The National Severe Storms Laboratory Mesocyclone Detection Algorithm for the WSR-88D*

GREGORY J. STUMPF,[+#] ARTHUR WITT,[+] E. DEWAYNE MITCHELL,[+#] PHILLIP L. SPENCER,[+#] J. T. JOHNSON,[+] MICHAEL D. EILTS,[+] KEVIN W. THOMAS,[+#] AND DONALD W. BURGESS[@]

[+] *National Severe Storms Laboratory, Norman, Oklahoma*
[#] *Cooperative Institute for Mesoscale Meteorology Studies, University of Oklahoma, Norman, Oklahoma*
[@] *National Weather Service Operational Support Facility, Norman, Oklahoma*

(Manuscript received 5 March 1997, in final form 3 December 1997)

ABSTRACT

The National Severe Storms Laboratory (NSSL) has developed a mesocyclone detection algorithm (NSSL MDA) for the Weather Surveillance Radar-1988 Doppler (WSR-88D) system designed to automatically detect and diagnose the Doppler radar radial velocity patterns associated with storm-scale (1–10-km diameter) vortices in thunderstorms. The NSSL MDA is an enhancement to the current WSR-88D Build 9.0 Mesocyclone Algorithm (88D B9MA).

The recent abundance of WSR-88D observations indicates that a variety of storm-scale vortices are associated with severe weather and tornadoes, and not just those vortices meeting previously established criteria for mesocyclones observed during early Doppler radar studies in the 1970s and 1980s in the Great Plains region of the United States. The NSSL MDA's automated vortex detection techniques differ from the 88D B9MA, such that instead of immediately thresholding one-dimensional shear segments for strengths comparable to predefined mesocyclone parameters, the initial strength thresholds are set much lower, and classification and diagnosis are performed on the properties of the four-dimensional detections. The NSSL MDA also includes multiple range-dependent strength thresholds, a more robust two-dimensional feature identifier, an improved three-dimensional vertical association technique, and the addition of time association and trends of vortex attributes. The goal is to detect a much broader spectrum of storm-scale vortices (so that few vortices are missed), and then diagnose them to determine their significance. The NSSL MDA is shown to perform better than the 88D B9MA at detecting storm-scale vortices and diagnosing significant vortices.

Operational implications of the NSSL MDA are also presented. In light of the new WSR-88D observations of storm-scale vortices and their association with severe weather and tornadoes, it is clear that the operational paradigms of automated vortex detection require changes.

1. Introduction and background

The National Weather Service (NWS) Weather Surveillance Radar-1988 Doppler (WSR-88D) system (Crum and Alberty 1993) contains a suite of severe weather detection algorithms designed to detect various severe weather signatures in Doppler radar. This suite includes the WSR-88D Build 9.0 Mesocyclone Algorithm (hereafter called the 88D B9MA), which is designed to detect the Doppler radar radial velocity signature of significant rotation in severe thunderstorms. The 88D B9MA is based on the algorithm described by Zrnić et al. (1985). The 88D B9MA has been in use in NWS warning operations since the first WSR-88D was deployed. With the operational use of the 88D B9MA, along with the advent of the collection of large amounts of data from the newly deployed WSR-88D network, several shortcomings of the 88D B9MA have been discovered. This paper describes the efforts of the National Severe Storms Laboratory (NSSL) to develop an enhanced version of the 88D B9MA that addresses these shortcomings.

Traditionally, a *mesocyclone* has been defined as the Doppler radar velocity signature of a storm-scale (1–10-km diameter) Rankine vortex (Donaldson 1970; Burgess 1976), which corresponds to the rotating updraft–downdraft couplet of a supercell thunderstorm. The mesocyclone is the cyclonic vortex that is larger than and may contain the more intense tornado vortex.[1] About

* An electronic supplement to this article may be found on the CD-ROM accompanying this issue or at http://www.ametsoc.org/AMS.

Corresponding author address: Gregory J. Stumpf, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069.
E-mail: stumpf@nsslsun.nssl.noaa.gov

[1] A related WSR-88D algorithm, the NSSL Tornado Detection Algorithm (Mitchell et al. 1998) is designed to detect even smaller (≤2-

one-half of the mesocyclones observed during the Joint Doppler Operational Project (JDOP) (Burgess et al. 1979) produced tornadoes. Since JDOP, continued radar observations of mesocyclones have shown that the percentage of mesocyclones believed to be tornadic is much less than 50%, and perhaps as low as 30% (Burgess and Lemon 1991) or lower. As important, however, is that more than 90% of mesocyclones are associated with severe weather of some kind (i.e., tornadoes, hail $\geq$19-mm diameter, wind damage, and/or measured surface wind speeds $\geq$25 m s^{-1}). Detecting mesocyclones is clearly important to accurate and timely severe weather warning operations.

In the past decade, certain strength, spatial, and temporal criteria for mesocyclone recognition have been developed based on extensive Doppler radar observation (Burgess et al. 1979; Wieler 1986; Burgess et al. 1993). These criteria, along with various conceptual models of intense atmospheric vortices (e.g., Brown and Wood 1991) have led to the development of algorithms that automatically detect these phenomena. However, since these early studies, there has been some debate over the proper definition of a mesocyclone in terms of radar-derived parameters. Furthermore, a mesocyclone is also difficult to verify using in situ observations (e.g., spotter reports), as some of the rotation is obscured within the storm cloud. What degree of rotation is required in a thunderstorm in order for that rotation to be considered operationally significant? Consequently, what *defines* significant?

Early mesocyclone studies and algorithm tests were also restricted in the amount of radar data that was analyzed. In the ''pre-WSR-88D era,'' there were very few operational Doppler radars collecting data (with most mesocyclone cases being recorded in central Oklahoma). In addition, limited computing resources prohibited the analysis of large datasets with geographical and seasonal diversity. For example, Zrnić et al. (1985) and Desrochers and Donaldson (1992) (hereafter DD92) studied datasets on the order of a few hours and a few storms from only the central Oklahoma region. Although these early algorithm tests resulted in great success with high skill scores, the data used for testing were limited.

Recently, with the full deployment of the WSR-88D radar network across the United States, the amount of data collected has increased by several orders of magnitude; data are no longer restricted to a few areas of the country and are recorded year round, providing greater geographical and seasonal diversity. The current widespread operational use of the 88D B9MA in warning operations has led to the discovery of several shortcomings in the algorithm, including an overabundance of detections (high false alarm rate) in certain severe storm situations, such as squall lines and bow echoes (Burgess et al. 1993; Lee and White 1998).

The NSSL has developed a new mesocyclone detection algorithm that addresses these shortcomings and helps meet the needs of the meteorologist who has to make warning decisions. The first version of the new design addressed the false alarm problem by raising existing thresholds and creating new thresholds (including strength, spatial, and time continuity thresholds). Also, the new design attempted to overcome the difficulty of the 88D B9MA in detecting mesocyclones at far ranges. The result was a new mesocyclone detection algorithm (hereafter called the Early MDA) (Stumpf and Witt 1994), which improved upon the WSR-88D algorithm with a higher critical success index (CSI) (Wilks 1995). However, the detection capabilities of the Early MDA were too conservative and a large number of mesocyclones associated with severe weather and tornadoes were not being detected. The development of the Early MDA took place in the early deployment phase of the WSR-88D network, and testing on geographically and seasonally diverse datasets had not yet occurred. Most of the data used to develop both the 88D B9MA and the Early MDA were not even collected using the WSR-88Ds, but rather used data from other research radars with different characteristics. Also, computational resources were still very limited during the development of the 88D B9MA and the Early MDA.

As new WSR-88D systems were deployed at the NWS offices across the United States, resulting in a greater seasonal and geographical climatology of storm types, additional shortcomings of the 88D B9MA were found (as anticipated by Burgess et al. 1993). For example, the 88D B9MA (using its fielded default site-adaptable parameters) has a significant failure rate in detecting low-topped or minisupercells (Burgess et al. 1995). Also, many storm-scale vortices that failed to meet designated criteria for ''mesocyclone'' were not being detected, yet were still producing severe weather or tornadoes. Recent Doppler radar observations and theoretical studies are discovering that not all tornado and mesocyclone morphologies fit the original tornado and mesocyclone conceptual models (Burgess et al. 1982; Burgess et al. 1993; Doswell and Burgess 1993). For instance, not all supercell tornadoes are preceded by the development of a midlevel mesocyclone that builds to the ground [an example is shown by Watson et al. (1995)]. Also, although MDAs are not designed to detect nonsupercell tornadoes (those tornadoes not associated with a larger parent vortex) (Wakimoto and Wilson 1989), there are varieties of tornadoes that are sometimes classified as nonsupercell tornadoes, but are associated with radar signatures of vortices that are of a larger scale than the actual tornado [e.g., bow-echo tornadoes and boundary layer vortices/mesocyclones

km diameter), and more intense, vortices associated with tornado cyclones [a vortex between the diameter scales of a tornado and a mesocyclone, as defined by Rotunno (1986) and Wicker and Wilhelmson (1993)] and tornadoes.

(Stumpf and Burgess 1993) and tornadoes associated with tropical cyclones (McCaul 1987)].

With all these shortcomings in mind, NSSL has designed an upgrade to the Early MDA, hereafter named the NSSL Mesocyclone Detection Algorithm (NSSL MDA). The NSSL MDA development effort was primarily funded by the NWS Operational Support Facility (OSF) and the Federal Aviation Administration (FAA), with an initial goal of enhancing the 88D B9MA and incorporating some ideas from other algorithms [including those of DD92 and Johnson et al. (1998)]. The goal included the development of an algorithm that was tested using strictly WSR-88D data, as the algorithm is intended for use only with that radar for NWS and FAA warning operations. The NSSL MDA design takes advantage of greater computational power for analysis of larger radar datasets, and state-of-the-art color computer display systems for enhanced graphical output analysis products.

The mesocyclone definitions used in the 88D B9MA and the Early MDA have specific strength, spatial, and temporal thresholds. However, mesocyclone intensity is not always a fail-safe indicator of severity (Burgess and Lemon 1991). Some tornado or severe weather events occur from storms associated with vortices not reaching any of these thresholds. Also, some strong vortices (that do meet the criteria) occur without a report of a tornado or severe weather. Because of this, it became apparent that any new version of the mesocyclone detection algorithm should be able to detect an entire *spectrum* of storm-scale vortices (on the order of 1–10 km in diameter) and then later classify each detection with a robust set of rules. Additionally, the attributes of each storm-scale vortex could then be diagnosed to determine the probability that the storm-scale vortex will be associated with a tornado or severe weather occurrence.

With all of this in mind, the NSSL MDA builds upon the work of the Early MDA and DD92. Instead of defining the strength thresholds at the very first analysis step (as the 88D B9MA and the Early MDA does), the NSSL MDA attempts to detect *all* storm-scale vortices (1–10-km diameter) and then diagnose them to determine if they are significant. This necessitates a change in operational strategies of the warning forecasters who use mesocyclone algorithm output. Detecting all storm-scale vortices and providing the user a variety of diagnostic information and tools (e.g., graphical output products) should produce a more-informed severe weather warning decision. By detecting vortices from their incipient signatures, through maturity, and onto demise, the NSSL MDA can provide a complete time history of the vortices and produce more reliable diagnostic information to the warning forecaster.

2. Algorithm analysis procedures

a. Overview

The basis of all mesocyclone detection algorithms is the automated pattern recognition of an area of rotation in Doppler radar radial velocity data. Because only the component of airflow directly toward or away from the radar (along-radial) is measured, areas of rotation appear as a couplet of strong localized opposing flows (outbound velocities next to inbound velocities at approximately constant range).

The NSSL MDA processes the Doppler radar radial velocity data in four dimensions. WSR-88D data are collected in volume scans consisting of a series of 360° antenna sweeps at constant elevation. Each consecutive sweep is taken with increasing elevation angle. The algorithm first processes data at the one-dimensional (1D) level; shear segments of cyclonic azimuthal shear are detected. Next, the shear segments are horizontally associated to form two-dimensional (2D) features. The NSSL MDA then uses vertical association to create three-dimensional (3D) detections at the end of each volume scan. Finally, time association is employed to complete the process.

Most algorithms are designed with site-adaptable parameters that can be fine-tuned for better operational skill given varying near-storm environments or geographical regions. Although the NSSL MDA has a set of adaptable parameters to create 3D detections (Stumpf and Marzban 1995), fine-tuning these parameters may not be necessary. With an algorithm designed to detect *all* storm-scale vortices and provide diagnostic information (and improved data display formats, to be discussed later), the user can make more informed decisions about which vortices are considered significant for that particular near-storm environment and geographical region.

b. Data preprocessing

In order for any MDA to function properly, high quality Doppler velocity data are required. If data artifacts are present, such as noisy or improperly dealiased velocities, algorithm performance can be adversely affected. As a first step at eliminating noisy data, velocities whose corresponding reflectivity values are below a preset threshold (typically 0–20 dBZ) are deleted.[2] The velocity dealiasing scheme used to preprocess all the data analyzed here is the WSR-88D system (Eilts and Smith 1990) modified using adaptable parameter optimization and new logic configuration (Conway et al. 1995) to correct some existing deficiencies in the WSR-88D dealiasing algorithm.

c. Analysis across adjacent radials (1D)

Automated vortex detection starts by searching for cyclonic azimuthal shear patterns in Doppler velocity

[2] WSR-88D base data are initially thresholded for low signal-to-noise ratio (6 dBZ for precipitation mode) prior to base data archiving.

