

Validation of Polarimetric Hail Detection

PAMELA L. HEINSELMAN* AND ALEXANDER V. RYZHKOV*

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

(Manuscript received 14 June 2005, in final form 25 March 2006)

ABSTRACT

This study describes, illustrates, and validates hail detection by a simplified version of the National Severe Storms Laboratory's fuzzy logic polarimetric hydrometeor classification algorithm (HCA). The HCA uses four radar variables: reflectivity, differential reflectivity, cross-correlation coefficient, and "reflectivity texture" to classify echoes as rain mixed with hail, ground clutter-anomalous propagation, biological scatterers (insects, birds, and bats), big drops, light rain, moderate rain, and heavy rain. Diagnostic capabilities of HCA, such as detection of hail, are illustrated for a variety of storm environments using polarimetric radar data collected mostly during the Joint Polarimetric Experiment (JPOLE; 28 April-13 June 2003). Hail classification with the HCA is validated using 47 rain and hail reports collected by storm-intercept teams during JPOLE. For comparison purposes, probability of hail output from the Next-Generation Weather Radar Hail Detection Algorithm (HDA) is validated using the same ground truth. The anticipated polarimetric upgrade of the Weather Surveillance Radar-1988 Doppler network drives this direct comparison of performance. For the four examined cases, contingency table statistics show that the HCA outperforms the HDA. The superior performance of the HCA results primary from the algorithm's lack of false alarms compared to the HDA. Statistical significance testing via bootstrapping indicates that differences in the probability of detection and critical success index between the algorithms are statistically significant at the 95% confidence level, whereas differences in the false alarm rate and Heidke skill score are statistically significant at the 90% confidence level.

1. Introduction

Over the past 45 yr, numerous studies have explored the relationships between the conventional (i.e., single polarization) radar reflectivity factor (hereafter reflectivity) and hail occurrence at the ground. These studies usually employ radar reflectivity data only (e.g., Donaldson 1959; Geotis 1963; Rinehart and Staggs 1968; Waldvogel and Federer 1976; Dye and Martner 1978; Amburn and Wolf 1997), radar reflectivity data combined with observations from other sensors (e.g., Mather et al. 1976; Foote and Knight 1979; Waldvogel et al. 1979; Auer 1994; Hardaker and Auer 1994; Billet et al. 1997; Witt et al. 1998), or dual-wavelength radar

reflectivity data (e.g., Atlas and Ludlum 1961; Eccles and Atlas 1973; Féral et al. 2003).

Single-radar reflectivity approaches to hail diagnosis began in the late 1950s and early 1960s when Donaldson (1959) and Geotis (1963) sought to delineate areas of rain and hail using low-elevation reflectivity values exceeding a certain threshold. This technique capitalizes on the strong dependence of reflectivity on hydrometeor diameter and the effects of Mie scattering for hail-size particles (Dye and Martner 1978; Doviak and Zrnić 1993). Another reflectivity-centric approach employs vertically integrated liquid water (VIL) to detect hail (Greene and Clark 1972). Correlations of VIL with severe weather occurrence, first noted by Elvander (1977), led to the development of the VIL-based Weather Surveillance Radar-1988 Doppler (WSR-88D) severe weather potential (SWP) algorithm (Kitzmillier et al. 1995) and the use of "VIL of the day" and "VIL density" [ratio of VIL to echo-top height; Amburn and Wolf (1997)] by forecasters to assess hail potential. The limited success of the latter two approaches is due to variability in VIL values and echo-

* Additional affiliation: NOAA/National Severe Storms Laboratory, Norman, Oklahoma.

Corresponding author address: Pamela Heinselman, National Weather Center, 120 David L. Boren Blvd., Room 4905, Norman, OK 73072.

E-mail: pam.heinselman@noaa.gov

top heights with radar range (Edwards and Thompson 1998; Maddox et al. 1999).

Hail diagnosis using single-radar reflectivity data and observations from other sensors began in the 1970s when hail suppression experiments showed a relation between hail occurrence and the height of the 45-dBZ contour above the freezing level (Mather et al. 1976; Foote and Knight 1979; Waldvogel et al. 1979). Interestingly, this result corroborates Donaldson's (1959) earlier finding that high-reflectivity values at the upper levels of storms are related to hail fall at the ground. The relation of hail at the ground to vertical profiles of reflectivity, employed by Petrocchi (1982), constitutes the Next-Generation Weather Radar (NEXRAD) hail algorithm used by the National Weather Service (NWS) during the 1980s and 1990s. Later, Witt et al. (1998) used data presented by Waldvogel et al. (1979) to develop a simple relation where higher probabilities of hail correspond with maximum reflectivity values located higher above the melting level. This relation is used in the current NEXRAD hail detection algorithm (HDA) to forecast the probability of hail of any size. Because the HDA is cell based, problems with storm cell detection and tracking can make it difficult to quantify correctly the vertical storm structure. The addition of vertical temperature profile data makes it necessary to attain accurate and timely updates, especially in situations where the environment is evolving rapidly. A different multisensor approach to hail diagnosis seeks to discriminate between hail and heavy rain through a linear regression analysis of single-radar reflectivity data and infrared cloud-top temperatures from satellite imagery (e.g., Auer 1994; Hardaker and Auer 1994). A limitation of this technique is the relatively long update time of satellite data.

The last conventional hail diagnosis technique reviewed in this paper capitalizes on differences in dual-wavelength signals from hail observed by radars of different wavelength. Atlas and Ludlum (1961) showed that, due to effects of Mie scattering, the ratio of radar reflectivities at 10- and 3-cm wavelengths for hail is much larger than for rain and that such a ratio can be used for hail detection. Two distinct limitations of this dual-wavelength technique are the strong attenuation of radar returns at the shorter wavelength and the need to match antenna patterns. Although Eccles and Atlas (1973) successfully address the 3-cm-wavelength attenuation problem by computing the range derivative of the dual-wavelength power ratio, it is difficult to utilize their technique operationally because it requires collocated radar beams at two different radar wavelengths. Recently, Féral et al. (2003) proposed a more operationally viable dual-wavelength approach to hail

detection by using 10- and 5-cm wavelength radars located far from each other within France's radar network. For each storm observed by both radars, Féral et al. (2003) identify areas where the reflectivity is 40 dBZ and higher and compute the ratio of average reflectivity within the storm core (maximum $Z - 3$ dB) to the average reflectivity within a 3-km radius of the storm core (at each radar wavelength). Once spatially corresponding radar bins are identified, Féral et al. (2003) compute the dual-wavelength reflectivity hail ratio (DWHR). DWHR is defined as the ratio of the previously computed 10-cm ratio to the 5-cm ratio, multiplied by 100. Owing to differences in scattering between the two radars, values of DWHR higher than 100% denote hail. Because this approach uses radar reflectivity thresholds to define storm areas of interest, heavy rain may be misdiagnosed as hail when precipitation is driven by warm rain processes.

While each hail detection technique described above has its own limitations, conventional approaches to hail detection, in general, have common disadvantages compared to polarimetric approaches. Most importantly, dual-polarization radars measure supplemental radar variables that characterize the differences between radar returns at two orthogonal polarizations and, consequently, provide information regarding various microphysical properties of scatterers. Using these measurements, polarimetric approaches capitalize on the differences in shapes and orientations between hailstones and raindrops. Because hailstones tend to tumble and their orientation is more chaotic than that of raindrops, the differential reflectivity Z_{DR} and specific differential phase K_{DP} in hail are substantially lower than in rain with the same radar reflectivity factor Z . In contrast, linear or circular depolarization ratios (LDR and CDR, respectively) are higher in hail than in rain. The reader is referred to the monographs of Doviak and Zrnić (1993) and Bringi and Chandrasekar (2001) for definitions of basic polarimetric variables.

