WESTERN REGION TECHNICAL ATTACHMENT
NO. 98-27
JULY 21, 1998

AN EXAMINATION INTO THE UTILITY OF THE BUILD 9 HAIL DETECTION ALGORITHM (HDA) ACROSS NORTHERN UTAH

Randall A. Graham - NWSFO SLC
and
James A. Nelson - NWSFO SLC/ WRH-SSD

Introduction

The development and maturation of hailstones are very complex processes. Numerous factors impact the resultant size of the hailstone including updraft strength, storm scale wind profile, height of the freezing level, and the mean temperature and relative humidity of downdraft air (Johns and Doswell 1992). The complexities of hail formation and sub-cloud processes make utilizing Doppler radar data to forecast the occurrence of large hail difficult. Verification of hail events is also important, but is a cumbersome process due to the limited temporal and spatial distribution of the event. Verification is further complicated in areas with low population densities, such as the Intermountain West.

To help improve forecasts of large hail, the National Severe Storms Laboratory (NSSL) developed an improved Hail Detection Algorithm (HDA) which was introduced into the WSR-88D radar in software Build 9. Three products are produced by the algorithm: 1) the Probability of Hail (POH), 2) the Probability of Severe Hail (POSH), and 3) the Maximum Estimated Hail Size (MEHS). In a previous study for northern Utah (Vasiloff et al. 1997), the new HDA was found to perform well in large hail events, but tended to overestimate the POSH for weaker cells. It was also noted that large hail was associated with large CAPE and strong vertical wind shear. In response to similar reports of an over-warning bias from other high elevation radars, the NSSL made changes to the algorithm. In particular, NSSL adjusted the Warning Threshold Selection Model (WTSM) and POSH Offsets for high elevation radar sites (Witt 1998). The changes are detailed later in this Technical Attachment (TA).

The purpose of this study is to compare the skill of the Build 9 HDA (in particular the POSH) with the original WTSM Offset versus the new WTSM Offset suggested by NSSL. In addition, an attempt will be made to determine if the elevation difference between the Promontory Point (KMTX) radar and most valley locations across northern Utah has a detrimental effect on algorithm performance. In particular, results of the Build 9 HDA using the height of the 0C and -20 C levels above the RDA (using both the original and updated WTSM Offset) will be compared to computing the same environmental input with respect to the valley floor. The HDA output will be compared with ground truth reports for eight storm days from the 1996 and 1997 convective seasons. No attempt has been made to discriminate among storm type.

Topography of Northern Utah

Northern Utah marks the boundary between the western edge of the Rocky Mountain chain and the eastern portion of the Great Basin. The Promontory Point WSR-88D (KMTX), at an elevation of 6570 feet MSL, is located on a peninsula along the northern portion of the Great Salt Lake. The Wasatch Front extends roughly from Brigham City (BMC in Fig. 1), at the northern extent, to Provo (PVU) in the south. This 80 nm (~145 km) strip is home to approximately 85% of Utah's population. Most of the valley locations along this stretch are at an elevation of around 4200 feet MSL. Directly to the east of the Wasatch Front lie the Wasatch Mountains with peaks generally from 9500-11500 feet MSL.

Description of the WSR-88D Build 9 Hail Detection Algorithm

Hail growth occurs above the freezing level with most large hail growth occurring above the -20C level. Thus, the HDA focuses on high reflectivity values above the freezing level and -20C level. The WSR-88D Build 9 HDA was based primarily on empirical studies conducted on low-elevation radar sites in the eastern United States (Witt 1997). While the HDA performed reasonably well for low-elevation radars (Barjenbruch and Laplante 1997, Klimowski et. al. 1997), it had an over-warning bias for high elevation radars in the West (Witt 1998). For use in the original Build 9 HDA, the height of the 0C and -20C levels above radar level are used in the calculations (e.g., Klimowski et at. 1997 and Vasiloff et al. 1997). Essentially, the HDA is computing the POH, POSH, and MEHS for an elevation of 6570 feet MSL. The 2300 foot difference between the RDA and the majority of valley locations (and population) is not accounted for in the algorithm.

The Probability of Hail (POH) is determined simply by comparing the maximum height of the 45 dBZ echo to the freezing level. The greater the height of the 45 dBZ echo above the 0C level the greater the POH.

