Polarimetric Radar Characteristics of Melting Hail. Part III: Validation of the Algorithm for Hail Size Discrimination

Kiel L. Ortega, John M. Krause, and Alexander V. Ryzhkov

*Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR
National Severe Storms Laboratory, Norman, Oklahoma*

ABSTRACT

This study is the third part of a paper series investigating the polarimetric radar properties of melting hail and application of those properties for operational polarimetric hail detection and determination of its size. The results of theoretical simulations in Part I were used to develop a hail size discrimination algorithm (HSDA) described in Part II. The HSDA uses radar reflectivity Z, differential reflectivity Z_{DR}, and cross-correlation coefficient ρ_{hv} along with melting-level height within a fuzzy-logic scheme to distinguish among three hail size classes: small hail (with diameter $D < 2.5$ cm), large hail ($2.5 < D < 5.0$ cm), and giant hail ($D > 5.0$ cm). The HSDA validation is performed using radar data collected by numerous WSR-88D sites and more than 3000 surface hail reports obtained from the Severe Hazards Analysis and Verification Experiment (SHAVE). The original HSDA version was modified in the process of validation, and the modified algorithm demonstrates probability of detection of 0.594, false-alarm ratio of 0.136, and resulting critical success index (CSI) equal to 0.543. The HSDA outperformed the current operational single-polarization hail detection algorithm, which only provides a single hail size estimate per storm and is characterized by CSI equal to 0.324. It is shown that HSDA is particularly sensitive to the quality of Z_{DR} measurements, which might be affected by possible radar miscalibration and anomalously high differential attenuation.

1. Introduction

a. Hail identification and sizing by radar

The identification of hail is important for the U.S. National Weather Service (NWS) severe weather warnings, especially detection of hail exceeding the severe diameter threshold of 25 mm (1 in.; NWS 2014). Currently many hail detection and sizing algorithms and techniques for NWS operations are based upon single-polarization radar, such as vertically integrated liquid (VIL; Greene and Clark 1972) and VIL density (Amburn and Wolf 1997). Specific hail size estimates for a storm are provided through the hail detection algorithm (HDA; Witt et al. 1998a). The HDA combines a vertical profile of reflectivity and near-storm temperature profiles based upon the work of Waldvogel et al. (1979). Using manual interrogation of storms, Donavon and Jungbluth (2007) took a similar approach in combining thermodynamic information and reflectivity profiles to create warning threshold guidance using the 50-dBZ heights above the melting level. Doppler velocity data have also been explored for use in determining hail size, especially the relationship of storm-top divergence to the resulting hail size (Witt and Nelson 1991; Boustead 2008). NWS forecasters are also trained in using three-body scatter spikes (Lemon 1998) and updraft proxy signatures, like bounded weak echo regions, to infer severe hail within thunderstorms (http://www.wdtd.noaa.gov/courses/rac/severe/svr-hail/presentation.html). Heinselman and Ryzhkov (2006) provide an overview of other hail detection techniques and research using single-polarization radars.

The United States' WSR-88D network recently underwent an upgrade in capabilities via implementation of dual polarization. This upgrade allows NWS operational forecasters to use polarimetric variables to assist in forecast and warning duties. For hail detection, there is currently only one operational algorithm in place: the hydrometeor classification algorithm (HCA; Park et al. 2009). The HCA does not discriminate hail size and only identifies regions of rain and hail mixture. However,

forecasters are given training in the identification of hail using polarimetric variables and manual interrogation of thunderstorms (http://www.wdtd.noaa.gov/courses/rac/severe/svr-hail/presentation.html). Further, the integration of dual-polarization radar data into operations was shown to increase forecaster confidence in hail identification during the Joint Polarization Experiment (JPOLE; Ryzhkov et al. 2005; Scharfenberg et al. 2005).

Heinselman and Ryzhkov (2006) summarize previous studies that investigated hail detection using polarimetric capabilities. More recent studies have investigated the properties of melting hail and applications of those findings, including hail size discrimination. Ryzhkov et al. (2013a, hereinafter Part I) investigated hail using 1D and 2D cloud models combined with T-matrix calculations to simulate polarimetric variables. The 1D model was based upon the work of Rasmussen and Heymsfield (1987a,b) and was used to investigate the impact of hail size distributions, hail densities, and thermodynamic profiles on polarimetric variables. The 2D model, the Hebrew University of Jerusalem Cloud Model (HUCM; Khain et al. 2004, 2011), was used to investigate the impact of size sorting, which can have strong impact on polarimetric variables. The findings were applied to create an initial hail size discrimination algorithm (HSDA; Ryzhkov et al. 2013b, hereinafter Part II), which is designed to augment the HCA with hail size estimates. The HSDA is a fuzzy-logic scheme that determines one of the three hail size classes using reflectivity at horizontal polarization Z_h, differential reflectivity Z_{DR}, and cross-correlation coefficient ρ_{hv}. The HSDA accounts for the degree of hail melting by utilizing separate membership functions for Z_h, Z_{DR}, and ρ_{hv} for six different height layers with respect to the melting level.

b. Reports of surface hail fall and sizing

The most common source of hail reports at the surface is *Storm Data*, which is the official publication of the National Oceanic and Atmospheric Administration for the documentation of rare, significant, or damaging storms and meteorological phenomena (NWS 2007). However, limitations of the data have been documented (Amburn and Wolf 1997; Witt et al. 1998b; Blair and Leighton 2012). These limitations of the hail reports within *Storm Data* are due to the fact that the reports are primarily for NWS severe weather warning verification and thus the temporal/spatial spacing between the reports can be quite coarse (on the order of tens of minutes and tens of kilometers). Further, since *Storm Data* is primarily intended for the verification of warnings, it rarely has events of nonsevere magnitude. These characteristics severely limit *Storm Data*'s application for algorithm verification.

Only two previous studies have extensively verified polarimetric hail detection with reports not originating from *Storm Data*. Heinselman and Ryzhkov (2006) used hail reports from storm-intercept vehicles operating during JPOLE. Depue et al. (2007) used door-to-door ground surveys of impacted populations to determine hail size. Another study used a storm-intercept vehicle to make in situ measurements of hail for a single storm in eastern Colorado and compared the observations with polarimetric data (Hubbert et al. 1998). The National Severe Storms Laboratory, in response to the need for reports with higher spatial resolution and fairly large geographic coverage, began conducting the Severe Hazards Analysis and Verification Experiment (SHAVE) in 2006 (Ortega et al. 2009). The primary goal for SHAVE was to collect hail reports—including information on maximum and average hail sizes, and the start and end time of the hail fall—via telephone surveys, in the wake of thunderstorms at spacings near 1.6 km,[1] including reports of nonsevere hail and no hail along and near the thunderstorm's path. The inclusion of nonsevere and null events within the SHAVE dataset makes those reports more efficient for algorithm evaluation than those within *Storm Data*. The current study will make exclusive use of SHAVE hail reports (http://www.nssl.noaa.gov/projects/shave/archive.php) for HSDA verification.

2. Hail size discrimination algorithm

a. HSDA description

The initial version of HSDA has been introduced in Part II as an extension of the existing WSR-88D HCA, which identifies the class "hail mixed with rain" (Park et al. 2009). The HSDA splits this class designation into three categories of hail size: small (diameters < 25 mm), large (diameters between 25 and 50 mm), and giant (diameters > 50 mm). These classes were chosen on the basis of current NWS thresholds for severe-hail criteria.

The HSDA operates after the HCA and only for pixels identified as "rain/hail." The HSDA uses Z_h, Z_{DR}, and ρ_{hv} within a fuzzy-logic scheme to calculate an aggregation value A_i for three hail classes as in Park et al. (2009):

$$A_i = \frac{\sum\limits_{j=1}^{3} W_j Q_j P^i(V_j)}{\sum\limits_{j=1}^{3} W_j Q_j}, \tag{1}$$

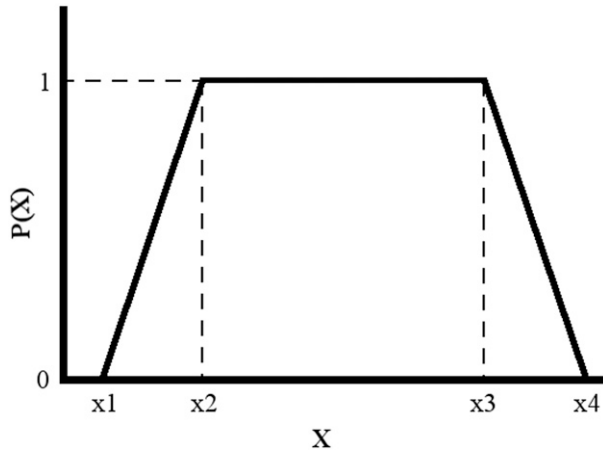[1] Approximate spacing of road networks in rural areas of the central plains of the United States.

FIG. 1. Trapezoidal membership function, where x can be Z_h, Z_{DR}, or ρ_{hv} and x_1, x_2, x_3, and x_4 are bounds defined in Table 2.

TABLE 1. Weights of polarimetric variables for each height layer used in the $\mathbf{W}$ vector.

Height layer	Z_h	Z_{DR}	ρ_{hv}
$H_b \geq H(T_w = -25°C)$	1.0	0.3	0.6
$H(T_w = -25°C) < H_b \leq H(T_w = 0°C)$	1.0	0.3	0.6
$H(T_w = 0°C) < H_b \leq H(T_w = 0°C) - 1\,km$	0.8	0.5	0.6
$H(T_w = 0°C) - 1\,km < H_b \leq H(T_w = 0°C) - 2\,km$	0.7	0.8	0.6
$H(T_w = 0°C) - 2\,km < H_b \leq H(T_w = 0°C) - 3\,km$	0.7	1.0	0.6
$H_b < H(T_w = 0°C) - 3\,km$	0.7	1.0	0.6

where $P^{(i)}(V_j)$ is a membership function that characterizes the distribution of the jth radar variable for the ith hail size class, and Q_j and W_j are weights between 0 and 1 assigned to the jth variable. The classification decision is based on the maximal aggregation value. The membership functions for the three radar variables ($V_1 = Z_h$, $V_2 = Z_{DR}$, $V_3 = \rho_{hv}$) are assumed to have trapezoidal shape with a maximal value of 1 and a minimal value of 0 (Fig. 1). The elements of the confidence vector $\mathbf{Q}$ characterize the accuracy of radar measurements, which can be biased and noisy. The $\mathbf{Q}$ vector used within the HSDA is the same as the $\mathbf{Q}$ vector in the HCA (Park et al. 2009). The $\mathbf{Q}$ vector takes into account several factors causing data quality degradation, such as attenuation, partial beam blockage, and nonuniform beamfilling. Some of these effects intensify with range, which results in a general decrease of the elements of $\mathbf{Q}$ with increasing range. The vector of weights $\mathbf{W}$ characterizes the discrimination efficiency of each variable at each height layer (Table 1). All elements of $\mathbf{W}$ were set to 1 in the HSDA of Part II.

