within storms is inferred here using a kinematic approach based on divergence observations. Kinematic approaches for inferring upward motion involve vertical integration of the horizontal wind divergence through a column with the assumptions of an incompressible

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

or anelastic atmosphere (e.g., O'Brien 1970). Strong upper-level divergence located at altitudes above low-level convergence within convection (i.e., a two-layer divergence profile) implies strong upward motion due to the conservation of mass in the atmosphere. While the radar and satellite observations can only measure winds within and atop storms, respectively, the upper-level divergence alone can (with assumptions) serve as a proxy for updraft strength in deep convection.

The utility of upper-level divergence as a proxy for updraft strength is primarily limited by variations in the depth of analyzed storms and coarse vertical sampling. If all storms spanned the same depth in the atmosphere and had equivalent divergence profile shapes, differences in the upper-level divergence (or low-level convergence) maxima would be proportional to differences in vertical velocity. Since the vast majority of storms analyzed in this study reach the tropopause and the tropopause altitude varies by <3 km across the 27 cases analyzed, it is assumed that the differences in storm depth have a minor impact on the use of upper-level divergence as a proxy for updraft strength. In a scenario where two storms had equivalent maxima in upper-level divergence but differed by 3 km in depth, the inferred updraft speed for the deeper storm would be 25% larger than that of the shallower storm. Errors could be larger if the divergence profile shapes differed considerably between storms, which is not possible to adequately assess with the data used in this study. Single-radar estimates of divergence at high elevation angles (i.e., those obtaining measurements in the upper troposphere) contain additional error due to contributions from the vertical component of the wind and hydrometeor fall speeds to the measured V_R, but these errors are expected to be relatively small (or potentially helpful for diagnosing relative differences in updraft strength given the relationship between vertical velocity and the horizontal divergence). Others have had success assuming upper-level divergence is related to updraft strength, for example, in hail size nowcasting (e.g., Witt and Nelson 1991; Boustead 2008; Blair et al. 2011).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

h. Performance Evaluation

As outlined in Section 3a, an evaluation of the ability of a simple objective technique to identify storms capable of producing tornadoes before they occur was performed using the product of two radar-derived kinematic fields: divergence and rotation. To avoid being overly restrictive with an arbitrary altitude threshold, the column-maximum divergence is used in the product calculation. The rotation in the divergence-rotation product (maximum divergence multiplied with maximum rotation) is the maximum at upper- and mid-levels (i.e., the largest value found anywhere at and above 4 km). Storms that exceed a single threshold value of this product (i.e., [divergence $\times$ rotation] $\geq$ threshold) for a specified time period are flagged as potentially tornadic and the time at which the condition is met is recorded. For a predictive model, the resulting probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and bias forecast skill metrics for the storm population are computed using Equations 1 through 4.

$$\mathrm{POD} = \frac{\text{No. correctly flagged storms}}{\text{No. tornadic storms}} \tag{1}$$

$$\mathrm{FAR} = \frac{\text{No. incorrectly flagged storms}}{\text{No. storms flagged}} \tag{2}$$

$$\mathrm{CSI} = \left(\frac{1}{1-\mathrm{FAR}} + \frac{1}{\mathrm{POD}} - 1\right)^{-1} \tag{3}$$

$$\mathrm{Bias} = \frac{\mathrm{POD}}{\text{1 - FAR}} \tag{4}$$

A perfect forecast has a 100% POD, 0% FAR, and a CSI and bias of 1 (e.g., Roebber 2009). Correctly flagged storms are tornadic storms identified prior to the occurrence of the first tornado and incorrectly flagged storms are those flagged that never produce a tornado. Mean and median

Accepted for publication in *Journal of Applied Meteorology and Climatology.* DOI 10.1175/JAMC-D-18-0241.1.

lead times of the potentially-tornadic identification relative to the first occurrence of a tornado (hereafter the flag lead time) within each storm are also computed. Flag lead times reported in this study are computed only for correctly flagged storms (i.e., missed tornadic storms are not included in lead time calculations as having lead times of 0). Tornadic storms with 0 or negative lead times are considered to be missed storms, which is accounted for in the POD. For evaluation purposes, the first instance of a tornado warning for a storm from the NWS served as a baseline potentially-tornadic identification for comparison with the objective threshold exceedance method. Apart from the difference in storm identification method, the performance of the objective method and NWS tornado warnings is evaluated in the same way. Thus, calculation of lead times for these metrics may favor the objective approach given the fact that NWS warnings are commonly issued for a finite duration of 30 or 45 minutes, but the corresponding POD, FAR, and CSI calculations do not favor either method.

Performance evaluations can also be made for varying storm environments, which is done here using the number of tornadic storms for a given day when the primary storm mode was discrete convection (i.e., supercells and ordinary cells). All cases for which the primary mode was multicellular convection (typically mesoscale convective systems or MCSs) are analyzed separately because the environments in which they occur often differ considerably from supercells (e.g., see Flournoy and Coniglio 2019, and references therein). The primary modes were subjectively evaluated, where the mode that is dominant during the actively tornadic period was chosen. MCSs are the primary storm mode for five of the 27 cases (Table 1). Events for which the dominant storm mode was discrete convection are grouped into those having 1-5, 6-15, or 16+ tornadic storms.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

3. Results

The analysis of 27 severe weather day cases, based on both kinematic and physical metrics, shows that significant severe and tornadic storms generally have greater inferred upward motion and rotation than severe and non-severe non-tornadic storms (Figs. 3 & 4). The maximum divergence estimated from both radar and satellite is substantially stronger for significant severe non-tornadic and tornadic storms compared to that found in non-tornadic non-severe storms, especially when there is a tornado on the ground (Figs. 3A-3B). Severe non-tornadic storms show intermediate divergence magnitudes relative to the significant severe and non-severe storm populations. Divergence for significant severe non-tornadic storms is similar to that observed in tornadic storms prior to tornadogenesis, suggesting little to no ability to distinguish between the two storm types before a tornado has occurred. The difference in median values between the significant severe or tornadic storms (especially leading up to the first tornado) and the non-tornadic storms is greater for the radar-estimated divergence than the satellite divergence, with clear and consistent differences prior to first tornado occurrence. Divergence estimates from the radar and satellite sources here do not account for density changes in the atmosphere with height (i.e., differences in storm depth); thus, inferring a stronger updraft within storms containing larger divergence involves an incompressible atmosphere assumption. Though not shown, using an anelastic assumption (where base state density varies with height) and deriving mass-flux divergence instead provides consistent results with those shown here.

Differences between the divergence estimated from ground-based radar and satellite imagery are likely due to both the limited information detected by satellite (i.e., at cloud top only) and the differences in the spatial resolution of the two datasets. It is also possible that some of the difference can be due to the limitations of the radar-derived divergence due to the previously discussed

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

issue with the radar beam inclination. Although the number of cases differs from the satellite to the radar data, the cases where 1-min GOES-14 imagery was available were previously analyzed separately for the radar divergence with nearly identical results to the 27 cases in this study (not shown), indicating that the differences between radar and satellite divergence are not due to a sampling issue. Fig. 4A, which shows the maximum upper-level divergence, is nearly identical to the column-maximum divergence in Fig. 3A, implying that column-maximum divergence typically occurs at altitudes above 8 km.

Physical metrics of strong updrafts show behavior consistent with that observed from radar and satellite divergence. Specifically, radar-observed 40-dBZ echo-top altitudes (the maximum altitude reached by radar-indicated precipitation of considerable size – e.g., large rain drops or ice particles such as hail) imply that significant severe non-tornadic and tornadic storms have stronger updrafts than weaker severe and non-severe non-tornadic storms (Fig. 3C). This is not a surprising result and is due to the fact that larger precipitation particles have faster fall speeds, meaning stronger in-cloud vertical motion is required to loft them to higher altitudes. Identifying cloud-top altitudes from satellite is challenging when storms reach the tropopause (commonly the case for storms analyzed in this study) due to the dependence of the relationship between cloud top temperature and altitude in the stratosphere on both the resolution of the IR imager and the assumed environmental temperature profile, which can vary greatly in the extratropical lower stratosphere (e.g., Griffin et al. 2016). Alternatively, it is possible to measure the visible texture of the cloud top from satellite to indicate the tropopause-relative depth of OTs (Bedka and Khlopenkov 2016). A high visible texture rating implies a more complex texture, which is shown here to be correlated with stronger upward motion and higher tropopause-relative cloud tops (Fig. 5). Indeed, the visible texture rating is also highest in the tornadic storms examined here during tornadoes, providing further evidence of stronger upward motion than that in non-tornadic non-severe storms (Fig. 3D).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Tropopause-relative IR cloud-top temperatures show similar characteristics, but less contrast. Reduced contrast in IR is likely due to the 16 times poorer spatial resolution (compared to the visible) of the GOES imagery used in this study (Fig. A1B). As observed for divergence, the differences between physical characteristics of tornadic and non-tornadic storms are reduced when considering observations for the most intense periods in significant severe non-tornadic storms (expected to be the most extreme non-tornadic storms).

Three additional metrics that are related to upward motion in storms are shown to provide further evidence of a unique relationship between both significant severe non-tornadic and tornadic storms and updraft strength. First, column-maximum V_R spectrum width from radar is shown (Fig. 3E) due to its dependence on turbulence that increases as the updraft strength increases (Doviak and Zrnić 1993; Feist et al. 2019). Spectrum width shows similar contrast between large values in significant severe non-tornadic and tornadic storms and much lower values in non-tornadic non-severe storms to that observed for column-maximum divergence, further supporting the inference that significant severe non-tornadic and tornadic storms are characterized by stronger upward motion than weaker severe and non-severe non-tornadic storms.

Second, stronger upward motion has implications for lightning activity. Data from ENTLN show that flash density is greatest in significant severe non-tornadic storms and similarly high in tornadic storms during the time a tornado is occurring (Fig. 3F). This result is comparable to the so-called "lightning jump" feature discussed in previous studies and linked to severe weather (Williams et al. 1999; Schultz et al. 2009), although this study evaluates the absolute value of flash density rather than how rapid the lightning activity is increasing over time. Despite the large flash rates observed within tornadic storms, the lightning data also show considerable overlap between the severe non-tornadic and tornadic storm populations prior to the first tornado, which indicates

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

that this metric is better at discriminating between severe and non-severe rather than tornadic and non-tornadic storms.