The calculation of the Probability of Severe Hail (POSH) is more complicated. A vertical integration is done in which reflectivities greater than 40 dBZ that are above the freezing level are evaluated. A weighting factor is used in which the greater the reflectivity and the higher it is above the freezing level, the greater the weight given to that echo. Reflectivities which are greater than 50 dBZ and above the -20C level are given the most weight. The result of this vertical integration is the Severe Hail Index (SHI). The WTSM, which is a function of the height of the 0C level and several empirically derived values, and the SHI are then used to generate the POSH. Due to this method of derivation, the POSH is dependant on the height of the 0C and -20C levels, as well as the elevation of the radar.

Maddox et al. (1997) showed how the POSH for the same storm varied among radars at different altitudes. In order to address these differences, the NSSL administered the following changes to the WTSM and POSH Offsets.

The equation for POSH is:
$\mathrm{POSH} = 29\ln(\mathrm{SHI}/\mathrm{Wt}) + \mathrm{POSH\ offset}$

where SHI is the severe hail index. Wt is defined as:
$\mathrm{Wt} = 57.5(\mathrm{H0}) - \mathrm{WTSM\ offset}$

where H0 is the height of the freezing level (km) above the radar.

There were two changes to the HDA:

1) The 0C height above the radar level is still used (actually the input value is for the station height and then the radar height is used to make the final height relative to the radar). The WTSM (or Wt) offset is changed effectively making the HDA relative to height above sea level; and

2) POSH offset is reduced to 30% from 50% during "summertime" storms.

Let's use KMTX as an example for how the new WTSM offset will affect POSHs.

Assume 0C height is 3 km above KMTX
Assume the Severe Hail Index ~100

OLD METHOD: $\mathrm{Wt} = 57.5(\mathrm{H0}) - 121$, where H0 is the freezing level in km above the radar (ARL); -121 is the WTSM offset.

$\mathrm{WT} = 51.5$ and $\mathrm{POSH} = 29\ln(100/51.5) + 50 = 70\%$

NEW METHOD: $\mathrm{Wt} = 57.5(\mathrm{H0}) - 5.8$, where H0 is the freezing level in km ARL; -5.8 is the new WTSM offset.

$\mathrm{WT} = 166$ and $\mathrm{POSH} = 29\ln(100/166) + 50 = 35\%$

Note that for the same storm there is a difference in POSH of 35%. This does not account for the further reduction that will be caused by lowering the POSH offset from 50 to 30.

Methodology

Using the WATADS software, the HDA was evaluated for eight storm days. Three different ways of calculating the POH and POSH were examined for each event. The first run used the original WTSM Offset. In this run, the height of the 0C and -20C levels are input into the algorithm with respect to the RDA. These will be referred to as ARL (Above Radar Level) runs. In the second run, the POSH was calculated as if the RDA were at valley level instead of on Promontory Point (4200 feet instead of 6570 feet MSL). Thus, the 2370 foot difference between the valley floor and the RDA would effect the freezing level influence the calculation of POH and POSH as detailed in Fig 2. These will be referred to as AVL (Above Valley Level) runs. The final run was the same as the original (ARL) run except with the new WTSM Offset recommended by NSSL was used in the algorithm. These will be referred to as WTSM (Warning Threshold Selection Model) runs.

For each of these runs the POH and POSH were recorded for cells for which there was reliable ground truth. It was assumed that the hail reports were accurate with respect to time and size. It was also assumed that the largest hail size reported was the largest hail that reached the surface. The majority of the cells used in this study moved over highly populated areas and produced multiple hail reports. For the null/no hail cases only those cells that moved over the densely populated regions of the Wasatch Front were used.

Results

Eight storm days were examined in which algorithm output was compared to ground truth reports across northern Utah. The POH was compared amongst the three runs (ARL, AVL, and WTSM) for selected cells. Only those volume scans in which the POH was 50% or greater in the original ARL runs were used (approximately 300 volume scans). The 50% threshold was chosen in order to eliminate weak, nondescript cells. The POH in the AVL runs was always less than or equal to the POH in the ARL runs (due to the addition of 2300' onto the height of the freezing level). The difference between the POH for the ARL runs and AVL runs was 20% or less 81% of the time. Since the new WTSM Offset does not affect the POH values, there was virtually no change in the POH between the ARL runs and the WTSM runs. In general, the HDA output for POH did not change significantly for the three different runs.

