INVESTIGATION OF RADAR VARIABLES AND NEAR SURFACE ENVIRONMENTS FOR DEVELOPING A SURFACE HAIL FALL PRODUCT

Sarah K. Mustered[1, 2] and Kiel Ortega[3]

[1]National Weather Center Research Experiences for Undergraduates Program
and
[2]Valparaiso University, Valparaiso, Indiana

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

ABSTRACT

In 2006, the Severe Hazards Analysis and Verification Experiment (SHAVE) was formed to collect high resolution severe weather reports. The resulting dataset is a detailed and accurate record of surface hail fall. Currently, the exact processes that determine the size and distribution of hail at the surface is relatively unknown. While there are numerous products that address the presence of hail cores aloft, a gridded surface product is missing. The benefit of a gridded surface product is a more accurate understanding of surface hail size at any particular location. In this study, we incorporated the near surface environment (NSE) from the 20 km RUC analysis and existing radar products with SHAVE hail reports to determine if the NSE could be a beneficial component of a surface hail fall product. We found that the current resolution and reliability of the NSE data is too low for the addition of NSE variables to add significantly to the accuracy of the existing radar products.

1. INTRODUCTION

Each year, hailstorms affect thousands of people throughout the country, leaving behind significant impacts on life and property. Despite the widespread impacts of hail fall and theoretical knowledge of hail growth, relatively little is known about the complex processes that determine the actual hail size recorded at the surface. Even more of an unknown is the distribution of hail at the surface as opposed to a single maximum value expected from a particular storm. Existing algorithms developed for predicting maximum hail size have several limitations. Sounding based products such as HAILCAST (Brimelow et al. 2002) only produce a single value of maximum hail size for the day. While this value can be useful for forecasters, it provides no guidance in the distribution of the hail fall. Another sounding product commonly available to forecasters is the AWIPS hail algorithm. As explained by Jewell and Brimelow (2009), this particular version of the product was particularly inaccurate in high CAPE scenarios, sometimes forecasting hail as large as 86 inches. Radar based products for hail prediction include the Maximum Expected Size of Hail (MESH) and the Severe Hail Index (SHI), which is a component of MESH (Witt et al. 1998). This original algorithm was adapted from single-radar to a multi-radar gridded product by Stumpf et al. (2002). This product is useful in real-time hail forecasting, but is limited in that it is primarily represents the hail core aloft.

While the National Weather Service (NWS) has regularly collected hail reports for years, this source of hail data (*Storm Data*) is notoriously inaccurate. Jewell and Brimelow (2009) describe the problems in depth, particularly noting the massive bias for quarter-size (1”), golf ball size (1.75”), and baseball-size hail (2.75”). The public tendency to report hail as one of three categories combined with the low resolution of *Storm Data* (often only one report per storm) has been a long standing difficulty for hail research

[1] *Corresponding author address:*
Sarah Mustered, 812 S Race St, Urbana, IL 61801
Email: sarah.mustered@valpo.edu

(Witt et al. 1998, Ortega et al. 2006). These problems were the driving force behind the formation of the Severe Hazards Analysis and Verification Experiment (SHAVE; Ortega et al. 2009) in 2006. During the summer months, undergraduate students make phone calls to homes within the path of a target storm shortly after it passes with the goal of collecting high-resolution data of hazards including wind and flooding reports in addition to extensive hail data. The resulting reports are of fine resolution and encompass not only severe reports, but also reports of no hail and non-severe hail. This project seeks to make use of this high resolution data to better understand the conditions that affect surface hail size. The original goals of this research were to:

- Investigate the relationship of the near surface environment (NSE) of storms and existing radar products to the reported hail size
- Create a surface hail fall product with a spatial grid of expected hail size for any specific location

2. DATA AND METHODOLOGY

2.1 *Data*

This project made extensive use of pre-processed SHAVE data, spanning the summers of 2006 through August 2009. This period yielded 140 distinct storm cases with a total of 8,716 hail reports. Each case could vary in temporal and physical scale from a 60 minute period over Connecticut to a 5 hour sweep over several states. Single cases could also contain multiple storms as long as they existed in the same storm environment. The hail reports were spread throughout the continental United States, with a primary focus over the Central Plains and Midwest regions. Storm modes were varied, with many examples of short-lived, unorganized convection with minimally severe hail alongside more the larger hail sizes from supercells. However, it is important to note that an implicit limitation of SHAVE data is that it is limited to the summer months, since the project is not active during the rest of the year. The merged radar products used in this study are on an approximately 1 km x 1 km grid and are output from a real-time severe weather application, the Warning Decision Support System of Integrated Information (WDSSII; Lakshamanan et al. 2007). The NSE variables are derived from the 20 km RUC analysis.

