P2.12 An Examination of the Hail Detection Algorithm over Central Alabama

Kevin B. Laws*, Scott W. Unger and John Sirmon
National Weather Service Forecast Office
Birmingham, Alabama

1. Introduction

With the wide variety of decision tools available to forecasters for use during severe weather warning operations, the accuracy and reliability of certain products in the NWS Advanced Weather Interactive Processing System (AWIPS) are often questioned. The hail detection algorithm (HDA; Witt et al., 1998) is one of these tools. Originally deployed along with the storm cell identification and tracking algorithm (SCIT; Johnson et al., 1998), the HDA uses the vertical profile of reflectivity created by the SCIT to determine the hail potential of a cell-based thunderstorm. The HDA algorithm output provides the warning meteorologist with three parameters: the maximum estimated hail size (MEHS), the probability of hail (POH) and the probability of severe hail (POSH). Though this algorithm has been employed for 10 years, the only study of its effectiveness in the southeastern United States was completed in 1998 and is specifically noted for its use along the immediate Gulf Coast (Lenning and Fuelberg, 1998).

Another analysis tool utilized by warning forecasters is vertically integrated liquid (VIL; Greene and Clark, 1972) and the associated VIL of the Day (VOTD; Paxton and Shepherd, 1993). VIL was originally thought to gather *in situ* measurements to calculate the liquid water content of an averaged 4 x 4 km gridded cell, but it was quickly determined that the resultant estimation represented integrated reflectivity. Furthermore, since it is assumed that VIL only estimated liquid water, column reflectivity values less than 18 dBZ are removed from the calculation and values greater than 56 dBZ are truncated to 56 dBZ. The higher reflectivities are assumed to be ice contaminated and are thereby removed by the original calculation.

While the VIL product is often used by warning forecasters to make a determination of hail, its most useful feature represents the locations of significant reflectivity cores. Utilizing the VIL product exclusively to justify hail storms can often be misleading due to the known limitations. Being that the product consists of integrated reflectivity, higher reflectivity that exists below the freezing level will result in higher VIL output, while the maximum precipitation core is well below the hail growth zone. Also, severe hail thresholds are air mass dependent and need to be modified when using VIL in the warning decision process as the depths of the growth zones can vary greatly. VIL of the day (described in more detail in Section 2), is thought to be an estimation of severe hail by assuming the hail growth zone exists between the 400 and 500 mb layer. Again, this calculation is air mass dependent as the growth zone may change vertically between storm systems and certainly between seasons.

Due to these and other limitations, individual warning decision tools should never be used unilaterally, but each in conjunction with the other to assist the forecaster in making a

* Corresponding author address: Kevin B. Laws, National Weather Service, 465 Weathervane Rd, Calera, AL 35040-5427; email: kevin.laws@noaa.gov

more informed decision when issuing warnings.

The purpose of this paper is to assess the output of the HDA using a statistical comparison to storm data. This is accomplished by examining two things: first, a comparison of the HDA output to numerous hail reports followed by a second comparison of the traditionally used VOTD. This paper compares the two warning decision tools to 368 local storms reports (LSR) within the Birmingham, Alabama county warning area.

2. Data

Much thought was put into how to collect the data needed to successfully complete this study. It is the policy of the National Weather Service to issue warnings based on storm structure and attributes rather than based on location, i.e. if severe reports are likely to be received. This study is based solely on reports received, no matter where the storm occurred.

The comparison reexamines archived data between February 2005 and April 2007 from the WSR-88D radar (KBMX) located in Calera, Alabama. This data was initially compared to 821 LSRs of hail that were collected over the period of study. Based on the time and location of individual LSRs, each storm cell was analyzed for VIL and each of the HDA outputs, maximum estimated hail size (MEHS), probability of hail (POH) and probability of severe hail (POSH). A noted comparison is shown in Figure 1 from the tornadic outbreak across central Alabama on March 1, 2007. Each LSR available was used for this initial step. There was no consideration given to population in a given area.