A much greater change was seen in the values of the POSH from run to run. Again, only those volume scans in which the POH in the original HDA runs (ARL runs) was 50% or greater were used for comparison. Running the HDA, with respect to the valley floor and using the new WTSM Offset both, resulted in substantially lower values of POSH than the original ARL runs (Fig. 3). The POSH decreased by an average of 30% per volume scan between the ARL and the AVL runs. Over half of the volume scans in the sample (55%) experienced a decrease of 30% or more. The change between the ARL and the WTSM runs was even more dramatic. The POSH dropped by an average of 43% per volume scan in the WTSM runs compared to the ARL runs. Over half (53%) of the scans showed a decrease of 50% or greater.

Average changes in the values of POH and POSH were only an initial measuring stick of the affects on the performance of the algorithm. The true utility of the HDA, and of most importance to the warning forecaster, is found in the algorithms' skill in differentiating cells which are likely to produce severe hail at the surface versus those which will produce small hail or no hail at all.

The POSH values were recorded for each volume scan of the three runs of the HDA. The maximum POSH recorded within thirty-five minutes of the largest hail report for an individual cell was used to compare the HDA to ground truth reports. For storms that produced no hail, the maximum POSH within fifteen minutes of the cell being over a highly populated area was used. Great care was taken in the null (no hail) cases to use only those cells which were mature, and subsequently collapsed, over densely populated areas where the Salt Lake City National Weather Service Forecast Office has numerous spotters.

Severe Hail Cases

For the severe hail cases (hail diameter 1.9 cm or greater), the average maximum POSH for the ARL runs was 94% (Fig. 4). For the AVL runs, the maximum POSH averaged 71% while the WTSM runs averaged just 50%. The ARL runs generated high POSH estimates for the strongest cells (maximum POSH 70% or greater in all severe cases). The AVL runs also did a good job of generating high POSH values in the severe hail cases (POSH of 70% or greater in 75% of the cases). However, the WTSM runs drastically underestimated the POSH for the severe events generating a POSH of 70% or greater in only 13% of the severe cases (Table 1).

Small Hail Cases

For the small hail cases (hail diameter less than 1.91 cm) the ARL runs averaged a maximum POSH of 71% (Fig 4). The AVL runs averaged 39% and the WTSM runs averaged just 23% (Fig. 4). One of the major complaints with the Build 9 HDA was that, while it had high POSH values for the majority of severe cells, it tended to over-warn on weaker cells. This is certainly evident in this data set. The ARL maximum POSH values were 80% or greater in half of the small hail cases and 70% or greater in 79% of the cases. Thus, the ARL runs resulted in a high probability of detection for the severe hail cases it also generated a substantial false alarm ratio (POSH of 70% or greater with no severe hail) in the small hail cases.. The AVL runs on the other hand generated high POSH values for the severe hail case and produced values of 70% or greater in only 21% of the small hail cases (36% had a POSH of 60% or greater), a substantial improvement over the ARL runs (Table 1).

The AVL runs had a difference of 32% in the average maximum POSH between the severe hail cases and the small cases (severe 71% - small 39%), compared to a 23% difference for the ARL runs (severe 94% - small 71%). The AVL runs also resulted in significantly fewer false alarms. Thus, the AVL runs exhibited more skill than the ARL runs in separating the large hail cases from the small hail cases. The WTSM runs had a difference of 27% (severe 50% - small 23%) from the severe hail cases to the small hail cases, but dramatically under forecast the severe cases resulting in a low probability of detection (Fig. 4).

Null Cases

For the null cases, the ARL runs averaged a maximum POSH of 50% with several false alarms (42% had a POSH greater than or equal to 70%). The AVL runs averaged a maximum POSH of 20% while WTSM runs fell to 7% (Fig. 4). The ARL runs' tendency to overestimate is evident even in the weak cells that produced no hail. The AVL and WTSM runs both generated only low POSH values for the null cases with no false alarms.