The membership functions vary for each of the three radar variables depending upon the height of the radar sampling volume with respect to six height intervals (see Fig. 1 within Part II). The six height intervals are defined by location of the center of the radar sampling volume (H_b) with respect to the heights of wet-bulb temperature, $H(T_w)$, equal to 0° and −25°C:

1) $H(T_w = 0°C) - 3\,km > H_b$,
2) $H(T_w = 0°C) - 3\,km \leq H_b < H(T_w = 0°C) - 2\,km$,
3) $H(T_w = 0°C) - 2\,km \leq H_b < H(T_w = 0°C) - 1\,km$,
4) $H(T_w = 0°C) - 1\,km \leq H_b < H(T_w = 0°C)$,
5) $H(T_w = 0°C) \leq H_b < H(T_w = -25°C)$,   and
6) $H_b \geq H(T_w = -25°C)$.

b. Modifications of the Part II version of the HSDA

The HSDA presented in Part II (hereinafter referred to as original HSDA) was modified during the evaluation process for the current study. The changes made during the course of this study include 1) modifying the membership functions, 2) defining some Z_{DR} membership function bounds as functions of Z_h, 3) adding a tunable ΔZ_{DR} parameter to the Z_{DR} membership functions, and 4) modifying the $\mathbf{W}$ vector with different weights for each variable per height interval. The new weights $\mathbf{W}$ are listed in Table 1, and the new membership functions are summarized in Table 2, where

$$f_1 = -0.5 + 2.5 \times 10^{-3} \times Z_h$$
$$+ 7.5 \times 10^{-4} \times Z_h^2 + \Delta Z_{DR}, \quad (2)$$

$$f_2 = 0.1 \times (Z_h - 50.0) + \Delta Z_{DR}, \quad (3)$$

$$f_3 = 0.1 \times (Z_h - 60.0) + \Delta Z_{DR}, \quad (4)$$

$$g_1 = -0.9 + 1.5 \times 10^{-2} \times Z_h$$
$$+ 5.0 \times 10^{-4} \times Z_h^2 + \Delta Z_{DR}, \quad (5)$$

$$g_2 = 0.075 \times (Z_h - 50.0) + \Delta Z_{DR}, \quad \text{and} \quad (6)$$

$$g_3 = 0.075 \times (Z_h - 60.0) + \Delta Z_{DR}. \quad (7)$$

The changes to the membership functions were the result of differences in the variable distributions from observations compared to those from the modeling results presented in Part I. Further, observations of Z_h and Z_{DR} at the warmer height layers revealed that a two-dimensional membership function for Z_{DR}, with its parameters based upon Z_h, is appropriate within those layers. A ΔZ_{DR} parameter was added to Eqs. (2)–(7) to handle possible Z_{DR} calibration deficiencies or forecaster preference for larger hail size designations by the HSDA. Changes in the relative importance of the different radar variables in discriminating hail sizes within different height layers required the use of varying weights per height layer for each input variable. The quantity Z_{DR} is the most informative parameter in the interval of heights lower than 2 km below the melting

TABLE 2. Membership functions (MF) for height layers and hail classes for modified HSDA.

Height layer	MF bound	Small hail			Large hail			Giant hail		
		Z_h	Z_{DR}	ρ_{hv}	Z_h	Z_{DR}	ρ_{hv}	Z_h	Z_{DR}	ρ_{hv}
$H_b \geq H(T_w = -25°C)$	x_1	45	−0.50	0.92	48	−0.50	0.92	50	−8.75	−1.00
	x_2	50	−0.30	0.96	58	−0.30	0.96	60	−7.75	0.00
	x_3	60	0.30	0.99	63	0.30	0.99	100	0.30	0.99
	x_4	65	0.50	1.00	68	0.50	1.00	101	0.50	1.00
$H(T_w = -25°C) < H_b \leq H(T_w = 0°C)$	x_1	45	−0.50	0.92	48	−0.50	0.86	50	−8.75	−1.00
	x_2	50	−0.30	0.96	58	−0.30	0.90	60	−7.75	0.00
	x_3	60	0.30	0.99	63	0.30	0.96	100	0.20	0.93
	x_4	65	0.50	1.00	68	0.50	0.98	101	0.50	0.98
$H(T_w = 0°C) < H_b \leq H(T_w = 0°C) - 1 \, km$	x_1	45	−0.10	0.93	48	−0.30	0.80	50	−8.75	−1.00
	x_2	50	0.30	0.96	58	0.10	0.91	60	−7.75	0.00
	x_3	60	0.70	0.99	63	0.50	0.97	100	0.20	0.94
	x_4	65	1.20	1.00	68	1.00	0.98	101	0.70	0.98
$H(T_w = 0°C) - 1 \, km < H_b \leq H(T_w = 0°C) - 2 \, km$	x_1	45	$g_2 - 0.3$	0.94	50	$g_3 - 0.3$	0.80	52	−8.75	−1.00
	x_2	52	g_2	0.96	60	g_3	0.91	62	−7.75	0.00
	x_3	62	g_1	0.98	65	g_2	0.97	100	g_3	0.96
	x_4	67	$g_1 + 0.3$	1.00	70	$g_2 + 0.3$	0.98	101	$g_3 + 0.3$	0.98
$H(T_w = 0°C) - 2 \, km < H_b \leq H(T_w = 0°C) - 3 \, km$	x_1	45	$f_2 - 0.3$	0.91	50	$f_3 - 0.3$	0.8	50	−8.75	−1.00
	x_2	49	f_2	0.94	57	f_3	0.90	59	−7.75	0.00
	x_3	59	f_1	0.96	62	f_2	0.96	100	f_3	0.93
	x_4	64	$f_1 + 0.3$	0.99	67	$f_2 + 0.3$	0.99	101	$f_3 + 0.3$	0.98
$H_b < H(T_w = 0°C) - 3 \, km$	x_1	45	$f_2 - 0.3$	0.91	50	$f_3 - 0.3$	0.8	50	−8.75	−1.00
	x_2	47	f_2	0.94	55	f_3	0.90	57	−7.75	0.00
	x_3	57	f_1	0.96	60	f_2	0.96	100	f_3	0.93
	x_4	62	$f_1 + 0.3$	0.99	65	$f_2 + 0.3$	0.99	101	$f_3 + 0.3$	0.98

level where contributions of hailstones and raindrops originating from melting hail to Z_{DR} are generally comparable. The Z_{DR} is the least important variable compared to Z_h and ρ_{hv} at higher levels, and Z_h is the most informative radar parameter at higher altitudes.

New limitations were placed upon the algorithm to develop spatial coherency for the HSDA designations: 1) if the membership function value for any of the three polarimetric parameters was less than 0.2, the aggregation value for the associated hail size designation was set to zero; 2) if the largest aggregation score for the hail size designations did not exceed 0.6, the HSDA assigned small hail to the pixel; 3) if the designation was for large or giant hail and $Z_{DR} \geq 2 \, dB$, the designation was changed to small hail; and 4) a "despeckle" method along each radial downgraded isolated, single pixels of giant hail to large hail and isolated, single pixels of large hail to small hail. If no designation could be made following the above rules the default designation was small hail. Output from the new HSDA is presented in Figs. 2 and 3.

c. Radar data and HSDA configurations

Heights of $T_w = 0°C$ and $T_w = -25°C$ were obtained from the hourly RAP or RUC analysis for the radar's location. Radar data were processed through the WSR-88D preprocessor (Istok et al. 2009), which applies corrections to Z_{DR} for attenuation and system calibrations and smooths the polarimetric variables along the radial. The smoothed fields from the preprocessor are then used within the HCA and HSDA and for the matching to hail reports. In all, eight versions of the HSDA were tested: the original HSDA, original HSDA with the added weights (Table 1), and six versions of the modified HSDA described in Tables 1 and 2 with ΔZ_{DR} set to values between −0.5 and 0 dB at 0.1-dB increments, where increasingly positive values of ΔZ_{DR} generally result in more designations of large and giant hail by the HSDA (Fig. 4).

3. Hail reports and matching techniques

SHAVE data collection operations within 120 km of polarimetric WSR-88D units were considered for this study. The maximum size from the SHAVE reports was used for comparisons to radar parameters and HSDA output. The HSDA was not available during SHAVE operations, thus the collection of the reports was independent of the HSDA output. The primary guidance for where to make verification phone calls was the multiradar, multisensor maximum expected size of hail product (Lakshmanan et al. 2006). Operations resulting in report swaths that exhibited high spatial resolution between the reports and good completeness of the
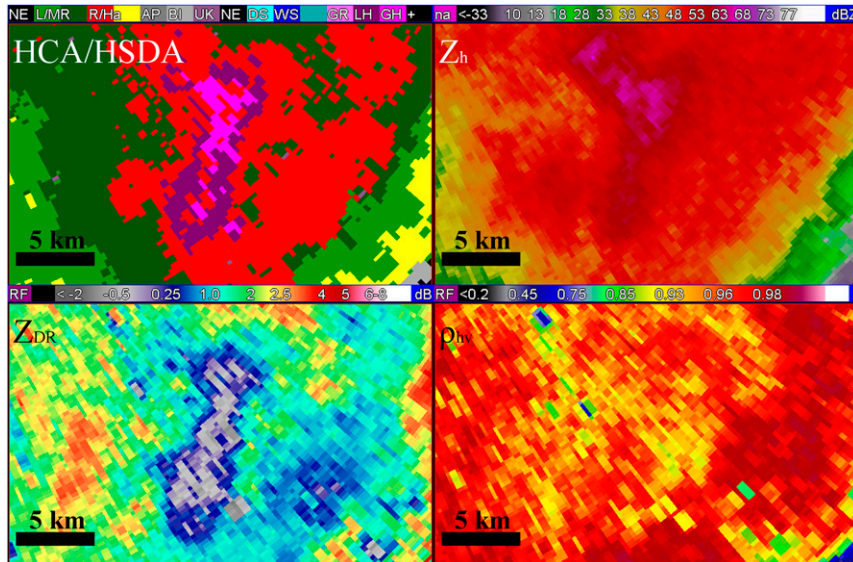
FIG. 2. (top left) HCA and HSDA output, (top right) Z_h, (bottom left) Z_{DR}, and (bottom right) ρ_{hv} for the storm producing giant hail approximately 47 km from the KDMX (Des Moines, IA) WSR-88D at 1856 UTC 30 Jun 2014. The key in the top-left panel corresponds to the categories listed in Fig. 3, below.

report swath (e.g., no large areas without reports) were evaluated. In this study, 79 SHAVE operations spanning the years 2010 through 2014 were used. These operations yielded 3257 total reports (1115 no-hail reports, 1150 small-hail reports, 786 large-hail reports, and 206 giant-hail reports) across the contiguous United States (Fig. 5). Seven of the operations were conducted on storms that produced only small-hail or no-hail reports, and 26 operations yielded at least a single giant-hail report. The median spacing between the reports was 2.15 km, and the interdecile and interquartile ranges of the hail diameter differences between neighboring reports were approximately ±20 and ±10 mm, respectively.

The reported hail fall times within SHAVE data are problematic because SHAVE operators do not ask for the time when hail with the maximum size fell (instead, they ask for the start and end times of the hail fall in general). In this study, the size of hail and its location are compared to the polarimetric variables and HSDA output. Thus, it is difficult to match precisely the maximum hail size to radar volumes during the time when the hail with maximum size was falling. Further, reports of "no hail" cannot have an actual time applied as the reports serve as verification of no event occurring. However, SHAVE points are precisely located since the location of the person called is known. Four different matching techniques were employed to obtain vertical "matching" profiles of Z_h, Z_{DR}, and ρ_{hv} corresponding to the given SHAVE hail report. These techniques relied upon maximum Z_h at the exact point location of the

report and within a 4 km × 4 km search window. The matching techniques used for this study, and illustrated in Fig. 6, included the following:

1) For "point match," the volume with the maximum Z_h at the 0.5° tilt at the exact report location was used to create the matching vertical profiles of the three radar variables. The Z_h, Z_{DR}, and ρ_{hv} values come from the pixel at the exact report location at each tilt for the matched volume. The corresponding values of Z_h in the matching volume marked as V_2 are shown with open circles in the diagram in Fig. 6. Two primary limitations are associated with this method: the first is imprecise matching due to offset of time of radar volume and actual time of observed hail fall (which would lead to spatial discrepancies); the second is imprecise matching to higher elevations due to storm motion or storm tilt.