Third, as an updraft intensifies within a rotating storm, stretching of vertical vortex tubes within provides increased vertical vorticity relative to storms with weaker updrafts (Markowski and Richardson 2009), which is demonstrated well in the radar observations of rotation at all altitudes (Figs. 4B, 4D, and 4F). Increased lightning activity and low-to-mid-altitude rotation are currently being used as variables of interest for probabilistic forecasts of tornadoes (Smith et al. 2016). Here, of the three altitude layers of rotation analyzed, mid-level rotation (Fig. 4D) shows the greatest potential for discriminating between significant severe and non-severe (and tornadic and weakly severe or non-severe non-tornadic) storms, with similar separation between categories to that found for radar-derived divergence. The lack of separation in low-level rotation between tornadic and non-tornadic storm categories deserves some explanation here. Considering the methods used to calculate rotation outlined in Section 2b (smoothing via 3×3 median filter and 5-gate running-mean and centered differencing), there are minimum scales of rotation that can be resolved and retained in the merged multi-radar volumes. In addition, because the native radar data have higher azimuthal sampling in the lowest elevation scans, the minimum scales of rotation that can be resolved are smaller at low levels and larger at mid and upper levels. In most cases, these minimum resolvable scales are 2-3 km at low levels and 3-6 km at mid and upper levels. Thus, since mesocyclone diameters are commonly between 1 and 10 km (Stumpf et al. 1998), the smallest mesocyclones will not be detected in these data. Low-level observations here have an advantage in the scales (and magnitudes) of rotation that can be retained due to the enhanced resolution there compared to higher altitudes, so a lack of mesocyclone detection doesn't explain the differences between low-level and mid-level rotation. The minimum threshold of $Z_H \geq 30$ dBZ applied to analyses of rotation could also be a source of reduced discrimination at low levels, since

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

strong rotation can often be found within weaker echoes at such altitudes. Thus, we did evaluate rotation using a weaker threshold of $Z_H \geq 10$ dBZ, which did show some increases in low-level rotation for tornadic storms overall, but also an increase in the spread of rotation values for all storm populations (not shown).

a. Evaluation of a Simple Objective Short-Term Tornadic Storm Forecast Product

While the statistical evaluations in Figures 3 and 4 show that radar-derived divergence and rotation provide the largest separation between tornadic and weakly severe or non-severe non-tornadic storms prior to tornadogenesis, they do not evaluate the potential usefulness of the variables for real-time discrimination. The figures also demonstrate that tornadic and significant severe non-tornadic storms show little separation, but both populations are small in number compared to the more prevalent weakly severe and non-severe storms. Given these results and the societal relevance of tornadoes, an evaluation of the ability of a simple objective technique based on the product of radar-derived rotation and divergence to identify storms capable of producing tornadoes before they occur is warranted. Although low-level rotation shows significant differences between the non-tornadic categories and the tornadic periods, the limited number of observations available compared to that for mid- to upper-level rotation (see Table 3) leads to the exclusion of low-level rotation in the product of rotation and divergence here. To provide context for this objective threshold method for storm discrimination, performance results (i.e., the ability to identify observed tornadoes) are compared with the first tornado warning given to each storm by the responsible National Oceanic and Atmospheric Administration (NOAA) NWS forecast office, which serves as a metric of the first public recognition that a storm was potentially tornadic by forecasters. Note that the first warning is used here as a short-term forecast of a storm's potential to become tornadic for context only, not to be confused with the evaluations conducted by the NWS of the performance of

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

all individual warnings, which aim to evaluate whether or not a warning encompassed the time of an observed tornado. The tornado warnings are linked to the storm tracks generated for this study, so the exact same storms are analyzed for both the radar-based and warning-based methods.

As outlined in Section 2h, the divergence-rotation product is based on column-maximum divergence and the maximum of rotation from mid- and upper-levels. It was found that a rotation-divergence product threshold of $42 \cdot 10^{-6}$ s^{-2} is comparable to the cumulative performance of the NWS warning-based potentially tornadic storm flag over all 27 severe weather days (see Figs. 4E and 6). This decision was arbitrarily made to facilitate direct comparison between the objective threshold method and the NWS warning-based method. From Fig. 6, a 5-min time period of the divergence-rotation product exceeding the threshold is deemed sufficient for the objective threshold technique, since the product did not appear to be greatly affected by random time variations (i.e., noise).

For objective divergence-rotation thresholds ranging from $5 \cdot 10^{-6}$ s^{-2} to $80 \cdot 10^{-6}$ s^{-2} applied to data from all 27 severe weather days, the CSI largely varies between 0.1 and 0.2 (Fig. 6). In comparison, the CSI of the NWS warning-based method is $\sim$0.13 (indicated by the black circle in Fig. 6). The objective threshold method achieved a comparable CSI to the NWS method at a POD of approximately 58.3% and an FAR of approximately 85.9%, while the POD and FAR based on the NWS method are approximately 51.7% and 84.9%, respectively. The mean flag lead time is 43 min using the objective threshold method, while the median flag lead time is 35 min. Similar performance (skill) with positive lead time by the objective method indicates that the divergence-rotation product provides a comparable ability to discriminate between tornadic and non-tornadic storms prior to tornadogenesis.

The single-value divergence-rotation threshold calculated from the cumulative performance of all 27 days is applied to groupings based on the number of tornadic storms for a given case (Table

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

4). The two lower-impact groupings (1-5 and 6-15 tornadic storms) showed both higher POD and FAR than the overall performance, with slightly lower skill. Though the POD decreases from ∼70% to ∼60% for the high-end days (those with 16+ tornadic storms), the FAR also decreases by a considerable amount, which in turn increases the skill of the objective method to 0.19. The performance decreases for the objective threshold method when the dominant storm mode is an MCS. Namely, the lowest POD and highest FAR values are found in these cases, with the objective threshold method showing the poorest performance. However, the median flag lead times from the objective threshold method are still the same as the overall median flag lead times from the 27 cases. These results reveal that the ability of the objective threshold method to discriminate between tornadic and non-tornadic storms is greatest in discrete cases (i.e., supercell storms) and the lead time of discrimination is relatively insensitive to the variation in event type.

In order to illustrate the spatial appearance of the objective threshold evaluation, maps of instantaneous fields at 20-min intervals from the 31 May 2013 event are shown in Fig. 7. Areas exceeding the single-value divergence-rotation threshold are shown in purple in each map. Storms 1 and 4 exceed the threshold for extended periods of time and are each responsible for producing several tornadoes (times indicated in each map), while storms 2 and 3 briefly exceeded the single-value threshold and never or only once produced a tornado, respectively. All four storms were tornado warned by the NWS for some time during their life cycles. The southern storm (labeled 1) produced an EF3 tornado near El Reno, OK at 23:03 UTC, as well as an EF0 tornado shortly prior to the EF3 tornado. The first exceedance of the divergence-rotation product for storm 1 was observed at 21:50 UTC and the divergence-rotation product exceeded the threshold value over a larger area for storm 1 than the remaining storms, both prior to and especially during the EF3 tornado.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

4. Summary and Discussion

This study employed radar, satellite, and lightning observations from a large dataset of more than 7000 storms to examine the ability of modern, high-resolution remote sensing data to objectively discriminate between severe and non-severe storms, with an additional focus on severe storms that produce tornadoes. It was found that radar-observed/derived physical and kinematic characteristics routinely enable discrimination between significant severe or tornadic and non-severe non-tornadic storms, with indications from all datasets that inferred upward motion is strongest and rotation is fastest in tornadic storms during the occurrence of a tornado (see Figs. 3 & 4). Significant severe non-tornadic storms were found to broadly overlap with tornadic storms in most observations, but the size of the significant severe non-tornadic population is relatively small. While the tornadic and non-tornadic discrimination results are broadly consistent in both radar and satellite-derived flow observations, larger differences were seen between the storm categories in the radar observations. The separation between the tornadic and non-tornadic storm characteristics was found to be large enough such that a simple objective threshold method based on the product of radar-derived storm divergence and rotation was able to provide early indication of potentially tornadic storms with comparable performance to indications based on NWS tornado warnings (see Fig. 6).

Previous studies have shown somewhat similar separation between storm categories using environmental measurements from numerical model analyses and forecasts, such as the significant tornado parameter (e.g., Thompson et al. 2003). These studies typically isolate environments based on the most intense storm within close proximity to the model grid point. However, as outlined in Section 1, it is common to find both tornadic and non-tornadic (or severe and non-severe) storms within very similar environments, which makes it challenging to use these metrics for objective storm discrimination. Analysis of such environmental variables was conducted during the

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

course of this research, but greater overlap, and thus weaker discrimination, between storm categories was found compared to that provided by the radar-observed/derived physical and kinematic characteristics (not shown).

With respect to tornadic versus non-tornadic storms, the results of this study agree with the current understanding of the three-step process for tornadogenesis within supercells (Markowski and Richardson 2009; Davies-Jones 2015). Namely, the first step in a storm's evolution to become tornadic is the development of a strong mid-level circulation, which is found routinely in the radar observations at long lead times to tornadogenesis (see Fig. 4D). The second step for a tornadic storm is the development of a strong near-surface circulation as a result of processes occurring as air descends through the low-level outflow. The third and final step to becoming a tornadic storm is having this near-surface rotation come into alignment with in-storm perturbation pressure gradients associated with rotation aloft, that lift the air and contract it to tornado strength (Markowski and Richardson 2014). The maximum values observed in almost all physical and kinematic metrics evaluated here being associated with time periods during observed tornadoes is evidence of the extreme and deep rotating updrafts associated with tornadogenesis in the conceptual model.

Given the extensive knowledge base that exists for severe, non-severe, tornadic and non-tornadic storms and the discussion given in the previous paragraph, it is not surprising to find that, on average, significant severe non-tornadic and tornadic storms have stronger inferred updrafts and greater rotation than non-severe non-tornadic storms. These findings are in agreement with a similar argument for tornadic storms that has recently been made for an association between the strength of a storm's mesocyclone and the width of the updraft, which Trapp et al. (2017) tied to tornado strength based on numerical simulations of tornadic storms. As shown in the example maps of the divergence-rotation product (Fig. 7), the storm responsible for the 2013 El Reno, OK EF3-tornado was associated with a higher area of divergence-rotation threshold exceedance than

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

nearby storms with weaker tornadoes, which may be an indication of a broader updraft within the El Reno storm. Future studies should investigate the relationship between metrics of updraft width and tornado strength in observations.