Based on the time of the hail report and corresponding radar scan, the maximum VIL and each of the HDA components were recorded. The maximum grid-based VIL for a corresponding volume scan was only applied to an LSR if the reported time was after cell passage. If the report was before the cell passage over the reported location, the LSR was discarded. The LSRs were also discarded if there was no usable archived data. This process created nearly 500 useable data sets.

The next data set that was collected was the upper air data for each of the days examined. Upper air data was collected from the National Oceanic and Atmospheric Administration's (NOAA) radiosonde database, located at the website http://raob.fsl.noaa.gov. Data collected included the temperature of the 400 and 500 mb levels. If upper air data was not available for that day, the LSR was discarded. This process reduced the data set to 368 useable LSRs. VOTD was then calculated for each day. If there were two upper air runs for a particular day, both were calculated and applied to the appropriate LSR. VOTD is calculated using the following equation from Paxton and Shepherd (1993):

$$VOTD = 750 / [(T500 + T400) / 2],$$

where T500 and T400 are the environmental temperatures at 500 and 400 mb, respectively.

The data were then compiled and sorted by significant event. At this point in the analysis, it was decided that only two parameters from the HDA should be included in the statistical output; from here forward there will be no mention of POH, as much of the data show that this parameter is very likely to be 100% when any size hail is present. The HDA, including both the MEHS and POSH, and the VOTD were compared to the reported hail size of each LSR.

3. Analysis Highlights

When each of the 368 LSR's was analyzed by a trained meteorologist, several key parameters were collected from each event. As mentioned, POSH and MEHS were gathered along with the sounding parameters that were instrumental in making the calculations in the HDA – the freezing level, -20 degree level and