The most commonly used polarimetric method for hail detection utilizes combined measurements of Z at horizontal polarization and Z_{DR} . In this method, rain and hail are separated in the $Z-Z_{DR}$ plane by a boundary defined with relations $Z_{DR} = f(Z)$ or $Z = g(Z_{DR})$ as specified by Leitao and Watson (1984) and Aydin et al. (1986). Similar relations $Z = y(K_{DP})$ were used by Balakrishnan and Zrnić (1990a), Ryzhkov and Zrnić (1994), and Smyth et al. (1999) for the delineation of rain and hail in the $Z-K_{DP}$ plane. Sometimes all three radar variables, Z , Z_{DR} , and K_{DP} , are involved in the designation of hail. Because the three radar parameters are interdependent in rain medium [i.e., they are "consistent" in rain; e.g., Scarchilli et al. (1996)] lack of such

a consistency may point to the presence of hail (Smyth et al. 1999; Brandes and Ryzhkov 2004).

LDR was used for identification of hail by Bringi et al. (1986), Holler et al. (1994), Kennedy et al. (2001), and Zeng et al. (2001), among others. It is widely accepted that radar echoes from rain are characterized by LDRs of less than -25 dB, whereas hail, melting snowflakes, and nonmeteorological scatterers may have much larger LDRs.

Balakrishnan and Zrnić (1990b) suggested using the cross-correlation coefficient between horizontally and vertically polarized signals ρ_{hv} in addition to other polarimetric variables to better recognize hail. At 10-cm wavelength, ρ_{hv} in rain rarely drops below 0.98. Lower values of ρ_{hv} combined with high Z might indicate either pure hail or a mixture of hail and rain.

Numerous experimental studies indicate that quite often rain and hail partially overlap in the Z - Z_{DR} or Z - K_{DP} planes. Such an overlapping might be exacerbated by measurement errors in the radar polarimetric variables and by biases in Z and Z_{DR} . Biases in Z and Z_{DR} caused by attenuation and differential attenuation are difficult to correct in media containing hail. In addition, specific differential phases are prone to errors due to the large gradients of the total differential phase (Ryzhkov and Zrnić 1998) that are commonly observed in hailstorms. The inherent “fuzziness” of rain-hail boundaries is more adequately addressed in the framework of a fuzzy logic approach than by deterministic or “rigid” boundaries. Several fuzzy logic hydrometeor classifiers based on polarimetric measurements have become increasingly popular in recent years (Vivekanandan et al. 1999; Zrnić and Ryzhkov 1999; Straka et al. 2000; Liu and Chandrasekar 2000; Zrnić et al. 2001; Keenan 2003; Lim et al. 2005). Hail is one of several classes of meteorological and nonmeteorological scatterers that can be distinguished using fuzzy logic classification routines.

Because hail diagnosis by polarimetric measurements is based on the bulk properties of hydrometeors rather than on reflectivity alone or reflectivity combined with other data platforms, it is likely that the former method outperforms the latter methods. Indeed, a recent comparison of algorithm performance showed that a simplified version of the National Severe Storms Laboratory’s (NSSL’s) fuzzy logic hydrometeor classifier outperformed the NEXRAD HDA (probability of hail indicator) during the recent Joint Polarization Experiment (JPOLE; Ryzhkov et al. 2005). To date, this is the only study in the literature that compares directly the statistical performance of polarimetric and conventional hail diagnosis algorithms. The importance of

such comparative studies is supported by the anticipated polarimetric upgrade of the WSR-88D network.

The purpose of this study is to describe a simplified version of the NSSL’s fuzzy logic hydrometeor classification algorithm (HCA) for rain and hail discrimination, illustrate diagnostic uses of the HCA, and review and expand upon hail validation results reported by Ryzhkov et al. (2005). Polarimetric data for this study were collected with NSSL’s research and development WSR-88D radar in Norman, Oklahoma (KOUN hereafter), to which polarimetric diversity was added in March 2002. The radar can operate in two different modes. In its simultaneous horizontal and vertical (SHV) transmission and reception mode, the radar simultaneously transmits and receives horizontally (H) and vertically (V) polarized waves. The following variables are measured in the SHV mode: radar reflectivity factor Z at horizontal polarization, Doppler velocity V , spectral width σ_w , differential reflectivity Z_{DR} , differential phase Φ_{DP} , and the magnitude of the cross-correlation coefficient ρ_{hv} between two copolar components of the radar signal. In the linear depolarization ratio (LDR) mode, only horizontally polarized radiation is transmitted but both copolar and cross-polar components of the radar return are received. Hence, LDR can be estimated in the LDR mode at the expense of Z_{DR} and ρ_{hv} . During data collection, the KOUN radar was operated mostly in the SHV mode and thus LDR was not available. Therefore, an additional objective of this study is to evaluate the quality of polarimetric hail detection in the absence of LDR.

2. Algorithm description

Among several polarimetric classification routines used at NSSL (Ryzhkov et al. 2005), we selected a simplified version of a fuzzy logic classifier tailored for hail detection from the data collected mostly at the lowest PPI scan during the JPOLE Intense Operation Period (IOP) in April–June 2003. The algorithm distinguishes between echoes caused by 1) ground clutter/anomalous propagation (AP), 2) biological scatterers (insects, birds, and bats), 3) “big drops” (rain with drop spectra characterized by the presence of big drops and a deficit of small drops), 4) light rain, 5) moderate rain, 6) heavy rain, and 7) rain mixed with hail. The median volume drop diameter for rain with big drops typically exceeds 2.5 mm. Most of light rain is characterized by rain rates lower than 5 mm h^{-1} , moderate rain has rates between 5 and 30 mm h^{-1} , and rates exceeding 30 mm h^{-1} are common for heavy rain. The structure of the algorithm and choice of classes were dictated by the conditions and timing of the validation experiment. Because only

data from the lowest elevation scan were used for comparison with ground truth and observations were limited to within 150 km of the radar, none of the snow categories or graupel were included in the list of classes.

Four radar variables, Z , Z_{DR} , ρ_{hv} , and the “reflectivity texture” $SD(Z)$, are used for classification. The parameter $SD(Z)$ characterizes the magnitude of small-scale fluctuations of Z along the radar ray. To obtain $SD(Z)$, we average Z data (sampled every 0.267 km) along the radial, using a 1-km running average, and subtract the smoothed estimates of Z from their original values. Because $SD(Z)$ is much higher for nonmeteorological radar echoes than for echoes associated with hydrometeors, this parameter is very useful for distinguishing between ground clutter (GC)/AP and hail echoes that may have very similar Z , Z_{DR} , and ρ_{hv} values. In other versions of the NSSL fuzzy classifier, the texture parameter of the differential phase [$SD(\Phi_{DP})$] is utilized together with $SD(Z)$ (Ryzhkov et al. 2005). However, we prefer not to use the standard deviation of Φ_{DP} in hail-bearing storms because it increases rapidly as the cross-correlation coefficient ρ_{hv} drops due to hail presence and/or nonuniform beam filling (NBF), which is very common in this type of storm. Hence, it can confuse the classification algorithm by producing false designations of clutter/noise within the storm.