One caveat of this study is that only 27 events from a period spanning 5 years were evaluated, with most events occurring during the April-June time period. Thus, to demonstrate that our methods for case selection were not inherently biased, an analysis based on 22 additional severe weather days that were randomly selected from a single year (2011) is included in the Appendix. The results from these cases are generally consistent with that presented above and further support the argument that our case selection for the events analyzed throughout the paper was not biased.

Another caveat of this analysis is the focus on single-polarization radar variables. While several dual-polarization variable extrema were investigated during the study summarized here, none of this analysis yielded statistically significant differences between severe categories and was therefore not reported. A lack of significant differences could be due to insufficient diagnosis of dual-polarization signatures associated with tornadoes using extrema alone. It is also possible these signatures are suffiecently small in scale such that they are smoothed out in the gridded radar dataset. Nevertheless, dual-polarization radar observations and their utility for severe and tornadic storm discrimination should be investigated further in future work.

In the tornado warning process, the NWS forecaster faces two primary challenges: timely identification of tornadic storms, and production of spatially concise warnings that sufficiently identify locations likely to be affected by a tornado. The objective system developed here can only help with the former, as the physical connection between strong divergence near the storm-top and the eventual development of rotation near the surface is unknown and is required to confirm the eventual presence and approximate location of a tornado. Furthermore, this study has shown that while weakly severe and non-severe non-tornadic storms are often considerably different than tor-

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

nadic storms in radar and satellite observations, significant severe non-tornadic storms (those most likely to be non-tornadic supercells) do not differ considerably from tornadic storms prior to tornadogenesis. Thus, additional work is required to evaluate the utility of the physical and kinematic radar observations (especially those at middle and upper levels) for the tornado warning decision making process. Furthermore, while the objective method of potentially tornadic storm detection using the divergence-rotation product was found to perform at a skill similar to NWS warnings, the true value of this metric for the warning decision making process should be evaluated in greater detail in future studies. Namely, this work would benefit from increasing the number of cases to reduce uncertainty and include greater representation of observed seasonality and convective mode. Observing system simulation experiments (commonly referred to as OSSEs), which have been used to estimate radar multi-Doppler wind retrieval uncertainties (e.g., Potvin et al. 2012), may also be helpful for improving understanding of the limitations of and uncertainties in the divergence-rotation product.

One potential barrier to implementing the divergence-rotation product evaluated here in an operational setting is the necessary step of dealiasing radial velocity fields, which is the most crucial and time consuming element of the process. However, dealiasing is commonly executed in real-time within the software used by forecasters. In addition, while the divergence-rotation product was calculated from multi-radar composites, it could easily be implemented using single-radar observations. If similar methods to this study are used for computing divergence and rotation, differences between the magnitude of the product in single-radar fields and the multi-radar composites are expected to be minimal, but it is required to evaluate the product from multiple neighboring radars to achieve similar vertical sampling. One aspect that was not investigated in this study is the development of a variable divergence-rotation threshold for the objective method to account for potentially relevant factors such as seasonality, location, or storm mode. It is likely that the

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

threshold found here is not "one size fits all", but will vary based on such factors as evident by the variation in performance between case types (Table 4). Again, further research is needed to examine the best way to utilize these results in an operational setting.

In conclusion, these findings provide an opportunity for improving the early recognition of significant severe and potentially tornadic storms from modern radar, satellite, and lightning observations. Increases in the spatial and temporal resolution of visible and IR satellite imagery now available following the transition of GOES-16 to operations in January 2018 will likely demonstrate improved capability to infer updraft intensity in the future. There are ongoing efforts to investigate these metrics further using machine learning techniques, which will likely yield a greater performance than the simple objective threshold technique introduced here.

Acknowledgments. This work was supported by the National Aeronautics and Space Administration (NASA) under Award NNX15AV81G. The authors would like to thank Harold Brooks (NOAA National Severe Storms Laboratory) for essential guidance on the performance analyses conducted for this manuscript, Elisa Murillo (University of Oklahoma - OU) for assisting in the data management process, Corey Potvin (OU-CIMMS) and Alan Shapiro (OU) for valuable feedback on this manuscript, Christopher Jewett for gridding the ENTLN flash detection data, and Benjamin Scarino (NASA Langley Research Center) for his insight and software development support. Radar, satellite, and environmental data used in this study can be obtained from NCEI (`https://www.ncdc.noaa.gov/data-access`). The authors would like to thank Chris Sloop and Stan Heckman (Earth Networks) for access to the ENTLN data. Multi-dataset storm track files produced during this effort are available from the authors upon request.

APPENDIX

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

a. Tropopause-Relative Infrared Brightness Temperature

Since IR brightness temperatures serve as a proxy for cloud top height and can help to indicate the depth of overshooting tops (and thus, the strength of upward motion within convection), we did evaluate the potential of brightness temperatures to discriminate between severe and non-severe (and tornadic and non-tornadic) storms. IR brightness temperature did not show an ability to discriminate between severe and non-severe and tornadic and non-tornadic storms. A storm with a cold cloud top does not indicate that the storm will necessarily be severe or tornadic. Fig. A1A shows storms from the same example as in Fig. 2. The northeastern-most storm was producing a tornado at the time that map was valid for, while the other storm with a cold cloud top (deep blue shading) on the KS-NE border only produced a few $\sim$1-inch hail reports and a 52-kt wind report. One of the southern storms also produced a tornado at a later time, but its minimum brightness temperature was only 3 K colder than at this time, which was considerably warmer than the KS-NE border storm that never produced a tornado. Note also that anvil regions well removed from precipitation echoes are comparably cold to the strong convective cores, which complicates the use of IR temperature thresholding for severe storm discrimination.

The minimum IR brightness temperature from GOES within a storm is calculated and compared with the temperature at the tropopause in order to investigate the minimum tropopause-relative temperature of the cloud tops. Generally, if the tropopause-relative cloud-top temperature is negative, the storm is penetrating into the stratosphere.

The tropopause temperature was extracted from Rapid Update Cycle (RUC) or Rapid Refresh (RAP) hourly output (Benjamin et al. 2004, 2016). The RUC/RAP models have a horizontal resolution of 13 km and 50-51 vertical levels, and were retrieved from the National Centers for Environmental Prediction (NCEP; National Oceanic and Atmospheric Administration, Earth Sys-

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

tem Research Laboratory 2012). The temperatures from RAP/RUC were linearly interpolated to the radar-based storm tracks in space and time for analysis.

Although there is considerable overlap between the categories for the minimum IR brightness temperature, it is worth noting that the 5th percentile of the tornadic category is much colder than the other categories (Fig. A1B). This implies that tornadoes seldom occur when cloud tops are warm relative to the tropopause, which differs from the severe and significant severe non-tornadic categories. While GOES-13/14 data cannot resolve the coldest temperatures due to its spatial resolution, the data indicate that cloud tops are relatively cold for tornadic storms on average. After updraft intensification while the storm is tornadic, there is a lot of cold outflow generated, resulting in a smaller range of values after the tornado dissipates.

b. Additional Cases

An analysis of 22 randomly selected severe weather cases from 2011 (Table A1) supports the general result from this study, though the upper-level variables show less separation between the most intense period of the non-tornadic significant severe storms and the tornadic storms (Fig. A2). These cases were tornado days that were randomly picked throughout the year. At least one case from each month of the year is included to account for potential seasonality in tornadic and non-tornadic storm characteristics. This analysis suggests that there is an important seasonality to the divergence-rotation threshold, as the separation between storm populations remains similar to the results presented in the main text, but all the boxes are shifted downward (i.e., magnitudes of each metric are smaller).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

References

Amburn, S. A., and P. L. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12**, 473–478.

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2016: Investigation of near-storm environments for tornado events and warnings. *Wea. Forecasting*, **31 (6)**, 1771–1790.

Apke, J. M., J. R. Mecikalski, K. Bedka, E. W. McCaul Jr, C. R. Homeyer, and C. P. Jewett, 2018: Relationships between deep convection updraft characteristics and satellite based super rapid scan mesoscale atmospheric motion vector derived flow. *Mon. Wea. Rev.*, **(2018)**.

Apke, J. M., J. R. Mecikalski, and C. P. Jewett, 2016: Analysis of mesoscale atmospheric flows above mature deep convection using super rapid scan geostationary satellite data. *J. Appl. Meteor. Climatol.*, **55 (9)**, 1859–1887, doi:10.1175/JAMC-D-15-0253.1.

Bedka, K. M., and K. Khlopenkov, 2016: A probabilistic multispectral pattern recognition method for detection of overshooting cloud tops using passive satellite imager observations. *J. Appl. Meteor. Climatol.*, **55 (9)**, 1983–2005.

Bedka, K. M., and J. R. Mecikalski, 2005: Application of satellite-derived atmospheric motion vectors for estimating mesoscale flows. *J. Appl. Meteor.*, **44 (11)**, 1761–1772.

Bedka, K. M., C. Wang, R. Rogers, L. D. Carey, W. Feltz, and J. Kanak, 2015: Examining deep convective cloud evolution using total lightning, WSR-88D, and GOES-14 super rapid scan datasets. *Wea. Forecasting*, **30**, 571–590, doi:10.1175/WAF-D-14-00062.1.

Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132 (2)**, 495–518.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144 (4)**, 1669–1694.

Blair, S. F., D. R. Deroche, J. M. Boustead, J. W. Leighton, B. L. Barjenbruch, and W. P. Gargan, 2011: A radar-based assessment of the detectability of giant hail. *E-Journal of Severe Storms Meteorology*, **6 (7)**.

Boustead, J. M., 2008: Using maximum storm-top divergence and the vertical freezing level to forecast hail size. Preprint. *Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., P6.6.

Brooks, H. E., and J. Correia Jr., 2018: Long-term performance metrics for national weather service tornado warnings. *Wea. Forecasting*, **33 (6)**, 1501–1511.