2) For "point match (max tilts)," the volume with the maximum Z_h at the 0.5° tilt and exact report location served as the match for the 0.5° tilt and as a reference point (again, V_2 in Fig. 6.) Two additional volumes at times preceding (V_1) and succeeding (V_3) the matched volume at V_2 were also considered for finding maximal values of Z_h at higher elevations than 0.5°. For each tilt above 0.5°, the volume with the maximum Z_h at the exact report location from any of the three successive volumes (at V_1, V_2, and V_3) was used as the match (dashed open circles in Fig. 6) The Z_h, Z_{DR}, and ρ_{hv} values come from the
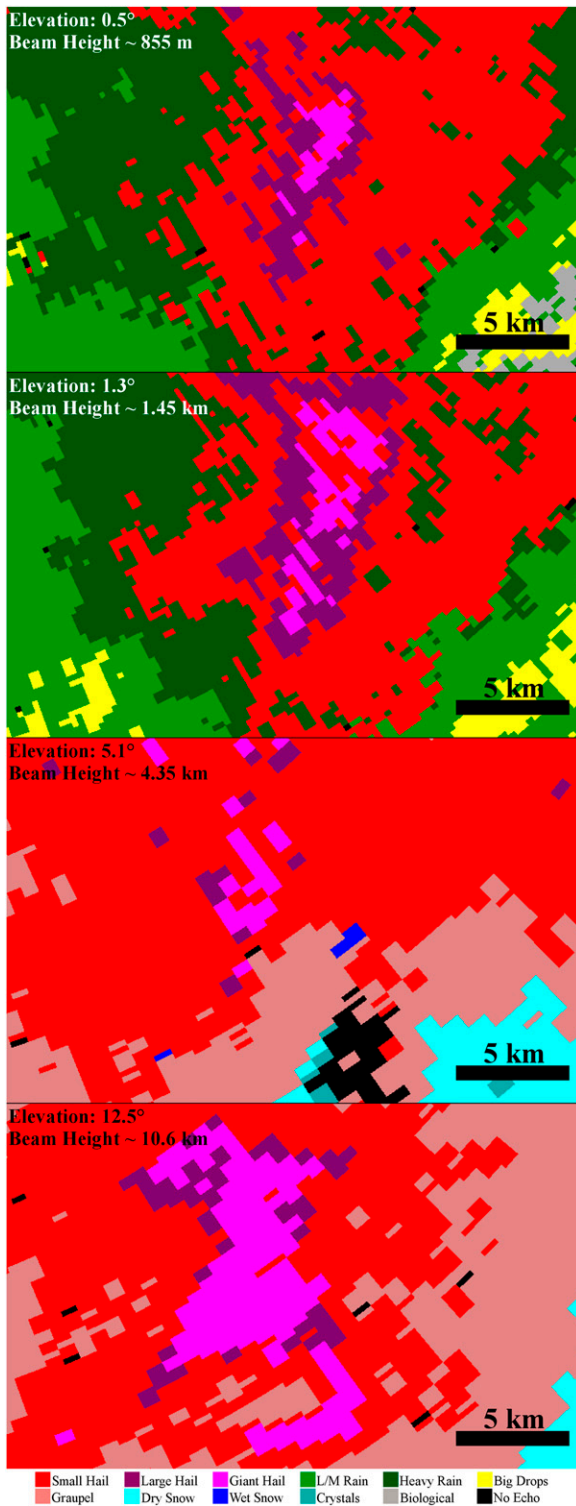
FIG. 3. HCA and HSDA output at multiple antenna elevations for the same volume as in Fig. 2; $H(T_w = 0°C)$ was 3.82 km, and $H(T_w = -25°C)$ was 8.23 km. All heights are above mean sea level. The storm was approximately 45 km from the radar.

exact report location for the matching volume. This method may allow for more accurate matching aloft; however, precise matching is limited because of storm motion and horizontal advection of hydrometeors.

3) For "windows (max Z_h)," the volume with the maximum Z_h at the 0.5° tilt within a 4 km × 4 km search box centered on the report was selected as the matching volume. For each tilt within the matched volume, the pixel with the maximum Z_h within the 4 km × 4 km search box was used for extracting the Z_h, Z_{DR}, and ρ_{hv} values. This method might help correct the difference between the time of the radar volume and the time of the actual hail fall. However, it still may suffer from potentially inaccurate matching at higher antenna elevations because of storm motion or storm tilt.

4) For "windows (max Z_h–max tilts)," the volume with the pixel having maximum Z_h at the 0.5° tilt within a 4 km × 4 km search box centered on the report served as the match for 0.5° and also as a reference point. The tilts above 0.5° from the volumes preceding and succeeding the matched volume were also considered for matching. For each tilt above 0.5° from any of the three successive volumes, the pixel with the maximum Z_h within the 4 km × 4 km search box is used for extracting the Z_h, Z_{DR}, and ρ_{hv} values. Hence, the values for the vertical matching profiles of the radar variables may come from different locations and different volumes of radar data. This method allows for the most uncertainty with the matching, thus allowing highly displaced reports (e.g., reports whose locations are between or away from the highest reflectivity regions within the three volumes because of storm motion) to match to the radar volumes appropriately. A limitation of this methodology is matching to higher-than-expected Z_h, and the corresponding Z_{DR} and ρ_{hv}, values than would be expected for different hail sizes due to reflectivity gradients within the storm.

These matching techniques were selected after several iterations of report matching conducted during the course of this study. The 4 km × 4 km search box was chosen because of typical SHAVE report spacing, usually around 1.7 km. The 4 km × 4 km box limits the number of radar pixels that overlap between multiple reports.

The HSDA was compared with current operational, single-polarization hail size discrimination algorithm, the HDA. HDA data were obtained from the Severe Weather Data Inventory (SWDI, available online at http://www.ncdc.noaa.gov/swdi; Ansari et al. 2006). The SWDI archives level-III data from the WSR-88D network, including output from the HDA. There are some
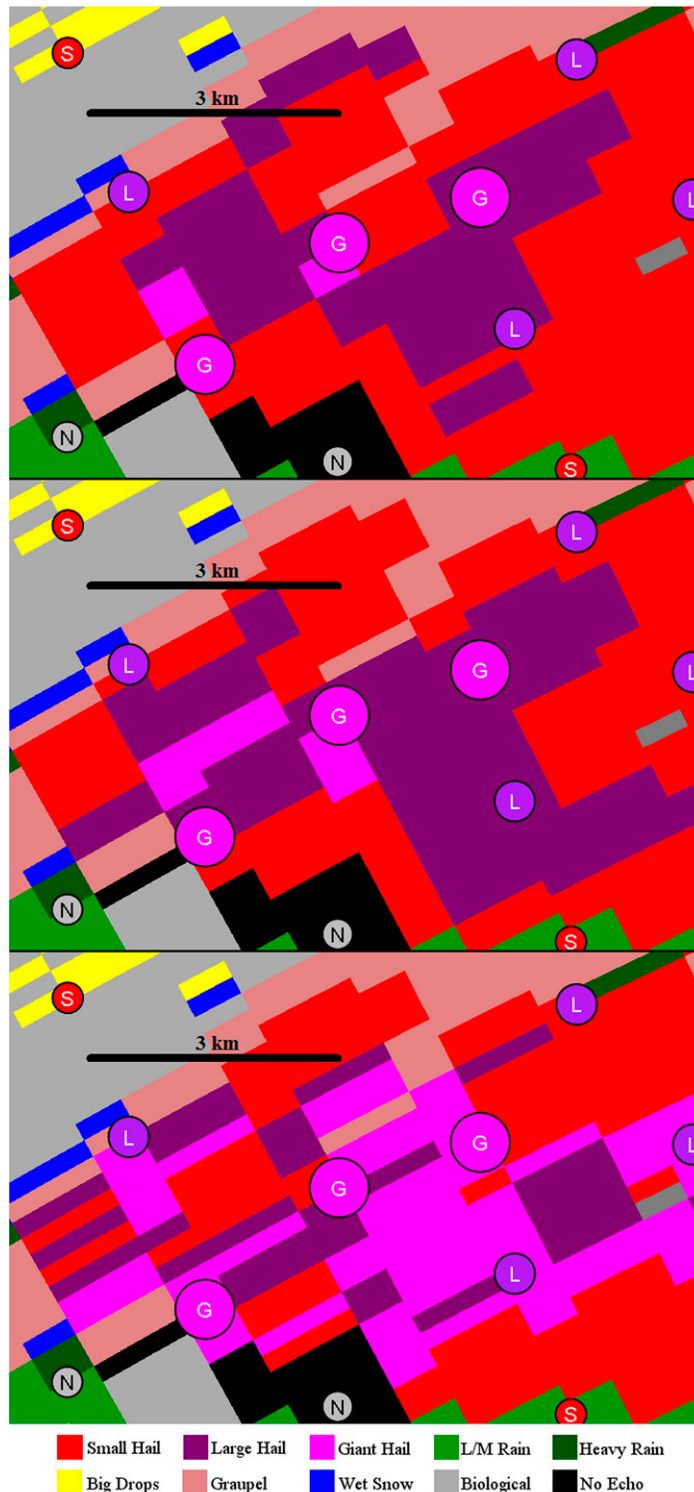
FIG. 4. HCA and HSDA outputs from the different versions of the HSDA run on the KDTX (Detroit, MI) WSR-88D data collected at 1830 UTC 27 Jul 2014. The storm was approximately 120 km from the radar. The panels correspond to the modified version of the HSDA, with the ΔZ_{DR} parameter equal to (top) -0.5 dB, (middle) -0.2 dB, and (bottom) original HSDA. All SHAVE hail reports from the entire case (and not specifically from the radar time displayed) are shown by the colored circle icons: gray N (no hail), red S (small hail), purple L (large hail), and magenta G (giant hail).
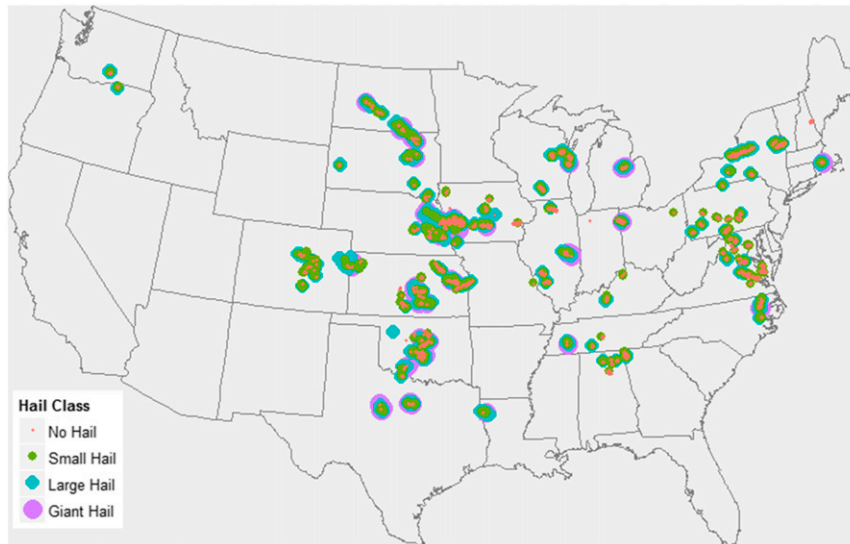
FIG. 5. Map of SHAVE hail reports used in the study.

missing data within the SWDI database, and 15 of the 79 cases examined in this study had no SWDI output. However, 1018 HDA hail size estimates were still available for evaluation. The HDA detections were matched to the largest hail report within a 5 km radius. Since the HDA was designed to estimate exact sizes of hail, its output was reclassified to the small-, large-, and giant-hail categories used within the HSDA.

4. Results

a. Observed radar profiles

Using the four matching techniques, vertical matching profiles of Z_h, Z_{DR}, and ρ_{hv} were created for each hail report, and the distributions of the three radar variables were generated for each of the six height intervals and four hail size classes (no hail, small, large, and giant) used within the HSDA (Figs. 7–9). The distributions of the variables obtained from the different matching techniques exhibit subtle but important differences. In general, adding uncertainty to the matching by using search windows and multiple volumes to create matching vertical profiles of radar variables resulted in more conceptually acceptable vertical profiles and ranges of variables at hail locations (e.g., Z_{DR} is closer to 0 dB for larger hail sizes at lower altitudes), narrower distributions, and better separation of the distributions for different hail size classes. However, the addition of uncertainty to the matching does shift the distributions for no-hail reports toward more unreasonable values (e.g., Z_h values are commonly near 60 dBZ), which is expected.

The following discussion will focus on the results obtained using the windows (max Z_h_max tilts) matching technique since these most closely matched theoretical expectations and minimized the variability within the distributions. The median value of Z_h for giant hail is 4–5 dBZ higher than the median for small hail. The vertical profiles of Z_h (Fig. 7) reveal very little change in Z_h with height for giant hail, while there is a 3–4-dB increase of Z_h with height up to the level where shedding of raindrops from the surface of melting hailstones starts, which for large hail is in the layer between 2 and 3 km below the melting level. This is in agreement with theory (Part I). There is a significant overlap of the Z_h distributions for large and giant hail within the layers at or below the melting level (height intervals 1–4 in Fig. 7). Some overlap of the Z_h distributions for small and large hail is also present; however, small-hail distributions are shifted toward smaller Z_h relative to those of large hail.