Conclusions

The POH and POSH were calculated three different ways (ARL, AVL, and using the new WTSM Offset) for eight storm days from the convective seasons of 1996 and 1997. The differences among values of POH and POSH for the three runs were calculated for volume scans in which the cell of interest had a POH of greater than or equal to 50% (from the ARL runs). When compared to the ARL runs, the POH in the AVL runs dropped by 20% or less in 81% of the volume scans. There was essentially no change in the POH between the ARL and WTSM runs. The POSH estimates differed significantly in the three runs. The POSH values dropped by an average of 30% for the AVL runs and 43% for the WTSM runs when compared to the original ARL runs. Thus, the AVL runs showed greater skill than the ARL and WTSM runs in differentiating severe hail storms from those which contained only small hail. The ARL runs had a significant number of false alarms for both the small hail and null cases.

The new WTSM Offset resulted in large decreases in the POSH values and significant underestimates for the severe hail cases resulting in a low probability of detection. Thus, the WTSM Offset appears to have little skill and simply resulted in significant underestimates, as opposed to the overestimates that the high elevation radars have experienced using the ARL method. The most skill was established when using the freezing level and -20C level relative to the valley floor.

This study was conducted on a limited data set (eight storm days and approximately thirty storm cells) and for only one radar (Promontory Point - KMTX). However, the results demonstrate the ability to tailor the HDA to a specific radar and terrain situation in the Intermountain West.

Acknowledgments

The authors would like to thank Steven Vasiloff, NSSL/WRH-SSD, for his many helpful and constructive comments. We would also like to thank him for reviewing this TA to improve its quality and effectiveness.

References

Barjenbruch, K. M., R. Laplante, 1997: An Assessment of the WSR-88D Hail Detection Algorithm's Performance Over Northern Ohio and Northwest Pennsylvania. Submitted to *National Weather Digest.*

Johns, R. H., C. Doswell III, 1992: Severe Local Storms Forecasting. *Wea. Forecasting*, **7**, 588-612.

Klimowski, B. A., S. Vasiloff, and A. Witt, 1997: On the Nature of the WSR-88D Build 9 Hail Detection Algorithm Part I: Review of the Algorithm and Operational Considerations. WR-Technical Attachment 96-25.

Vasiloff, S., R. Graham, J. Nelson, B. Klimowski, and A. Witt, 1997: On the nature of the WSR-88D Build 9 Hail Detection Algorithm Part II: Performance of the Algorithm over Utah. WR-Technical Attachment 96-26.

Vasiloff, S. and J. Nelson, 1998: Recent performance of the WSR-88D with new offsets. WR-Technical Attachment 98-26.

Witt, A., 1997: An enhanced hail detection algorithm for the WSR-88D. Submitted to *Wea. Forecasting.*

____, 1998: Evaluation of the Hail Detection Algorithm for High-Elevation WSR-88D Sites. Report prepared for the WSR-88D Operational Facility.

Wyatt, A., and A. Witt, 1997: The effects of population density on ground-truth verification of reports used to score a hail detection algorithm. Preprints, *28th Conf. On Radar Meteorology*, Amer. Meteor. Soc.