Browning, K. A., and R. J. Donaldson, 1963: Airflow and structure of a tornadic storm. *J. Atmos. Sci.*, **20 (6)**, 533–545.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and A. K. Heidinger, 2013: Evolution of severe and nonsevere convection inferred from GOES-derived cloud properties. *J. App. Met. Clim.*, **52**, 2009–2023, doi:10.1175/JAMC-D-12-0330.1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29 (3)**, 639–653, doi:10.1175/WAF-D-13-00113.1.

Cintineo, J. L., and Coauthors, 2018: The NOAA/CIMSS ProbSevere Model: Incorporation of total lightning and validation. *Wea. Forecasting*, **33 (1)**, 331–345.

Coffer, B. E., and P. M. Markowski, 2018: Comments on “the regulation of tornado intensity by updraft width”. *J. Atmos. Sci.*, **75 (11)**, 4049–4056, doi:10.1175/JAS-D-18-0170.1.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Coffer, B. E., M. D. Parker, J. M. L. Dahl, L. J. Wicker, and A. J. Clark, 2017: Volatility of tornadogenesis: An ensemble of simulated nontornadic and tornadic supercells in VORTEX2 environments. *Mon. Wea. Rev.*, **145 (11)**, 4605–4625, doi:10.1175/MWR-D-17-0152.1, URL https://doi.org/10.1175/MWR-D-17-0152.1, https://doi.org/10.1175/MWR-D-17-0152.1.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteorol. Soc.*, **74 (9)**, 1669–1687.

Davies-Jones, R., 2015: A review of supercell and tornado dynamics. *Atmos. Res.*, **158**, 274–291.

Deierling, W., and W. A. Petersen, 2008: Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res.*, **113**, D16 210, doi:10.1029/2007JD009598.

Doviak, R. J., and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. 2nd ed., ISBN 0-486-45060-0, Dover Publications, Inc., 31 East 2nd Street, Mineola, NY, 11501.

Feist, M. M., C. D. Westbrook, P. A. Clark, T. H. Stein, H. W. Lean, and A. J. Stirling, 2019: Statistics of convective cloud turbulence from a comprehensive turbulence retrieval method for radar observations. *Quart. J. Roy. Meteor. Soc.*, **in press**, doi:10.1002/qj.3462.

Flournoy, M. D., and M. C. Coniglio, 2019: Origins of vorticity in a simulated tornadic mesovortex observed during pecan on 6 july 2015. *Mon. Wea. Rev.*, **147 (1)**, 107–134, doi: 10.1175/MWR-D-18-0221.1.

Fujita, T., 1958: Mesoanalysis of the Illinois tornadoes of 9 April 1953. *J. Meteor.*, **15 (3)**, 288–296.

Goodman, S. J., and Coauthors, 2013: The GOES-R geostationary lightning mapper (GLM). *Atmospheric research*, **125**, 34–49.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Gravelle, C. M., J. R. Mecikalski, W. E. Line, K. M. Bedka, R. A. Petersen, J. M. Sieglaff, G. T. Stano, and S. J. Goodman, 2016: Demonstration of a GOES-R satellite convective toolkit to "bridge the gap" between severe weather watches and warnings: An example from the 20 May 2013 Moore, Oklahoma, tornado outbreak. *Bull. Amer. Meteor. Soc.*, **97 (1)**, 69–84.

Greene, D. R., and R. A. Clark, 1972: Vertically integrated liquid water—a new analysis tool. *Mon. Wea. Rev.*, **100 (7)**, 548–552.

Griffin, S. M., K. M. Bedka, and C. S. Velden, 2016: A method for calculating the height of overshooting convective cloud tops using satellite-based IR imager and *cloudsat* cloud profiling radar observations. *J. Appl. Meteor. Climatol.*, **55**, 479–491, doi:10.1175/JAMC-D-15-0170.1.

Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34 (1)**, 3–15.

Helmus, J., and S. Collis, 2016: The Python ARM Radar Toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *Journal of Open Research Software*, **4 (1)**.

Homeyer, C. R., J. D. McAuliffe, and K. M. Bedka, 2017: On the development of above-anvil cirrus plumes in extratropical convection. *J. Atmos. Sci.*, **74 (5)**, 1617–1633.

Iowa Environmental Mesonet, 2017: Archived NWS Watch/Warnings. Iowa Environmental Mesonet, accessed February 2017 to May 2017, URL http://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml.

Krzywinski, M., and N. Altman, 2014: Points of significance: Visualizing samples with box plots. *Nature Methods*, **11 (2)**, 119–120.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47 (7)**, 1940–1961.

Lemon, L. R., R. J. Donaldson, D. W. Burgess, and R. A. Brown, 1977: Doppler radar application to severe thunderstorm study and potential real-time warning. *Bull. Amer. Meteor. Soc.*, **58 (11)**, 1187–1193.

Lemon, L. R., and C. A. Doswell III, 1979: Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Wea. Rev.*, **107 (9)**, 1184–1197.

Line, W. E., T. J. Schmit, D. T. Lindsey, and S. J. Goodman, 2016: Use of geostationary super rapid scan satellite imagery by the Storm Prediction Center. *Wea. Forecasting*, **31 (2)**, 483–494.

Liu, C., and S. Heckman, 2011: The application of total lightning detection and cell tracking for severe weather prediction. *WMO Technical Conference on Instruments and Methods of Observation (TECO-2010)*.

Markowski, P. M., and Y. P. Richardson, 2009: Tornadogenesis: Our current understanding, forecasting considerations, and questions to guide future research. *Atmos. Res.*, **93 (1)**, 3–10.

Markowski, P. M., and Y. P. Richardson, 2014: What we know and don't know about tornado formation. *Phys. Today*, **67**, 26–31, doi:10.1063/PT.3.2514.

McCann, D. W., 1983: The Enhanced-V: A satellite observable severe storm signature. *Mon. Wea. Rev.*, **111**, 887–894.

Mecikalski, J. R., and K. M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, **134 (1)**, 49–78.

Menzel, W. P., and J. F. W. Purdom, 1994: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. *Bull. Amer. Meteorol. Soc.*, **75 (5)**, 757–781.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

National Centers for Environmental Information, 2017: NOAA's Severe Weather Data Inventory. National Centers for Environmental Information, accessed January 2015 to May 2017, URL https://www.ncdc.noaa.gov/swdi.

National Oceanic and Atmospheric Administration, Earth System Research Laboratory, 2012: Rapid Refresh (RAP). National Centers for Environmental Prediction, accessed August 2016 to July 2018, URL https://rapidrefresh.noaa.gov/.

NOAA, 1994: NOAA Geostationary Operational Environmental Satellite Imager Data. NOAA Comprehensive Large Array-data Stewardship System, accessed January 2015 to May 2017, URL http://www.class.ncdc.noaa.gov.

NOAA/NWS/ROC, 1991: NOAA Next Generation Radar (NEXRAD) Level II Base Data. NOAA National Centers for Environmental Information, accessed January 2015 to May 2017, doi: 10.7289/V5W9574V.

O'Brien, J. J., 1970: Alternative solutions to the classical vertical velocity problem. *J. Appl. Meteor.*, **9 (2)**, 197–203.

OFCM, 2005: Federal Meterological Handbook No. 11 – Doppler Radar Meteorological Observations, Part B: Doppler Radar Theory and Meteorology. FCM-H11B-2005 (Available online at http://www.ofcm.gov/publications/fmh/allfmh2.htm).

OFCM, 2006: Federal Meterological Handbook No. 11 – Doppler Radar Meteorological Observations, Part C: WSR-88D Products and Algorithms. FCM-H11C-2006 (Available online at http://www.ofcm.gov/publications/fmh/allfmh2.htm).

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 (1)**, 45–68.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Parker, M. D., 2014: Composite VORTEX2 supercell environments from near-storm soundings. *Mon. Wea. Rev.*, **142 (2)**, 508–529.

Potvin, C. K., D. Betten, L. J. Wicker, K. L. Elmore, and M. I. Biggerstaff, 2012: 3dvar versus traditional dual-doppler wind retrievals of a simulated supercell thunderstorm. *Mon. Wea. Rev.*, **140 (11)**, 3487–3494, doi:10.1175/MWR-D-12-00063.1.

Przybylinski, R. W., 1995: The bow echo: Observations, numerical simulations, and severe weather detection methods. *Wea. Forecasting*, **10 (2)**, 203–218.

Rasmussen, E. N., and D. O. Blanchard, ????: *Wea. Forecasting*.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, doi:10.1175/2008WAF2222159.1.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, and D. S. Zrnić, 2005: Polarimetric tornado detection. *J. Appl. Meteor.*, **44 (5)**, 557–570.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98 (4)**, 681–698.

Schmit, T. J., and Coauthors, 2013: Geostationary Operational Environmental Satellite (GOES)-14 super rapid scan operations to prepare for GOES-R. *J. Appl. Remote. Sens.*, **7 (1)**, 073 462–073 462.

Schultz, C. J., L. D. Carey, E. V. Schultz, and R. J. Blakeslee, 2017: Kinematic and microphysical significance of lightning jumps versus nonjump increases in total flash rate. *Wea. Forecasting*, **32 (1)**, 275–288.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *J. Appl. Meteor. Climatol.*, **48 (12)**, 2543–2563.

Smith, T. M., and K. L. Elmore, 2004: The use of radial velocity derivatives to diagnose rotation and divergence. *Preprints, 11th Conf. on Aviation, Range, and Aerospace, Hyannis, MA, Amer. Meteor. Soc. P*, Vol. 5.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97 (9)**, 1617–1630.

Starzec, M., C. R. Homeyer, and G. L. Mullendore, 2017: Storm labeling in three dimensions (SL3D): A volumetric radar echo and dual-polarization updraft classification algorithm. *Mon. Wea. Rev.*, **145 (3)**, 1127–1145.

Stensrud, D. J., J. V. Cortinas Jr, and H. E. Brooks, 1997: Discriminating between tornadic and nontornadic thunderstorms using mesoscale model output. *Wea. Forecasting*, **12 (3)**, 613–632.

Stensrud, D. J., and Coauthors, 2009: Convective-scale Warn-on-Forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90 (10)**, 1487–1499.

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16.