2.2 *Radar Swaths*

In order to compare the existing radar products to the NSE data and hail size, we used WDSSII to create temporal and spatial swaths of three radar products:

- Low-Level Composite Reflectivity (LLCR), which is the maximum 0-3 km dBZ
- Maximum Expected Size of Hail (MESH; Stumpf et al. 2004)
- Vertically Integrated Liquid (VIL)

These swaths were then overlaid with the SHAVE hail reports. An example of one such swath is shown in Figure 1.

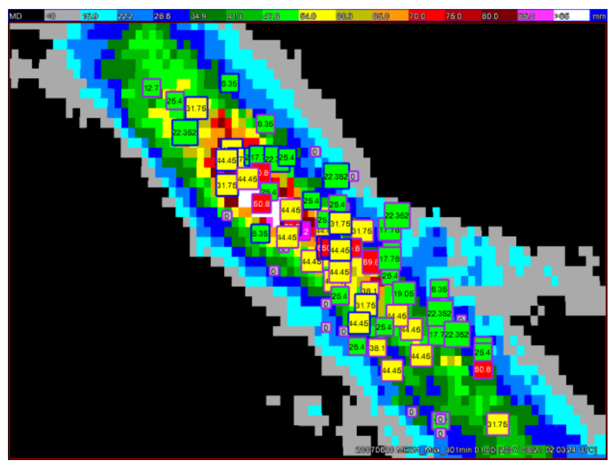

Figure 1*: An example of a MESH swath overlaid with SHAVE hail reports in mm.*

Each of the analyzed variables was chosen with the *surface* hail size in mind. Low-level composite reflectivity was selected in order to get an overall picture of the storm intensity near the ground. This would theoretically better capture the distribution of the surface hail reports than traditional composite reflectivity, which considers the entire column. MESH is a well established radar-based product that has many useful applications for studying hail. VIL was chosen due to its long-lived association with nowcasting hail and its accessibility to the meteorological community. However, several studies, including Edwards and Thompson (1998), suggest that VIL is not actually skillful at predicting hail size.