Some hydrometeor classification algorithms reported in the literature (e.g., Zrnić et al. 2001; Keenan 2003; Lim et al. 2005) utilize specific values for the differential phase K_{DP} and temperature as input variables. In our validation study, hail identification is performed below freezing level and temperature information is not crucial.

As far as K_{DP} is concerned, its high-resolution estimation in localized severe storms may suffer from extra noisiness caused by the low magnitudes of ρ_{hv} in hail and from biases due to NBF as shown in the study by Ryzhkov (2005). We believe that any quantitative use of K_{DP} (either for hail detection or rainfall estimation) should be preceded by an evaluation of the NBF-related perturbations of the radial profiles of Φ_{DP} using the estimates of the cross-beam gradients of Φ_{DP} and Z according to Ryzhkov (2005). In the areas where such perturbations are too high, the estimates of K_{DP} are not reliable. Given this reasoning and for the sake of simplicity, we do not utilize K_{DP} in this study.

The seven classes are described by 28 one-dimensional membership functions $P^{(i)}(Y_j)$ that characterize the distribution of a radar parameter Y_j [e.g., Z , Z_{DR} , ρ_{hv} , $SD(Z)$] for the i th class. It is assumed that one-dimensional membership functions have an asym-

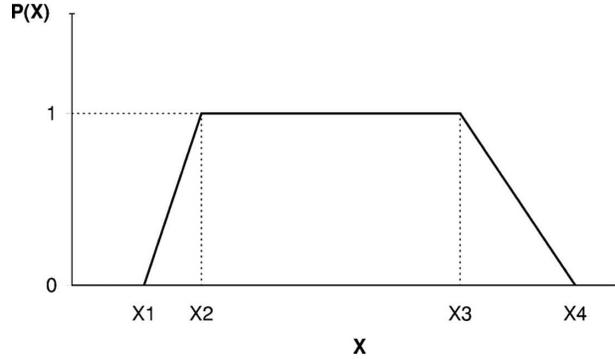


FIG. 1. Trapezoidal membership function.

metric trapezoidal shape with a maximal value of one and a minimal value of zero (Fig. 1). These trapezoidal functions are described by four parameters: $X1$, $X2$, $X3$, and $X4$, as shown in Fig. 1. Before these functions are applied to the data, Z and Z_{DR} are corrected for attenuation and differential attenuation (Bringi et al. 1990), and the measured radar variables, which are obtained from a polar grid, are estimated on a Cartesian grid by averaging values within $1 \text{ km} \times 1 \text{ km}$ grid cells.

An aggregation value for each class of radar echo is defined as

$$Q_i = \frac{\sum_{j=1}^M W_j P^{(i)}(Y_j)}{\sum_{j=1}^M W_j}, \quad (1)$$

where W_j is a weight between 0 and 1 assigned to the j th variable and $M = 4$ is the number of variables. The type of radar echo is identified by the maximal aggregation value.

The parameters $X1$ – $X4$ describing 28 one-dimensional membership functions for seven classes are presented in Table 1. In Table 1, fl , fh , and fb are functions of the radar reflectivity,

$$fl(Z) = -0.50 + 2.50 \cdot 10^{-3} Z + 7.50 \times 10^{-4} Z^2, \quad (2)$$

$$fh(Z) = 0.08 + 3.64 \cdot 10^{-2} Z + 3.57 \times 10^{-4} Z^2, \quad \text{and} \quad (3)$$

$$fb(Z) = -0.20 + 0.108 Z - 6.43 \times 10^{-4} Z^2, \quad (4)$$

where Z is expressed in dBZ. Equal weights are currently given to all four radar variables in Eq. (1). Zero weights can be given to any variable that is not useful for classification. The class with the highest aggregation score is designated as the most likely scatterer type. If at any given location GC/AP is identified as the most likely source of the radar echo and the absolute value of the mean Doppler velocity V is larger than 1 m s^{-1} , then the designation is made for a class with the second-

TABLE 1. Parameters of membership functions for the following seven classes: ground clutter and anomalous propagation (GC/AP), biological scatterers (BS), big drops (BD), light rain (LR), moderate rain (MR), heavy rain (HR), and rain mixed with hail (RH).

	$P(Z)$						
	GC/AP	BS	BD	LR	MR	HR	RH
X1 (dB)	15	5	15	5	30	40	45
X2 (dB)	20	10	20	10	35	45	50
X3 (dB)	70	20	45	35	45	55	75
X4 (dB)	80	30	50	40	50	60	80
	$P(Z_{DR})$						
	GC/AP	BS	BD	LR	MR	HR	RH
X1 (dB)	-4	0	$fh(Z) - 0.3$	$fl(Z) - 0.3$	$fl(Z) - 0.3$	$fl(Z) - 0.3$	-0.3
X2 (dB)	-2	2	$fh(Z)$	$fl(Z)$	$fl(Z)$	$fl(Z)$	0.0
X3 (dB)	1	10	$fb(Z)$	$fh(Z)$	$fh(Z)$	$fh(Z)$	$fl(Z)$
X4 (dB)	2	12	$fb(Z) + 1.0$	$fh(Z) + 0.3$	$fh(Z) + 0.3$	$fh(Z) + 0.3$	$fl(Z) + 0.3$
	$P(\rho_{hv})$						
	GC/AP	BS	BD	LR	MR	HR	RH
X1	0.5	0.3	0.94	0.95	0.95	0.95	0.85
X2	0.6	0.5	0.97	0.98	0.98	0.98	0.97
X3	0.9	0.8	1.0	1.0	1.0	1.0	1.0
X4	0.95	0.83	1.01	1.01	1.01	1.01	1.01
	$P[SD(Z)]$						
	GC/AP	BS	BD	LR	MR	HR	RH
X1 (dB)	2	1	0	0	0	0	0
X2 (dB)	4	2	0.5	0.5	0.5	0.5	0.5
X3 (dB)	10	4	3	3	3	3	3
X4 (dB)	15	7	6	6	6	6	6

highest aggregation score. This helps to prevent the misclassification of ground clutter/AP as hail. The parameters X1–X4 in Table 1 were obtained by using various thresholds published in the literature (see a review by Straka et al. 2000) or by examining histograms and scatterplots of the measured radar variables within the areas with a priori known sources of radar echo.

As shown in Table 1, the fuzzy logic approach allows the four rain categories (BD, LR, MR, and HR) to overlap each other and to overlap with hail possibly mixed with rain (RH), in terms of all four radar variables. A boundary separating areas of rain and hail in the Z – Z_{DR} plane [where $P(Z_{DR})$ is equal to 1 for rain and hail, respectively] is defined by the dependence $Z_{DR} = fl(Z)$ from Eq. (2). Figure 2 illustrates the position of this demarcation line with respect to the ones determined by functions $Z = g(Z_{DR})$ as determined by Leitao and Watson (1984),

$$\begin{aligned} Z &= -4Z_{DR}^2 + 19Z_{DR} + 37.5 & \text{if } 0 < Z_{DR} < 2.5\text{dB} \\ Z &= 60 & \text{if } 2.5 \leq Z \leq 4.0\text{dB}, \end{aligned} \tag{5}$$

and Aydin et al. (1986),

$$\begin{aligned} Z &= 27 & \text{if } Z_{DR} \leq 0\text{dB} \\ Z &= 19Z_{DR} + 27 & \text{if } 0 \leq Z_{DR} \leq 1.74\text{dB} \\ Z &= 60 & \text{if } Z_{DR} > 1.74\text{dB}. \end{aligned} \tag{6}$$