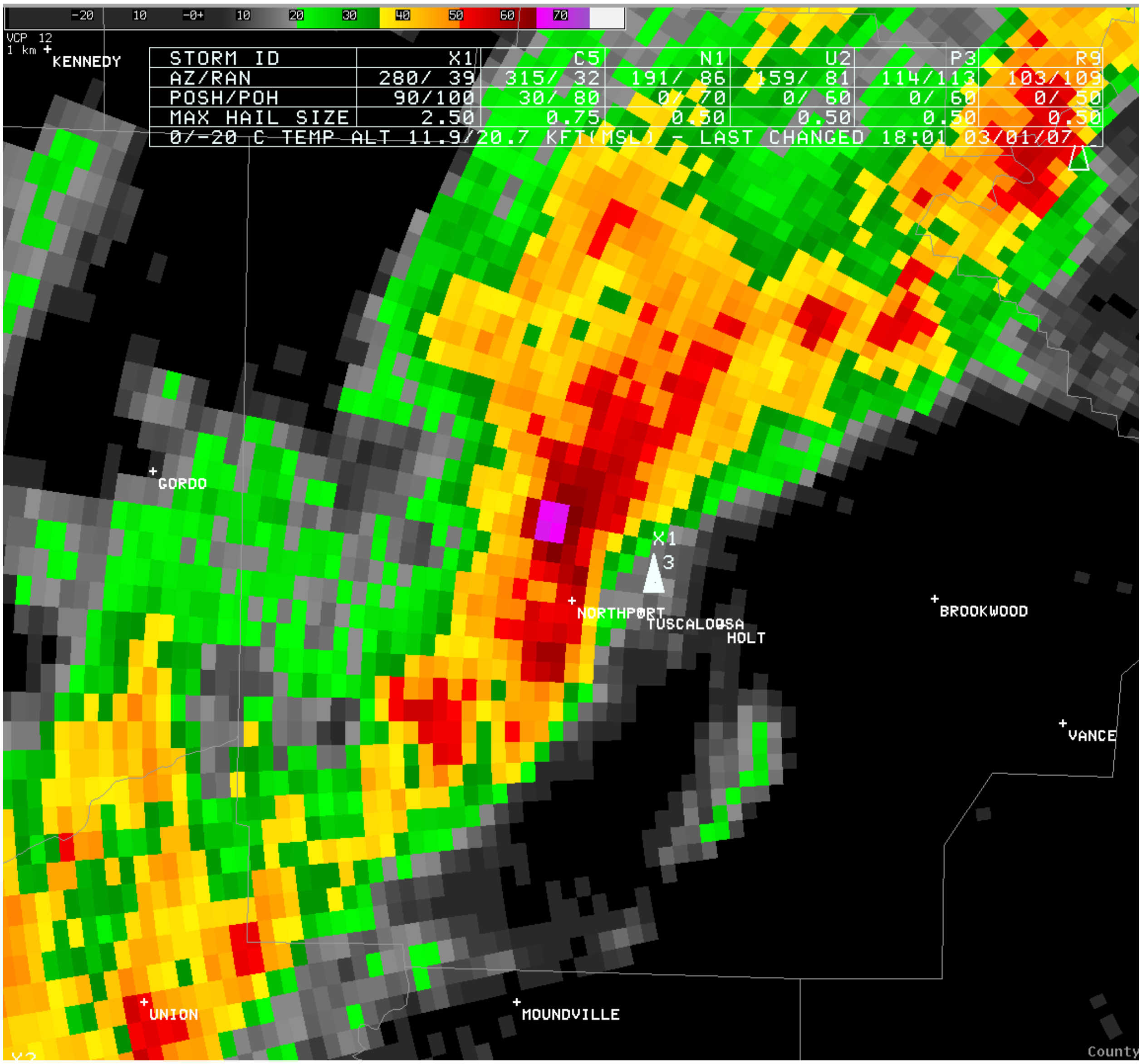

Figure 1. Example of the hail detection algorithm being applied to a supercell on March 1, 2007.

Table 1. Example of the data collected from each LSR or event. This sample was obtained from the storm shown in Figure 1.

Date	Time UTC	Time CST	City	County	Magnitude	Lat	Lon	VIL	D-VIL
3/1/2007	22:44	4:44 PM	1E NORTHPORT	TUSCALOOSA	0.88 in	33.24	-87.57	40	80

POSH	POH	Max Hail Size	0 degree Height	-20 degree Height	Environmental Update	T400mb	T500mb	VOTD	LI	CAPE	PW (mm)	PW (in)
90%	100%	2.5 in	11.9	20.7	18Z	-21.9 C	-10.7 C	46	-2.47	965	40.42	1.59

the corresponding sounding time that was used as initial input into the algorithm within the Radar Product Generator (RPG). For comparison purposes the maximum grid-based VIL along with the calculated VOTD were also collected for each event.

Products such as Digital VIL, and several parameters from a nearby sounding such as lifted index, convective available potential energy and precipitable water values were also collected for each event. Although not analyzed in this study, these parameters will prove useful in future studies. Table 1 shows an example of the parameters that were collected per event or LSR from the event mentioned in Figure 1.

Table 2 shows the number of severe hail reports, which is defined by the National Weather Service as greater than or equal to 0.75 inches. As shown, the number of severe reports (84% of the events) greatly outweighs the non severe events. This is often the case as non severe hail is usually underreported or if reported, LSRs are typically not generated for small hail.

Also shown in Table 2 are the distributions of severe and non severe reports as they relate to POSH. It was decided for the purposes of our study any POSH greater than 50% would be the most viable candidate to contain severe hail so the table is categorized as such. As expected, as the POSH increases from 50 to 100%, the frequency of severe reports increases from 75 to 94%, respectively. The false alarm ratio (FAR) or when the events were reported less than 0.75 inches is also shown to decrease significantly as POSH increases. It is notable that a large distribution of the events occurred when POSH is 100% (67 of 368 reports), which is not discussed within the scope of this study.

When incorporating the VOTD qualifier or when the VOTD criteria is met for each event, Table 2 shows that when POSH values of 80% are exceeded, little to no useful information is added to the frequency of severe to non severe reports. However, when the POSH is equal to the marginal percentages (50 to 70%), the data set shows some improvement in the false alarm ratio.

There are numerous ways to depict hail size distributions when compared to POSH. For this study, it was decided that the frequency of occurrence of each reported size would be compared based on non severe hail (<0.75 inches), marginally severe hail (0.75 – 1.25 inches), medium hail (1.25 – 1.75 inches), and large hail (>1.75 inches).