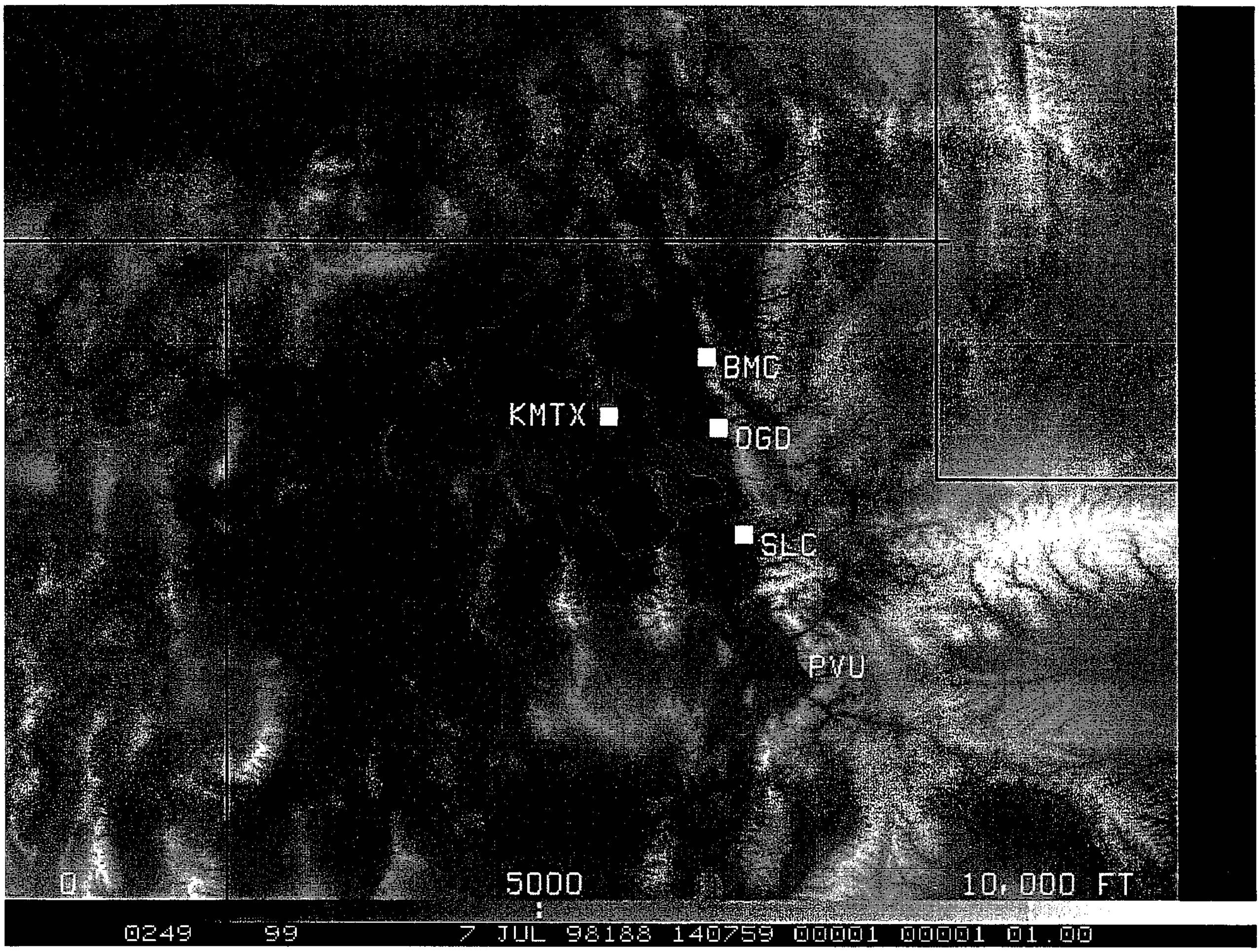

Figure 1. Close-up view of northern Utah. The location of the Promontory Point radar (KMTX) is denoted by a radar symbol at the north end of the Great Salt Lake. The Wasatch Front extends from Brigham City south through Provo (BMC - Brigham City, OGD - Ogden, SLC - Salt Lake City, PVU - Provo). The spine of the Wasatch Mountains is noted by the shaded area east of the Wasatch Front.

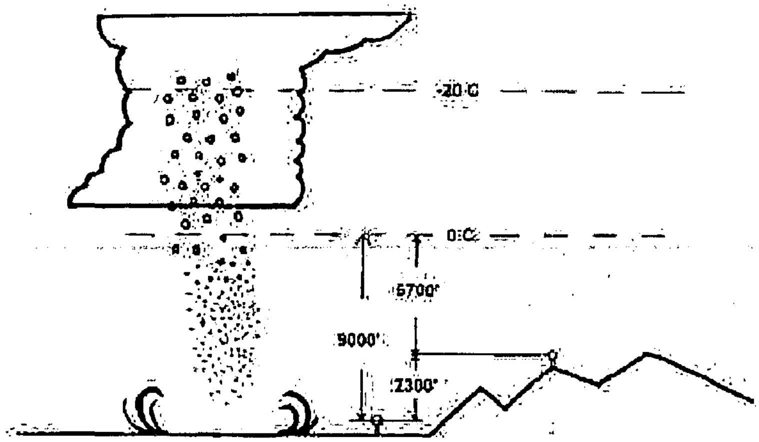

Figure 2. Schematic showing the difference in the relative height of the 0C level input into the HDA using the ARL (Above Radar Level) versus the AVL (Above Valley Level) method. In this case, instead of using the 6700' difference between the RDA and the 0C height, a value of 9000' would be used (6700' plus the 2300' difference between the the RDA and the average valley floor).