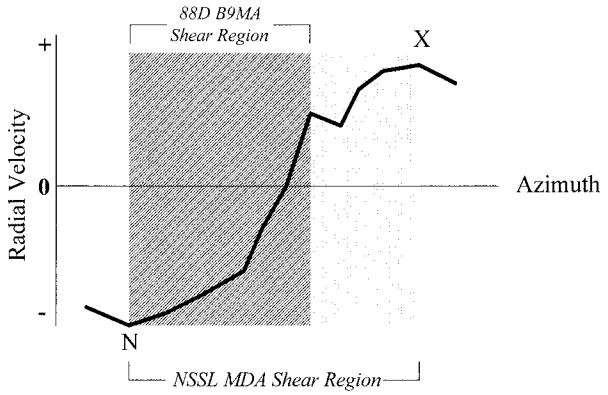

FIG. 1. An azimuthal shear segment represented on an azimuth versus velocity plot (range is constant). Darker-shaded area denotes the region of cyclonic shear (increasing velocity with azimuth) that would be detected using the 88D B9MA. The darker plus the lighter shaded areas denote the region of cyclonic shear that would be detected using the NSSL MDA. Here, X and N denote the maximum outbound and inbound velocities in the NSSL MDA shear segment, respectively.

data. Cyclonic shear (in the Northern Hemisphere) is characterized as increasing radial velocity with increasing azimuth (the WSR-88D collects data for each elevation scan in a clockwise manner). The procedure begins by comparing two adjacent velocity values at constant range from the radar. If cyclonic shear exists, an ''open'' or incomplete *shear segment* is created at that range. This procedure continues for the next two radials, and so on, until cyclonic shear is no longer detected at that range.

In the 88D B9MA, a shear segment is characterized by a run of velocities whose gate-to-gate azimuthal shear is continuously positive.[3] However, since the velocity field within a mesocyclone signature often contains small-scale perturbations, the profile of velocity as a function of azimuth through the signature will not always be characterized by continuously positive azimuthal shear. Thus, the 88D B9MA method would often result in the detection of only part of the total mesocyclone signature. To overcome this problem, a range-dependent look-ahead capability, similar to that used by Hermes et al. (1993), is employed in the NSSL MDA when building shear segments. Once cyclonic shear is no longer detected across a specified number of radials, then the shear segment is ''closed.'' Figure 1 illustrates a run of velocity values at a constant range and the resulting shear segments produced by the 88D B9MA and the NSSL MDA.

For each NSSL MDA shear segment, a number of attributes are computed. The velocity difference [or delta-V (Δv)] is the difference between the maximum inbound velocity and the maximum outbound velocity (the distance between these two velocities is the length of the shear segment). The shear (S) is the velocity difference divided by the length of the shear segment. The maximum gate-to-gate velocity difference (GTGVD) is the greatest velocity difference between two adjacent velocity values within the shear segment. Referring back to Fig. 1, note the perturbation in the velocity profile. The 88D B9MA would close its shear segment at this perturbation (the darker-shaded area), resulting in a shorter segment with a lower Δv and different shear calculation. The NSSL MDA, through its look-ahead capability, would extend the length of the shear segment through the lighter-shaded region.

The 88D B9MA and the Early MDA use one set of thresholds to discard shear segments [and the 88D B9MA uses momentum (length of shear segment multiplied by the velocity difference) thresholds]. Both algorithms are restrictive in that they attempt to apply predetermined strength criteria for mesocyclones at the initial analysis phase of the algorithm, that is, the shear segments. Unlike the NSSL MDA, neither algorithm attempts to detect weaker storm-scale vortices, which may also be significant. The lack of identification of a single shear segment, simply because it falls slightly below a threshold, may cause the algorithm to entirely miss a storm-scale vortex that may in fact meet the operational strength criteria used to detect the vortex.

The 88D B9MA also does not account for radar sampling limitations (Wood and Brown 1995) that affect the resolution of a vortex measured at increasing range from the radar. Wieler (1986) was the first to attempt a resolution-correction to the velocity data, followed by DD92. The NSSL MDA instead uses *range-dependent* velocity thresholds, rather than the resolution-correction method, which requires adjustments to the actual measured velocity data. At far ranges, a ''small and strong'' vortex could be similarly measured as a ''large and weak'' vortex, and a resolution correction would be inappropriate, especially if the vortex being measured is, in fact, a large and weak vortex. The range-dependent threshold method has limitations, as the measurement of a vortex still can be affected by the varying azimuthal separation between the vortex centroid and the radar beam center (Wood and Brown 1995).

In the Early MDA (Stumpf and Witt 1994), the range-dependent thresholds are set at the minimum values typically observed with mesocyclones, as shown in Table 1. These values were derived from the analysis of central Oklahoma mesocyclone data in the years following formal mesocyclone studies in the 1970s and 1980s (Burgess 1976; Burgess et al. 1979; Burgess et al. 1982; Zrnić et al. 1985), and preceding the development of

[3] ''Gate-to-gate'' defines azimuthally adjacent sample volumes at a constant range. Although radar sample volumes do not cover discrete spaces with finite boundaries, and they overlap with their neighboring sample volumes, the colloquial term gate-to-gate is used throughout this paper. This term is widely (although inappropriately) used within operational radar meteorology circles and is also used in various radar display software systems (including NSSL's display system, briefly described later in this paper).

TABLE 1. The range-dependent shear segment strength thresholds for the early MDA.

Range (km)	Velocity difference (Δv) (m s^{-1})	Shear (m s^{-1} km^{-1})
0–100	30	6
100–200*	30 $\rightarrow$ 22	6 $\rightarrow$ 3
>200	22	3

* The values of the thresholds decrease linearly between the ranges of 100 and 200 km.

TABLE 2. NSSL MDA strength rank thresholds. The values shown are used for shear segments within the range 0–100 km. The thresholds decrease linearly to 75% for Δv and 50% for shear between the range of 100 and 200 km. For ranges greater than 200 km, the thresholds remain constant at the decreased values valid for 200 km. For example, a shear segment at a range of 75 km from the radar, and whose Δv = 22.3 m s^{-1}, shear = 3.82 m s^{-1} km^{-1}, and GTGVD = 24.5 m s^{-1}, will be assigned a strength rank = 3. Only the first 9 of the 25 strength ranks are shown.

Strength rank	Velocity difference (Δv) (m s^{-1})	Shear (m s^{-1} km^{-1})
1*	10.0	3.00
2	15.0	3.75
3	20.0	4.50
4	25.0	5.25
5**	30.0	6.00
6	35.0	6.75
7	40.0	7.50
8	45.0	8.25
9	50.0	9.00

* Phillips Laboratory criteria (DD92).
** Early-MDA criteria.

the Early MDA. The Δv threshold is used for both the velocity difference and the maximum GTGVD. A shear segment is saved if the maximum GTGVD is above the Δv threshold for its given range, or if the velocity difference and shear are above their respective range-dependent thresholds.

Since a goal of the NSSL MDA is to detect all storm-scale vortices (including weaker vortices), the first step in the design of the NSSL MDA was to reduce the Early MDA shear segment strength thresholds (Table 1). By simply reducing these thresholds, the subsequent 2D and 3D analyses can fail [this will be shown in section 3c(2)]. For example, lowering the strength thresholds can cause the detection of long azimuthal shear regions associated with gust fronts (i.e., false vortex detection). Therefore, multiple thresholding of shear segments and feature core extraction [similar to what is used in the NSSL Storm Cell Identification and Tracking (SCIT) algorithm (Johnson et al. 1998)] are employed in the NSSL MDA. This new method of shear segment thresholding includes assigning a *strength rank* (a nondimensional number from 1 through 25) to each shear segment. The strength ranks are designed so that the original strength thresholds from the Early MDA correspond to a strength rank of 5 (rank 5). The lowest rank used to save shear segments (rank 1) corresponds to the thresholds used by the Phillips Laboratory MDA (DD92). Threshold values for the other ranks (2–4, 6–25) are determined using linear interpolation and extrapolation for the values used for rank 1 and rank 5. Only extremely weak shear segments whose attributes are below the rank 1 thresholds are discarded. Table 2 shows the new range-dependent thresholds and associated strength ranks. Values beyond 200 km are kept constant because if strength thresholds were reduced further, the velocities would be too close to background wind fields; there is not much velocity data beyond 200 km, and any velocity data in that range increment is not worth much for definitive vortex detection. The strength rank of a shear segment is taken to be the higher strength rank corresponding to either the 1) GTGDV, or 2) the smaller of the two strength ranks computed for both Δv and S.

Another limitation resulting from detecting shear segments with much lower strength thresholds is that some shear segments with large azimuthal extent are produced (and whose lengths are greater than the scale of the vortex being detected). Although velocity magnitude varies in a linear fashion across a Rankine vortex core, this is not always observed in nature (due to vortex asymmetries). Therefore, if the length of a shear segment exceeds a maximum segment-length threshold (typically 10 km), the strength thresholds are increased by one strength rank at a time and are applied to the run of velocity data inside the shear segment until a core segment less than the maximum segment-length threshold is detected. The shear segment attributes saved are based on the core segment. If a core segment cannot be found, the entire shear segment is discarded. The final output of the 1D analysis is a list of shear segments of varying strength ranks.

d. Analysis at the end of an elevation scan (2D)

Because the NSSL MDA produces a very large number of 1D shear segments, it is necessary that the horizontal 2D feature-association technique be robust enough to extract 2D vortex patterns without a large number of false detections. To accomplish this, feature core extraction [method similar to Johnson et al. (1998)] is employed. This method uses a multiple threshold technique to isolate a vortex core from broader regions of 2D azimuthal shear [which could be the Rankine ''potential vortex'' (Burgess 1976) or a broad region of azimuthal shear along a gust front or squall line]. By using overlap rules and a range versus azimuth aspect ratio threshold, 2D features should be fairly compact and round, emulating the 2D pattern associated with a Rankine vortex core.

Individual shear segments are combined into 2D *shear features* (Fig. 2) based on spatial proximity. Those segments that overlap in azimuth, and are separated in

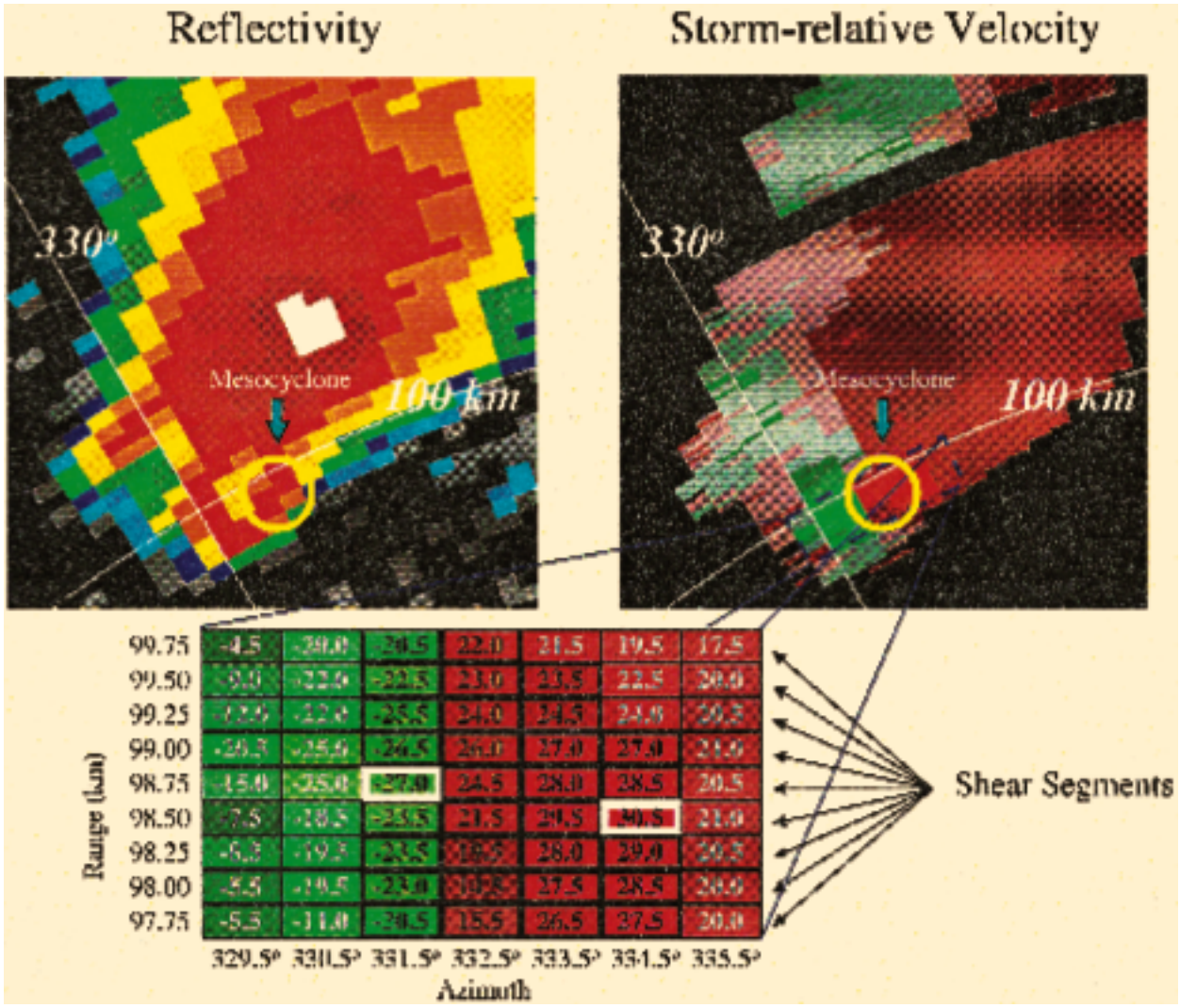

FIG. 2. An example of a 2D vortex signature in actual radar data (2107 UTC 7 May 1995, 1.5° elevation, Fort Worth, TX). Expanded insert depicts the radial velocity values for individual sample volumes in m s^{-1}. The shear segments are highlighted with black numbers and bold outlines. Other velocity values are depicted in gray. The maximum outbound velocity (30.5 m s^{-1}) and maximum inbound velocity (-27.0 m s^{-1}) of the resulting 2D feature are highlighted with white borders.

range by no more than 1 km, are combined into a common 2D feature. The multiple thresholds in Table 2 are applied to combine shear segments into 25 groups of 2D features corresponding to each of the 25 different strength ranks. The analysis begins by creating a group of 2D features using all available shear segments (rank 1 and greater). Next, a second group of 2D features is created using those shear segments whose rank is 2 or greater. Each additional group of 2D features is generated using successively larger strength ranks. Many shear segments are used more than once, and features from different groups of strength rank often overlap. Figure 3 shows an example of a 2D feature containing seven shear segments of ranks 3, 4, and 5. This particular 2D feature is initially saved as a rank 3 feature (all shear segments are rank 3 or greater).