Stumpf, G. J., A. Witt, E. D. Mitchell, P. L. Spencer, J. T. Johnson, M. D. Eilts, K. W. Thomas, and D. W. Burgess, 1998: The national severe storms laboratory mesocyclone detection algorithm for the wsr-88d. *Wea. Forecasting*, **13 (2)**, 304–326.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

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 (6)**, 1243–1261.

Thompson, R. L., 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 (5)**, 1136–1154.

Trapp, R. J., G. R. Marion, and S. W. Nesbitt, 2017: The regulation of tornado intensity by updraft width. *J. Atmos. Sci.*, doi:10.1175/JAS-D-16-0331.1.

Vicente, G., J. Davenport, and R. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *Int. J. Remote Sens.*, **23 (2)**, 221–230.

Williams, E., and Coauthors, 1999: The behavior of total lightning activity in severe Florida thunderstorms. *Atmos. Res.*, **51 (3)**, 245–265.

Witt, A., and S. P. Nelson, 1991: The use of single-doppler radar for estimating maximum hailstone size. *J. Appl. Meteor.*, **30 (4)**, 425–431.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

LIST OF TABLES

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

TABLE 1. Dates, number of storms, number of tornadic storms, number of tornadoes, dominant storm mode (discrete or mesoscale convective system), and the longitude-latitude coordinates of the analysis domain for the 27 severe weather days analyzed in this study. Dates in bold represent days where GOES-14 and ENTLN data were available, with one exception (ENTLN data were not obtained for the 4 June 2015 case).

Date	No. Storms (No. Severe)	No. Tornadic (No. Tornadoes)	Dominant storm mode	Analysis Domain Coordinates [lon_0, lat_0, lon_1, lat_1] (°W and °N)
22 May 2011	469 (68)	21 (59)	Discrete	[95.5, 35.5, 87.0, 46.5]
24 May 2011	450 (88)	24 (64)	Discrete	[101.5, 32.0, 92.5, 39.0]
9 April 2012	30 (4)	1 (6)	Discrete	[101.0, 33.5, 95.0, 37.5]
13 April 2012	97 (6)	3 (14)	Discrete	[100.5, 34.5, 95.0, 37.0]
14 April 2012	313 (30)	23 (96)	Discrete	[101.0, 36.0, 95.5, 41.5]
20 May 2013	246 (67)	16 (35)	Discrete	[99.0, 31.5, 93.0, 40.0]
31 May 2013	391 (63)	14 (36)	Discrete	[99.0, 34.5, 87.0, 40.5]
12 June 2013	555 (126)	10 (21)	MCS	[96.0, 38.0, 80.0, 45.0]
27 April 2014	223 (57)	8 (21)	Discrete	[99.0, 34.0, 91.5, 42.0]
10 May 2014	112 (40)	2 (5)	Discrete	[99.0, 36.0, 90.0, 43.0]
11 May 2014	330 (63)	10 (41)	Discrete	[102.0, 36.0, 92.0, 44.5]
21 May 2014	54 (10)	2 (5)	Discrete	[106.0, 37.5, 101.0, 41.0]
16 June 2014	406 (66)	10 (40)	Discrete	[100.0, 41.0, 89.0, 44.0]
17 June 2014	155 (22)	7 (16)	Discrete	[106.0, 41.5, 94.5, 48.0]
18 June 2014	79 (8)	5 (13)	Discrete	[100.0, 43.5, 98.0, 46.5]
13 October 2014	707 (80)	17 (24)	MCS	[95.5, 29.5, 84.5, 40.5]
6 May 2015	202 (53)	23 (52)	Discrete	[100.0, 32.5, 95.5, 41.5]
19 May 2015	329 (32)	13 (36)	Discrete	[103.0, 29.0, 94.0, 37.0]
24 May 2015	123 (16)	1 (10)	MCS	[105.0, 36.0, 97.0, 41.0]
25 May 2015	669 (64)	18 (28)	MCS	[105.0, 25.0, 89.0, 41.0]
27 May 2015	387 (48)	8 (18)	Discrete	[104.0, 29.5, 96.0, 41.5]
4 June 2015	290 (42)	3 (23)	Discrete	[108.0, 34.0, 93.0, 43.0]
23 December 2015	137 (34)	7 (26)	MCS	[92.5, 33.5, 84.0, 42.0]
15 April 2016	160 (28)	4 (12)	Discrete	[104.0, 34.5, 99.0, 40.5]
9 May 2016	199 (64)	10 (26)	Discrete	[100.0, 33.0, 94.0, 41.5]
24 May 2016	150 (35)	11 (44)	Discrete	[104.0, 35.5, 97.0, 41.0]
25 May 2016	17 (6)	2 (6)	Discrete	[99.5, 35.5, 95.0, 40.0]
Total	**7280 (978)**	**273 (777)**	–	–

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

TABLE 2. All variables presented in this study, categorized by their source and type (physical or kinematic).

	Radar	Satellite	Lightning
Kinematic	Rotation Extrema	Cloud Top Vorticity Extrema	N/A
	Divergence Extrema	Cloud Top Divergence Extrema	
	Velocity Spectrum Width Extrema		
Physical	Echo Top Altitude (at a 40-dBZ Z_H threshold)	Visible Texture Rating	Total Flash Density

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

TABLE 3. The number of 1-min observations contributing to box plots in this study.

Figure	Non-tornadic	Non-tornadic severe	Non-tornadic significant severe	30 min before first tornado	15 min before first tornado	During tornado	15 min after last tornado	30 min after last tornado
3A	125615	16373	3256	961	1164	6207	1053	817
3B	15307	2215	558	112	138	541	93	54
3C	128713	16585	3318	961	1164	6210	1058	821
3D	21372	3221	735	117	232	977	168	119
3E	125671	16355	3281	961	1164	6210	1053	821
3F	36212	5284	1177	185	226	844	206	130
4A	122414	16260	3272	955	1164	6135	1041	796
4B	120454	15739	3287	953	1159	6134	1040	792
4C	96926	13586	2672	825	983	5243	863	672
4D	121654	16056	3220	951	1149	6123	1049	817
4E	121151	16143	3338	953	1159	6134	1040	792
4F	44429	6710	1431	378	465	3243	424	316
A1B	24699	2658	608	223	269	1086	231	170
A2A	231963	15850	2648	275	345	878	260	205
A2B	103295	13164	2289	224	290	745	195	158
A2C	220861	15224	2589	255	331	844	253	195
A2D	231293	15595	2653	275	345	878	260	205
A2E	232414	15790	2644	275	345	878	260	205
A2F	220677	15282	2588	260	331	845	253	195
A2G	238546	16080	2648	275	345	878	260	205
A2H	229796	15515	2648	275	345	878	260	205

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

TABLE 4. Values of the performance metrics for the rotation-divergence product using a threshold of $42 \cdot 10^{-6}$ s^{-2} [where NWS skill (CSI) was matched using data from all 27 cases]. Here, values are shown for the performance when the threshold was used for all 27 cases, cases dominated by MCSs, and cases grouped by the number of tornadic storms that occurred.

	POD (%)	FAR (%)	Mean flag lead time (min)	Median flag lead time (min)
27 cases	58.30	85.85	43.0	35
MCS	32.00	92.52	38.4	35
1-5 tornadic storms	69.57	87.30	54.4	47
6-15 tornadic storms	69.23	87.20	39.5	34
$\geq$ 16 tornadic storms	58.88	77.90	44.8	34

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Table A1. Similar to Table 1. Dates, number of storms, number of tornadic storms, number of tornadoes, dominant storm mode (discrete or mesoscale convective system), and the longitude-latitude coordinates of the analysis domain for the randomly selected 2011 severe weather days.

Date	No. Storms (No. Severe)	No. Tornadic (No. Tornadoes)	Dominant storm mode	Analysis Domain Coordinates [lon_0, lat_0, lon_1, lat_1] (°W and °N)
25 January 2011	665 (18)	6 (7)	MCS	[84.0, 25.0, 80.0, 31.0]
24 February 2011	547 (97)	10 (19)	Discrete	[95.0, 31.0, 82.0, 38.5]
5 March 2011	501 (8)	4 (7)	Discrete	[96.0, 28.5, 86.5, 34.5]
19 March 2011	102 (10)	2 (2)	Discrete	[103.5, 32.0, 78.5, 35.5]
29 March 2011	1101 (45)	2 (3)	Discrete	[94.5, 28.5, 84.5, 36.5]
09 April 2011	296 (34)	7 (23)	Discrete	[98.5, 40.5, 90.0, 45.5]
21 April 2011	113 (7)	3 (5)	Discrete	[104.5, 28.5, 96.5, 35.5]
26 May 2011	1415 (192)	12 (14)	MCS	[91.5, 29.5, 74.0, 43.0]
29 May 2011	346 (47)	3 (4)	MCS	[94.0, 40.0, 81.5, 45.5]
1 June 2011	351 (48)	4 (7)	Discrete	[80.5, 39.5, 67.0, 46.5]
10 June 2011	1016 (74)	1 (1)	Discrete	[97.0, 37.0, 80.5, 42.0]
12 June 2011	557 (41)	9 (12)	Discrete	[108.0, 38.0, 74.5, 47.5]
27 June 2011	216 (8)	2 (2)	Discrete	[85.5, 36.5, 79.5, 42.0]
29 June 2011	199 (12)	2 (2)	Discrete	[117.0, 43.0, 104.0, 49.0]
17 July 2011	241 (17)	4 (8)	Discrete	[104.0, 44.0, 93.5, 49.0]
26 July 2011	1416 (50)	4 (7)	Discrete	[104.0, 40.0, 69.5, 47.0]
2 August 2011	212 (1)	1 (1)	Discrete	[84.5, 25.0, 80.0, 30.5]
17 September 2011	488 (15)	1 (3)	Discrete	[102.5, 31.5, 94.0, 37.5]
7 October 2011	425 (21)	3 (4)	Discrete	[104.0, 34.0, 95.5, 43.0]
7 November 2011	552 (24)	2 (15)	MCS	[103.0, 32.0, 92.5, 38.0]
21 December 2011	101 (1)	1 (1)	MCS	[89.0, 30.0, 82.5, 35.5]
22 December 2011	632 (21)	10 (18)	MCS	[93.5, 29.0, 83.0, 35.0]
Total	**11492 (791)**	**93 (165)**	–	–