The vertical profiles and ranges of the Z_{DR} distributions (Fig. 8) correspond well to their theoretical predictions for all hail size categories. The largest changes in Z_{DR} with decreasing height are for smaller hail sizes. As was observed with the Z_h distributions, the biggest separation is between no-hail and giant-hail reports. At the lowest height interval (labeled 1 in Figs. 6–8), the median value of Z_{DR} changes from 2.1 dB for small hail to 0.9 dB for giant hail. These are very close to the centers of the trapezoids representing the membership functions for Z_{DR} (2.3 and 1.2 dB for small and giant hail, respectively). Such consistency of the observations and theory attests to the validity of basic models of melting hail utilized in Part I, which were used for designing the membership functions of the HSDA. The
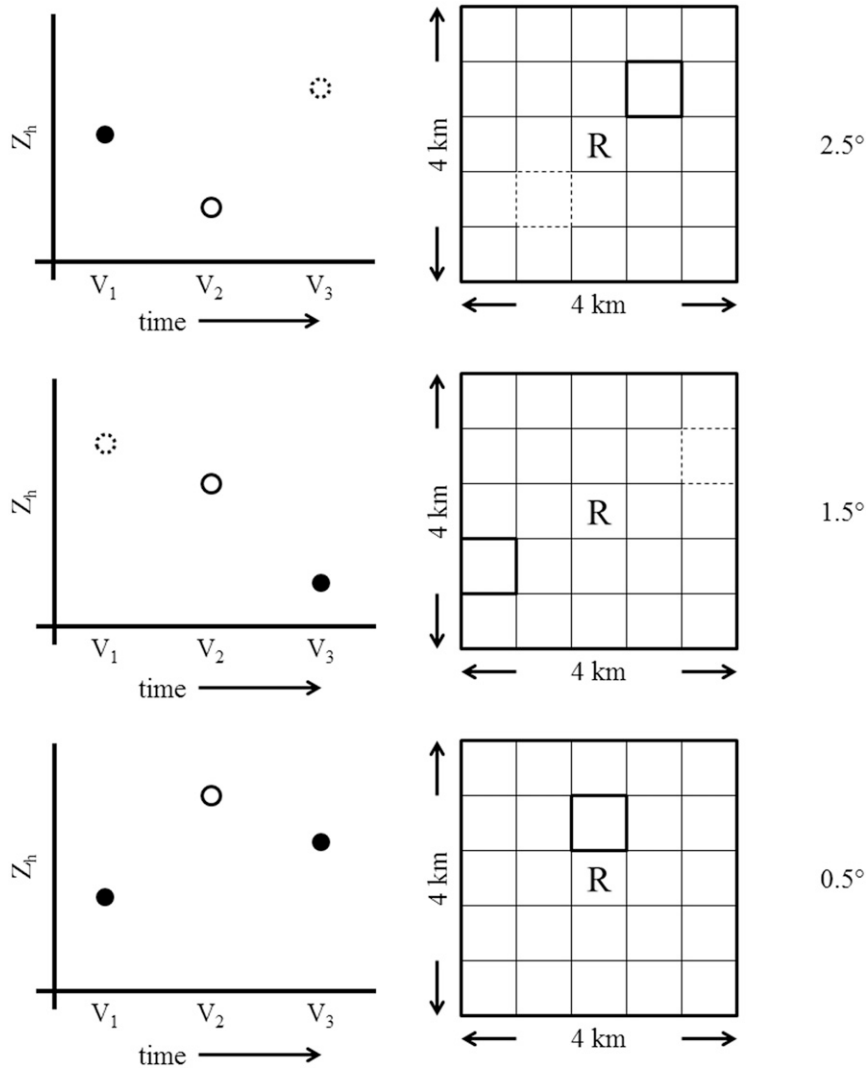
FIG. 6. Diagrams of the matching techniques used to extract the Z_h, Z_{DR}, and ρ_{hv} values for the vertical profiles. Each row is a different tilt for a given volume. (left) Example Z_h values for three volumes, and (right) the 4 km × 4 km search window. The R designates the report location. For methods 1–4, the Z_h, Z_{DR}, and ρ_{hv} values were obtained from: the volume V_2 (solid open circle) and the pixel labeled R, the constructed volume marked with the dashed open circle and also from the pixel labeled R, the volume V_2 and the pixel highlighted with a thick outline, and the constructed volume marked with the dashed open circle and the pixel highlighted by a dashed outline (except for the 0.5° tilt). For the 0.5° tilt, the same volume (V_2) was used in all methods; only the matching pixel changed according to the prescribed methods.

Z_{DR} differences between different hail sizes gradually decrease with height and practically disappear at the melting layer and above. Above the melting level, Z_{DR} loses its discrimination power, which justifies minimization of the weight W_2 for Z_{DR} at and above the melting level (Table 1).

Hail with larger sizes exhibits smaller ρ_{hv} values compared to smaller sizes at all heights (Fig. 9) as was first predicted by Balakrishnan and Zrnić (1990). An exception

is for no-hail reports near and above the melting level. The no-hail reports tend to be associated with ρ_{hv} values lower than those with small hail and similar to large hail. The distributions also show that ρ_{hv} values less than 0.9 tend to be within the tails of the distributions, even for giant hail. A raw count of the number of pixels within the search windows with ρ_{hv} values less than 0.9 at the height layer $H(T_w = 0°C) - 3 \, \mathrm{km} \leq H_b < H(T_w = 0°C) - 2 \, \mathrm{km}$ revealed only 10% of all pixels are associated with such
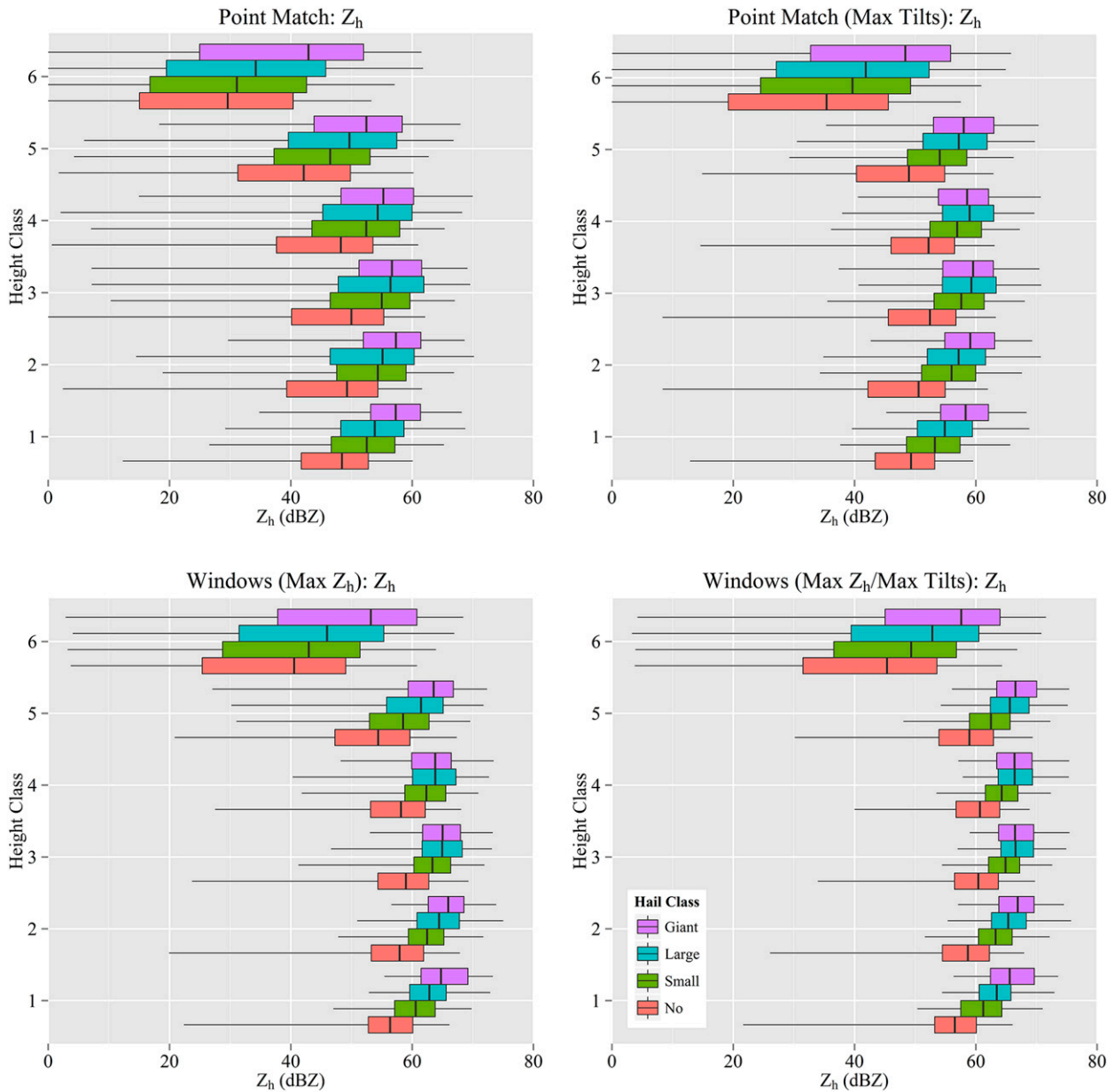
FIG. 7. Distributions of Z_h in different height intervals within 1–6 km for four categories—point match, point match (max tilts), windows (max Z_h), and windows (max Z_h–max tilts)—of hail size and no hail obtained from radar observations. The whiskers mark the 95th percentile, the boxes mark the interquartile range, and the vertical line marks the median value. The distributions for the windows were formed by using the maximum Z_h value within the 4 km × 4 km search box centered on the report. Height class 1 corresponds to the $H(T_w = 0°C)$ − 3 km > H_b layer, and height class 6 corresponds to the $H_b ≥ H(T_w = −25°C)$ layer (as labeled in the text). Layers 1–4 are warm (with $T_w ≥$ 0°C), and layers 5 and 6 are cold ($T_w < 0°C$) (see Tables 1–4).

values. In general, when comparing the observed distributions to the membership functions, observed ρ_{hv} values tend to be higher. However, given algorithm performance (discussed in the next section) it does not seem prudent to adjust the membership functions to more closely match observations. Increasing the upper bounds of the ρ_{hv} membership functions could result in

more-frequent large-hail and giant-hail designations that could increase the number of false alarms.

The consistency of the Z_h-dependent Z_{DR} membership functions [Eqs. (2)–(7)] and the Z_h–Z_{DR} scatterplots obtained from observations were examined. Figure 10 shows the scatterplots for the Z_h–Z_{DR} pairs for the different hail size report categories at the
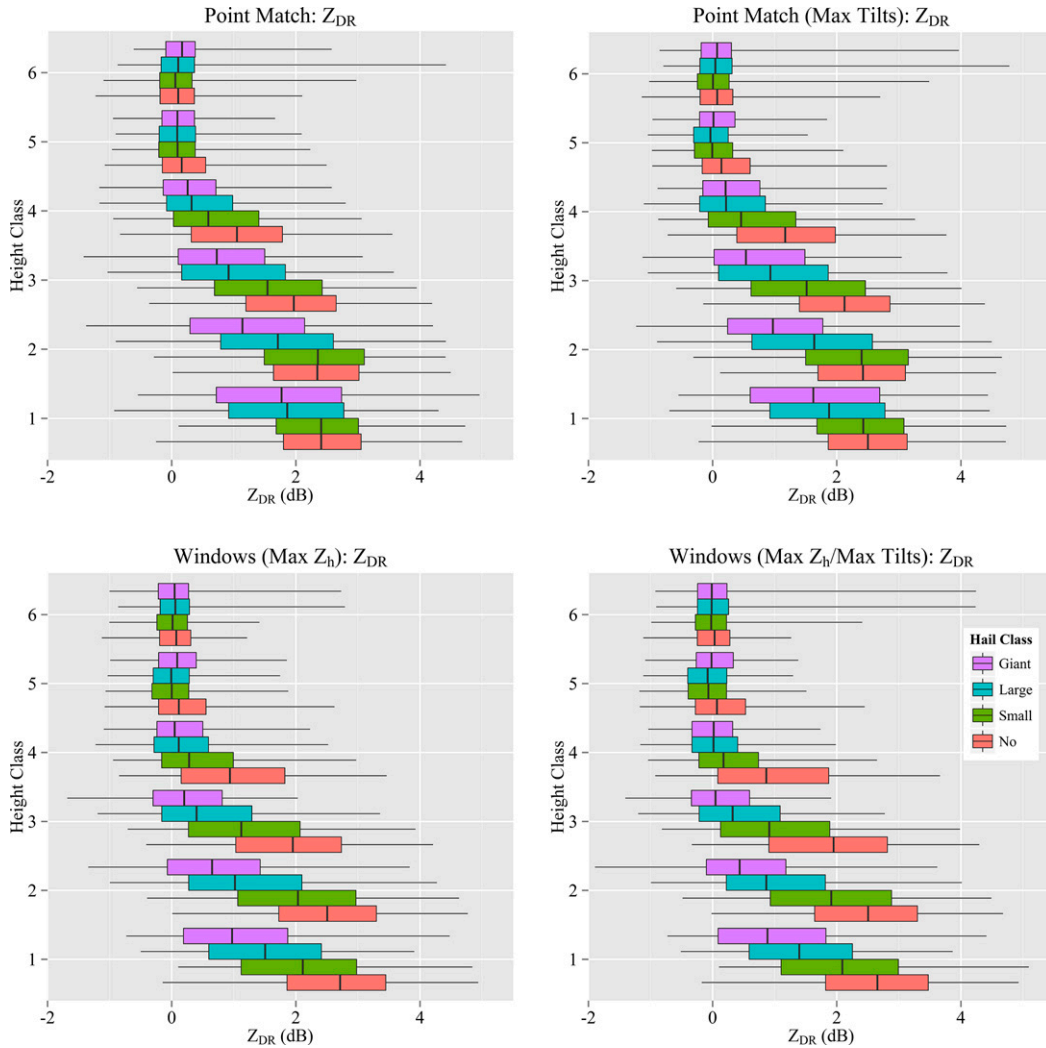
FIG. 8. As in Fig. 7, but for Z_{DR}. The Z_{DR} values used in the distributions were those paired to the Z_h values displayed in Fig. 7.