If the total number of segments in any completed 2D feature is less than four, or if any feature extends less than 1 km in the radial direction (again, to force a 1-km minimum diameter), that feature is discarded. In comparison, the 88D B9MA discards any 2D features comprising less than 10 shear segments (the default value[4]). This prohibits the 88D B9MA from finding smaller vortices (e.g., those associated with minisupercells). Any 2D features higher than 12 km above radar level (ARL) (where the airflow in a thunderstorm is often highly divergent) are also discarded. Finally, in an at-

[4] Lowering the value of this parameter increases the probability of detection of minisupercells but raises the overall false alarm ratio for all storm types (Lee and White 1998).

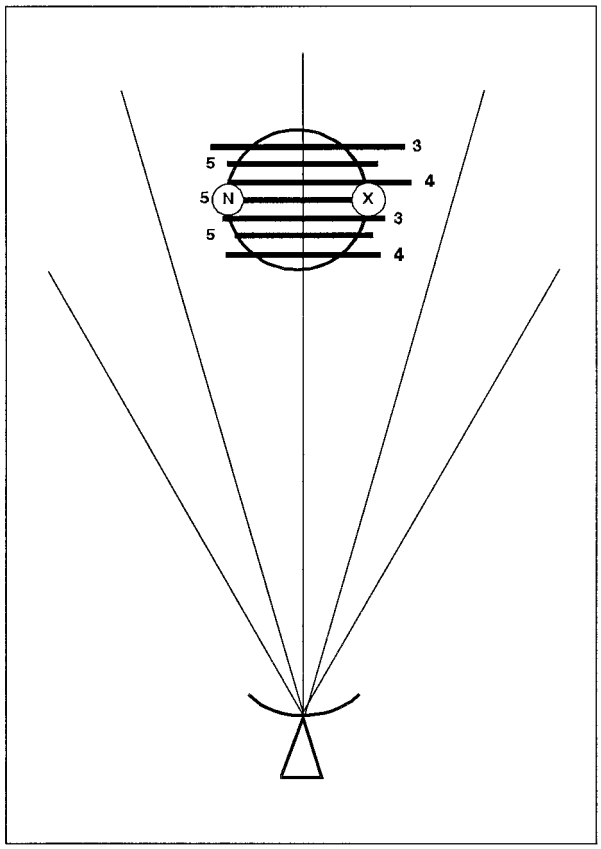

FIG. 3. A 2D feature, showing individual shear segments as horizontal line segments. The circled N and X show the locations of the maximum inbound and outbound velocities, respectively. The length of the line drawn between these two points equals the diameter (D) of the 2D feature. The large circle is the detected boundary of the vortex core. The radar is depicted at the bottom of the figure.

tempt to isolate roughly circular 2D cores of azimuthal shear, any 2D feature whose aspect ratio (radial distance divided by the azimuthal distance) is greater than two is discarded.

For each 2D feature, the following attributes are determined: 1) number of 1D shear segments, 2) centroid location, 3) centroid height ARL, 4) diameter (D), 5) rotational velocity (V_r), 6) shear (S), 7) maximum GTGVD for all shear segments, 8) aspect ratio, 9) 2D strength rank (which is recomputed later), 10) minimum azimuth, 11) maximum azimuth, 12) minimum range, 13) maximum range, and 14) elevation angle.

The rotational velocity is calculated using the maximum outbound velocity (V_{max}) and the maximum inbound velocity (V_{min}) determined from all of the shear segments in the 2D feature, with this equation:

$$V_r = \frac{V_{max} - V_{min}}{2}. \quad (1)$$

The center azimuth and range are determined as the midpoint of a line connecting the locations of V_{max} and V_{min}. The diameter (D) is the distance between V_{max} and

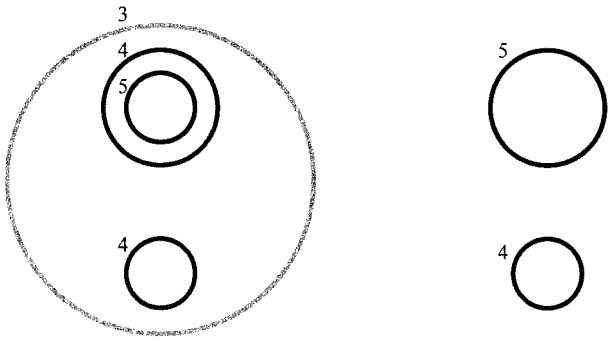

FIG. 4. Overlapping 2D features and feature cores on the left. On the right are the resulting two core regions detected. The upper-core region is reassigned rank 5 (using the maximum inbound and outbound velocities in the original rank 5 feature), though it contains rank 4 shear segments.

V_{min}. The shear is calculated as $S = V_r/D$. The GTGVD is the largest GTGVD for all the shear segments in the 2D feature. These values are used to recompute a 2D strength rank for the feature.

Radial convergence values are also computed for each 2D feature. This is done by averaging the radial velocity difference for all radial convergence shear segments detected within a 2-km buffer zone outside the diameter of the 2D feature. Hermes et al. (1993) describes the method used in the NSSL MDA to produce radial convergence shear segments.

To extract the azimuthal shear cores, each group of 2D features (corresponding to the 25 strength ranks) are checked for overlap and thresholded for aspect ratio. Analysis begins by comparing the 2D features with the highest strength rank with 2D features of the next lower rank. If any shear segments are shared between two overlapping 2D features, or if one feature centroid is within the minimum and maximum range and azimuthal extent of the other feature (and vice versa), then the higher-ranked feature is discarded and replaced with the lower-ranked feature (to retain all shear segments within the signature of the vortex core). However, if a lower-ranking feature overlaps more than one higher-ranking feature, then the lower-ranking feature is discarded. This prevents the inclusion of broad azimuthal shear regions and isolates the vortex cores. If a feature does not overlap any feature from a higher rank, it is saved. Figure 4 shows an example of overlapping 2D features and feature cores.

Once all 2D vortex core features are established, the strength rank of each 2D feature is recomputed (Table 2) by comparing the new maximum inbound and outbound velocities for the feature, and computing new values of rotational velocity, shear, and GTGVD. Typically, the strength attributes of the 2D feature will be the same as the highest-rank 2D feature that overlapped the final core feature. The feature core extraction technique allows for 2D feature creation using shear segments that may fall slightly below strength thresholds

and not meet the minimum number of shear segments per 2D feature criteria without the addition of those weaker shear segments.

e. Analysis at the end of a volume scan (3D)

After the analysis of all the elevation scans in the current volume scan is completed, vertical association is attempted for 2D features. Since the NSSL MDA detects a large number of 2D features, some of which are in close proximity to one another, a robust method (compared to the 88D B9MA and the Early MDA) of vertical association is used by the algorithm.

The process of vertical association is an iterative one, involving detections at two adjacent elevation scans. Vertical association results in a group of *3D features*, each of which consists of two or more 2D features at adjacent elevation angles. Because many more weaker 2D features are being detected, it is unnecessary to allow for vertical gaps in the 3D features as the 88D B9MA does. For each 2D feature from the higher of the two elevation scans, a set of candidate 2D features from the lower elevation scan is prepared. A first attempt is made to associate features whose centers are horizontally within 2 km of one another. When more than one association is found, this set of candidates will be sorted by strength rank. The search radius is then increased by 1 km, and the process is repeated for all the unassociated features. If more 2D features are associated with the feature from the higher elevation scan, they are also added to the bottom of the sorted set of candidates. This procedure continues until the search radius reaches a maximum of 8 km.

The next step in the analysis is to check to see if any 2D features from the higher elevation scan share the same best candidate 2D feature (the feature at the top of the sorted set of candidates) from the lower elevation scan. If so, then the best candidate becomes associated with the upper 2D feature whose centroid is closest in horizontal distance. Any other features from the higher elevation scan that shared that best candidate now use the second-best candidate in the list. This checking process continues until each feature from the higher elevation scan is uniquely associated with a feature from the lower elevation scan. If any feature from the higher elevation scan cannot be vertically associated with another feature, or if it loses all candidates to other features from the higher elevation scan, it remains unassociated. This process is repeated for the remaining pairs of adjacent elevation scans. Figure 5 illustrates this increasing radius search process.

Because the radar beamheight increases with range, it is possible that a distant vortex could be inadequately sampled within the volume scan (i.e., part of the 3D vortex may exist below the radar horizon). Thus, any 2D features from the lowest elevation scan that have not been vertically associated, and are located greater than 175 km from the radar, are tested against a set of

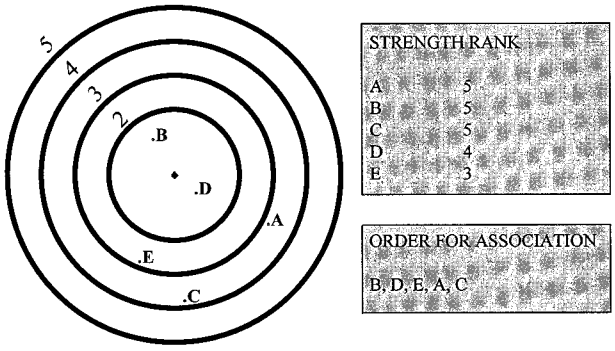

FIG. 5. The vertical association ''increasing search radius'' process. The diamond represents the centroid of the upper-elevation feature. The lettered dots (A–E) represent lower-elevation feature centroids. The top table on the right indicates the strength rank of each lower-elevation feature. The bottom table shows the order of the candidate lower-elevation features. This illustration also shows the method for time association, where the diamond represents a first-guess forecast position of a 3D feature from the previous volume scan, and the lettered dots as 3D features from the current volume scan. Only the first 5 km of all search radii are shown.

long-range strength thresholds (shear, rotational velocity, and gate-to-gate velocity difference). If the single 2D feature passes these thresholds, it is saved as a 3D feature. For the purposes of classification, these are called ''long-range 2D features.'' No similar procedure exists in the 88D B9MA.

For classification and diagnosis purposes, the following attributes are stored for each 3D feature: 1) center location using an average of all 2D centroids below 6 km; 2) center location of the 2D feature at the lowest elevation scan that it is detected; 3) two diameters (one from lowest-altitude 2D feature and the maximum diameter of all 2D features below 12 km); 4) rotational velocity (lowest-altitude and maximum); 5) shear (lowest-altitude and maximum); 6) GTGVD; (lowest altitude and maximum); 7) the heights of the maximum diameter, rotational velocity, shear, and GTGVD; 8) base (height ARL of lowest-altitude 2D feature used); 9) top (height ARL of highest-altitude 2D feature used); 10) depth (calculated by adding the half-power beamwidth to both the top and the base of the 3D feature); and 11) storm-relative depth [the percentage of the depth of the storm cell, as provided by either the NSSL SCIT algorithm (Johnson et al. 1998) or sounding data]. Low-altitude (0–2 km ARL) and midaltitude (2–4 km ARL) convergence are also calculated for each 3D feature. These are calculated by averaging the radial convergence shear segment velocity differences for all the NSSL MDA 2D features that are within these two respective layers.

f. Time association and tracking

When one or more 3D features are detected on two consecutive volume scans (that are less than 15 min apart), an attempt is made to associate the features in time. The 88D B9MA does not contain a time-associ-

TABLE 3. Example of 3D feature output: a) summary of the 3D detection information, and b) information about the 2D features that composed the 3D detection. In (b), the vertical core (rank 5 in this example) is boldfaced.

(a) 3D detection summary

Centroid azimuth (°)	Centroid range (km)	Base (km)	Top (km)	Total depth (km)	Core depth (km)	Relative depth (%)	Vortex type	MSI
106.4	115.1	1.7	9.2	9.6	3.9	39.2	MESO	3082

(b) 2D feature information

Elevation angle (°)	Centroid azimuth (°)	Centroid range (km)	Height (km)	Diameter (km)	V_r (m s^{-1})	S (m s^{-1} km^{-1})	GTGVD (m s^{-1})	Strength rank
0.4	106.4	115.1	1.7	5.9	13.5	4.6	23.5	4
1.4	**105.0**	**109.4**	**3.5**	**2.0**	**18.8**	**18.7**	**17.5**	**5**
2.3	**103.0**	**103.0**	**5.3**	**3.8**	**18.5**	**9.6**	**20.5**	**6**
3.3	102.0	102.4	7.3	4.1	12.2	6.0	20.0	3
4.2	101.2	101.8	9.2	3.9	10.8	5.5	16.5	3

V_r = rotational velocity.
S = shear.
GTGVD = Gate-to-gate velocity difference.

ation routine. The first step in the time-association procedure is to generate a first guess of the expected centroid location for the previously detected 3D features using either the previous motion vector of the MDA detection or a default motion vector (if the feature is first detected in the previous volume scan). The default motion vector is either an average of all the motion vectors from time-associated MDA detections in the previous volume scan, or an operator-defined value that is used when no MDA detections are present on the previous volume scan. This operator-defined value is typically taken to be the expected mesocyclone motion (the 0–6-km mean storm motion with a correction applied for the deviant motion of right-turning supercells), calculated from a near-storm sounding.

The time-association technique used is similar to the expanding search radius technique used for 3D vertical association (Fig 5). The first search radius begins at 2 km. The search radius then expands by 1-km annuli, until the search radius reaches the maximum distance threshold allowed. This maximum distance threshold is determined by computing the distance a vortex could travel from the previous volume scan time using a maximum speed threshold (33 m s^{-1}). For a 5- or 6-min. volume scan update rate, this distance is about 10 km. From the first-guess position of each previous feature, a set of candidate 3D detections from the current volume scan is created in the same fashion as the 3D vertical association (i.e., sorted by distance and strength rank). The best candidates for each first-guess position are chosen in the same manner as in 3D vertical association.