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

LIST OF FIGURES

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

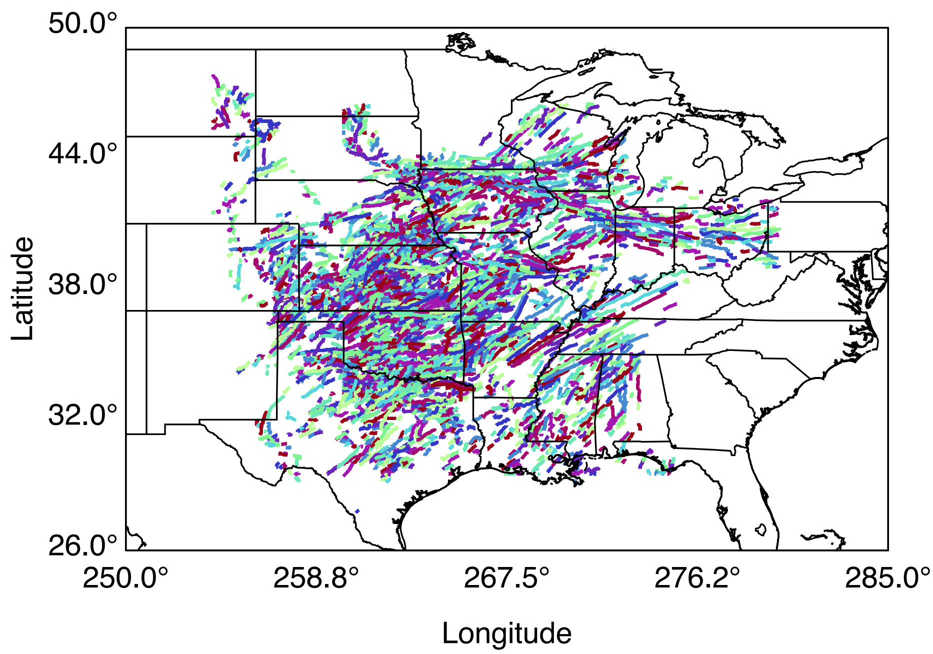


FIG. 1. Storm tracks of at least 30 min in length from all 27 cases. Variation in color is arbitrary and meant to improve interpretation of overlapping storms.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

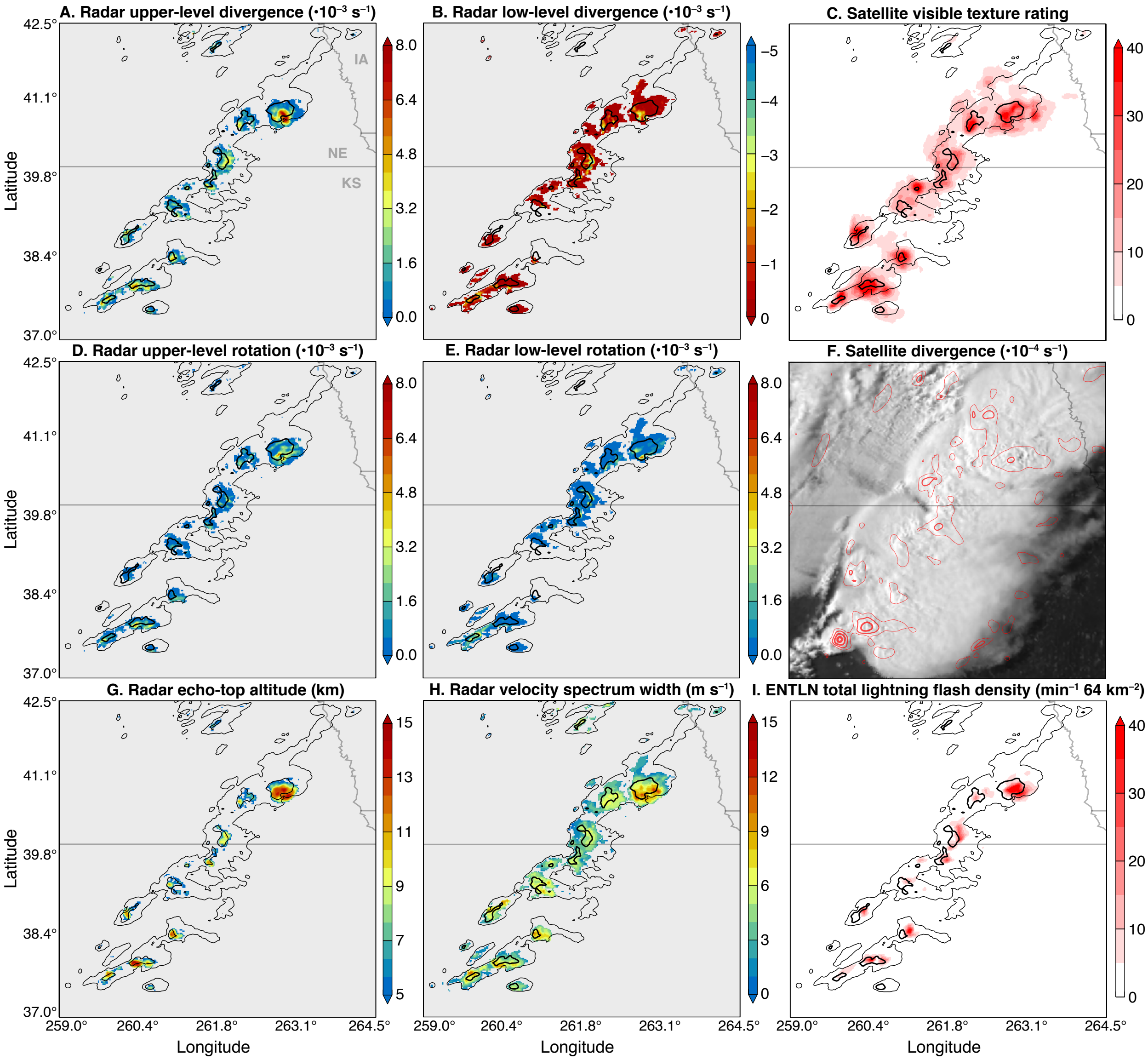
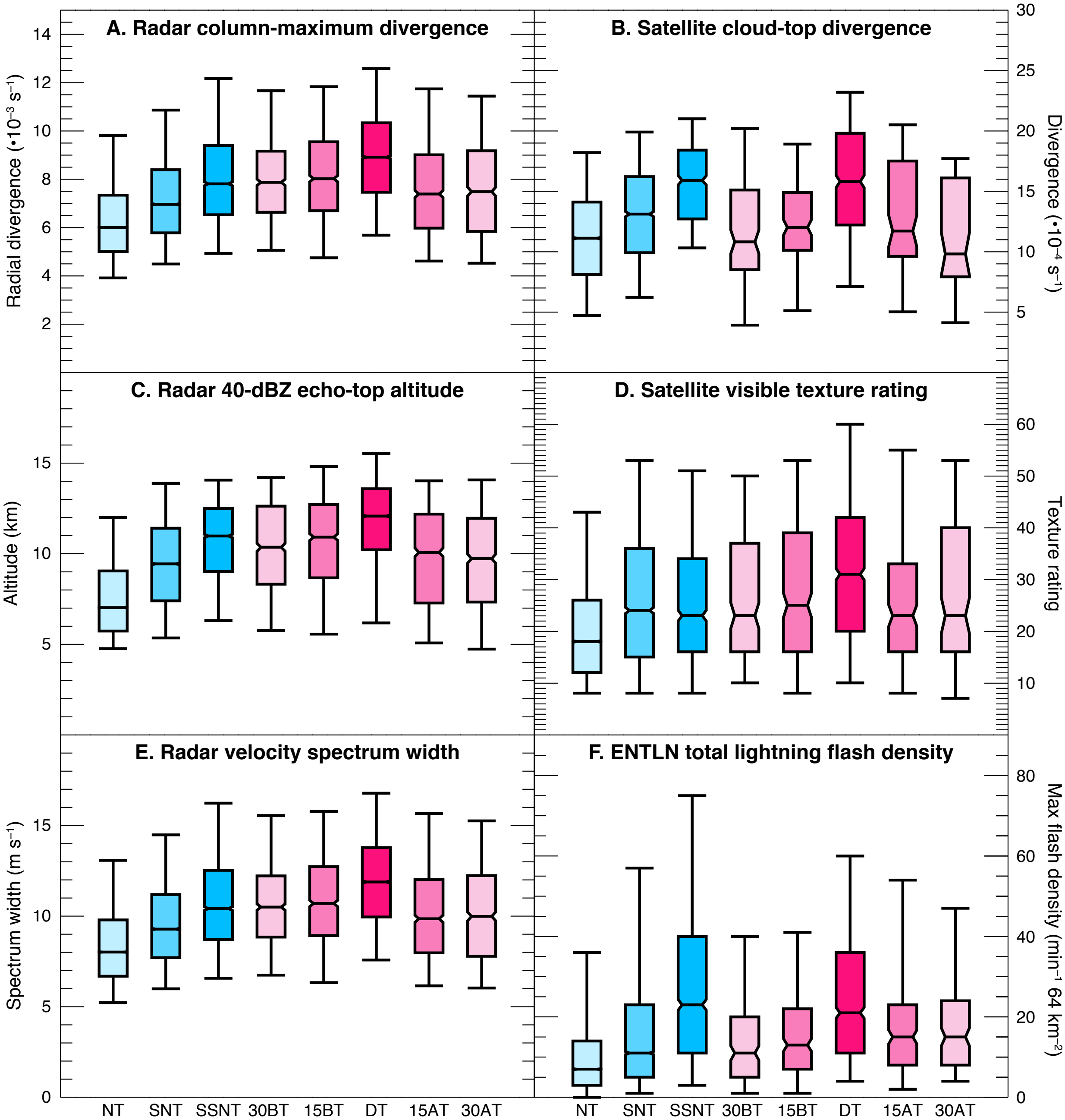


FIG. 2. Example maps of variables valid 11 May 2014 at 22:30 UTC in Kansas and Nebraska. 15- and 45-dBZ 0-5 km column-maximum reflectivity values are contoured in black (increasing thickness for increasing reflectivity) and superimposed in most panels. Red contours in panel F signify satellite mAMV divergence in $5 \cdot 10^{-4}$ s^{-1} increments with increasing thickness starting from $5 \cdot 10^{-4}$ s^{-1}. Satellite products were subject to the parallax correction method described in Section 2c.