different height levels using the windows (max Z_h–max tilts) matching method with the bounds of the membership functions for Z_h and Z_{DR} are also displayed. Limitations of the matching method are quite evident with a tendency for the smaller hail size categories to have higher Z_h values than the Z_h membership function bounds. However, the trend of increasing Z_h with decreasing Z_{DR} at the lower height levels is evident. Further, the collapse of the Z_{DR} dependence on Z_h near and above the melting level is also obvious, which justifies the elimination of the Z_{DR} membership functions dependence on Z_h in Table 2. It is well known that Z_{DR} monotonically increases with Z_h in pure rain, hence the inverse dependence of Z_{DR} on Z_h for the no-hail category in Fig. 10 does not seem correct. The likely reason for this is that using the windows (max Z_h–max tilts)

matching method inevitably incorporates many true hail pixels into the no-hail class. In other words, a good portion of hail-matching profiles becomes artificially associated with no-hail ground reports.

More realistic Z_h–Z_{DR} scatterplots for the no-hail category were obtained from the point-match scheme (Fig. 11). This scheme produces lower Z_h that better fits the membership function bounds but at the same time tends to generate higher Z_{DR} mostly outside the Z_{DR} membership bounds. In other words, the point-match schemes tend to produce too many rainlike matching vertical profiles in situations when SHAVE reports indicate hail. Such ambiguity is attributed to the difference between spatial resolution of the SHAVE and radar data (1.6 and 0.25 km, respectively), ambiguity between radar and report times, and the use of a 4 km × 4 km window.
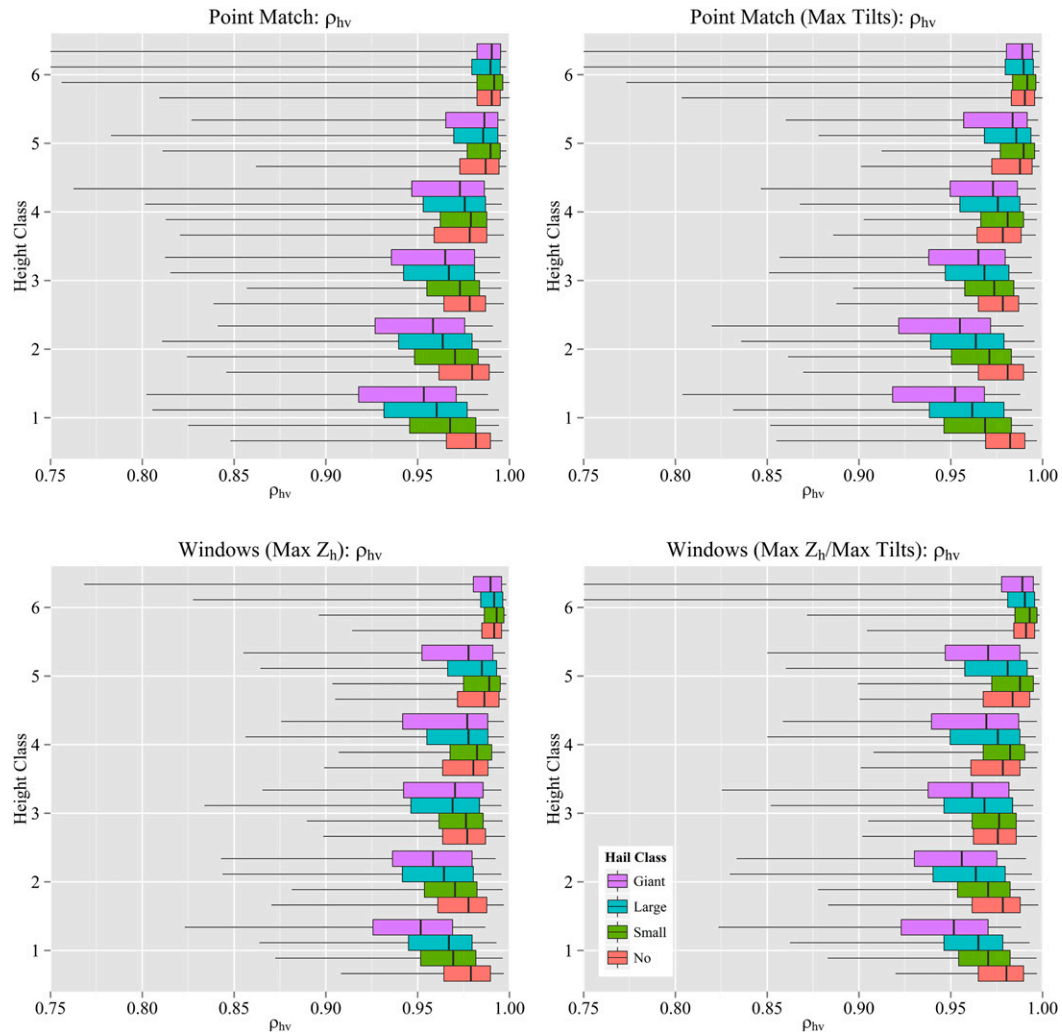
FIG. 9. As in Fig. 8, but for ρ_{hv}.

Despite these inherent uncertainties of the matching methodologies, the distributions of the three radar variables for different hail sizes at different height intervals presented in Figs. 7–9 provide strong evidence that the use of Z_{DR} and ρ_{hv} along with Z_h increases the ability to discriminate hail size compared to the use of Z_h alone. The vertical distributions also agree well with previous observational studies (e.g., Kaltenboeck and Ryzhkov 2013).

b. HSDA performance

Probability of detection (POD), false-alarm ratio (FAR), and critical success index (CSI) were calculated (Doswell et al. 1990; Schaefer 1990) for all eight versions of HSDA. The matched volumes used for the performance evaluation were from the windows (max Z_h–max windows) matching method. These volumes were selected because the method provided the best separation

of the polarimetric variables for different hail size classes. The no-hail reports were not used for the HSDA evaluation. These reports would be useful for an evaluation of the HCA; however this study is exclusively focused on the HSDA performance. Hail reports with no designations of rain/hail within the 4 km × 4 km search boxes were also excluded. These reports are associated with failures by the HCA and not the HSDA. Such failures of detection by the HCA were rare, with fewer than 10% of all tilts with small-hail reports and fewer than 5% of all tilts for large-hail and giant-hail reports. Many of these tilts with no rain/hail designations were aloft—above the hail core and even storm top. This confirms high probability of hail (or rain/hail) detection by the HCA in hail-producing storms, which was also reported by Heinselman and Ryzhkov (2006).

The maximum reported hail size from each SHAVE report was paired with the counts of the number of pixels
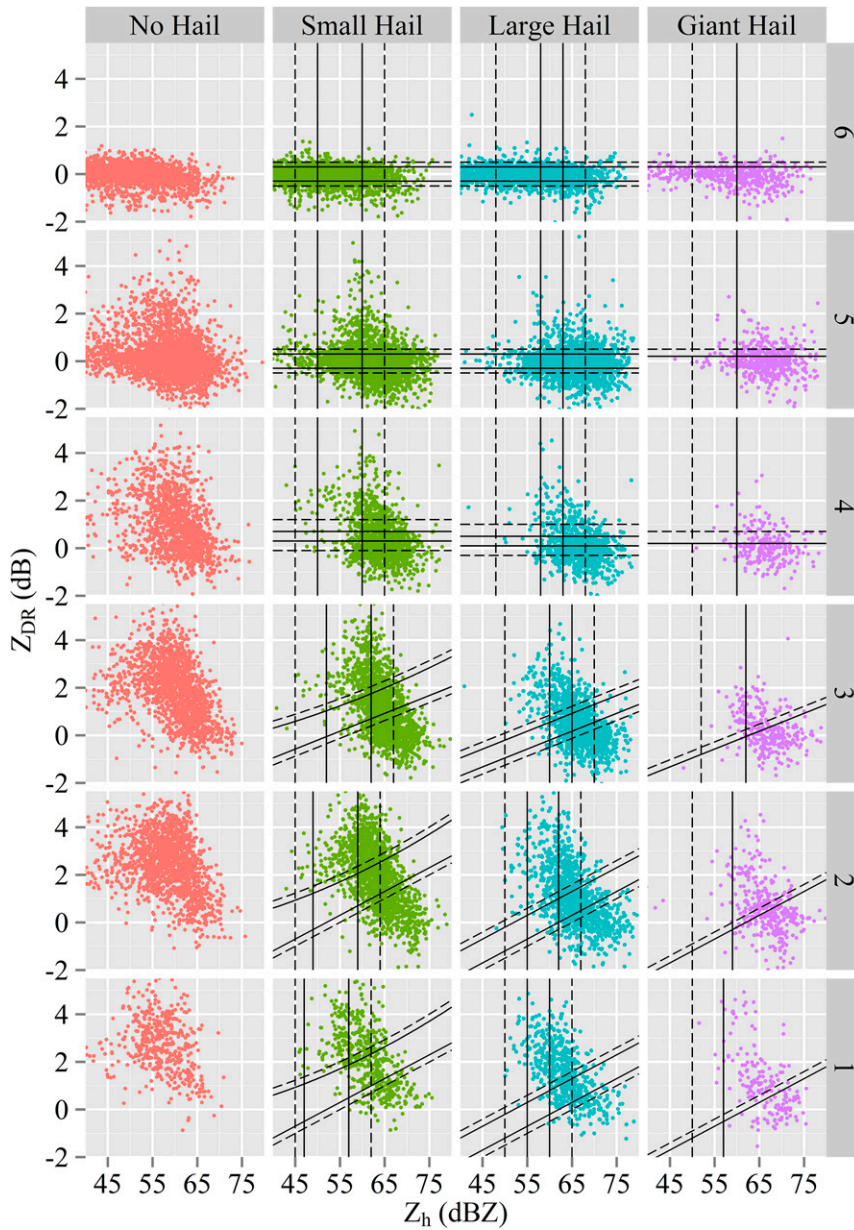
FIG. 10. Scatterplot of Z_h and Z_{DR} pairs for the windows (max Z_h–max tilts) matching technique. The columns are for each hail size category, while the rows are for each height class. The height classes are as labeled in Fig. 7 and as discussed in the text. For the small-, large-, and giant-hail categories, the membership function bounds for Z_h (vertical lines) and Z_{DR} (slanted and horizontal) are shown by the dashed and solid lines. The dashed lines mark x_1 and x_4, while the solid lines mark x_2 and x_3. The $\Delta Z_{DR} = -0.2$ dB for the Z_{DR} membership bounds.

with small-, large-, and giant-hail designations within the $4 \, km \times 4 \, km$ search window. The maximum HSDA designation and the most common (i.e., the modal) HSDA designation were used to quantify HSDA performance. Simple point matches were found to perform very poorly (POD < 0.2), which is most likely the result

of the difficulties matching the radar volumes to the reports as discussed earlier.