The last step involves calculating a new motion vector for each time-associated 3D feature. This is simply an average of the previous motion vector (if there is one) and the motion determined from the current and previous center locations. Using this method gives greatest weight to the most recent vortex locations and can account for the right-turning motion sometimes common to supercell thunderstorms. The motion vector is saved and used to forecast the future positions (every 5 min., up to 30 min. ahead of the current volume scan time) of the 3D feature. Up to 10 past positions (if available) for each detection are saved. The age of the circulation is also saved as a detection attribute for classification and diagnosis purposes.

g. Vortex diagnostic measurements

1) THREE-DIMENSIONAL (3D) STRENGTH RANK

Each 3D feature is assigned a 3D strength rank by finding the strongest continuous vertical core of 2D features whose 2D strength ranks are greater than or equal to a given strength rank. This core must be at least 3 km in half-beamwidth depth, and the base of the core must be below 5 km ARL. An example is shown in Table 3. In this example, although a rank 6 2D feature is part of the 3D feature, the 3D feature has a 3D strength rank of rank 5. This is because the rank 6 core is both less than 3 km deep and has a base above 5 km ARL. The base and depth of the core are saved as detection attributes.

2) MESOCYCLONE STRENGTH INDEX (MSI)

A vertically integrated vortex strength index, called the Mesocyclone Strength Index (MSI), is calculated. This is done by vertically integrating the strength ranks (multiplied by 1000) of all the 2D features used to create the 3D feature [the vertical integration is similar to those used by DD92 and Witt et al. (1998a)]. The strength rank for each 2D feature is weighted by the average air density (from the standard atmosphere profile used by DD92) across the half-power beamwidth depth at the

height of the 2D feature so that more weight is given to the 2D features at lower altitudes. Integration is done from the base (plus the half-beamwidth depth) of the 3D feature to the top (plus the half-beamwidth depth) of the 3D feature or 8 km, whichever is lower in altitude. The MSI is divided by the total depth of the 3D feature to normalize the MSI values for a variety of mesocyclone depths (from shallow to deep). MSI has an inherent range correction (via the strength ranks) to account for radar sampling limitations with increasing range.

3) NEURAL NETWORK PROBABILITY FUNCTION

A neural network (NN) is employed to diagnose the likelihood of tornadoes or severe winds with each vortex detection. The use of an NN in the NSSL MDA represents an important leap in the level of analysis for a severe storm algorithm and the first foray into the use of neural networks in any WSR-88D algorithm. Marzban and Stumpf (1996, 1998) provide a detailed description of the NN used in the NSSL MDA. Only a brief description is given here.

The NN is designed to be *trained* with a very large dataset containing a large number of vortex detections and their corresponding attributes, or ''input nodes.'' The software designed to train the NN is not contained within the NSSL MDA. This separate NN training program is designed to produce an NN function containing the NN weight function matrices. Instead, what the NSSL MDA software includes is the NN function that uses the vortex detection attributes to diagnose each vortex.

The NN has been trained to predict the existence of tornadoes and severe winds using reports contained in *Storm Data* (NCDC 1992–95). The NN can be trained on a variety of datasets, including datasets from particular near-storm environments, geographical regions, or climatological seasons. In this paper, we will describe the results of an NN trained on only one particular storm type (section 3a).

For training, a *tornadic vortex* is a vortex detection that occurs during a volume scan containing a reported tornado. Vortices are also considered tornadic if they are detected within 20 min. prior to the beginning time of the tornado report, or within one volume scan (5 or 6 min.) after the ending time of the report. All other vortex detections are considered *nontornadic*. *Severe wind* and *nonsevere wind* detections are also classified in this manner for training using reports of both high ($\geq$25 m s^{-1}) or damaging winds, and tornadoes. Training results in an NN consisting of a nonlinear function that uses a matrix of weight functions derived from the training dataset.

Twenty-three radial-velocity attributes of each 3D vortex detection are used as input into the NN (Marzban and Stumpf 1996). The NN is trained on these attributes to determine the probability that a vortex is associated with tornadoes or severe winds at the surface. The output of the NN are two numbers between 0 and 1 transformed into probability functions, the probability of tornado (POT) and probability of severe wind (POSW) [the latter is called probability of severe weather by Marzban and Stumpf (1996)].

h. Vortex classification

The next step in vortex analysis is 3D classification based on several rules defining the strength, spatial, and age characteristics of each detection. The definition of a mesocyclone used by the NSSL MDA is the one used in the early-MDA and described in Stumpf and Witt (1994): 1) 3D vortex with a continuous half-beamwidth depth of at least 3 km, 2) 3D vortex base at or below 5 km ARL, 3) all 2D features within the 3D vortex must meet the strength thresholds given in Table 1, and 4) the vortex must be detected during at least two consecutive volume scans (5 or 6 min.). The first three conditions are satisfied only if the 3D strength rank is greater than or equal to 5. Therefore, if a 3D feature has a 3D strength rank $\geq$5, and it can be time associated with a detection from the previous volume scan (of any 3D strength rank), then it is classified as a ''mesocyclone'' (MESO). If a 3D feature has a 3D strength rank $\geq$5, but it cannot be time associated with any detection from the previous volume scan, it is classified as a ''couplet'' (CPLT). If the 3D strength rank of any 3D feature is less than 5, then the vortex is classified as ''circulation'' (CIRC).

The architecture of the NSSL MDA allows for many different classification schemes. For example, ''low-topped mesocyclones'' (LOWTOP), which are associated with minisupercells (Burgess et al. 1995), are vortices that do not meet the criteria for mesocyclone because the vertical core depth is less than 3 km. Instead, the core depth of a low-topped mesocyclone is measured relative to the depth of the storm cell containing the vortex detection [using the storm-cell height as provided by either the NSSL SCIT algorithm (Johnson et al. 1998) or sounding data]. The core depth must be at least 25% of the storm depth, and the vortex detection must meet all other criteria for a mesocyclone. This classification was developed so that the warning forecaster could recognize and anticipate the sampling limitations in environments characterizing low-topped convection (e.g., at far ranges from the radar, the mesocyclones may be unobservable).

A ''shallow vortex'' (SHALLO), which is sometimes associated with strong gust fronts/squall lines or a boundary layer vortex (Stumpf and Burgess 1993), is classified as such if the top of the vertical feature core is below 3 km ARL, the base is at the 0.5° elevation scan, the core depth is at least 1 km but less than 25% of the storm depth, and the 3D strength rank of this shallow core is at least 5.

Unlike the 88D B9MA, the NSSL MDA does not

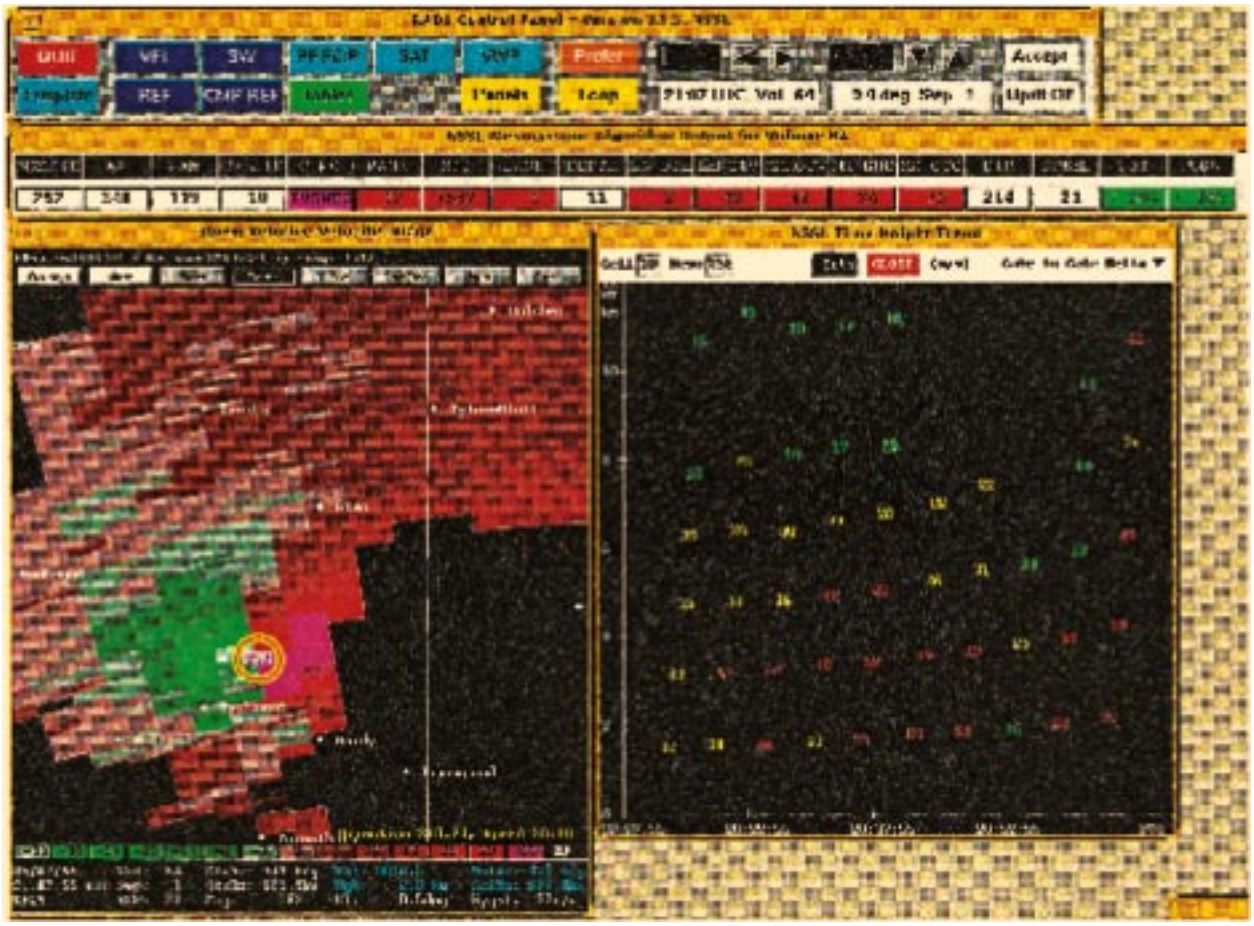

FIG. 6. NSSL MDA output in RADS (Sanger et al. 1995). The top box is the RADS control panel. The second-to-the-top box is the mesocyclone attribute table. The lower left-hand box is the radial velocity (m s^{-1}) display, zoomed and centered on a supercell storm (yellow and red circle labeled 252) from 2107 UTC 7 May 1995, 0.5° elevation, Fort Worth, TX. The lower right-hand box displays time–height trends of GTGVD for the mesocyclone detection 252 for the last hour ending 2107 UTC.

diagnose the detections for tornadic vortex signatures (TVSs). Instead, a separate algorithm, the NSSL Tornado Detection Algorithm (NSSL TDA) (Mitchell et al. 1998) is used to detect the gate-to-gate intense shears associated with the tornado vortex. Thus a NSSL TDA detection does not require the presence of a mesocyclone detection. If any TVS detections are produced by the NSSL TDA, then a proximity check is performed between the NSSL TDA and NSSL MDA detections. If any TVS is found within the diameter (plus 2 km) of a mesocyclone, that mesocyclone is classified as a ''TVS–mesocyclone'' (TVSMES).

Since scientists have yet accept a single mesocyclone definition, the NSSL, MDA can allow for many definitions. For example, vortices can be classified using any number of mesocyclone definitions, including those prescribed by Phillips Laboratory (that uses resolution-corrected shear; (DD92) and the NWS (OFC 1991). With the statistical analysis of NSSL MDA output (including the NN), a more refined mesocyclone definition (that could include more than just Doppler velocity attributes) might be attainable.

i. Graphical product output

The NSSL MDA outputs a variety of information that is displayable using the NSSL. Radar and Algorithm Display System (RADS). Figure 6 shows an example of RADS output of the NSSL MDA. The reader is asked to refer to Sanger et al. (1995) for more details about RADS.

1) MDA OVERLAYS

Three types of overlay icons are displayed. The first is a thick yellow circle that represents vortices classified as mesocyclones. TVS mesocyclones, and low-topped mesocyclones (all have 3D strength ranks ≥ 5). A thick yellow circle with an inlaid thin red circle is used to denote one of these types of detection whose base is

TABLE 4. Vortex classification hierarchy (with most significant vortex type at the top of the table) and RADS mesocyclone table presentations.

Vortex classification	RADS table identifier	RADS table color
TVS–mesocyclone	TVSMES	Magenta
Low-topped mesocyclone	LOWTOP	Red
Mesocyclone	MESO	Red
Couplet	CPLT	Orange
Circulation	CIRC	White
Long-range 2D feature	2DFT	Cyan
Shallow vortex	SHALLO	Yellow

detected by all the lowest elevation scan (0.5°) in WSR-88D data, or whose base is below 1 km ARL. Finally, thin yellow circles are used to denote all other types of vortices. The plotted center of the meso icon is the center of the lowest-altitude 2D feature, and the diameter of the icon is proportional to the diameter of the actual vortex.

The NSSL MDA also provides tracking information (whereas the 88D B9MA does not). Past centroids (up to 10) and 30 min. of forecast information (at 5-min. intervals) are displayed.

2) MDA ATTRIBUTE TABLE

The attributes of the 15 strongest vortices from the present volume scan are presented in a mesocyclone table. The detections are sorted with the most significant detection at the top of the table. The order of the sort is based on vortex classification and various diagnostic strength parameters. This sorting allows the warning forecaster to identify the most significant vortices first. Also, the information is color coded within the table based on either severity (using a progressively more significant ''green → yellow → red'' scheme) or color coded by vortex classification (with ''hot'' colors denoting more significant vortices).

The initial sort is based on vortex classification. Table 4 shows the order of these classifications and the respective representations in the mesocyclone table. Secondary and tertiary sorts are based on 3D strength rank and MSI.