In (5) and (6), Z is expressed in dBZ and Z_{DR} is expressed in dB. The Z – Z_{DR} boundary used in our classification algorithm (marked as line 3 in Fig. 2) is shifted slightly upward with respect to line 2 [corresponding to Eq. (6)] and is significantly different from

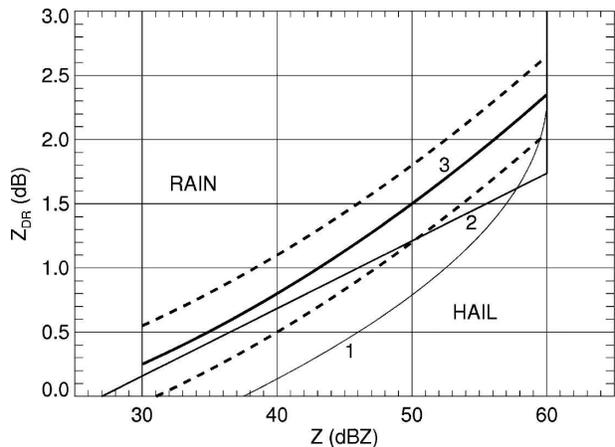


FIG. 2. Boundaries between rain and hail in the Z – Z_{DR} plane used by Leitao and Watson (1984) (curve 1), Aydin et al. (1986) (curve 2), and in the NSSL’s fuzzy logic classification algorithm (curve 3). Dashed lines encompass the area for $P(Z_{DR})$ where hail and rain overlap according to the fuzzy logic algorithm.

TABLE 2. Storm intercept data used in the statistical evaluation including the date, storm type, radar range, range of observed hail sizes, and number of observations. Here, LP denotes low precipitation.

Date	Type	Range (km)	Hail size (cm)	No.
1 May 2003	LP supercell	120–140	0.64–4.5	7
14 May 2003	Multicells	40–60	0.50–0.65	8
19 May 2003	Supercells	30–50	0.50–4.50	21
		80–100		
10 Jun 2003	Multicells	30–50	0.50–0.88	11
		110–130		

line 1 [determined by Eq. (5)]. Note that for $P(Z_{DR})$ the fuzzy logic classification routine allows for overlapping of rain and hail $Z-Z_{DR}$ pairs within the “fuzzy” boundary stretching 0.6 dB along the Z_{DR} axis between the dashed lines in Fig. 2, whereas curves 1 and 2 depict deterministic boundaries.

3. Data collection

The polarimetric radar data used to illustrate the diagnostic capability of the HCA or to validate its performance (Ryzhkov et al. 2005) were collected mostly during JPOLE. During the course of this experiment (28 April 2003–13 June 2003), two vehicles intercepted thunderstorm cores that had the potential to produce hail at the surface. While meteorologists noted rainfall and hail fall characteristics within storm cores, a global positioning system recorded vehicle position. Storms intercepted included an isolated low-precipitation (LP) supercell storm (Bluestein and Parks 1983) on 1 May, a classic supercell storm on 19 May, two linearly aligned LP supercell storms on 11 June, and lines of convective storm cells on 14 May and 10 June, all within 150 km of KOUN (Table 2). The ensuing dataset included 47 reports of rain intensity ($N = 14$) and hail size ($N = 33$), with hail diameters ranging from 0.5 to 4.5 cm (Table 2). These data were used for comparative validation of the NEXRAD HDA (Witt et al. 1998) and the fuzzy logic polarimetric classifier described in the previous section. This validation dataset excluded 11 June 2003 because data from the conventional radar at Norman/Twin Lakes, Oklahoma (KTLX), used to validate the NEXRAD HDA (Ryzhkov 2005), were not recorded. Other polarimetric radar cases examined in this study include a developing supercell on 8 May 2003 and an isolated severe storm on 1 July 2004. We use the latter cases, and the 19 May 2003 supercell, to illustrate the diagnostic capability of the HCA.

4. Diagnostic capability

To be an operationally viable algorithm, the HCA must be able to diagnose regions of severe hail (with

hailstone diameter larger than 1.9 cm) within a thunderstorm. Another desirable capability is the diagnosis of polarimetric signatures indicative of hail at any radar elevation within a storm, under a variety of environmental conditions. These capabilities are illustrated by polarimetric signatures and HCA output for a supercell that produced 2.5-cm hail (19 May 2003), a developing supercell with hail aloft (8 May 2003), and a severe storm embedded in AP (1 July 2004). The latter event emphasizes the HCA’s capability to distinguish between meteorological and nonmeteorological echoes.

Figure 3 illustrates the diagnosis of regions of severe hail within the 19 May 2003 supercell, which produced hail of 0.5–4.5-cm size. At 2330 UTC, the 2.5-cm hail reported on the ground at a location of approximately $X = -45.5$ km and $Y = -82$ km was associated with Z of 55 dBZ, Z_{DR} of 0.8 dB, and ρ_{hv} of 0.92 at 0.5° elevation [~ 1.2 km above radar height (ARH); Figs. 3a–c, respectively]. When these values are run through the HCA, the target is classified as a rain–hail mixture (Fig. 3d).

Besides classifying hail at the lowest elevation, the HCA is capable of tracking the development of hail aloft and assessing the potential for hailfall at the ground—especially for storms located within about 120 km of the radar. In Fig. 4, this situation is illustrated for a developing storm on 8 May 2003 (at ~ 2115 UTC) that later evolved into a tornadic supercell. While polarimetric signatures within the broad storm core at 0.5° elevation (~ 0.6 km ARH) are characterized by a relatively weak reflectivity core (40–50 dBZ; Fig. 4a), moderate Z_{DR} (2–3 dB; Fig. 4c), and relatively high ρ_{hv} (≥ 0.97 ; Fig. 4e) indicative of rain (Fig. 4g), polarimetric signatures within the more compact storm core ($X = -45$ km and $Y = -4$ km) at 3.5° elevation (~ 2.9 km ARH) are characterized by high reflectivity (≥ 60 dBZ; Fig. 4b), relatively low Z_{DR} (0.5–1.0 dB; Fig. 4d), and a nearby minima in ρ_{hv} (0.9–0.95; Fig. 4f) indicative of hail (Fig. 4h). About 25 min later, *Storm Data* (NCDC 2003) reported 2.2-cm hail falling from this storm.

In addition to detecting regions of rain and hail, the HCA detects regions of nonmeteorological echo: biological scatterers (BS) and GC/AP. Owing to the similarity of reflectivity characteristics among storms and GC/AP, this capability is particularly important when severe storms develop within a region of GC/AP. This situation arose on the morning of 1 July 2004 (1425 UTC) as a hail-producing storm located between 60–80 km west and 0–10 km south of KOUN became embedded in AP (Fig. 5). Near that time (1200 UTC), vertical profiles of temperature and humidity (estimated from a sounding taken at Norman, Oklahoma) were favorable for the superrefraction of the radar beam (not shown).