The original HSDA with weights (Table 1) was found to perform similarly to the original HSDA without additional weights. The HSDA versions with different ΔZ_{DR} values had similar performance and, in general,
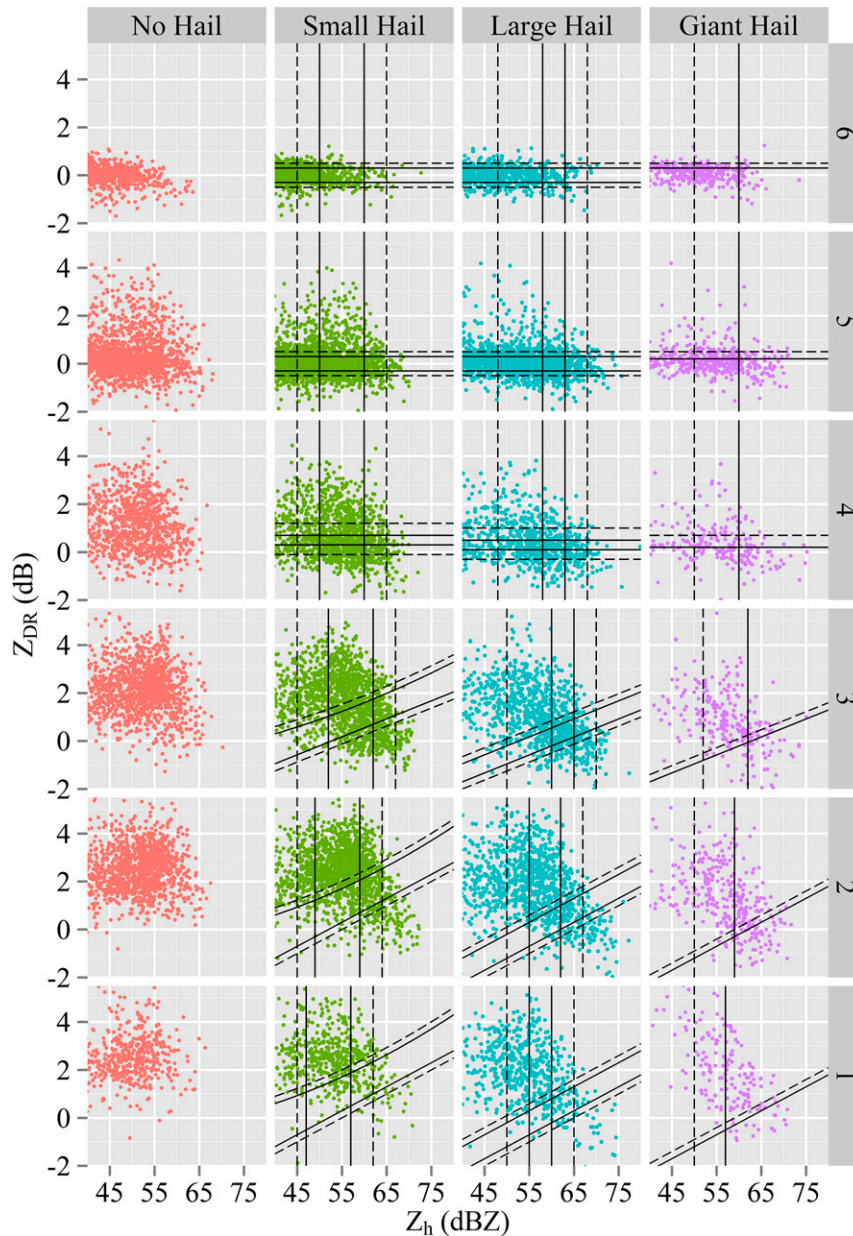
FIG. 11. As in Fig. 10, but for the point-match method.

increasing (decreasing) ΔZ_{DR} decreased (increased) overall skill, as measured by CSI. To evaluate the impact of ΔZ_{DR}, CSI scores for each of the different ΔZ_{DR} values for each case were compared. The resulting hits, misses, and false alarms for the ΔZ_{DR} value that resulted in the maximum CSI for each case were used for the aggregate skill score calculations of all cases. The skill scores of the original and modified HSDA are summarized in Tables 3 and 4. Table 3 summarizes the skill scores for all height layers if HSDA hail size designations at all radar elevations are used. There is

obviously a loose connection between radar observations and designations at higher tilts aloft, and surface hail reports. However, from a warning decision-making perspective, the skill scores provide forecasters information on the value of the algorithm designations aloft for quantifying hail size at the surface, thus giving forecasters some lead time prior to surface hail fall. As opposed to Table 3, only the HSDA classification at the lowest 0.5° tilt is taken into account to calculate skill scores in Table 4. It is not surprising that comparing surface hail reports with radar designations at the lowest tilt results in

TABLE 3. Skill scores for the HSDA: original (orig) and new version with the best ΔZ_{DR} selection per case for each height layer. The two scoring techniques are 1) maximum HSDA designation within the 4 km × 4 km window and 2) the most common HSDA designation within the 4 km × 4 km search box. The no-hail reports and reports with no HSDA designations within the 4 km × 4 km search box were excluded.

Height layer	Scoring method	POD		FAR		CSI	
		New	Orig	New	Orig	New	Orig
$H_b \geq H(T_w = -25°C)$	Maximum	0.601	0.604	0.414	0.368	0.422	0.447
	Common	0.484	0.484	0.102	0.059	0.459	0.470
$H(T_w = -25°C) < H_b \leq H(T_w = 0°C)$	Maximum	0.695	0.760	0.433	0.621	0.454	0.339
	Common	0.557	0.580	0.076	0.113	0.533	0.540
$H(T_w = 0°C) < H_b \leq H(T_w = 0°C) - 1\,km$	Maximum	0.768	0.846	0.578	0.789	0.374	0.203
	Common	0.584	0.606	0.131	0.179	0.537	0.536
$H(T_w = 0°C) - 1\,km < H_b \leq H(T_w = 0°C) - 2\,km$	Maximum	0.807	0.891	0.596	0.807	0.369	0.188
	Common	0.584	0.641	0.118	0.264	0.542	0.521
$H(T_w = 0°C) - 2\,km < H_b \leq H(T_w = 0°C) - 3\,km$	Maximum	0.797	0.928	0.484	0.807	0.456	0.190
	Common	0.538	0.709	0.091	0.507	0.511	0.410
$H_b < H(T_w = 0°C) - 3\,km$	Maximum	0.672	0.949	0.423	0.774	0.450	0.223
	Common	0.429	0.716	0.069	0.500	0.416	0.418

higher scores. The CSI combined for all height layers for the modified (new) HSDA shown in Table 4 is higher than for the original HSDA version: 0.543 versus 0.432 if the common scoring method is used and 0.465 versus 0.205 if the maximal scoring method is utilized.

In general, the original HSDA has a higher FAR than the newer version, while the newer version has a lower POD than the original. Overall, however, the modified version seems to outperform the original version in terms of CSI if the radar data from the lowest antenna tilt are utilized for hail size designations. This is generally confirmed by the feedback from forecasters participating in the Experimental Warning Program at NOAA's Hazardous Weather Testbed in Norman, Oklahoma (Calhoun et al. 2014). The forecasters notice that the original version of HSDA tends to overestimate hail size. This is also illustrated in Figs. 4, 12, and 13 where the fields of the HSDA classification results are compared for the two versions of the algorithm. As illustrated by the 27 July 2014 example in Fig. 4 and the

15 May 2013 example in Fig. 12, the original HSDA identified giant hail over a large area, which matches well to the reports on 15 May 2013 but not well to the reports on 27 July 2014. Further, the areas of giant hail identified by the original HSDA are scattered and lack any spatial coherency—giant hail was essentially designated anywhere within the hail core. This was a very common result for all of the cases investigated, even for storms not producing giant hail. The 5 April 2012 example in Fig. 13 shows a large area of giant hail detected by the original HSDA for a marginally severe storm. The newer HSDA in all cases reduces the area of giant-hail designations and adds spatial coherency to the designations.

To evaluate possible impact of the radar data quality on the performance of HSDA in most challenging situations, the cases with giant-hail reports were more closely examined. Figure 14 illustrates the range of Z_{DR} values for selected cases with giant-hail reports. The middle example, a giant-hail-producing storm observed

TABLE 4. As in Table 3, but for the 0.5° elevation tilt only. The combined scores are aggregated scores for all height layers intersected by the 0.5° elevation tilts.

| Height layer	Scoring method	POD		FAR		CSI	
		New	Orig	New	Orig	New	Orig
Combined	Maximum	0.782	0.929	0.465	0.792	0.465	0.205
	Common	0.594	0.765	0.136	0.501	0.543	0.432
$H(T_w = 0°C) < H_b \leq H(T_w = 0°C) - 1\,km$ (65 reports)	Maximum	0.905	0.833	0.698	0.922	0.292	0.077
	Common	0.824	0.813	0.250	0.304	0.646	0.600
$H(T_w = 0°C) - 1\,km < H_b \leq H(T_w = 0°C) - 2\,km$ (590 reports)	Maximum	0.859	0.876	0.558	0.816	0.412	0.180
	Common	0.680	0.758	0.203	0.352	0.580	0.537
$H(T_w = 0°C) - 2\,km < H_b \leq H(T_w = 0°C) - 3\,km$ (864 reports)	Maximum	0.823	0.949	0.427	0.806	0.510	0.192
	Common	0.619	0.797	0.108	0.567	0.576	0.390
$H_b < H(T_w = 0°C) - 3\,km$ (535 reports)	Maximum	0.661	0.954	0.375	0.727	0.473	0.269
	Common	0.448	0.717	0.060	0.574	0.436	0.364
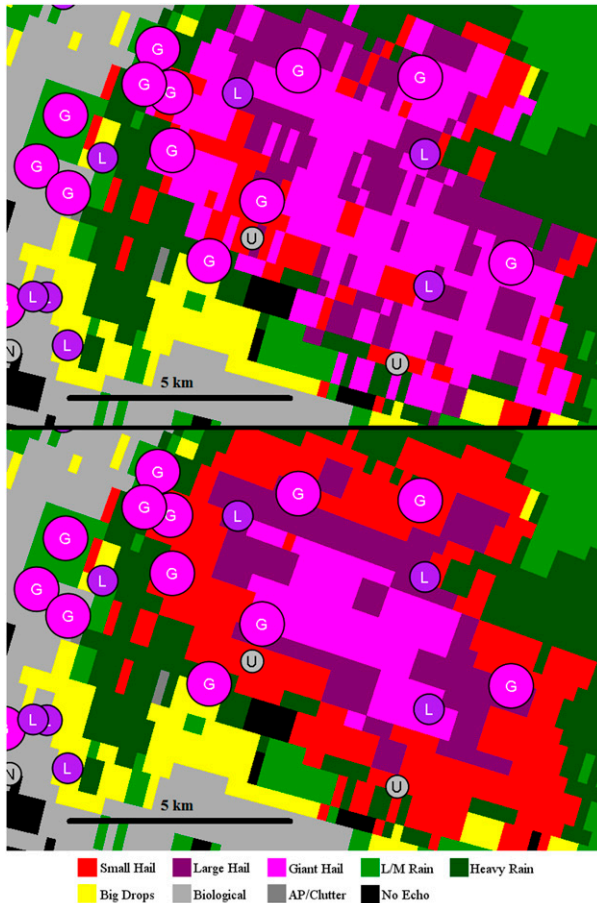
FIG. 12. As in Fig. 4, but for the KFWS (Fort Worth, TX) radar at 2328 UTC 15 May 2013. (top) The original HSDA output, and (bottom) the modified HSDA output with $\Delta Z_{DR} = -0.2$ dB.