In Table 2, the greatest distribution of non severe hail is located below 40% POSH, however there were scattered reports above 50% including 7% of the events with a calculated POSH of 100%.

The greatest number of reports was generated from the marginally severe hail events (58%). It was hoped that this category would show the most significant improvement as POSH increases as most warning decisions are made at this categorical threshold. Table 2 reveals that little to no improvement was shown as POSH increases per these marginal events. It is notable here that many of the events (32%) indicated a POSH of less than 40%.

The most encouraging data set was obtained from the large hail cases, even though it contained the least amount of events (3%). As POSH reaches 100%, half of the large hail reports are received. Likewise, it was shown that as POSH decreases the number of large hail reports significantly decreases with only 1 case below a POSH of 70%.

When looking at the distribution of severe and non severe reports compared to the MEHS, the results are shown in Figure 2. Similar to the POSH, as the MEHS increases, the reports of severe hail also increase, with MEHS greater than 4 inches showing 100% of the severe distribution. It was encouraging once again to note that when the MEHS was less than severe (<0.75 inches), the number of severe reports decreased somewhat to 68%.

Table 2. Comparison of severe hail reports to probability of severe hail from the hail detection algorithm. The reported hail size distribution is also compared to POSH.

			POSH <= 40%		POSH 50%		60%		70%		80%		90%		100%	
	Frequency	Percentage	Frequency	Percentage												
Severe Hail Reports (>= 0.75 Inches)	310	84%	78	80%	27	75%	31	78%	46	85%	37	86%	28	93%	63	94%
Severe Hail Reports with the VOTD Qualifier			89	78%	25	74%	29	83%	44	86%	35	88%	26	93%	62	94%
Non Severe Hail Reports (< 0.75)	58	16%	20	20%	9	25%	9	22%	8	15%	6	14%	2	7%	4	6%
Non Severe Hail Reports with the VOTD Qualifier			25	22%	9	26%	6	17%	7	14%	5	12%	2	7%	4	6%
Hail Size Reports <0.75"	58	16%	20	34%	9	16%	9	16%	8	14%	6	10%	2	3%	4	7%
Hail Size Reports 0.75 - 1.25"	215	58%	68	32%	21	10%	26	12%	26	12%	22	10%	17	8%	35	16%
Hail Size Reports > 1.25 - 1.75"	81	22%	10	12%	5	6%	5	6%	18	22%	13	16%	9	11%	21	26%
Hail Size Reports > 1.75"	14	3%	0	0%	1	7%	0	0%	2	14%	2	14%	2	14%	7	50%