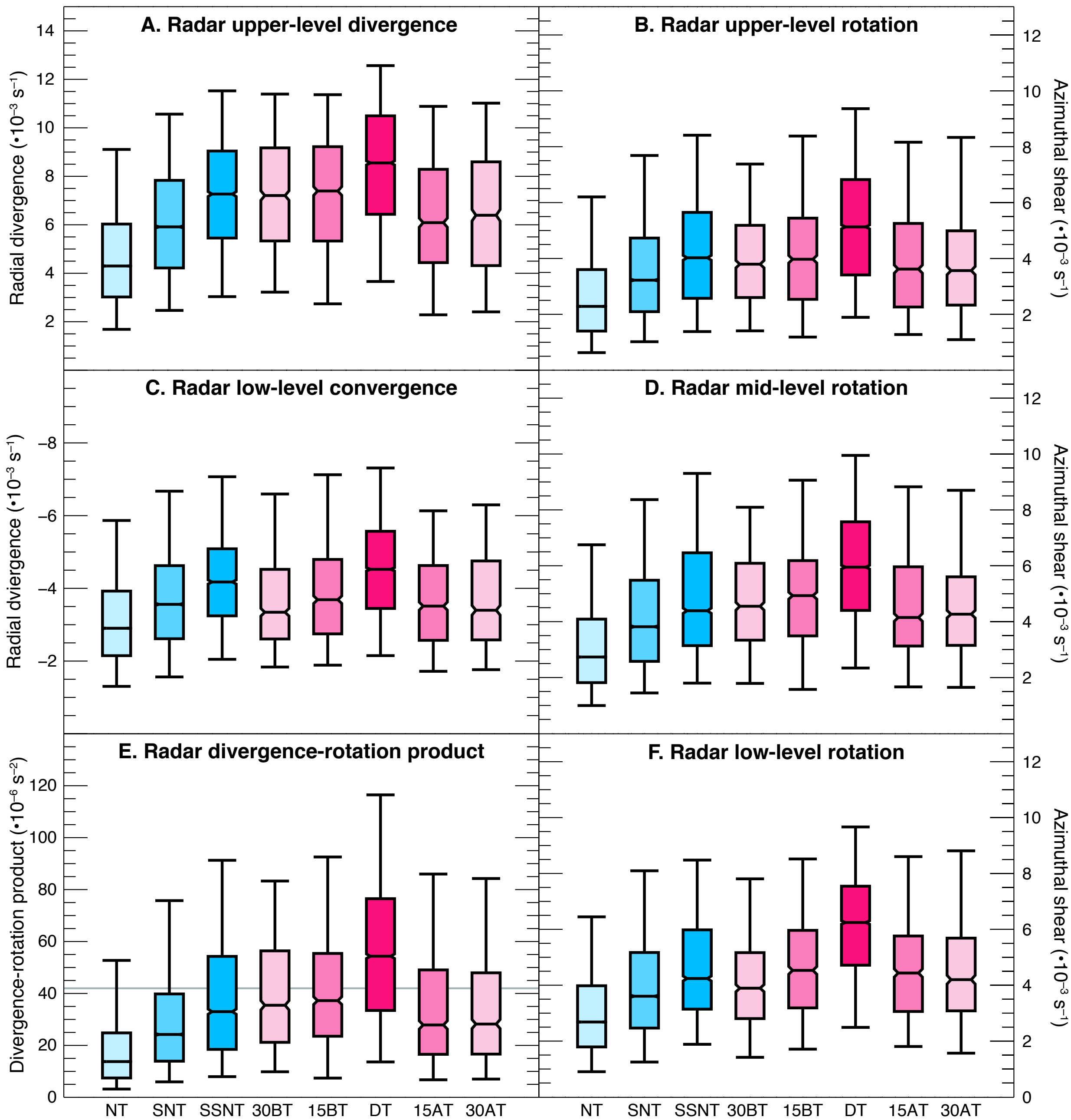
Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.



Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

FIG. 3. Box plots for kinematic and physical metrics of upward motion derived from radar, satellite, and lightning data. The notched box-and-whiskers show the 5th, 25th, 50th, 75th, and 95th percentiles of each metric for all severe weather days for which data are available. Notches in the boxes emanating from the median values represent the 95% confidence interval for the median values. When the notches of different boxes within the same subplot do not overlap, the medians are taken to be significantly different (Krzywinski and Altman 2014). The three leftmost boxes in each subplot show distributions based on the 30-min periods around the maximum of a given variable for all non-tornadic storms (NT), severe non-tornadic storms (SNT), and significant severe non-tornadic storms (SSNT). The five remaining boxes show distributions for tornadic storms at 30 and 15 min prior to the first tornado (30BT and 15BT), during the lifecycle of all tornadoes (DT), and 15 and 30 min after the last tornado (15AT and 30AT). The number of observations contributing to each box in every figure, as well as the values for the 5 percentiles for each box, can be found in the supplemental tables.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.



1001 FIG. 4. As in Fig. 3, but for divergence and rotation variables. The gray horizontal line in panel E represents
1002 the threshold used for the the objective tornadic storm identification method evaluated in Section 3a.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

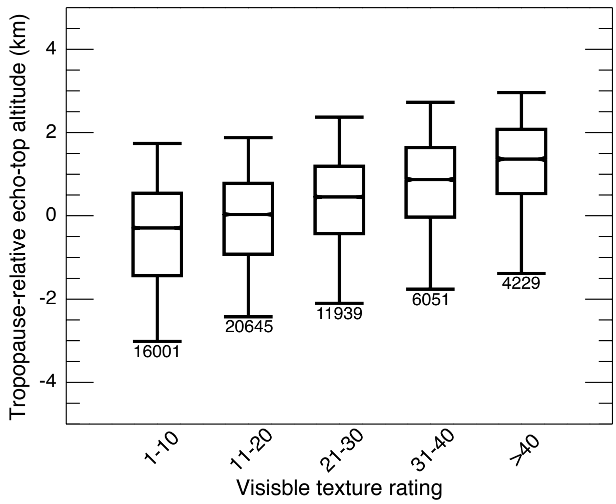


FIG. 5. Tropopause-relative Z_H = 10 dBZ echo-top altitude from radar versus visible texture rating from satellite imagery for all 27 severe weather days. Each box-and-whisker represents the 5th, 25th, 50th, 75th, and 95th percentiles of the echo-top altitude distribution at a specified range of visible texture ratings. Numbers at the bottom of each box represent the number of contributing 1-min observations.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

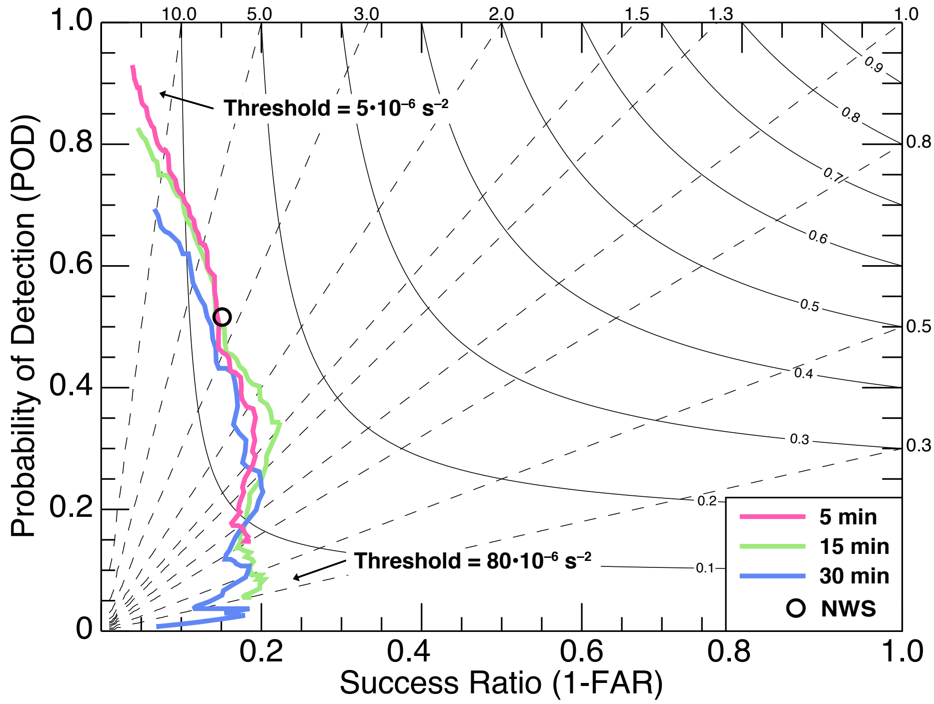
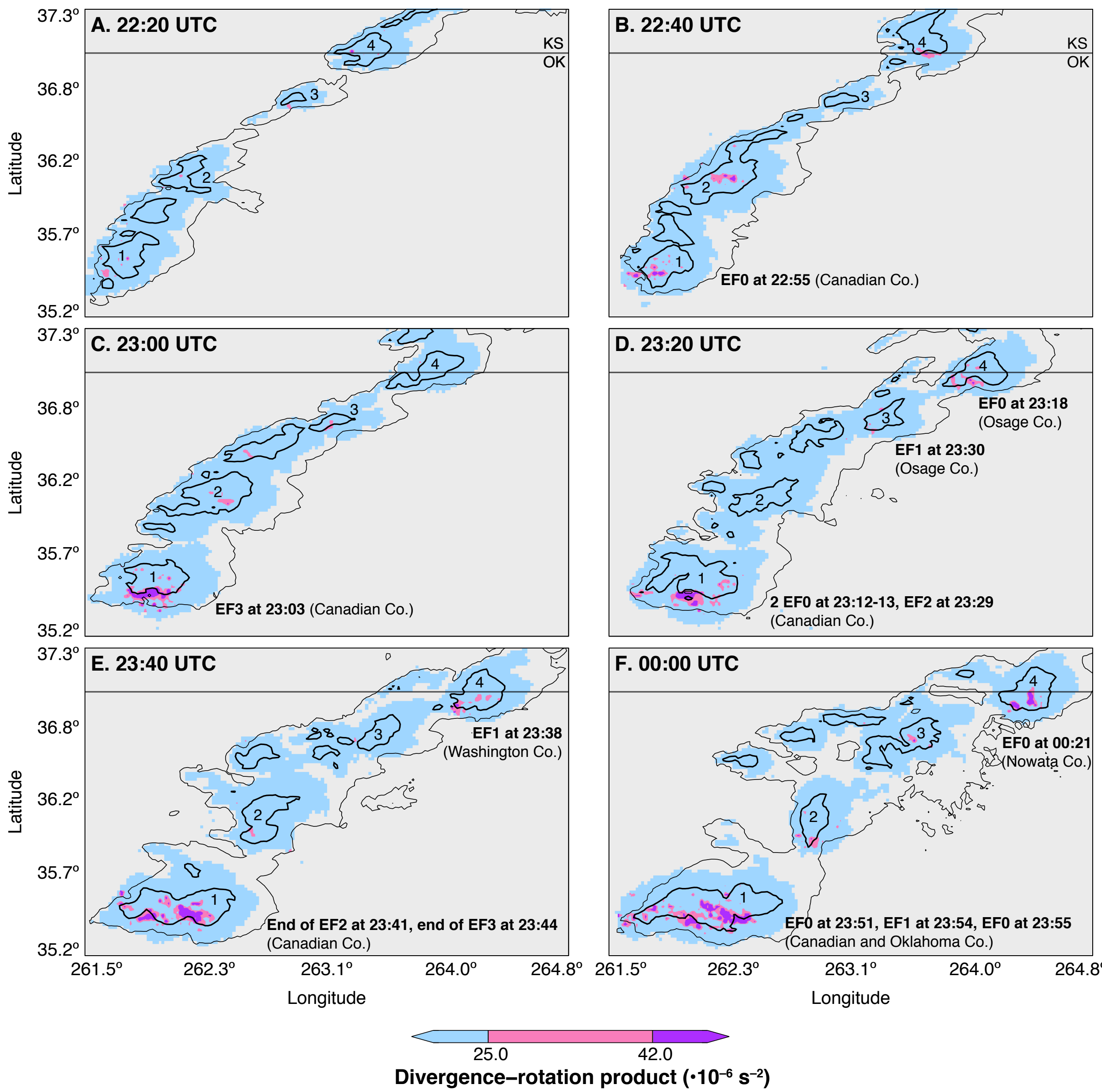