FIG. 13. As in Fig. 12, but for the KHTX (Huntsville, AL) radar at 0011 UTC 6 Apr 2012.

by the KABR (Aberdeen, South Dakota) WSR-88D on 21 June 2013, is of particular interest (Fig. 14c). The highly negative Z_{DR} values were caused by very high differential attenuation that was not fully corrected within the preprocessor. For the top two cases illustrated (Figs. 14a,b), Z_{DR} was positively biased, while the fourth case (Fig. 14d) seems to demonstrate good Z_{DR} calibration and the fifth case (Fig. 14e) exhibits a slightly negative Z_{DR} bias.

Of the 206 giant-hail reports used in this study, only 2 were not associated with any HCA rain/hail designations within the 4 km × 4 km search box at the 0.5° tilt because the reports were on the outside edge of a storm path. For the 204 giant-hail reports that had at least one pixel of hail detection within the search box at the lowest tilt, 119 had no HSDA designations of giant hail for the newer algorithm. Distributions of Z_h, Z_{DR}, and ρ_{hv} for each group are compared in Fig. 15. The largest separation of the groups occurs for the Z_{DR} distributions, with reports associated with giant-hail designations
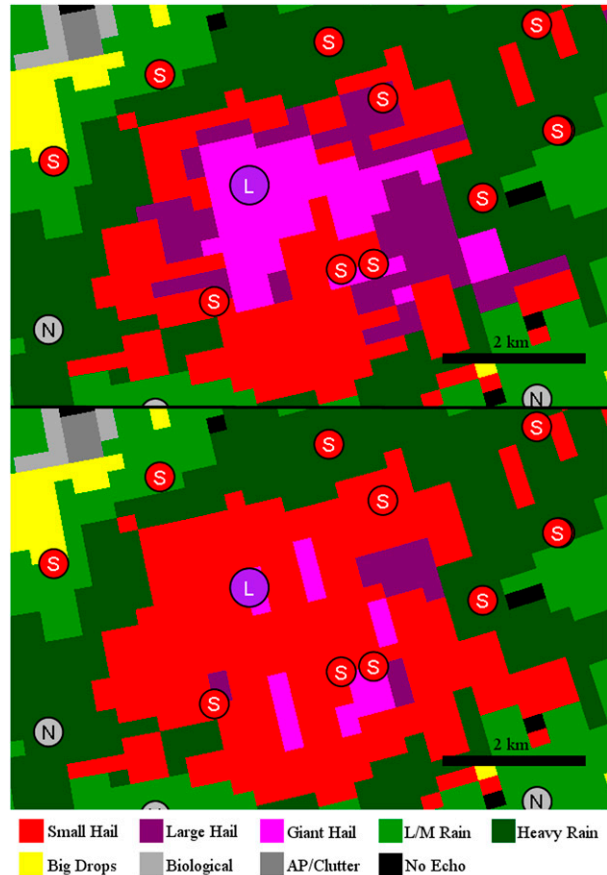
having much lower Z_{DR} values. There is also some separation of the Z_h distributions for the two groups. It is important to note that, for giant-hail designations, the use of Eq. (3) allows for a narrower Z_{DR} range for giant hail in pixels with lower Z_h values. Thus, for giant-hail-producing storms with moderate Z_h values even a moderate, positive Z_{DR} bias may cause deficiencies in giant-hail detection, as Z_{DR} membership function for giant hail would be bounded by lower Z_{DR} values relative to storms with larger Z_h values (Figs. 10 and 11).

c. HSDA performance compared with the current operational hail sizing algorithm

The evaluation for 1018 HDA detections yielded POD = 0.793, FAR = 0.646, and CSI = 0.324. This is as compared with all 0.5° elevation tilt HSDA scores of POD = 0.594, FAR = 0.136, and CSI = 0.432. With respect to CSI and overall skill, the HDA is worse than both of the scoring methodologies employed to evaluate the HSDA at all height levels. With respect to FAR, the HDA performs worse than the new HSDA for any height layer and either scoring methodology.
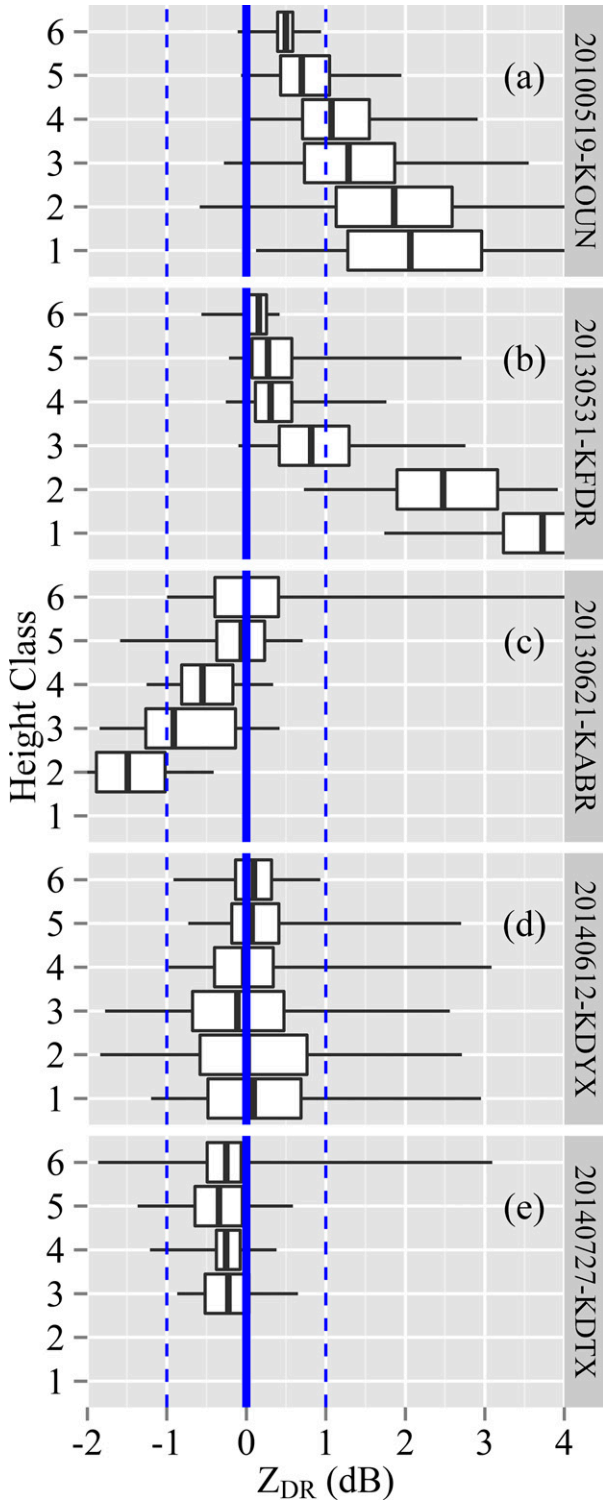
FIG. 14. The Z_{DR} distributions in different height intervals for selected cases of giant hail. The height intervals are the same as those in Fig. 7 and are as labeled in the text. For reference, the −1- and 1-dB locations are marked with the dashed lines and the 0-dB location is marked with the thick solid line.
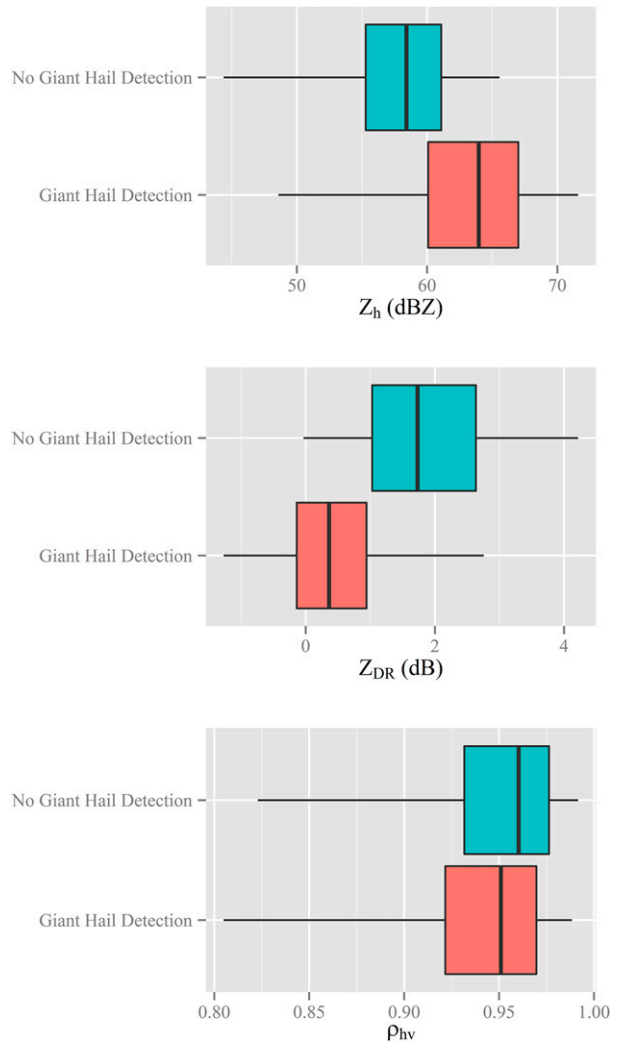

FIG. 15. Box plots for (top) Z_h, (middle) Z_{DR}, and (bottom) ρ_{hv} for giant-hail reports that had no giant-hail designations (teal boxplots) and at least one giant-hail designation (salmon boxplots) within the search box. The whiskers are the 95th percentile, the box outlines the interquartile range, and the vertical line is the median.

The HDA only provides a single estimate of hail size per storm, while the HSDA pinpoints locations of hail belonging to three different hail size categories. Further, the HDA is a volumetric product that only runs at the conclusion of radar volume scans. The HSDA runs on a pixel-by-pixel basis and thus is available at the conclusion at each individual tilt within the volume.

5. Discussion

As mentioned before, the performance of HSDA is particularly sensitive to the errors in the estimate of Z_{DR}, which can be biased from radar miscalibration or anomalously high differential attenuation. The skill

scores reported in this study reflect the current state of the quality of absolute calibration of Z_{DR} across the fleet of WSR-88Ds. As reported by Cunningham et al. (2013), the magnitude of Z_{DR} bias exceeds 0.2 dB for about 40% of the WSR-88D sites. It is expected that the quality of Z_{DR} calibration on the radar network will be continuously improved via utilization of the new methodologies for databased calibration and via better procedures for implementing needed correction in the radar system (Cunningham et al. 2013). There are also plans for further improvement of the algorithms for differential attenuation correction. At the moment, the attenuation-related bias of Z_{DR} is corrected by adding the term proportional to the total differential phase Φ_{DP} with a constant proportionality factor

$$\Delta Z_{DR}(r) = \beta[\Phi_{DP}(r) - \Phi_{DP}^{(sys)}],$$

where $\Phi_{DP}^{(sys)}$ is the system differential phase and the factor β is equal to 0.004 dB (°)$^{-1}$. In fact, factor β can be much higher in the case of anomalously high differential attenuation usually caused by melting hail (Part II). The procedure for automatic determination of β in the presence of hail is described in Part II (section 3) and is planned to be implemented on the WSR-88D network in the near future. These measures will mitigate the biases in Z_{DR}, which are illustrated in Fig. 14, and, it is hoped, help to improve overall HSDA performance.

Another potential source for anomalously high Z_{DR} values associated with giant hail is large drops formed by partially melted hail. As discussed in Part I, very large raindrops with diameters between 8 and 9 mm mainly originate from hailstones with diameters between 8 and 14 mm. Assuming these drops are also present in the mixture with giant hail with small mass water fraction, the expected Z_{DR} depression from the giant hail may be offset. While an interesting topic for future research, the distributions of Z_{DR} (Fig. 8) suggest this may be a relatively rare occurrence.