3) TRENDS

An array that contains the attributes for each NSSL MDA vortex detection from the current and nine previous volume scans (roughly 1 h in the WSR-88D precipitation mode) is created for the display of time series trends in RADS. These attributes include

- base, top, and depth (km);
- low-altitude diameter (km), rotational velocity (m s^{-1}), shear (m s^{-1} km^{-1}), and gate-to-gate velocity difference (m s^{-1}) (these values are taken from the 2D feature representing the lowest elevation angle of the 3D detection);
- maximum diameter, rotational velocity, shear, and gate-to-gate velocity difference (in some instances, these values will be the same as the low-altitude values), and the height ARL of these values; and,
- 3D strength rank, MSI, POT, and POSW.

Some of these attributes are also displayable in RADS as *trend sets* of five individual trends.

Time–height trend information is also created (see Fig. 6). The attributes shown in time–height trends include diameter, rotational velocity, shear, and maximum GTGVD. Up to 1 h of time–height trend information is stored, along with the elevation scan times and heights for each 2D feature used to plot the time–height trend. These are useful to analyze the vertical evolution of the vortex attributes over time.

3. Performance evaluation

a. The WSR-88D data used for testing

The performance evaluation consists of a comparative evaluation of both the NSSL MDA and the 88D B9MA, and an independent evaluation of the NSSL MDA. The various classification techniques and diagnostic parameters of each algorithm are evaluated using these data. The data processed by both algorithms are composed of WSR-88D level-II archive data (Crum et al. 1993). Preprocessing of the WSR-88D data is performed as described in section 2a. Although these datasets represent the largest mesocyclone dataset ever used for mesocyclone algorithm testing as reported in the literature, it still only represents a fraction of the data that could be available for testing.

Three groups of test data are used in the algorithm performance evaluation. The first group of test data consists of six cases in the *algorithm comparison dataset* (Table 5), and is used for the comparison scoring between the 88D B9MA and the NSSL MDA. This dataset consists of data from a variety of storm types (e.g., isolated supercells, low-topped minisupercells, squall lines) from various regions (e.g., south-central United States, northeast United States, Florida). The second dataset, the *NSSL MDA scoring dataset* (Table 6), is used for continued evaluation of just the NSSL MDA and the NN (but not the 88D B9MA). This dataset is also geographically and seasonally diverse, and contains a variety of storm types. The third group of data consists primarily of isolated tornadic supercell cases from southern United States WSR-88D recording sites. This dependent dataset was used to train the NSSL MDA NN (Marzban and Stumpf 1996) and is called the *NN training set* (Table 7). This dataset is not used for algorithm comparison purposes, nor is the NN scored on these data.

The NN requires a large dataset to be trained properly. At the time of the development of the NN used in this paper, the most prevalent mesocyclone storm type in the WSR-88D level-II database was the large isolated su-

TABLE 5. Algorithm comparison dataset.

Date	Radar location	Number of tornado reports	Number of severe wind reports	Number of volume scans	Comments
25 Mar 1992	Melbourne, FL	0	4	42	Nontornadic supercells
5 May 1993	Dodge City, KS	12	0	42	Isolated cyclic supercells
15 Apr 1994	St. Louis, MO	8	19	51	Tornadoes along leading edge of squall line
30 Apr 1994	Sterling, VA	3	0	70	Low-topped minisupercells
19 Apr 1995	Fort Worth, TX	18	30	129	Isolated supercells (near range); squall line
16 May 1995	Dodge City, KS	5	4	81	Isolated supercells
Total		46	57	415	

percell-type storm common to the U.S. southern plains. To avoid ''diluting'' the NN with other types of storms, the NN evaluated in this paper was developed to diagnose just isolated-supercell types of storms. The NN is highly flexible and can be trained on *any number and variety* of data (e.g., an NN can be trained to detect only tornadoes in low-topped minisupercells). Because of the nature of the NN training set used in our study, this particular NN performs better on the isolated supercell cases.

b. Verification (ground truth)

The algorithms are scored for their ability to correctly detect and classify operationally significant storm-scale vortices. The algorithms are evaluated versus a ground verification database consisting of NWS severe weather reports from *Storm Data* (NCDC 1992–95). This type of analysis avoids ambiguities associated with the subjective nature of determining vortex signatures from radar data and relates mesocyclones to criteria that the NWS uses to verify severe storm warnings.

The source of ground verification data used in this study has a number of shortcomings (Witt et al. 1998b). Algorithm scoring requirements are quite demanding, as the radar data involve small time and space scales, and more detail is needed than *Storm Data* usually provides.[5] Prior to use in algorithm scoring, the tornado and wind reports had to be validated. For those reports deemed to be in error, the report times and/or report locations were modified to make them consistent with the radar data. Just over half of the reports required at least a 5-min and/or at least a 5-km adjustment to the times and/or locations, respectively.[6] Reports are deleted when there is no supporting radar evidence to validate the reports (e.g., no storm echoes were present at the time and location of the report). Also, some reports are modified using several higher-confidence verification sources, such as NSSL field project observations (Rasmussen et al. 1994), NSSL damage survey information, and some of the lead author's personal storm intercept logs. We feel it is important to mention that the modification of the *Storm Data* reports are done *prior* to any knowledge of algorithm detection output. Ground truth modifications are established using base radar data alone.

Operationally significant vortices are defined as those

[5] *Storm Data* contains severe weather reports that have primarily been used to verify NWS severe weather warnings. These reports are often on timescales of 30–60 min., and space scales of partial or entire counties (10–60 km or more).

[6] Complete statistics on the modification of the tornado ground truth dataset are available in Witt et al. (1998b).

TABLE 6. Additional cases used for NSSL MDA scoring.

Date	Radar location	Number of tornado reports	Number of severe wind reports	Number of volume scans	Comments
27 May 1995	Des Moines, IA	6	2	58	Isolated supercells
12 May 1995	Goodland, KS	5	2	67	HP supercells
22 Jun 1995	Pueblo, CO	1	0	68	High plains minsupercell
13 Feb 1995	Phoenix, AZ	1	0	41	''Very'' minisupercell
21 Jul 1995	Minneapolis, MN	12	1	56	Isolated supercells
28 May 1994	Amarillo, TX	1	0	65	LP supercell
16 Nov 1993	Houston, TX	9	2	81	Close proximity; some minisupercells
15 Nov 1994	Melbourne, FL	5	0	123	Tropical cyclone minisupercells
22 May 1995	Amarillo, TX	8	0	51	Isolated supercells
5 May 1995	Fort Worth, TX	0	20	71	Hail and flash flood fatalities; supercell
Total		48	27	681	

TABLE 7. NN training set (all isolated supercell types).

Date	Radar	Number of tornado reports	Number of severe wind reports	Number of volume scans
11 May 1992	Norman, OK	22	4	78
7 May 1993	Dodge City, KS	9	1	44
6 May 1994	Tulsa, OK	3	14	100
27 Nov 1994	Memphis, TN	7	2	77
7 May 1995	Fort Worth, TX	12	20	120
2 Jun 1995	Lubbock, TX	9	4	107
Total		62	45	526

that produced a report of surface wind damage—either damage reported as a tornado, or damage reported from straight-line winds. Any ground truth tornado reports that are listed as ''short-and-narrow, no damage reported'' are deleted, provided there is not a significant storm-scale vortex visually observed in the radar data.[7] Any damaging wind reports that are not associated with a visually observed radar vortex (i.e., pure gust front or pure microburst events) are also deleted. Hail data are not used, as there are other reflectivity based parameters [and a hail detection algorithm (Witt et al. 1998a)] that should provide a more reliable indication of hail within a storm (hail production does not require the presence of a vortex within a storm). For the purposes of algorithm scoring (detailed in the next section), an observed event is either a tornado report or a report of severe wind (depending on what type of event is being scored). The term nonevent is used to describe the lack of any ground truth report.[8]

For each ground truth report, any volume scans encompassing the time of the report are identified. For each ground truth report, a ''time window'' (Witt et al. 1998b) encompassing that report is also created. The time window includes any volume scans occurring up to 20 min (three or four volume scans) prior to the beginning time of the event, and one volume scan (5 or 6 min) after the ending time of the event. The leading 20 min of the time window is chosen to allow the algorithm to diagnose the detections with sufficient lead time. This value is based on JDOP findings (Burgess et al. 1979). The trailing 5 or 6 min of the time window is to account for any small reporting errors in the ending time of the events. Certain storms produce multiple vortex events with overlapping time windows.

Once the time windows are established, all algorithm vortex detections are manually associated with the modified ground truth events. If there are any storm-scale vortices detected that coincide in space and time with an event or the event time window, then that vortex detection is considered significant. This is done to quantify whether a given ''snapshot'' of a vortex is currently (or will be in 20 min.) associated with a ground truth event. This association is a rigorous and lengthy process. Occasionally, the associations are very straightforward. In some instances, there are storms with multiple detections, and determining which vortex detection should be associated with the ground truth report is challenging. Typically, when there is more than one detection to choose from, the strongest and deepest detections are chosen for the association. In the cases where there are two or more vortex detections with roughly comparable strengths and depths, each vortex detection is tagged as significant. Vortex detections associated only with tornadoes are considered tornadic (the others are nontornadic). Those vortex detections associated with damaging winds (including tornadoes) are considered severe wind vortex detections (the others are nonsevere-wind vortex detections).

Ground truth events that are not associated with any algorithm detections (algorithm misses) are also tallied into three categories: 1) a vortex is visually observed in the radar data, but the algorithm fails to automatically detect it; 2) a vortex is *not* visually observed in the radar data because of sampling limitations (i.e., there is no signature in the data that the algorithm, or human, can detect); and 3) a potential vortex may exist in the radar data, but the data are obscured by radar artifacts (e.g., range or velocity aliasing).

c. Scoring

1) SCORING METHODOLOGY

All six storm cases listed in Table 5 are used to score the mesocyclone classification schemes of the NSSL MDA and 88D B9MA, and several of the diagnostic parameters (Table 8). The classification schemes are scored as a categorical decision of a ''no'' prediction or a ''yes'' prediction. Therefore, all detections that are classified as a TVMES (or a LOWTOP in the case of the NSSL MDA) are considered yes predictions, and all other detections are no predictions.

To score any diagnostic parameter (e.g., MSI), a threshold is varied from the minimum extreme of all cases to the maximum extreme. This threshold is used to discriminate between no predictions and yes predictions. This style of scoring is used to determine an appropriate threshold at which a particular diagnostic

[7] We define a ''visually observed radar vortex'' as a couplet of opposing radial velocities approximately along a constant range (but not exactly, owing to convergent, divergent, and elliptical vortices) and showing some azimuthal shear. The area of azimuthal shear should not be elongated (instead it should be roughly compact). No strength thresholds are imposed, as a storm-scale vortex could theoretically be of any strength, and radar sampling limitations may underestimate the true strength of a vortex.

[8] The absence of a ground truth report in *Storm Data* is assumed to mean that no severe wind or tornado occurred. Obviously, this is not always the case as some events go unreported. Without a sufficient database of actual reports of nonevents (e.g., ''The observed storm did not produce a tornado at that time''), this assumption must be made.

TABLE 8. Various vortex classification schemes and diagnostic measurements tested for both the NSSL MDA and 88D B9MA. The codes in the first column are used in Tables 14–16.

Code	Algorithm	Type
MESO	NSSL MDA	NSSL MESO and LOTOP classification
RANK	NSSL MDA	3D strength rank
MSI	NSSL MDA	Mesocyclone strength index
POT	NSSL MDA	NN probability of tornado
POSW	NSSL MDA	NN probability of severe wind
MES88	88D B9MA	WSR-88D MESO classification
IRS	88D B9MA	Integrated rotational strength index (Lee and White 1998)

quantity can be maximized for predictive skill. For example, when using MSI as the diagnostic measure to score, a threshold is selected such that all vortices with an MSI greater than or equal to that threshold are considered yes predictions, and all other detections are deemed no predictions. The threshold is then varied to determine the best skill score. Figure 7 shows the variation in skill scores (explained in following paragraphs) with varying MSI thresholds.

Figure 8 shows an example of a possible detection and ground truth scenario from a single volume scan. In this example, there are three tornado events in the ground truth (a, c, and e), and there are four algorithm detections (a, b, c, and d). Two of the four algorithm detections are yes predictions (a and b) and the other two are no predictions (c and d). One yes prediction is associated with an event a, and one no prediction is associated with an event c. This leads to (as labeled in Fig. 8) (a) correct yes prediction (hit), (b) incorrect yes prediction (false alarm), (c) incorrect no prediction (miss), or (d) correct no prediction. To score algorithm performance, the first four categories (a–d) are tallied and entered into a contingency table (Table 9) (Wilks 1995).

The event labeled as (e) on Fig. 8 is an event not associated with any algorithm detection. This type of event is counted as a miss (c) only when a vortex can be visually observed in the radar data. If the algorithm did not detect a vortex either because one is not visually observed in the radar data, or a vortex is obscured by data artifacts (range or velocity aliasing), then no penalty is assessed and the event is not counted as a miss.

Four scoring parameters (from Wilks 1995) can be computed from the contingency table:

$$\text{Probability of detection (POD)} = a/(a + c) \quad (2)$$

$$\text{False alarm rate (FAR)} = b/(a + b) \quad (3)$$

$$\text{Critical success index} = a/(a + b + c) \quad (4)$$

$$\text{Heidke skill score (HSS)} = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}. \quad (5)$$

Besides Eqs. (2)–(5), there are a variety of other scoring parameters that could be computed (Marzban 1998). Since Marzban (1998) has shown that HSS is one of the better scoring measures to use, and since CSI is a traditional scoring parameter used in weather forecasting studies, the algorithms are evaluated using only CSI and HSS. Note that CSI does not take into account the correct no predictions (d) in the contingency table.