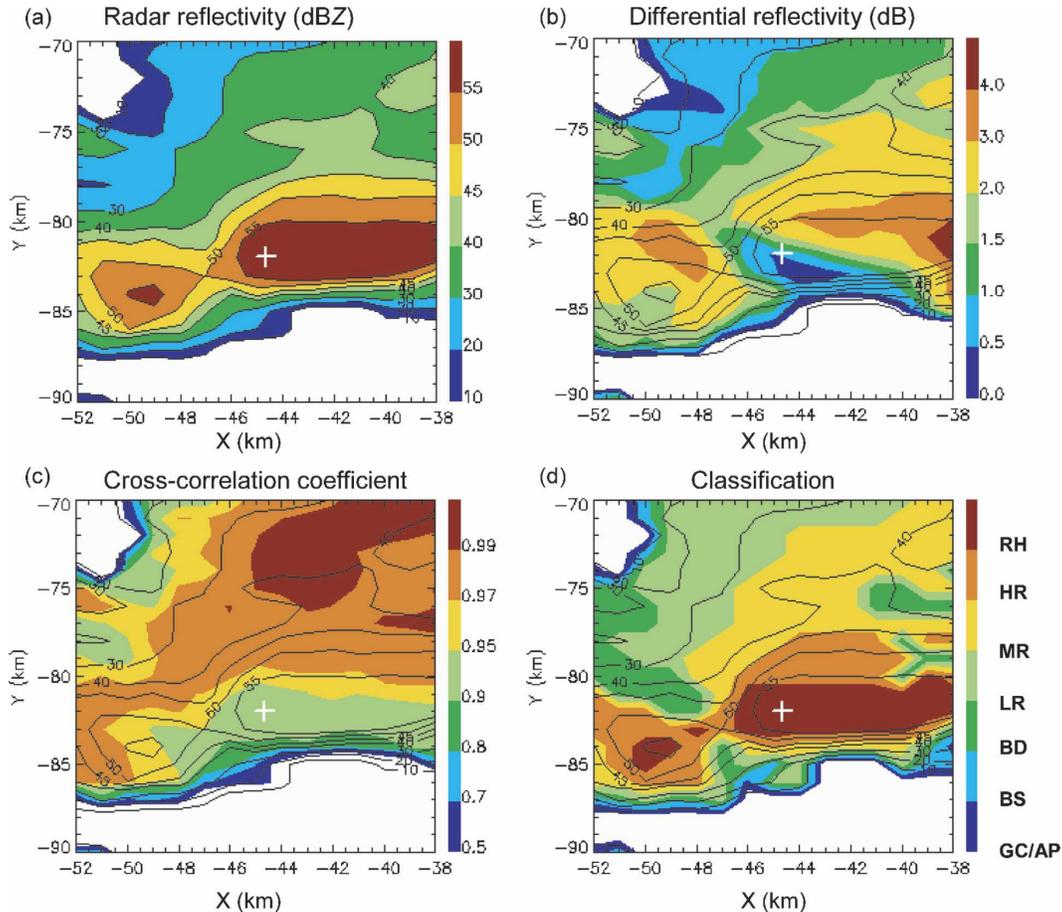


FIG. 3. At 0.5° elevation, the (a) Z , (b) Z_{DR} , (c) ρ_{hv} , and (d) classifications at 2330 UTC on 19 May 2003. Reflectivity contours overlay each panel. The white plus sign shows the location of 2.5-cm (1.0 in.) hail reported by the storm intercept team at approximately the same time.

Because reflectivity values within this region are quite high (≥ 55 dBZ), especially just east of the aforementioned cell, differentiating regions of rain and hail from regions of AP is difficult (Fig. 5a). As AP is essentially stationary (radial velocity ~ 0.0 m s $^{-1}$) and has robust polarimetric signatures—near-zero or negative Z_{DR} (Fig. 5b) and particularly low ρ_{hv} (0.7–0.9; Fig. 5c)—the HCA can correctly identify most areas that have polarimetric signals that are consistent with AP and rain. In this case, the HCA also identifies areas with polarimetric signals consistent with hail. One of these hail areas was associated with 2.5-cm (1 in.) hail reported at $X = -75$ km and $Y = -5$ km by *Storm Data* (NCDC 2004; Fig. 5d).

5. Validation of the HCA

The three cases described above illustrate the HCA's capability to diagnose polarimetric signatures indicative of severe hail at various radar elevations within a storm,

under a variety of environmental conditions. Another desirable attribute of the HCA is high statistical accuracy and skill. In an overview paper on JPOLE, Ryzhkov et al. (2005) report that the HCA not only verified well against hail reports attained during JPOLE, but attained higher accuracy and skill scores than did classifications based on the NEXRAD HDA's probability of hail output. In this section, we provide a more detailed description of the validation methodology and results presented by Ryzhkov et al. (2005).

To validate hail detection with the conventional (HDA) and polarimetric (HCA) algorithms, ground truth data collected during JPOLE (Table 2) were compared with the HDA and HCA output from the lowest radar elevation (0.5°). The HDA ingested data collected with KTLX, the nearest operational WSR-88D radar (20 km northeast of KOUN), whereas the HCA ingested data collected with KOUN. Because the HDA is a cell-based algorithm, it was necessary to determine which storm cell was located closest to the ground truth.