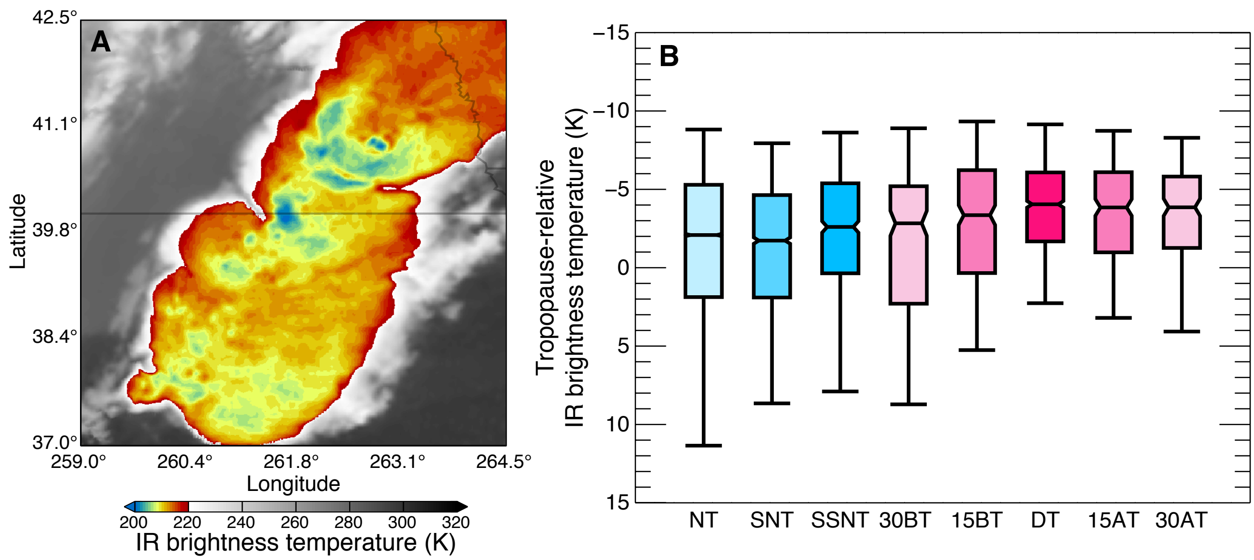
FIG. 6. Performance diagram for the objective threshold and NWS warning-based methods. Solid black lines are lines of constant CSI. The dashed lines represent bias, where values >1 signify over-forecasting and values <1 signify under-forecasting. Colored lines show the performance of the objective threshold method at multiple time periods (5, 15, and 30 min) of exceedance for divergence-rotation product threshold values ranging from $5 \cdot 10^{-6}$ to $80 \cdot 10^{-6}$ s^{-2}. The open black circle shows the cumulative performance for the 27 severe weather days for the NWS warning-based method.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.



1013 FIG. 7. Example maps of the radar divergence-rotation product valid 31 May 2013 in Oklahoma and Kansas
1014 in the time window from 22:20 UTC to 00:00 UTC on 1 June 2013 (at 20-min intervals). All values above
1015 the threshold of $42 \cdot 10^{-6}$ s^{-2} are colored purple, values between 25 and 42 $\cdot 10^{-6}$ s^{-2} are shown in pink, and
1016 any values below $25 \cdot 10^{-6}$ s^{-2} are colored blue. Storms of interest are labeled in the different panels and
1017 tornado reports (and the state counties in which they occurred) are noted in each map. 15- and 45-dBZ 0-5 km
1018 column-maximum reflectivity values are contoured in black (increasing thickness for increasing reflectivity).

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.



1019 Fig. A1. Panel A shows an example map of the IR brightness temperature for the same case as in Fig. 2. Panel
1020 B depicts box plots for minimum tropopause-relative IR brightness temperature similar to those in Fig. 3.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.

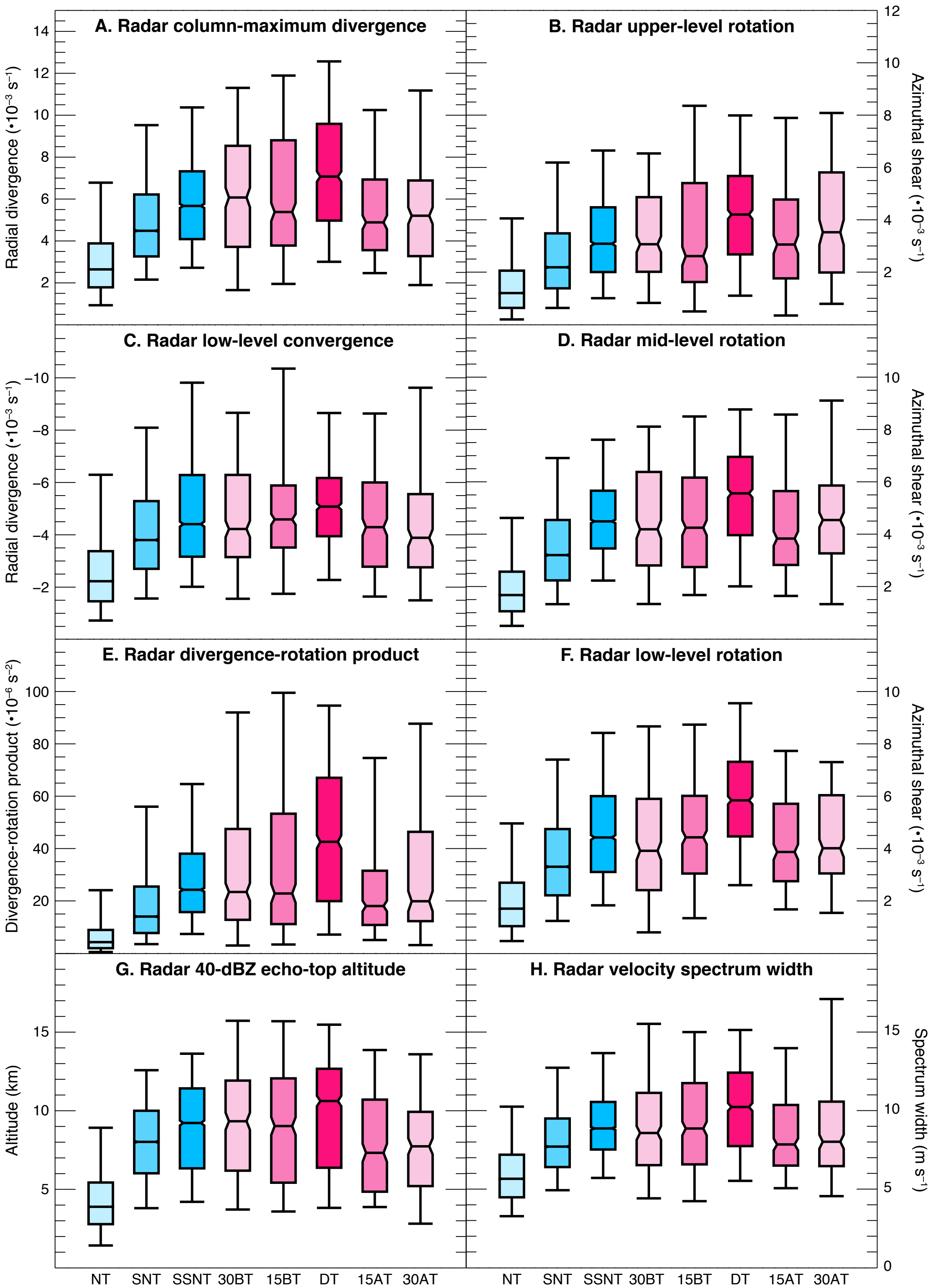


Fig. A2. Similar to Fig. 3. Box plots for the 22 randomly selected cases from 2011.

Accepted for publication in *Journal of Applied Meteorology and Climatology*. DOI 10.1175/JAMC-D-18-0241.1.