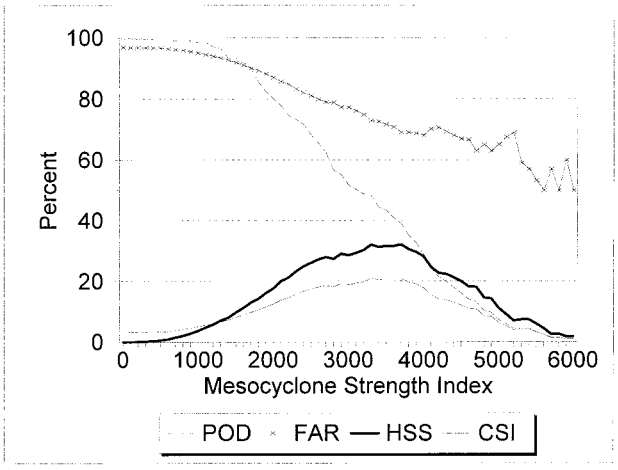

FIG. 7. Plot of skill scores (POD, FAR, HSS, CSI), for scoring tornado (only) events versus various mesocyclone strength index thresholds for no and yes predictions for the algorithm comparison dataset (Table 5).

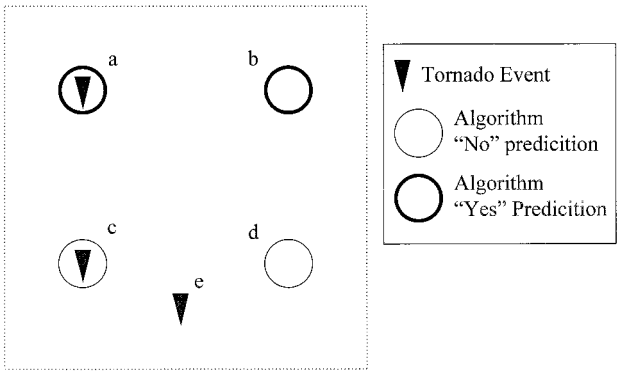

FIG. 8. Detections (including yes and no predictions) and ground truth (the triangles depict tornadoes) for a hypothetical volume scan. Labels a–e are described in the text of the paper.

TABLE 9. Contingency table used to score MDAs (after Wilks 1995).

Algorithm prediction	Observed event	
	Yes	No
Yes	a (hit)	b (false alarm)
No	c (miss)	d (correct no prediction)

TABLE 10. 88D B9MA Adaptable parameter values.

	Default	Optimized	Low threshold
High momentum ($km^2\ h^{-1}$)	540	540	180
High shear (h^{-1})	14.4	7.2	7.2
Low momentum ($km^2\ h^{-1}$)	180	180	90
Low shear (h^{-1})	7.2	7.2	3.6
Number of shear segments	10	6	4

2) COMPARISON OF SCORES BETWEEN 88D B9MA AND NSSL MDA USING SIX CASES

The NSSL MDA is compared to the 88D B9MA using three different sets of parameters for the 88D B9MA. These three sets include 1) default parameters that are those currently fielded in the NWS 88D system, 2) optimized parameters determined through independent testing by the OSF Applications Branch to correctly identify a wider variety of vortex types (Lee and White 1998), and 3) low-threshold parameters designed to emulate the lowest threshold parameters used by NSSL MDA to produce rank 1 detections. The values used for the 88D default, optimized, and low-threshold parameters are listed in Table 10. Table 11 lists information about the vortex detections for each of the algorithm and adaptable parameter configurations, including the percentages of significant vortices and the total number of detections produced.

For the NSSL MDA, one classification scheme (MESO in Table 8, which includes both MESO and LOWTOP) and four diagnostic measures (3D strength rank, MSI, POT, and POSW) are tested for sensitivity. For the 88D B9MA, one classification scheme and one diagnostic measure are tested for sensitivity. The 88D B9MA classification tested is the WSR-88D definition of a mesocyclone (OFC 1991) (MES88 in Table 8). The integrated rotational strength (Lee and White 1998) (IRS in Table 8), an experimental vertically integrated quantity similar to MSI, is the 88D B9MA diagnostic measure tested. The IRS uses different strength thresholds than the MSI and does not normalize by the depth of integration.

Using the six-case comparison dataset (Table 5), the POD of *any* event is listed in Table 12. The NSSL MDA vortex-signature detection techniques are superior to those of the 88D B9MA. The NSSL MDA performs significantly better than the 88D B9MA (for all three sets of thresholds) in *locating* significant vortices. Also, the NSSL MDA detected *every* vortex signature associated with a reported F2 or greater intensity tornado in this dataset (POD = 100%).[9] The goal of designing an algorithm that detects almost every significant storm-scale vortex signature is mostly achieved. The remaining challenge is to provide significant skill in properly diagnosing these detections (i.e., discriminating between tornadic and nontornadic vortices, or severe wind and nonsevere-wind vortices).

Table 13 shows the tally of algorithm-missed detections (as opposed to misdiagnosed no predictions). For the cases when a vortex signature is visible in the radar data, the 88D B9MA has a significantly higher number of missed detections. The NSSL MDA missed only one vortex associated with any tornado event, and the NSSL MDA detected all vortices associated with F2+ tornado events (Table 12). Missed detections due to radar data artifacts are less common, but are still noteworthy and suggest the need for continued improvements in the radar data preprocessing schemes. A number of events go undetected simply because a vortex signature is not evident in the radar data. Warning meteorologists should be aware that not all tornadic vortices are large or strong enough to be well sampled by the WSR-88D.

Table 14 summarizes scoring for all four algorithm runs and for the various classification schemes and diagnostic parameters listed in Table 8. The MESO and MES88 rows denote the MESO classification schemes for both the NSSL MDA (including LOWTOP) and the three 88D B9MA runs, respectively. For the rest of the rows, the values of each parameter that produce the highest skill are shown. Maximum values of HSS are determined for each of the diagnostic measurements.

From Table 14, note that the addition of the IRS parameter to the 88D B9MA offers an improvement in skill [this is also shown in Lee and White (1998)]. Also note from Table 14 that the 88D B9MA, using the low thresholds designed to emulate NSSL MDA rank 1 thresholds, without an IRS parameter, scores very low (HSS = 5.3%). Simply reducing the 88D B9MA thresh-

[9] Although we have shown a few statistics for the class of tornadoes generally defined as ''strong and violent'' (F2 or greater), we hold the opinion that an evaluation of any tornado prediction versus tornado F ratings should be used with extreme caution. First, the F scale was designed to be a *damage* scale, and not a strength scale, and the F scale attempts to relate observable tornado damage with probable wind speeds. However, in areas where there are no structures or tall vegetation on the ground to receive damage (e.g., the treeless high plains of the United States), strong and violent tornadoes are frequently rated F0, since they inflict no damage. Second, the reported F scale of a tornado is the subjectively determined *maximum* damage intensity along the entire path of the tornado, and variations at and below this strength exist along the rest of the tornado path. Since we are scoring the algorithms using 5- and 6-min. radar snapshots (scoring on a volume-scan basis), we chose not to further investigate the scores for various F-scale classes of tornadoes.

TABLE 11. Number and type of vortices detected by each algorithm (and various site-adaptable parameter sets) with the algorithm comparison dataset (Table 5).

Algorithm	Parameters	Total number of vortices	Number of tornadic vortices (%)	Number of severe vortices (%)
NSSL MDA	Default	6548	204 (3.1)	328 (5.0)
88D B9MA	Default	707	97 (13.7)	147 (20.8)
88D B9MA	Optimized	2616	145 (5.5)	239 (9.1)
88D B9MA	Low threshold	10 208	152 (1.5)	269 (2.6)

olds results in many failures in the 2D and 3D processing routines causing overall poor algorithm performance.

From Table 14, the HSS for every NSSL MDA classification and diagnostic attribute outperforms each incarnation of the 88D B9MA, for both tornadoes alone, and for reports of severe wind and tornadoes combined. When scoring against only tornado events, the NSSL MDA NNs, POT, and POSW share the highest HSS (35.7%). When scoring wind and tornado events (severe wind reports), the POSW has the highest HSS (37.5%). *The HSS values for the NSSL MDA NN are significantly higher than the best scores for any 88D B9MA attribute.* The NN also offers an improvement over the other NSSL MDA attributes and the MESO (and LOWTOP) classification schemes for this dataset.

Finally, the algorithm comparison dataset is subdivided into two smaller datasets. One set comprises mainly isolated supercell cases, and the other set contains all other storm types. Each algorithm performs better on the isolated supercell storms (Table 15). The NSSL MDA NN probabilities perform the best for every case, except when scoring severe wind events associated with the isolated supercell cases. The algorithms perform about *two to three times better* for isolated supercell events than for the other storm types. More work will be needed to accurately diagnose these other supercell types.

3) SCORING NSSL MDA USING 10 ADDITIONAL CASES

The NSSL MDA is independently scored on the 10 cases in Table 6 (scores presented in Table 16). The MESO classification scheme performs the best on these 10 datasets. The MESO classification has better skill with detecting severe wind (tornadoes and straight-line wind) over just tornadoes alone.

While the scores for the MESO classification scheme outperformed the NN parameters (POT, POSW) for the 10-case dataset, it is important to remember that the NN was trained on specific data (isolated supercells). The NN had the best skill with most of the isolated supercell cases (27 May 1995 Des Moines IA; 12 May 1995 Goodland KS; 21 July 1995 Minneapolis MN; and 28 May 1994 Amarillo TX) with one exception (22 May 1995 Amarillo TX) (Table 17).

TABLE 12. Probability of detecting any event with the algorithm comparison dataset (Table 5). POD is defined by Eq. (2).

Algorithm	Parameters	Ground truth	POD (%)
NSSL MDA	Default	All tornadoes	99.5
88D B9MA	Default	All tornadoes	49.0
88D B9MA	Optimized	All tornadoes	73.6
88D B9MA	Low threshold	All tornadoes	76.4
NSSL MDA	Default	Tornadoes $\geq$ F2	100.0
88D B9MA	Default	Tornadoes $\geq$ F2	46.8
88D B9MA	Optimized	Tornadoes $\geq$ F2	67.8
88D B9MA	Low threshold	Tornadoes $\geq$ F2	86.7
NSSL MDA	Default	Severe wind	94.0
88D B9MA	Default	Severe wind	55.1
88D B9MA	Optimized	Severe wind	79.1
88D B9MA	Low threshold	Severe wind	89.1

4. Scoring summary

The NSSL MDA offers a significant improvement in skill over the 88D B9MA. The NSSL MDA also offers the detection of an entire spectrum of vortices, time association and tracking of vortices, time series trends and time–height trends of vortices throughout a large portion of their lifetime, and the flexibility to diagnose vortices using several classification schemes and diagnostic measures. The NSSL MDA (via RADS) also offers novel approaches in graphical diagnostic output. The scoring results from the information shown in Tables 12–17 is summarized as follows:

- The NSSL MDA detects almost every vortex signature associated with a severe weather event (the POD is nearly 100%). The 88D B9MA misses many of the operationally significant signatures.
- The 88D B9MA IRS parameter offers an improvement in the diagnostic skills of the 88D B9MA alone.
- Lowering the strength thresholds in the 88D B9MA (to match the NSSL MDA rank 1 thresholds) results in a decrease in skill.
- The NSSL MDA diagnoses the vortex detections with much greater skill than the 88D B9MA, even when the optimized thresholds are used along with the IRS parameter.
- Both algorithms perform best when diagnosing isolated supercell events versus other storm types.
- The NSSL MDA NN has the best skill of all the NSSL MDA diagnostic parameters when tested on isolated supercell cases.

TABLE 13. Summary of algorithm misses with the algorithm comparison dataset (Table 5).

Algorithm	Parameters	Ground truth	Reason for algorithm miss		
			Algorithm failure; vortex visually observed in radar data	Vortex not visually observed in radar data	Radar data artifact obscuration
NSSL MDA	Default	Tornado	1	6	4
88D B9MA	Default	Tornado	101	6	4
88D B9MA	Optimized	Tornado	52	6	4
88D B9MA	Low threshold	Tornado	47	6	4
NSSL MDA	Default	Severe wind	21	22	12
88D B9MA	Default	Severe wind	120	22	12
88D B9MA	Optimized	Severe wind	63	22	12
88D B9MA	Low threshold	Severe wind	33	22	12

5. Operational implications of the NSSL MDA and automated storm-scale vortex detection

The NSSL MDA is designed to detect *all* storm-scale vortices (that are observable in the radar data), and then diagnose them for their significance. This is a major change in the detection techniques of the automated vortex detection algorithms in use on the WSR-88D system. The 88D B9MA detects substantially fewer storm-scale vortices (especially those that do not meet the specific criteria defined in the 88D B9MA). Because of this, many vortices, including significant vortices, are not detected by the 88D B9MA. Also, the 88D B9MA currently offers no diagnostic output with each detection.[10] By contrast, the NSSL MDA provides diagnostic guidance for almost every operationally significant storm-scale vortex to make a more informed warning decision.

This change in automated vortex detection (and diagnosis) necessitates a change in operational paradigms of the meteorologists who use mesocyclone algorithm output for warning decisions (as well as broadcast meteorologists who have access to WSR-88D algorithm output made available by commercial radar data vendors). With the NSSL MDA, the warning forecasters will have at their disposal a much larger amount of information. The warning forecaster must understand and interpret the output correctly and efficiently. Operational users are recommended to consider the NSSL MDA as a warning *guidance* tool, complete with diagnostic tools, rather than an automatic decision maker.

The NSSL MDA output should be treated as an alert for the warning forecaster to do a detailed radar data interrogation of the detected vortex signatures because data artifacts, such as velocity dealiasing errors and range folding obscuration of rotation signatures, may compromise algorithm detections. The radial velocity data should always be examined along with the algorithm output to verify that the detection is legitimate and a vortex is present in the data.