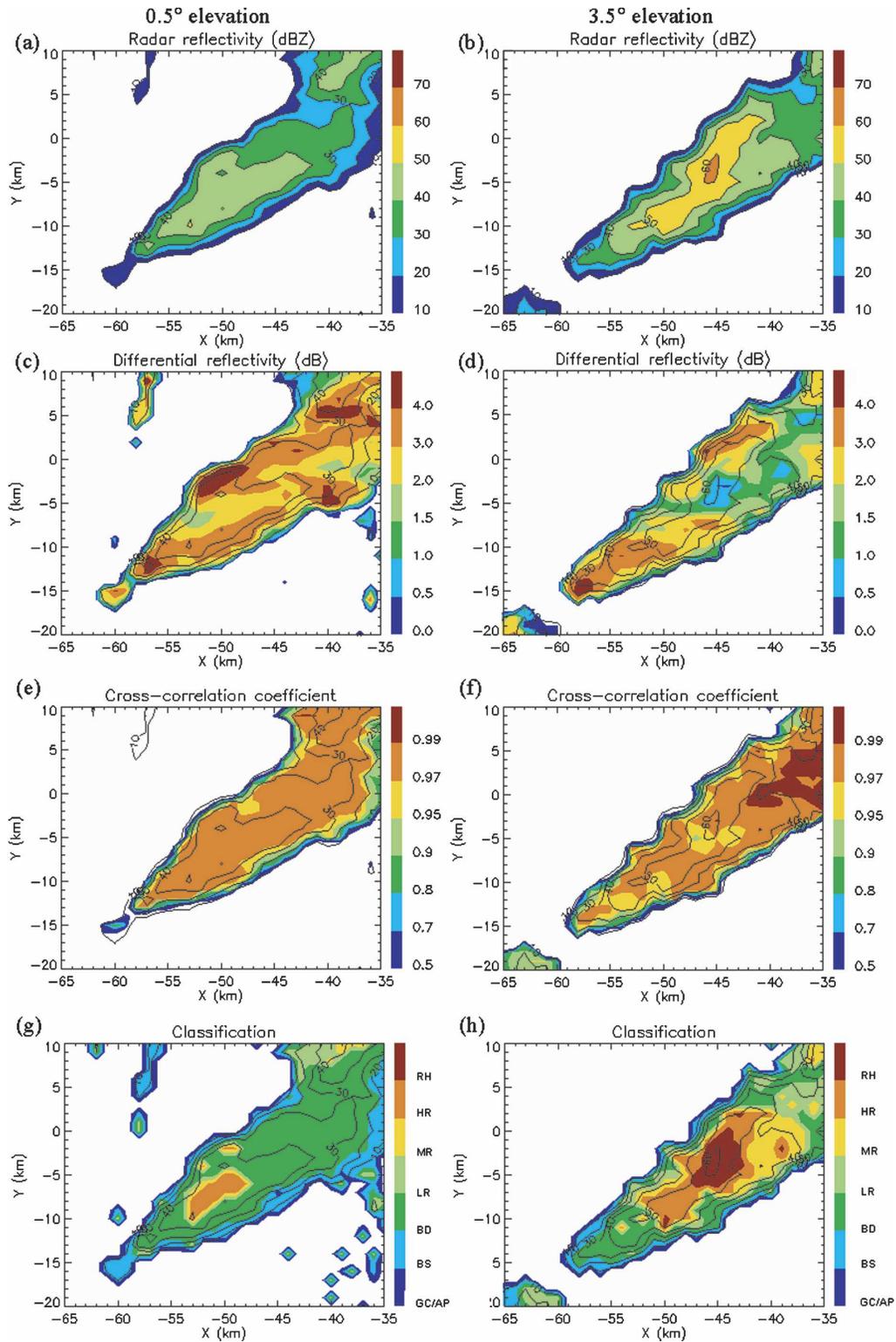


FIG. 4. At (left) 0.5° and (right) 3.5° elevations for (a), (b) Z , (c), (d) Z_{DR} , (e), (f) ρ_{hv} , and (g), (h) HCA classification at 2114 UTC 8 May 2003. Reflectivity contours overlay each panel.

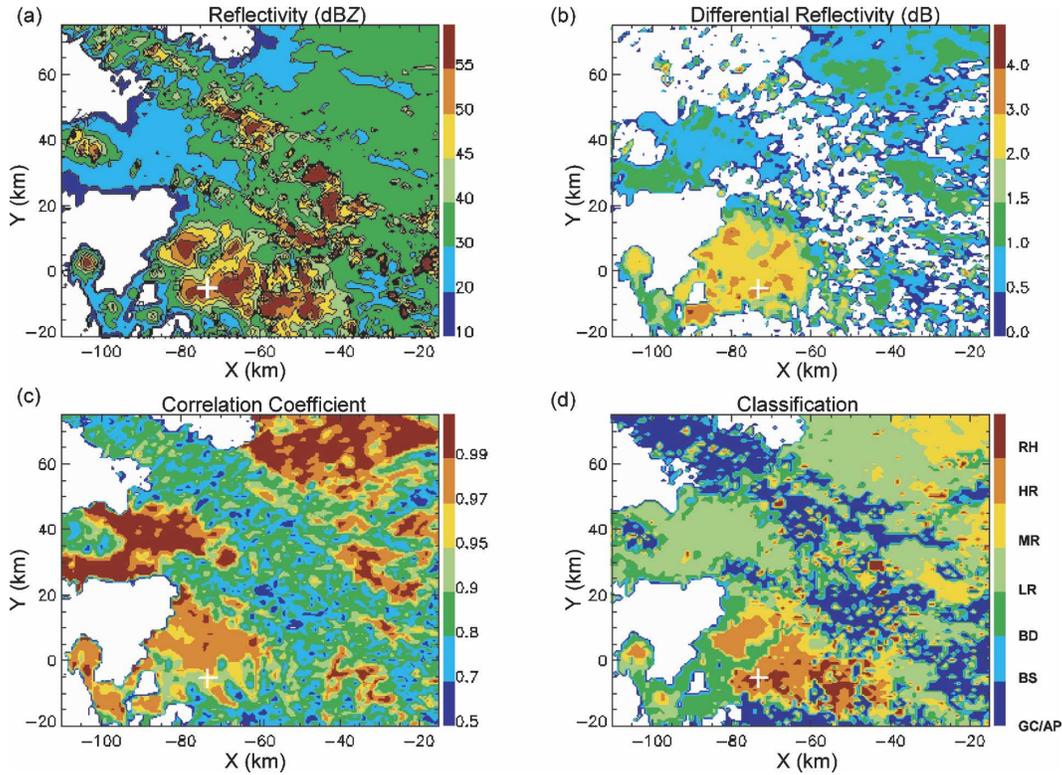


FIG. 5. The (a) Z , (b) Z_{DR} , (c) ρ_{hv} , and (d) classifications at 0.5° elevation at 1425 UTC on 1 Jul 2004. The white plus sign shows the location of 2.5-cm (1.0 in.) hail reported by *Storm Data* (NCDC 2004) at 1423 UTC.

Each hail report was matched with the storm cell whose 40-dBZ contour was located within the lowest acceptable distance and occurred within ± 6 min of the available radar data (both KOUN and KTLX). Acceptable distances varied from 3.2 to 5 km, depending on the speed of the storm movement. These criteria were used to consider the affects of storm movement, the possibility that the hail fell outside of the high-reflectivity core, and the high confidence in our ground truth.

Once each report was matched with a particular storm, the HDA and HCA output were validated. For the HDA, probabilities of hail 60% or higher were considered indicative of hail falling at the ground. Thus, HDA output was flagged as a hit when a hail probability 60% or higher corresponded with a hail report. The HDA output was flagged as a correct null when a hail probability less than 60% corresponded with a rain report. Because the HCA classifies hydrometeor type, HCA output was flagged as a hit when a region classified as hail was located within the acceptable distance of the hail report. The HCA output was flagged as a correct null when a region classified as rain was located within the acceptable distance of a rain report.

Using this rule base, a 2×2 contingency table was created for each day and all days combined, and was

used to quantify the accuracy and skill of the HCA and HDA. Within the contingency table (Table 3), a is a “hit,” b is a “false alarm,” c is a “miss,” and d is a “correct null” (Wilks 1995). These values are used to compute the following measures: probability of detection (POD),

$$\text{POD} = \frac{a}{a + c}; \tag{7}$$

false alarm rate (FAR),

$$\text{FAR} = \frac{b}{a + b}; \tag{8}$$

critical success index (CSI),

$$\text{CSI} = \frac{a}{a + b + c}; \text{ and} \tag{9}$$

and Heidke skill score (HSS),

$$\text{HSS} = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}. \tag{10}$$

Measures of accuracy and skill in Table 3b show that the performance of the HCA and the HDA varies

TABLE 3. (a) Example of a 2×2 contingency table, where a is a hit, b is a false alarm, c is a miss, and d is a correct null. (b) The HCA and HDA accuracy and skill measures.

(a)				
	Hail obs	Hail not obs		
Hail detected by the HCA or HDA	a	b		
Hail not detected by the HCA or HAD	c	d		

(b)				
Algorithm, date (2003)	POD	FAR	CSI	HSS
HCA, 1 May	100	17	94	88
HDA, 1 May	100	0	100	100
HCA, 14 May	100	0	100	100
HDA, 14 May	60	0	60	53
HCA, 19 May	100	0.1	94	88
HDA, 19 May	93	26	70	13
HCA, 10 Jun	100	0	100	100
HDA, 10 Jun	100	63	37	19

among the 4 days. Interestingly, the HDA attains perfect scores on 1 May 2003, while the HCA suffers from a few false alarms and hence, a higher FAR (17%) and lower CSI and HSS (94% and 88%, respectively; Table 3b). On this day, a chase vehicle intercepted an isolated LP supercell as it moved eastward from approximately 140 to 130 km west-northwest of KOUN. During this period (~1 h), the storm produced pea- to marble-size hail (Table 2) and rain at the ground. A different situation arose on the morning of 14 May 2003, when the HCA significantly outperforms the HDA (Table 3b). In this case, a chase vehicle intercepted a storm that produced very small hail (0.5–0.6 cm) within a west-to-east-oriented convective line 50–55 km west of KOUN. Owing to several misses, the HDA POD, CSI, and HSS are low compared to the corresponding HCA values (Table 3b).