Another resource for HSDA improvement is capitalizing on polarimetric radar signatures of convective updrafts where hail initially grows. The HSDA versions examined in this study are primarily based on polarimetric characteristics of melting hail below the melting level after hail has already been formed at higher altitudes. Designation of hail size above the melting level in the current HSDA is almost solely based on the radar reflectivity factor. It is well known that hail grows within strong convective updrafts stretching high above the melting level. Intense updrafts can be marked by bounded weak echo regions and be associated with Z_{DR} columns and, atop the Z_{DR} columns, depressions in the ρ_{hv} (e.g., Bringi et al. 1997; Hubbert et al. 1998;

Picca and Ryzhkov 2012; Snyder et al. 2013; Kumjian et al. 2014; Snyder et al. 2014). The mechanism of hail growth at the tops of the Z_{DR} columns is described in the recent paper of Kumjian et al. (2014) where it is shown that the depth of the Z_{DR} column is roughly proportional to the strength of the updrafts and, consequently, can be related to the potential size of hail growing in the updraft. Examining polarimetric radar signatures aloft will be instrumental in capturing the genesis of hail before it falls on the ground and likely will increase lead time for issuing warnings of severe hail at the surface.

6. Summary

An evaluation of the HSDA, a polarimetric radar algorithm for hail size classification for the operational WSR-88Ds, has been presented. The HSDA uses Z_h, Z_{DR}, and ρ_{hv} within a fuzzy-logic scheme to distinguish between small, large, and giant hail. Hail size classification is performed at all antenna elevations and different sets of the membership functions for the radar variables depending on the height of the radar resolution volume. The HSDA was validated using reports from the Severe Hazards Analysis and Verification Experiment. These hail reports are at high spatial resolution making them suitable to evaluate HSDA, which pinpoints locations of hail fall. The SHAVE operations yielded more than 3000 surface hail reports across the contiguous United States. One of the inherent limitations of such reports is the lack of exact timing of hail occurrence at the surface. Such uncertainty was addressed using various techniques to match surface hail reports and HSDA designations. Two techniques were also used for calculating skill scores for the HSDA. Despite methodological problems of matching radar data and surface observations, the observed ranges of Z_h, Z_{DR}, and ρ_{hv} and their median vertical profiles for different sizes of reported hail are found to be consistent with the HSDA membership functions, which were prescribed using theoretical simulations of radar variables based on cloud modeling of melting hail presented in Part I.

In the process of validation, the original version of HSDA described in Part II was modified to mitigate general overestimation of hail size by the original version. The new HSDA reduces the number of false alarms and quantifies hail size more accurately by adjusting membership functions within the fuzzy-logic scheme to more closely match observations, making the Z_{DR} membership functions dependent on Z_h values. Additionally, an adaptable parameter, ΔZ_{DR}, was added to adjust the Z_{DR} membership functions to

account for possible radar miscalibration and/or forecaster preference for increasing or decreasing warning thresholds for large and giant hail. The new HSDA also produces more coherent spatial patterns of hail size designations compared to the original algorithm.

The technique based on the matching of the most common HSDA hail size designation at lowest radar tilt within the $4\,km \times 4\,km$ area surrounding SHAVE hail report provides the most objective estimation of the algorithm skill. The corresponding probability of detection is 0.594, the false-alarm ratio is 0.136, and the critical success index is 0.543. The HSDA outperforms the traditional single-polarization algorithm (with CSI equal to 0.324) confirming the ability of polarimetric radars to better quantify hail size. Additionally, the HSDA pinpoints locations of hail of different sizes with the storm, whereas the single-polarization algorithm provides only a single hail size estimate per storm.

Further improvement of the HSDA performance is anticipated once existing problems with absolute calibration of Z_{DR} on the WSR-88D units are addressed, a more sophisticated scheme for differential attenuation correction is implemented, and polarimetric radar signatures of strong convective updrafts (such as BWER, Z_{DR} columns, and ρ_{hv} depressions above the melting level) are taken into account.

REFERENCES

Amburn, S. A., and P. L. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12**, 473–478, doi:10.1175/1520-0434(1997)012<0473:VDAAHI>2.0.CO;2.

Ansari, S., S. Del Greco, B. Nelson, and H. Frederick, 2006: The severe weather data inventory (SWDI): Spatial query tools, web services and data portals at NOAA's National Climatic Data Center (NCDC). *22nd Int. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 11.4. [Available online at https://ams.confex.com/ams/pdfpapers/100482.pdf.]

Balakrishnan, N., and D. S. Zrnić, 1990: Estimation of rain and hail rates in mixed-phase precipitation. *J. Atmos. Sci.*, **47**, 565–583, doi:10.1175/1520-0469(1990)047<0565:EORAHR>2.0.CO;2.

Blair, S. F., and J. W. Leighton, 2012: Creating high-resolution hail datasets using social media and post-storm ground surveys. *Electron. J. Oper. Meteor.*, **13** (3), 32–45.

Boustead, J. M., 2008: Using maximum storm-top divergence and the vertical freezing level to forecast hail size. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., P6.6. [Available online at https://ams.confex.com/ams/pdfpapers/142145.pdf.]

Bringi, V. N., K. Knupp, A. Detwiler, L. Liu, I. J. Caylor, and R. A. Black, 1997: Evolution of a Florida thunderstorm during the Convection and Precipitation/Electrification Experiment: The case of 9 August 1991. *Mon. Wea. Rev.*, **125**, 2131–2160, doi:10.1175/1520-0493(1997)125<2131:EOAFTD>2.0.CO;2.

Calhoun, K. M., G. S. Garfield, D. M. Kingfield, C. D. Karstens, W. Line, K. L. Ortega, T. M. Smith, and G. J. Stumpf, 2014: The 2013/2014 experimental warning program at the NOAA hazardous weather testbed. *27th Conf. Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 4B.1. [Available online at https://ams.confex.com/ams/27SLS/webprogram/Paper255935.html.]

Cunningham, J. G., W. D. Zittel, R. R. Lee, and R. L. Ice, 2013: Methods for identifying systematic differential reflectivity (Z_{DR}) biases on the operational WSR-88D network. *34th Conf. on Radar Meteorology*, Breckenridge, CO, Amer. Meteor. Soc., 9B.5. [Available online at https://ams.confex.com/ams/36Radar/webprogram/Manuscript/Paper228792/JCunningham_36thRadarConf_9B5.pdf.]

Depue, T. K., P. C. Kennedy, and S. A. Rutledge, 2007: Performance of the hail differential reflectivity (H_{DR}) polarimetric radar hail indicator. *J. Appl. Meteor. Climatol.*, **46**, 1290–1301, doi:10.1175/JAM2529.1.

Donavon, R. A., and K. A. Jungbluth, 2007: Evaluation of a technique for radar identification of large hail across the upper Midwest and central plains of the United States. *Wea. Forecasting*, **22**, 244–254, doi:10.1175/WAF1008.1.

Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585, doi:10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2.

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

Heinselman, P. L., and A. V. Ryzhkov, 2006: Validation of polarimetric hail detection. *Wea. Forecasting*, **21**, 839–850, doi:10.1175/WAF956.1.

Hubbert, J., V. N. Bringi, L. D. Carey, and S. Bolen, 1998: CSU–CHILL polarimetric radar measurements from a severe hail storm in eastern Colorado. *J. Appl. Meteor.*, **37**, 749–775, doi:10.1175/1520-0450(1998)037<0749:CCPRMF>2.0.CO;2.

Istok, M. J., and Coauthors, 2009: WSR-88D dual polarization initial operating capabilities. *25th Conf. IIPS*, Phoenix, AZ, Amer. Meteor. Soc., 15.5. [Available online at https://ams.confex.com/ams/pdfpapers/148927.pdf.]

Kaltenboeck, R., and A. Ryzhkov, 2013: Comparison of polarimetric signatures of hail at S and C bands for different hail sizes. *Atmos. Res.*, **123**, 323–336, doi:10.1016/j.atmosres.2012.05.013.

Khain, A., A. Pokrovsky, M. Pinsky, A. Seifert, and V. Phillips, 2004: Simulation of effects of atmospheric aerosols on deep turbulent convective clouds using a spectral microphysics mixed-phase cumulus cloud model. Part I: Model description and possible applications. *J. Atmos. Sci.*, **61**, 2963–2982, doi:10.1175/JAS-3350.1.

——, D. Rosenfeld, A. Pokrovsky, U. Blahak, and A. Ryzhkov, 2011: The role of CCN in precipitation and hail in a

mid-latitude storm as seen in simulations using spectral (bin) microphysics model in a 2D frame. *Atmos. Res.*, **99**, 129–146, doi:10.1016/j.atmosres.2010.09.015.

Kumjian, M., A. Khain, N. Benmoshe, E. Ilotoviz, A. Ryzhkov, and V. Phillips, 2014: The anatomy and physics of ZDR columns: Investigating a polarimetric radar signature with a spectral bin microphysical model. *J. Appl. Meteor. Climatol.*, **53**, 1820–1843, doi:10.1175/JAMC-D-13-0354.1.

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

Lemon, L. R., 1998: The radar "three-body scatter spike": An operational large-hail signature. *Wea. Forecasting*, **13**, 327–340, doi:10.1175/1520-0434(1998)013<0327:TRTBSS>2.0.CO;2.

NWS, 2007: *Storm Data* preparation. National Weather Service Instruction 10-1605, 97 pp. [Available online at https://www.ncdc.noaa.gov/stormevents/pd01016005curr.pdf.]

——, 2014: WFO severe weather products specification. National Weather Service Instruction 10-511, 35 pp. [Available online at http://www.nws.noaa.gov/directives/sym/pd01005011curr.pdf.]

Ortega, K. L., T. M. Smith, K. L. Manross, A. G. Kolodziej, K. A. Scharfenberg, A. Witt, and J. J. Gourley, 2009: The Severe Hazards Analysis and Verification Experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519–1530, doi:10.1175/2009BAMS2815.1.

Park, H. S., A. V. Ryzhkov, D. S. Zrinić, and K.-E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, **24**, 730–748, doi:10.1175/2008WAF2222205.1.

Picca, J., and A. V. Ryzhkov, 2012: A dual-wavelength polarimetric analysis of the 16 May 2010 Oklahoma City extreme hailstorm. *Mon. Wea. Rev.*, **140**, 1385–1403, doi:10.1175/MWR-D-11-00112.1.

Rasmussen, R. M., and A. J. Heymsfield, 1987a: Melting and shedding of graupel and hail. Part I: Model physics. *J. Atmos. Sci.*, **44**, 2754–2763, doi:10.1175/1520-0469(1987)044<2754:MASOGA>2.0.CO;2.

——, and ——, 1987b: Melting and shedding of graupel and hail. Part II: Sensitivity study. *J. Atmos. Sci.*, **44**, 2764–2782, doi:10.1175/1520-0469(1987)044<2764:MASOGA>2.0.CO;2.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrinić, 2005: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824, doi:10.1175/BAMS-86-6-809.

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

——, ——, ——, and P. Zhang, 2013b: Polarimetric radar characteristics of melting hail. Part II: Practical implications. *J. Appl. Meteor. Climatol.*, **52**, 2871–2886, doi:10.1175/JAMC-D-13-074.1.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, doi:10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Scharfenberg, K. A., and Coauthors, 2005: The Joint Polarization Experiment: Polarimetric radar in forecasting and warning decision making. *Wea. Forecasting*, **20**, 775–788, doi:10.1175/WAF881.1.

Snyder, J. C., H. B. Bluestein, V. Venkatesh, and S. J. Frasier, 2013: Observations of polarimetric signatures in supercells by an X-band mobile Doppler radar. *Mon. Wea. Rev.*, **141**, 3–29, doi:10.1175/MWR-D-12-00068.1.

——, A. V. Ryzhkov, H. B. Bluestein, and S. F. Blair, 2014: Polarimetric analysis of two giant-hail-producing supercells observed by X-band and S-band radars. *27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 166. [Available online at https://ams.confex.com/ams/27SLS/webprogram/Handout/Paper255455/giant_hail_storms_poster.pdf.]

Waldvogel, A., B. Federer, and P. Grimm, 1979: Criteria for the detection of hail cells. *J. Appl. Meteor.*, **18**, 1521–1525, doi:10.1175/1520-0450(1979)018<1521:CFTDOH>2.0.CO;2.

Witt, A., and S. P. Nelson, 1991: The use of single-Doppler radar for estimating maximum hailstone size. *J. Appl. Meteor.*, **30**, 425–431, doi:10.1175/1520-0450(1991)030<0425:TUOSDR>2.0.CO;2.

——, 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, doi:10.1175/1520-0434(1998)013<0286:AEHDAF>2.0.CO;2.

——, ——, ——, ——, ——, and ——, 1998b: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13**, 513–518, doi:10.1175/1520-0434(1998)013<0513:ETPOWS>2.0.CO;2.