The warning forecaster should also understand that not all NSSL MDA detections are going to be associated with significant weather at the ground. Storm-scale vortex detections that do not meet specified criteria for mesocyclone may be significant and be associated with severe weather, and vice versa. The three examples shown in Fig. 9 illustrate this point. In Fig. 9, each of these supercell storms is producing a tornado during the volume scan times shown. Note the relative size of each of the storms (the horizontal scale of each picture is the same). Each of these three storms are tornadic supercells, yet the two examples on the left are much smaller than the Great Plains supercell example on the right. Also, there is no visually observed radar vortex with the smallest of the storms (and subsequently, no algorithm vortex detection). Not all vortex detections are associated with severe weather, and those vortex detections should not be considered false alarms. Storm-scale vortices are only indirect evidence of potential severe weather. Therefore, the warning forecaster should interrogate all the diagnostic parameters associated with each detection. Finally, we stress again that radar meteorologists (and the media and public) should understand that not all storm-scale vortices (including tornadic vortices) are large or strong enough to be observed with our present operational Doppler radar system and, thus, cannot be detected using radar algorithms.

This interrogation of radar-observed storm-scale vortices is facilitated with novel approaches to the data product output. A mesocyclone attribute table (in RADS) provides the user with a sorted list of the significant vortex detections with color-coded (based on severity) diagnostic output. Effective management of algorithm output is facilitated with the table. Additional diagnostic tools in the form of new overlay icon types, and time series trends and time–height trends, provide the warning forecaster with additional analysis tools for making an informed decision.

We also stress the importance of the integration of information available from other sensors, especially

[10] The experimental IRS may be available in a future build of the WSR-88D Mesocyclone Algorithm (R. Lee 1997, personal communication).

TABLE 14. Comparative skill scores with the algorithm comparison dataset (Table 5) for the various vortex classification schemes and diagnostic measurements (Table 8) for the algorithm comparison dataset (Table 5). POD, FAR, CSI, and HSS are defined by Eqs. (2–5), respectively.

Algorithm	Parameters	Diagnostic measure	Tornadoes (%)				Diagnostic measure	Severe wind (%)			
			POD	FAR	CSI	HSS		POD	FAR	CSI	HSS
NSSL MDA	Default	MESO	52.7	72.6	22.0	33.3	MESO	39.0	65.5	22.4	32.8
NSSL MDA	Default	RANK ≥4	57.6	75.7	20.6	31.1	RANK ≥5	43.3	68.9	22.1	31.9
NSSL MDA	Default	MSI ≥3300	48.3	72.9	21.0	32.0	MSI ≥2600	49.0	74.4	20.2	28.6
NSSL MDA	Default	POT ≥18	46.3	67.7	23.5	35.7	POT ≥16	45.3	63.3	25.4	36.8
NSSL MDA	Default	POSW ≥20	50.7	69.3	23.6	35.7	POSW ≥16	49.9	64.6	26.1	37.5
88D B9MA	Default	MES88	40.9	85.0	12.3	−21.9	MES88	43.8	78.4	16.9	−25.4
88D B9MA	Default	IRS ≥7	32.8	79.9	14.3	−7.9	IRS ≥7	34.5	71.5	18.5	−6.6
88D B9MA	Optimized	MES88	64.5	90.2	9.3	4.7	MES88	65.9	84.7	14.2	7.9
88D B9MA	Optimized	IRS ≥9	36.5	73.9	18.0	23.8	IRS ≥8	39.4	66.9	22.0	27.0
88D B9MA	Low threshold	MES88	69.3	96.6	3.3	2.8	MES88	76.8	94.3	5.6	5.3
88D B9MA	Low threshold	IRS ≥10	30.7	76.5	15.3	24.9	IRS ≥8	40.7	76.6	17.4	26.9

near-storm environment information (from rawinsondes or mesoscale model output). The forecaster should also be cognizant that ''magic numbers'' or constant thresholds will not always define whether a vortex detection is significant or not, especially in varying near-storm environments. For example, for the low-topped supercell cases tested, the 3D strength rank that resulted in the highest skill score is rank 3 (detections with a 3D strength rank of 3 or greater are considered significant). On the other hand, for most of the isolated supercell cases, a 3D strength rank of 5 or greater resulted in the highest skill score. A strength rank filter was even developed for display purposes (in RADS) in which the forecaster can filter out (not display) any detections that fall below a specific 3D strength rank. For these reasons, operational diagnostic thresholds should vary with different near-storm environments, and with operational experience the warning forecaster will become attuned to these variations. This new operational paradigm also reduces the need for site-adaptable parameter studies designed to fine-tune algorithm thresholds to a specific region.

The NSSL MDA has been tested in a real-time operational mode at various NWS forecast offices since 1994 [Fort Worth, Texas, Peachtree City, Georgia (Atlanta area), Phoenix, Arizona, Pittsburgh, Pennsylvania, Minneapolis, Minnesota, Indianapolis, Indiana, Norman, Oklahoma (Oklahoma City area), Denver, Colorado, Salt Lake City, Utah, Melbourne, Florida, Charleston, South Carolina, Reno, Nevada, Albany, New York, Jackson, Mississippi, Pleasant Hill, Missouri (Kansas City area), and Sterling, Virginia (Washington, DC, area)]. Detailed results for one test (Fort Worth) are described in Stumpf and Foster (1996). These real-time tests provided an opportunity to test the NSSL MDA and associated new display concepts, and have provided additional valuable insight on the operational use of the NSSL MDA.

At the time of the publication of this article, the NSSL MDA will not be available operationally in the WSR-88D system. Although the impact of the information available with the NSSL MDA will not affect the warning forecasters immediately, it is important for them to understand these improvements that may soon be implemented. It is equally important that warning meteorologists understand the basis for the development of the NSSL MDA's detection strategy even while the 88D B9MA serves as the fielded algorithm.

6. Future work and concluding remarks

Although a considerable effort has been made to redesign the NSSL MDA and increase its skill, there are still several shortcomings that have yet to be addressed. The NSSL MDA is the result of an effort partially funded by the NWS OSF and the FAA, and time and funding constraints did not permit us to incorporate a number of additional ideas into the algorithm, nor were we able

TABLE 15. Best 88D B9MA (''optimized'' adaptable parameters) and NSSL MDA CSI and HSS for event types using the algorithm comparison dataset (Table 5). Diagnostic measure codes defined in Table 8. CSI and HSS are defined by Eqs. (4) and (5), respectively.

Algorithm	Diagnostic measure	Storm type	Tornado (%)		Severe wind (%)	
			CSI	HSS	CSI	HSS
88D B9MA	MES88	Isolated supercell	25.4	−26.6	41.6	−4.4
88D B9MA	IRS ≥7	Isolated supercell	35.1	44.4	46.7	55.9
NSSL MDA	POT ≥30	Isolated supercell	40.5	55.2	43.5	57.7
88D B9MA	MES88	Other storm types	6.9	−16.4	8.7	−31.2
88D B9MA	IRS ≥7	Other storm types	11.9	17.8	13.6	15.9
NSSL MDA	POT ≥30	Other storm types	13.2	21.2	14.3	21.1

TABLE 16. Composite scores for 10 cases listed in Table 6. Diagnostic measure code defined in Table 8. POD, FAR, CSI, and HSS are defined by Eqs. (2–5), respectively.

Diagnostic measure	Tornadoes (%)				Severe wind (%)			
	POD	FAR	CSI	HSS	POD	FAR	CSI	HSS
MESO	37.7	68.7	20.6	31.6	38.5	61.8	23.8	35.7
RANK ≥4	57.3	78.4	18.6	27.7	60.1	73.0	22.9	33.4
MSI ≥3000	41.0	81.8	14.4	21.4	43.0	77.2	17.5	25.8
POT ≥80	30.0	74.3	16.1	24.9	26.5	72.9	15.5	23.7
POSW ≥80	31.0	77.4	15.0	23.0	28.2	75.5	15.1	22.8

to carry out comparison testing between other MDAs besides the 88D B9MA. We encourage future work in both these areas. More of our suggestions for future work follow.

Although the NSSL MDA 2D feature detection and measurement technique has been significantly improved, further improvements are possible. The 2D feature detection still fails in areas of aliased data, and does not allow for the entire range of sampling limitations (Wood and Brown 1995), nor does it allow for the full range of elliptical vortices (Desrochers and Harris 1996) or for the detection of anticyclonic vortices. Estimating rotational and divergent characteristics of 2D features can be made more robust by applying techniques that compute these attributes using line integration about the edge of the vortex core (Davies-Jones and Stumpf 1997). All measures of height within the algorithm could be made storm relative rather than ground relative. We also encourage an evaluation of optimum lead times as DD92 and Witt et al. (1998a) have done.

Another limitation of the algorithm is its weakness in detecting certain size vortices at given ranges from the radar. Because vortex size and radar sampling size are related, tornado vortices at very close ranges will be sampled as mesocyclone-like vortices (with velocity peaks separated by more than one radial). At far ranges, mesocyclones resemble TVSs (with gate-to-gate velocity peak separation). Furthermore, the hierarchy of vortices that describe the rotation processes within a supercell thunderstorm is still not fully understood (Doswell and Burgess 1993; Mitchell and Stumpf 1996). A more robust method to detect *all* scales of vortices from the tornado to the mesocyclone needs to be developed. This could be accomplished by merging ideas from the NSSL MDA and NSSL TDA (Mitchell et al. 1998) into a single vortex detection and diagnosis algorithm. The NN could then be applied to diagnose an even greater spectrum of vortices related to tornadoes and severe weather at the ground.

Any storm-scale vortex detection algorithm adopted by the WSR-88D system should incorporate additional ideas from other development efforts. For example, DD92 showed that excess rotational kinetic energy (ERKE) could be used as a predictor of strong and violent tornadoes (F2 or greater F scale). Although we are cautious about scoring algorithms versus tornado F scale (see Footnote 9 in section 3 discussing these concerns), we feel that future work could include an evaluation of ERKE using the same dataset (one that is not restricted to Oklahoma storms) and the some of the scoring techniques we present in this paper.

Traditional mesocyclone diagnosis techniques by the warning forecaster include the interrogation of reflectivity (e.g., bounded weak echo regions) and near-storm environmental information. The neural network diagnoses of storm-scale vortices (Marzban and Stumpf 1996, 1998) will continue as new data are collected in the field and as new storm attributes (including trend information, reflectivity attributes, and near-storm environmental parameters) are integrated in the NN training datasets. Adding near-storm environmental param-

TABLE 17. The NN scores for the individual cases listed in Table 6. POD, FAR, CSI, and HSS are defined by Eqs. (2)–(5), respectively.

Date	Radar	Neural network measure	POD (%)	FAR (%)	CSI (%)	HSS (%)
27 May 1995	Des Moines, IA	POT	51.7	42.3	37.5	46.7
27 May 1995	Des Moines, IA	POSW	40.4	34.5	33.3	43.1
12 May 1995	Goodland, KS	POT	60.5	20.7	52.3	66.9
12 May 1995	Goodland, KS	POSW	57.1	25.0	48.0	62.7
22 Jun 1995	Pueblo, CO	POT	33.3	93.1	6.1	2.4
22 Jun 1995	Pueblo, CO	POSW	No severe wind reported			
13 Feb 1995	Phoenix, AZ	No detections are made				
21 Jul 1995	Minneapolis, MN	POT	37.3	57.6	24.8	37.0
21 Jul 1995	Minneapolis, MN	POSW	37.8	60.0	24.1	36.1
28 May 1994	Amarillo, TX	POT	100.0	50.0	50.0	61.2
28 May 1994	Amarillo, TX	POSW	No severe wind reported			
16 Nov 1993	Houston, TX	POT	14.6	62.5	11.8	18.8
16 Nov 1993	Houston, TX	POSW	14.0	60.0	11.5	18.5
15 Nov 1994	Melbourne, FL	POT	40.0	99.8	0.2	−0.7
15 Nov 1994	Melbourne, FL	POSW	No severe wind reported			
22 May 1995	Amarillo, TX	POT	96.4	90.4	9.5	9.2
22 May 1995	Amarillo, TX	POT	No severe wind reported			
5 May 1995	Fort Worth, TX	POT	No tornadoes reported			
5 May 1995	Fort Worth, TX	POSW	20.3	91.7	6.3	9.0

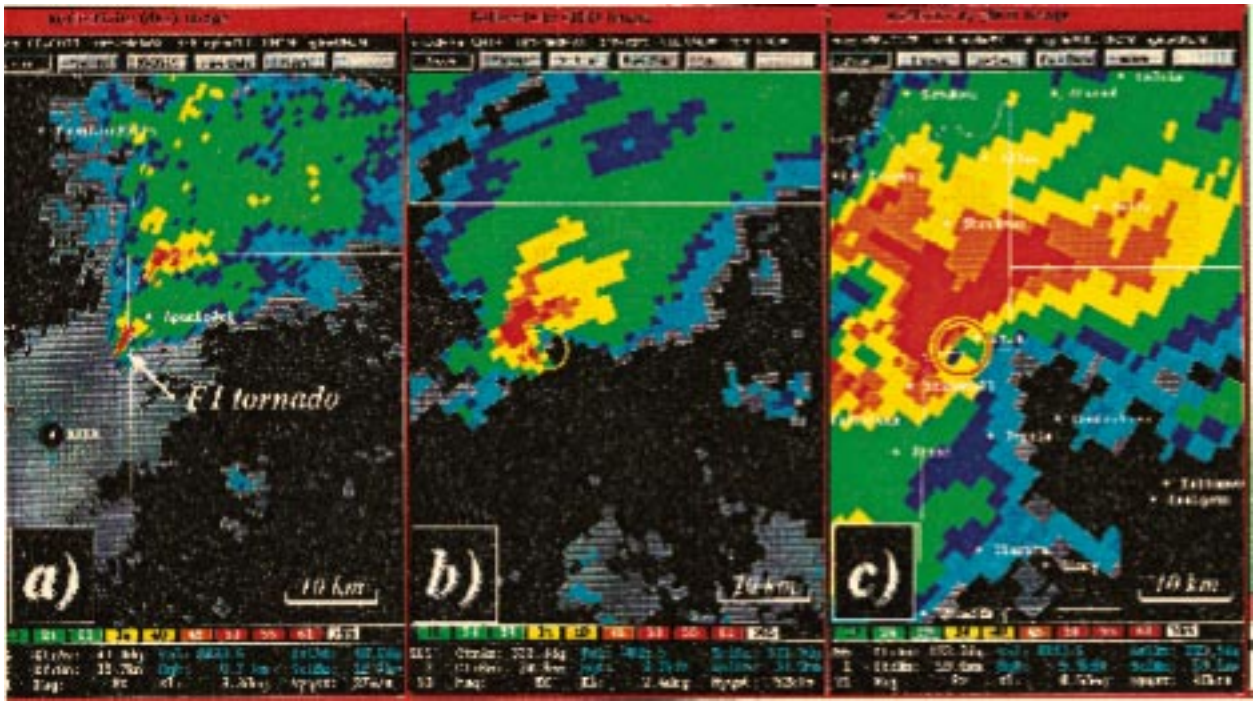

FIG. 9. Radar reflectivity plots of three tornadic supercells: (a) 2211 UTC 13 February 1995 near Mesa, AZ; (b) 2243 UTC 22 June 1995 near Falcon, CO; and (c) 1955 UTC 11 May 1992 near Stonewall, OK. Each plot is drawn to the same scale. NSSL MDA output is overlaid as circles on the radar data.

eters (from rawinsondes or from high-resolution mesoscale models) as NN inputs may help to eliminate some regional dependencies of the algorithm.