On 19 May 2003, chase vehicles intercepted a classic supercell located about 90 km south of KOUN that dropped up to 4.5-cm hail (1.75 in.) on the ground. This supercell produced hail cyclically over several hours as it moved eastward. Although the HCA had only one false alarm, the HDA had several, which resulted in a higher FAR (26% versus 0.1%) and a lower HSS (13% versus 88%; Table 3b). The 10 June event was similar to the 14 May 2003 in that storm cells formed a convective line and produced relatively small size hail (0.5–0.9 cm). Like that event, the HCA achieved perfect performance. In contrast, the HDA produced several false alarms that resulted in a high FAR (63%) and low CSI and HSS (37% and 19%, respectively; Table 3).

In terms of overall accuracy and skill, the HCA outperforms the HDA. Comparison of these statistical

measures suggests that the NEXRAD HDA suffers more false alarms (39.5% versus 11.5%) and more misses (88% versus 100%) than the HCA. As a result, the NEXRAD HDA also has a substantially lower CSI (56% versus 89%) and HSS (31% versus 80%). A question arising from these results, and unexplored in Ryzhkov et al. (2005), is whether these differences in performance are statistically significant.

To test the statistical significance, we bootstrapped values composing the contingency table and compared the resulting distributions (Wilks 1995). Bootstrapping is a nonparametric technique where contingency table data associated with each algorithm are resampled with replacements (5000 times) and the statistics are recomputed to yield a distribution of each statistic (i.e., POD, FAR, CSI, and HSS). The resulting distributions indicate the accuracy of each statistic given the available data. Statistical significance is determined by comparing the 90th and 95th percentiles; an absence of overlapping values indicates statistical significance at the 90% and 95% confidence levels, respectively. A comparison of these statistical distributions indicates that only the POD and CSI statistics are statistically significant at both confidence levels. However, FAR and HSS are statistically different at the 90% confidence level. While the superior performance of NSSL's classifier, compared to the HDA, demonstrates a statistical advantage of using polarimetric variables to discriminate hail from rain, that advantage is tempered by the dataset's relatively small size and limited geographic coverage. Furthermore, this statistical comparison did not consider hail size estimation or predictability. The planned polarimetric upgrade of the WSR-88D network will provide the opportunity to evaluate polarimetric and conventional hail classification algorithms in a more complete manner.

Regardless of these limitations, this study shows that high performance can be achieved without LDR. This assertion is supported by a recent study showing superior hail diagnosis using the hail parameter (H_{DR}) and the hail quadrature parameter (HQP), as compared to LDR (Depue and Rutledge 2003). The work of Depue and Rutledge is the only other study that validates these statistically polarimetric hail diagnosis techniques. Because their validation dataset was collected during a different time (summers of 1992 and 1993) and over a different region (northeastern Colorado), and is larger than the current dataset, their statistical results are not directly comparable to ours.

6. Summary

This paper describes one of the versions of the NSSL's polarimetric fuzzy logic hydrometeor classifica-

tion algorithm (HCA) that is tailored for hail detection, illustrates the HCA's operational capabilities in terms of rain and hail discrimination, and presents a comparative validation of hail detection by the HCA and HDA. The HCA's capabilities and statistical validation are demonstrated using polarimetric data collected mostly during JPOLE (Ryzhkov et al. 2005; Scharfenberg et al. 2005). The paper's operational focus is motivated by the upcoming polarimetric upgrade of the WSR-88D network. Within the HCA, values of four variables Z , Z_{DR} , ρ_{hv} , and $SD(Z)$ are used within asymmetric membership functions to discriminate seven classes. These seven classes are ground clutter/anomalous propagation, biological scatterers (insects, birds, and bats), big drops, light rain, moderate rain, heavy rain, and rain mixed with hail.

The HCA's viability as an operational algorithm is illustrated for three cases. The first case (19 May 2003) shows the HCA's capability to detect regions of severe hail (2.5 cm) within a supercell at 0.5° elevation. The second case (8 May 2003) demonstrates the classification of hail below the melting layer in a developing storm (3.5° elevation) that later evolved into a tornadic supercell. The third case (1 July 2004) illustrates the HCA's ability to discriminate regions of rain and hail from regions of surrounding AP.

The HCA and HDA outputs are validated for four cases collected during JPOLE, including an LP supercell (1 May 2003), a classic supercell (19 May 2003), and two convective lines (14 May and 10 June 2003). Statistical measures computed from contingency tables show that the HCA outperforms the HDA. In comparison to the HDA, the HCA's overall higher POD (100% versus 88%) and lower FAR (11% versus 39%) indicates that the HCA suffers from fewer misses and false alarms in these cases. As a result, the HCA attains a higher overall CSI (89% versus 56%) and HSS (80% versus 31%). Furthermore, bootstrapping the data reveals that differences in the distributions of POD and CSI are statistically significant at the 95% confidence level, whereas differences in FAR and HSS are statistically significant at the 90% confidence level. These results indicate that a relatively simple version of the polarimetric classification algorithm provides significantly higher quality of hail detection than a conventional hail detection algorithm currently utilized on single-polarization WSR-88D radars.

Acknowledgments. Funding for this study was provided by the NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, the U.S. Department of Commerce, and the U.S. National

Weather Service, the Federal Aviation Administration, and the Air Force Weather Agency through the NEXRAD Products Improvement Program. We also thank contributions from storm-intercept crews and operators of KOUN. Additionally, we thank Kim Elmore for his advice and discussions about bootstrapping.

REFERENCES

- Amburn, S. A., and P. L. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12**, 473–478.
- Atlas, D., and F. H. Ludlum, 1961: Multi-wavelength radar reflectivity of hailstorms. *Quart. J. Roy. Meteor. Soc.*, **87**, 523–534.
- Auer, A. H. J., 1994: Hail recognition through the combined use of radar reflectivity and cloud-top temperatures. *Mon. Wea. Rev.*, **122**, 2218–2221.
- Aydin, K., T. A. Seliga, and V. Balaji, 1986: Remote sensing of hail with a dual linear polarization radar. *J. Climate Appl. Meteor.*, **25**, 1475–1484.
- Balakrishnan, N., and D. S. Zrnić, 1990a: Estimation of rain and hail rates in mixed-phase precipitation. *J. Atmos. Sci.*, **47**, 565–583.
- , and —, 1990b: Use of polarization to characterize precipitation and discriminate large hail. *J. Atmos. Sci.*, **47**, 1525–1540.
- Billet, J., M. DeLisi, B. G. Smith, and C. Gates, 1997: Use of regression techniques to predict hail size and the probability of large hail. *Wea. Forecasting*, **12**, 154–164.
- Bluestein, H. B., and C. R. Parks, 1983: A synoptic and photographic climatology of low-precipitation severe thunderstorms in the southern plains. *Mon. Wea. Rev.*, **111**, 2034–2046.
- Brandes, E. A., and A. V. Ryzhkov, 2004: Hail detection with polarimetric radar. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., CD-ROM, P5.10.
- Bringi, V. N., and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press, 636 pp.
- , J. Vivekanandan, and J. D. Tuttle, 1986: Multiparameter radar measurements in Colorado convective storms. Part II: Hail detection studies. *J. Atmos. Sci.*, **43**, 2564–2577.
- , V. Chandrasekar, N. Balakrishnan, and D. S. Zrnić, 1990: An examination of propagation effects in rainfall on radar measurements at microwave frequencies. *J. Atmos. Oceanic Technol.*, **7**, 829–840.
- Depue, T. K., and S. A. Rutledge, 2003: Ground truth and modeling verification of the hail quadrature parameter. M.S. thesis, Dept. of Atmospheric Science, Colorado State University, 154 pp.
- Donaldson, R. J. J., 1959: Analysis of severe convective storms observed by radar—II. *J. Meteor.*, **16**, 281–287.
- Doviak, R., and D. Zrnić, 1993: *Doppler Radar and Weather Observations*. Academic Press, 562 pp.
- Dye, J. E., and B. E. Martner, 1978: The relationship between radar reflectivity factor and hail at the ground for northeast Colorado thunderstorms. *J. Appl. Meteor.*, **17**, 1335–1341.
- Eccles, P. J., and D. Atlas, 1973: A dual-wavelength radar hail detector. *J. Appl. Meteor.*, **12**, 847–854.
- Edwards, R., and R. L. Thompson, 1998: Nationwide comparisons of hail size with WSR-88D vertically integrated liquid water