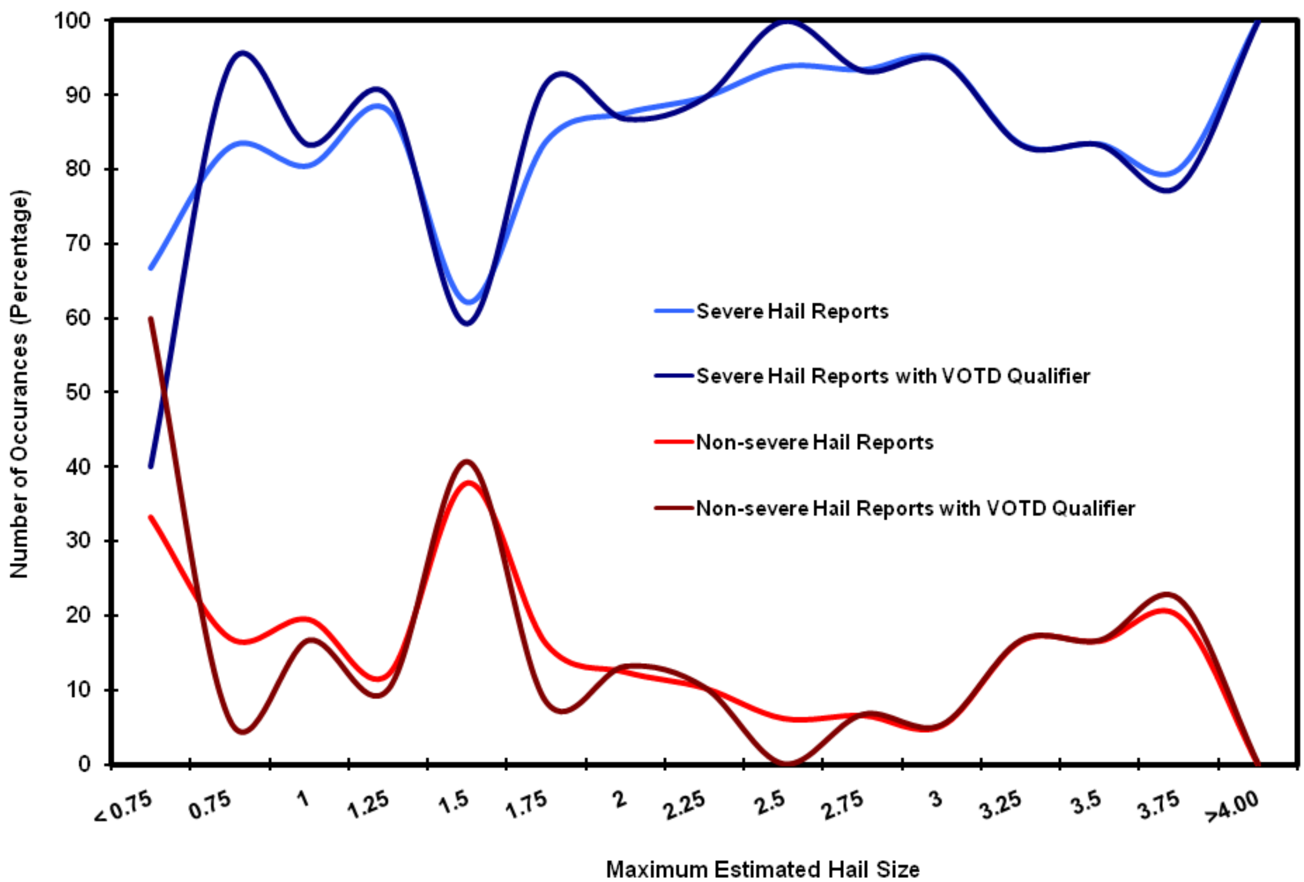

Figure 2. The maximum estimated hail size from the hail detection algorithm when compared to severe and non severe hail reports. Additional information is shown when the VIL of the Day qualifier is applied to the reports.

When VOTD is met, there is little to no improvement when the MEHS is above the marginally severe range (>1.25 inches). However, when the marginally severe hail is indicated, adding VOTD information to a warning decision is shown in the data set to increase the chances of receiving a severe hail report.

The most disappointing results shown in Figure 2 are when the MEHS output is not severe. When adding the VOTD qualifier, the FAR nearly doubles. This suggests that when the VOTD criteria is met, but the MEHS is showing hail less than 0.75 inches, the number of non severe reports received increases from 33 to 60%.

Acknowledgements

The authors would like to thank Kevin Pence and Paul Schlatter for their useful comments and suggestions, and for help in clarifying scientific concepts within this study. We would also like to express our appreciation to Kristina Sumrall for assistance in the writing process.

References

Greene, D.R. and R. A. Clark, 1972: Vertically integrated liquid water-A new analysis tool. Mon. Wea. Rev., **100**, 548-552.

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

Lenning, Eric and Henry E. Fuelberg, 1998: An Evaluation of WSR-88D Severe Hail Algorithms along the Northeastern Gulf Coast. Wea. Forecasting, **13**, 1029-1042.

Paxton, C.H. and J.M.Shepherd, 1993: Radar diagnostic parameters as indicators of severe weather in central Florida. NOAA Tech. Nemo. NWS SR-149, 12 pp. [Available from National Weather Service Southern Region Headquarters, 819 Taylor St., Room 10A26, Ft. Worth, TX 76102]

Witt, A., M.D.Elits, 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.