The skill scores shown in Tables 14 and 16 (in particular, the CSI) are much lower than those reported by Zrnić et al. (1985) in the evaluation of an early mesocyclone detection algorithm. However, the testing by Zrnić et al. (1985) was done using a very limited dataset comprising only a few hours of data, for a few select storms, all occurring in central Oklahoma. Although the diagnostic skill of the NSSL MDA is an improvement over the various incarnations of the 88D B9MA, the absolute skill is still quite low, partially owing to a still rather large FAR (as seen in Table 14). Burgess et al. (1993) estimates that correct mesocyclone detection will still give only a medium CSI for tornado warnings because even though POD could remain high, the FAR will always be medium (since many mesocyclones do not produce tornadoes). Further analysis of new diagnostic parameters, and future NN training on much larger datasets, may suggest whether a theoretical limit of skill simply based on mesocyclone climatology has been reached. On the other hand, since mesocyclone observation using Doppler radial velocity is limited in many respects (due to the radar's sampling characteristics), the integration of information from a variety of other sensors (e.g., near-storm environment, satellites, lightning, etc.) may provide important missing pieces to the puzzle, thereby distinguishing, with more skill, between significant and insignificant vortices (and thus reducing the FAR to a respectable low number).

Our understanding of tornadogenesis within thunderstorms is still limited. An extensive tornado field project (Rasmussen et al. 1994) completed a data collection phase in 1995, and it will be several more years before scientific hypotheses about tornadogenesis can be proven or refuted. Additionally, the understanding of the near-storm environment and its relationship to the development of severe weather or tornadoes from supercell thunderstorms is limited. Many of the processes leading to tornadogenesis can occur within the lowest few kilometers of the storm, altitudes that are often obscured by the radar horizon. If there are clues within the radar-observable portions of thunderstorms and about the near-storm environment that lead us to conclusions about what is happening at the ground, that information could lead to a more robust algorithm in the future.

Previous algorithm development (88D B9MA, early MDA) was limited by the lack of Doppler radar data with regional, seasonal, and climatological variations. Although development of the NSSL MDA is facilitated by a more robust WSR-88D radar dataset, WSR-88D data collection is still in its infancy, especially for nontraditional varieties of supercell thunderstorms (e.g., minisupercells, tropical mesocyclones). Many of these other varieties of severe storms are much rarer than the Great Plains supercells (in our current opinion; future

WSR-88D data collection may prove otherwise), and the development of a comparably sized database of nontraditional storm types will require many years (and good fortune, considering the events must first occur, be close to a radar and not obscured by radar data artifacts, and then also be recorded as WSR-88D level-II data). Until we have adequate observations of all types of supercell thunderstorms from across the country (and the world), the understanding of supercell evolution and morphology will remain insufficient. Future work must be focused on analyzing more WSR-88D data (and data from other sensors) from a myriad of storm environments and storm types, and these future datasets could be used to produce climatologies of storm-scale vortices.

Acknowledgments. The authors gratefully acknowledge the following NSSL individuals for their assistance: Vincent Wood, Chris Robbins, Todd Gladfelter, and Karen Cooper. Thoughtful comments were provided by two anonymous reviewers, Ralph Donaldson (formerly of Hughes STX Corporation), Dr. David Blanchard of NSSL, and Robert Lee of the NWS OSF (who also assisted in the scoring of the IRS parameter and the development of Fig. 7). Insight has been provided along the way by Drs. Erik Rasmussen and Rodger Brown of NSSL. Contributions from NSSL's Severe Weather Warning Applications and Technology Transfer (SWAT) team paved the path for the development and operational testing of the NSSL MDA. SWAT's members, past and present, have included the first five authors, along with Kim Elmore, Pamela MacKeen, Dr. Caren Marzban, Travis Smith, Steve Vasiloff, Amy Wyatt, and Daphne Zaras. Administrative assistance was provided by Kelly Lynn. Gratitude is extended to the many NWS meteorologists who participated in NSSL MDA real-time testing. This work was partially funded by the NWS OSF and the Federal Aviation Administration.

REFERENCES

Brown, R. A., and V. T. Wood, 1991: On the interpretation of single-Doppler velocity patterns within severe thunderstorms. *Wea. Forecasting,* **6,** 32–48.

Burgess, D. W., 1976: Single-Doppler radar vortex recognition. Part I: Mesoscale signatures. Preprints, *17th Conf. on Radar Meteorology,* Seattle, WA, Amer. Meteor. Soc., 97–103.

——, and L. R. Lemon, 1991: Characteristics of mesocyclones detected during a NEXRAD test. Preprints, *25th Int. Conf. on Radar Meteorology,* Paris, France, Amer. Meteor. Soc., 39–42.

——, R. J. Donaldson, T. Sieland, and J. Hinkelman, 1979: Final Report on the Joint Doppler Operational Project (JDOP 1976–1978). Part I: Meteorological Applications. NOAA Tech. Memo. ERL NSSL-86, NOAA, Boulder, CO, 84 pp. [NTIS PB80-107/88/AS.]

——, V. T. Wood, and R. A. Brown, 1982: Mesocyclone evolution statistics. Preprints, *12th Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 422–424.

——, R. J. Donaldson, and P. R. Desrochers, 1993: Tornado detection and warning by radar. *The Tornado: Its Structure, Dynamics, Prediction, and Hazards, Geophys. Monogr.,* No. 79, Amer. Geophys. Union, 203–221.

——, R. R. Lee, S. S. Parker, D. L. Floyd, and D. L. Andra Jr., 1995: A study of minisupercells observed by WSR-88D radars. Preprints, *27th Conf. on Radar Meteorology,* Vail, CO, Amer. Meteor. Soc., 4–6.

Conway, J. W., K. D. Hondl, M. J. Moreland, J. M. Cordell, and R. J. Harron, 1995: Improvements in the WSR-88D dealiasing algorithm: The pursuit of the final, most important gates. Preprints, *27th Conf. on Radar Meteorology,* Vail, CO, Amer. Meteor. Soc., 145–147.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. *Bull. Amer. Meteor. Soc.,* **74,** 1669–1687.

——, ——, and D. W. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.,* **74,** 645–653.

Davies-Jones, R., and G. J. Stumpf, 1997: On the detection and measurement of circulation and areal expansion rate with WSR-88D radars. Preprints, *28th Conf. on Radar Meteorology,* Austin, TX, Amer. Meteor. Soc., 313–314.

Desrochers, P. R., and R. J. Donaldson Jr., 1992: Automatic tornado prediction with an improved mesocyclone–detection algorithm. *Wea. Forecasting,* **7,** 373–388.

——, and F. I. Harris, 1996: Interpretation of mesocyclone vorticity and divergence structure from single-Doppler radar. *J. Appl. Meteor.,* **35,** 2191–2209.

Donaldson, R. J., 1970: Vortex signature recognition by a Doppler radar. *J. Appl. Meteor.,* **9,** 661–670.

Doswell, C. A., and D. W. Burgess, 1993: Tornadoes and tornadic storms: A review of conceptual models. *The Tornado: Its Structure, Dynamics, Prediction, and Hazards, Geophys. Monogr.,* No. 79, Amer. Geophys. Union, 161–172.

Eilts, M. D., and S. D. Smith, 1990: Efficient dealiasing of Doppler velocities using local environment constraints. *J. Atmos. Oceanic Technol.,* **7,** 118–128.

Hermes, L. G., A. Witt, S. D. Smith, D. Klingle-Wilson, D. Morris, G. J. Stumpf, and M. D. Eilts, 1993: The Gust front detection and wind-shift algorithms for the terminal Doppler weather radar system. *J. Atmos. Oceanic Technol.,* **10,** 693–709.

Johnson, J. T., P. L. MacKeen, A. Witt, E. D. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas, 1998: The Storm Cell Identification and Tracking (SCIT) algorithm: An enhanced WSR-88D algorithm. *Wea. Forecasting,* **13,** 263–276.

Lee, R. R., and A. White, 1998: Improvement of the WSR-88D Mesocyclone Algorithm. *Wea. Forecasting,* **13,** 341–351.

Marzban, C., 1998: Scalar measures of performance in rare-event situations. *Wea. Forecasting,* in press.

——, and G. J. Stumpf, 1996: A neural network for tornado prediction based on Doppler radar-derived attributes. *J. Appl. Meteor.,* **35,** 617–626.

——, and ——, 1998: A neural network for damaging wind prediction. *Wea. Forecasting,* **13,** 15–163.

McCaul, E. W., 1987: Observations of the Hurricane ''Danny'' tornado outbreak of 16 August 1985. *Mon. Wea. Rev.,* **115,** 1206–1223.

Mitchell, E. D., and G. J. Stumpf, 1996: The 19 April 1995 Fort Worth/Dallas tornado event: Implications for automated vortex recognition. Preprints, *18th Conf. on Severe Local Storms,* San Francisco, CA, Amer. Meteor. Soc., 574–576.

——, S. V. Vasiloff, G. J. Stumpf, A. Witt, M. D. Eilts, J. T. Johnson, and K. W. Thomas, 1998: The National Severe Storms Laboratory Tornado Detection Algorithm. *Wea. Forecasting,* **13,** 352–366.

NCDC, 1992–95: *Storm Data.* Vols. 34–37. [Available from the National Climatic Data Center, Federal Building, Asheville, NC 28801-2696.]

OFC, 1991: Doppler radar meteorological observations, Part C, WSR-88D products and algorithms. Federal Meteorological Handbook No. 11, FCM-H11C-1991, 208 pp. [Available from Office of the

Federal Coordinator for Meteorological Services and Supporting Research, Rockville, MD 20850.]

Rasmussen, E. N., J. M. Straka, R. Davies-Jones, C. A. Doswell III, F. H. Carr, M. D. Eilts, and D. R. MacGorman, 1994: Verification of the Origins of Rotation in Tornadoes Experiment: VORTEX. *Bull. Amer. Meteor. Soc.,* **75,** 995–1006.

Rotunno, R., 1986: Tornadoes and tornadogenesis. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., 414–436.

Sanger, S. S., R. M. Steadham, J. M. Jarboe, R. E. Schlegel, and A. Sellakannu, 1995: Human factors contributions to the evolution of an interactive Doppler radar and weather detection algorithm display system. Preprints, *11th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Dallas, TX, Amer. Meteor. Soc., 1–6.

Stumpf, G. J., and D. W. Burgess, 1993: Observations of lower-tropospheric misocyclones along the leading edge of a bow echo thunderstorm. Preprints, *26th Conf. on Radar Meteorology,* Norman, OK, Amer. Meteor. Soc., 215–217.

——, and A. Witt, 1994: The new NSSL mesocyclone detection algorithm functional description. Rep. for the NWS Operational Support Facility, 11 pp. [Available from G. J. Stumpf, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069.]

——, and C. Marzban, 1995: NSSL Build 2.0 Mesocyclone Detection Algorithm (MDA2) final documentation. Rep. for the NWS Operational Support Facility, 124 pp. [Available from G. J. Stumpf, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069.]

——, and M. P. Foster, 1996: The 1995 NSSL Warning Decision Support System test at the Fort Worth National Weather Service Forecast Office. Preprints, *18th Conf. on Severe Local Storms,* San Francisco, CA, Amer. Meteor. Soc., 570–573.

Wakimoto, R. M., and J. W. Wilson, 1989: Nonsupercell tornadoes. *Mon. Wea. Rev.,* **117,** 1113–1140.

Watson, A. I., T. R. Shepherd, E. N. Rasmussen, C. L. Ziegler, and R. J. Trapp, 1995: Mesocyclone evolution as seen by airborne pseudo-dual Doppler radar during VORTEX. Preprints, *27th Conf. on Radar Meteorology,* Vail, CO, Amer. Meteor. Soc., 522–524.

Wicker, L. J., and R. B. Wilhelmson, 1993: Numerical simulations of tornadogenesis within a supercell thunderstorm. *The Tornado: Its Structure, Dynamics, Prediction, and Hazards, Geophys. Monogr.,* No. 79, Amer. Geophys. Union, 75–88.

Wieler, J. G., 1986: Real-time automated detection of mesocyclones and tornadic vortex signatures. *J. Atmos. Oceanic Technol.,* **3,** 98–113.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. Mitchell, and K. W. Thomas, 1998a: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting,* **13,** 286–303.

——, ——, ——, E. D. Mitchell, J. T. Johnson, and K. W. Thomas, 1998b: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting,* **13,** 513–518.

Wood, V. T., and R. A. Brown, 1995: Effect of radar sampling on Doppler velocity mesocyclone signatures. Preprints, *27th Conf. on Radar Meteorology,* Vail, CO, Amer. Meteor. Soc., 403–405.

Zrnić, D. S., D. W. Burgess, and L. D. Hennington, 1985: Automatic detection of mesocyclonic shear with Doppler radar. *J. Atmos. Oceanic Technol.,* **2,** 425–438.