- and derived thermodynamic sounding data. *Wea. Forecasting*, **13**, 277–285.
- Elevander, R. C., 1977: Relationships between radar parameters observed with objectively defined echoes and reported severe weather occurrences. Preprints, *10th Conf. on Severe Local Storms*, Omaha, NE, Amer. Meteor. Soc., 73–76.
- Féral, L., H. Sauvageot, and S. Soula, 2003: Hail detection using S- and C-band radar reflectivity difference. *J. Atmos. Oceanic Technol.*, **20**, 233–248.
- Foote, G. B., and C. A. Knight, 1979: Results of a randomized hail suppression experiment in northeast Colorado. Part I: Design and conduct of the experiment. *J. Appl. Meteor.*, **18**, 1526–1537.
- Geotis, S. G., 1963: Some radar measurements of hailstorms. *J. Appl. Meteor.*, **2**, 270–275.
- Greene, D. R., and R. A. Clark, 1972: Vertically integrated liquid water—A new analysis tool. *Mon. Wea. Rev.*, **100**, 548–552.
- Hardaker, P. J., and A. H. Auer Jr., 1994: The separation of rain and hail using single polarisation radar echoes and IR cloud-top temperatures. *Appl. Meteor.*, **1**, 201–204.
- Holler, H., V. N. Bringi, J. Hubbert, M. Hagen, and P. F. Meischner, 1994: Life cycle and precipitation formation in hybrid-type hailstorm revealed by polarimetric and Doppler radar measurements. *J. Atmos. Sci.*, **51**, 2500–2522.
- Keenan, T. D., 2003: Hydrometeor classification with a C-band polarimetric radar. *Aust. Meteor. Mag.*, **52**, 23–31.
- Kennedy, P. C., S. A. Rutledge, W. A. Peterson, and V. N. Bringi, 2001: Polarimetric radar observations of hail formation. *J. Appl. Meteor.*, **40**, 1347–1366.
- Kitzmler, D. H., W. E. McGovern, and R. E. Saffle, 1995: The WSR-88D severe weather potential algorithm. *Wea. Forecasting*, **10**, 141–159.
- Leitao, M. J., and P. A. Watson, 1984: Application of dual linearly polarized radar data to prediction of microwave path attenuation at 10–30 GHz. *Radio Sci.*, **19**, 209–221.
- Lim, S., V. Chandrasekar, and V. N. Bringi, 2005: Hydrometeor classification system using dual-polarization radar measurements: Model improvements and in situ verification. *IEEE Trans. Geosci. Remote Sens.*, **43**, 792–801.
- Liu, H., and V. Chandrasekar, 2000: Classification of hydrometeors based on polarimetric radar measurements: Development of fuzzy logic and neuro-fuzzy systems, and in situ verification. *J. Atmos. Oceanic Technol.*, **17**, 140–164.
- Maddox, R. A., D. S. Zaras, P. L. MacKeen, J. J. Gourley, R. Rabin, and K. W. Howard, 1999: Echo height measurements with the WSR-88D: Use of data from one versus two radars. *Wea. Forecasting*, **14**, 455–460.
- Mather, G. K., D. Treddenick, and R. Parsons, 1976: An observed relationship between the height of the 45 dBZ contours in storm profiles and surface hail reports. *J. Appl. Meteor.*, **15**, 1336–1340.
- NCDC, 2003: *Storm Data*. Vol. 45, No. 5, 34 pp.
- , 2004: *Storm Data*. Vol. 46, No. 7, 36 pp.
- Petrocchi, P. J., 1982: Automatic detection of hail by radar. AFGL-TR-82-0277, Environmental Research Paper 796, Hanscom AFB, MA, 33 pp.
- Rinehart, R. E., and D. W. Staggs, 1968: Use of radar to delineate surface hail areas in weather modification experiments. *First National Conf. on Weather Modification*, Albany, NY, Amer. Meteor. Soc., 464–473.
- Ryzhkov, A. V., 2005: On the use of differential phase for polarimetric rainfall measurements. A new approach to K_{DP} estimation. Preprints, *32d Conf. on Radar Meteorology*, Albuquerque, NM, Amer. Meteor. Soc., CD-ROM, P9R.8
- , and D. S. Zrnić, 1994: Precipitation observed in Oklahoma mesoscale systems with polarimetric radar. *J. Appl. Meteor.*, **33**, 455–464.
- , and —, 1998: Beamwidth effects on the differential phase measurements of rain. *J. Atmos. Oceanic Technol.*, **15**, 624–634.
- , T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnić, 2005: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824.
- Sarchilli, G., E. Gorgucci, V. Chandrasekar, and D. Dobaie, 1996: Self consistency of polarization diversity measurements of rainfall. *IEEE Trans. Geosci. Remote Sens.*, **34**, 22–26.
- Scharfenberg, K. A., and Coauthors, 2005: The Joint Polarization Experiment: Polarimetric radar in forecasting and warning decision-making. *Wea. Forecasting*, **20**, 775–788.
- Smyth, T. J., T. M. Blackman, and A. J. Illingworth, 1999: Observations of oblate hail using dual polarization radar and implications for hail-detection schemes. *Quart. J. Roy. Meteor. Soc.*, **125**, 993–1016.
- Straka, J. M., D. S. Zrnić, and A. V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39**, 1341–1372.
- Vivekanandan, J., D. S. Zrnić, S. M. Ellis, R. Oye, A. V. Ryzhkov, and J. Straka, 1999: Cloud microphysics retrieval using S-band dual-polarization measurements. *Bull. Amer. Meteor. Soc.*, **80**, 381–388.
- Waldvogel, A., and B. Federer, 1976: Large raindrops and the boundary between rain and hail. Preprints, *17th Conf. on Radar Meteorology*, Seattle, WA, Amer. Meteor. Soc., 167–172.
- , —, and P. Grimm, 1979: Criteria for the detection of hail cells. *J. Appl. Meteor.*, **18**, 1521–1525.
- 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, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286–303.
- Zeng, Z., S. E. Yuter, R. A. House, and D. E. Kingsmill, 2001: Microphysics of the rapid development of heavy convective precipitation. *Mon. Wea. Rev.*, **129**, 1882–1904.
- Zrnić, D. S., and A. V. Ryzhkov, 1999: Polarimetry for weather surveillance radars. *Bull. Amer. Meteor. Soc.*, **80**, 389–406.
- , —, J. Straka, Y. Liu, and J. Vivekanandan, 2001: Testing a procedure for the automatic classification of hydrometeor types. *J. Atmos. Oceanic Technol.*, **18**, 892–913.