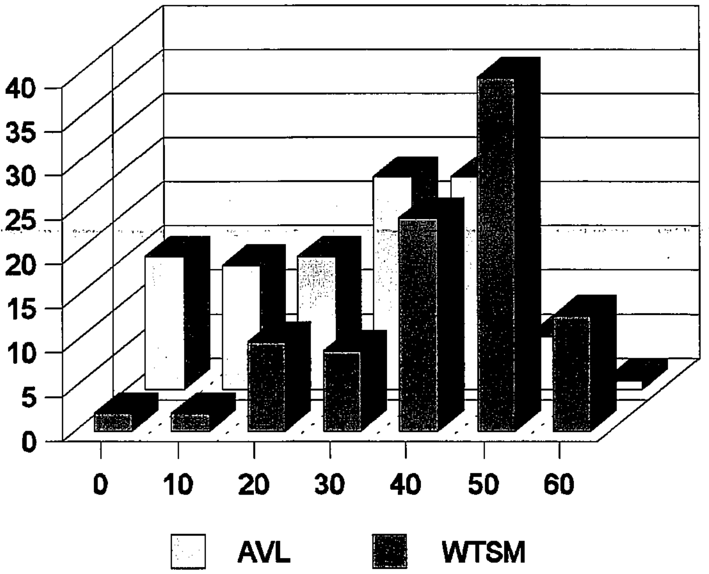

Figure 3. X-axis shows delta POSH (percent of change) for the AVL and WTSM runs versus the ARL runs. Y-axis shows the percent of volume scans in the study for which those values were recorded.

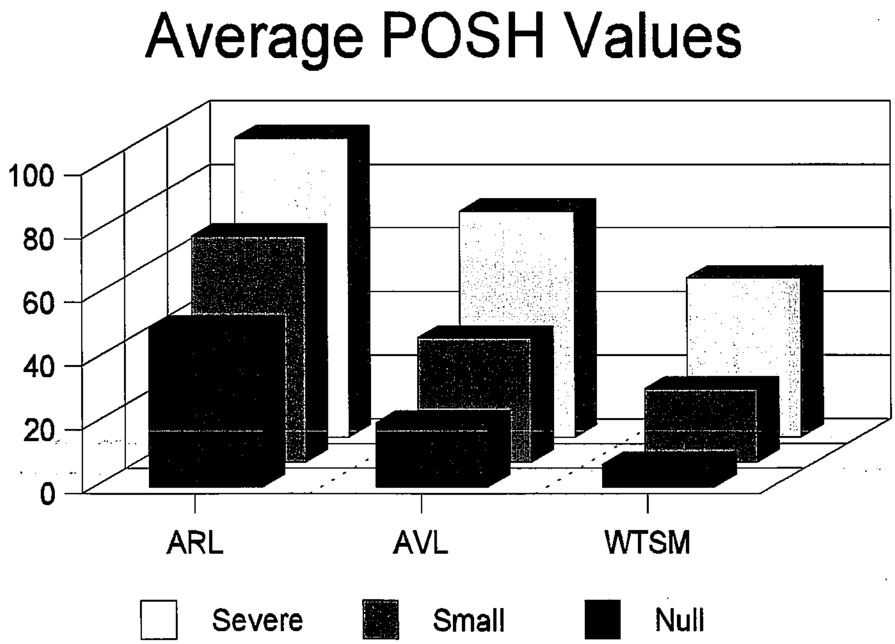

Figure 4. Average percent of maximum POSH values for the ARL, AVL, and WTSM runs for severe, small, and null hail cases.

	ARL	AVL	WTSM
Severe	100%	71%	13%
Small	79%	21%	0%
Null	43%	0%	0%

Table 1. Percentage of time that the three HDA runs produced POSH values greater than or equal to 70% for severe, small, and null hail cases