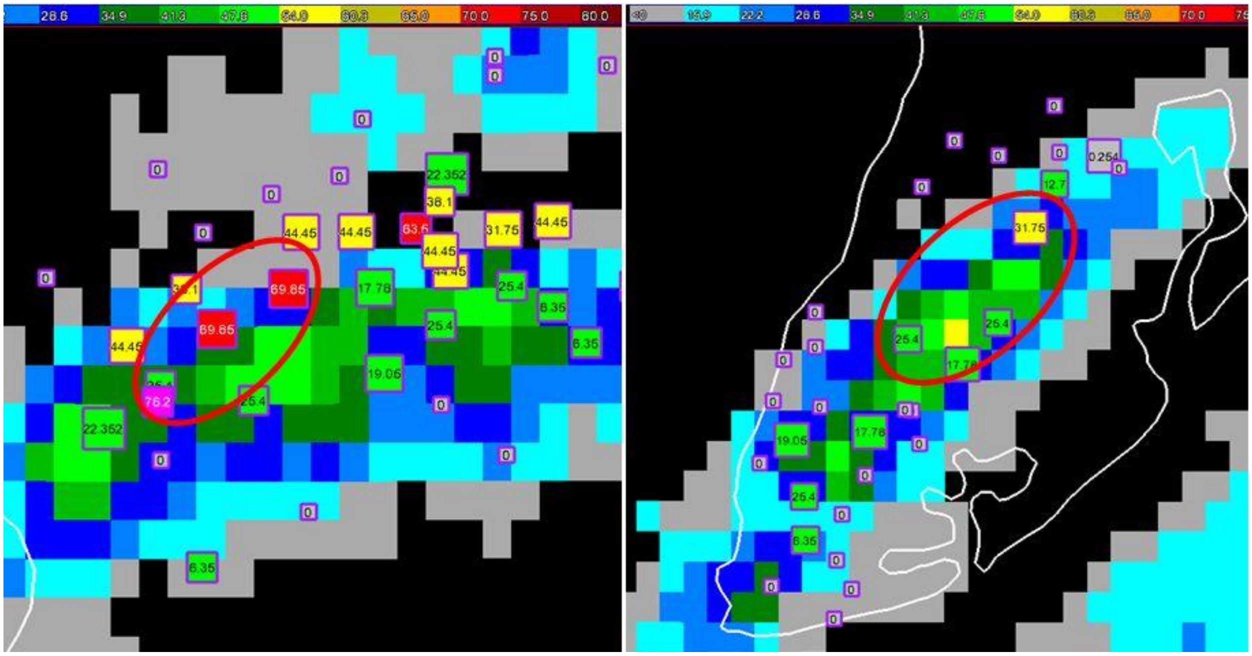

Figure 2*: Two similar MESH swaths from separate storm cases are displayed along with their associated hail reports. The areas of maximum hail size are circled. The storm on the left produced multiple reports of golf ball size hail (44.45 mm) with a max of 76.2 mm. In contrast, the storm on the right produced several marginal severe reports (25.4 mm) with a max of 31.75*

The processed radar swaths were part of an extensive manual investigation of the entirety of the processed dataset. For each case, maximum hail size was recorded along with verbal descriptions of the radar variables. This was done with the goal of creating a qualitative baseline for how current radar products compare to the high density SHAVE data. Particular attention was paid to similar swaths of MESH with drastically different surface hail falls. Figure 2 provides an example. These cases were instrumental in choosing which environmental parameters to investigate. This process included manual quality control. Due to time constraints, any cases with errors in mergers were completely eliminated from the analysis rather than repaired. Since this applied to fewer than 10 of the possible cases, this was considered a reasonable solution.

2.3 *Selection of the Near Surface Environments*

In order to see if the addition of variables from the near storm environment improved forecasts of hail size over predictions with radar data alone, we selected five potential variables that are commonly thought to be important to hail size.

- Surface Mixing Ratio (Stumpf et al. 2004)
- Surface Theta-E
- Height of Theta-E Minimum (H_{TEM})
- Height of the Wet-Bulb Zero (H_{WBZ}; Marzban and Witt 2001)
- Mean Relative Humidity From the Surface to 0°C (Witt et al. 1998)

Most of the selected variables are essentially different methods for expressing moisture, which is particularly important when considering hail size. The microphysical properties of ice make hail very sensitive to the moisture of its environment, but this relationship is very complex and varies based on hail size. By choosing several variables associated with the melting/evaporation process, we maximized our chances for finding a trend in the environmental data.

Due to time constraints and practical concerns in managing the large dataset, a single storm environment was chosen for each storm case. In other words, each case was assigned a single hourly RUC analysis. While this proved problematic for long-track storms, most cases were able to be assigned an environment that did not change drastically over the duration of the event. Analysis time were chosen by hand for each case by overlaying the hail reports over the MUCAPE (most unstable convective available potential energy) from the RUC analysis and looping the composite reflectivity. Based on the time that the storm passed over the reports, the most representative environment was subjectively selected. In cases where the best environment was not clear (such as a storm that was active from 2030-2130Z and had an equal claim on the 20 and 21Z environments), the time with the maximum MUCAPE was used. MUCAPE was chosen as a simple proxy for the overall storm environment, but is only an arbitrary cutoff.

In some cases, multiple distinct storms occurred within the same domain, but at different times. This was especially common in the older SHAVE data, which favored a single large area rather than several smaller cases per day. In order to account for this change in case divisions and provide a more accurate representation of the storms and their environments, a manual check was performed on any case with multiple distinct MESH tracks. If storms occurred more than 3 hours apart or if the NSE changed significantly (MUCAPE changed by over 500 J/kg) between the storms' durations, the case was split. Each split was then treated as a new, entirely separate case. In order to do this, we modified the WDSSII programs to analyze a region within a user-selected box.

2.4 *Data Analysis*

For this project, we chose to do a grid point by grid point analysis. The drawbacks of this method are discussed in later sections. The benefit of this approach is its simplicity. In WDSSII, an algorithm called "PointMatch" matches up the location of the SHAVE hail report, called "Truth", with the values of the selected radar products and NSE variables at that corresponding time and location within the grid cell. The algorithm matches the values of the grid directly, without taking the surrounding grid cells into consideration.

3. RESULTS AND DISCUSSION

Using the merged data, we created a series of scatter plots of the 5 NSE variables vs. the 3 radar products, leading to a total of 15 plots. Each point was color coded according to four possible categories for recorded hail size:

- "Null": No hail reported
- "Non-Severe": Hail was less than 1" (25.4 mm)
- "Severe": Hail was greater than or equal to 1" (25.4 mm) and less than 2" (50.8 mm)

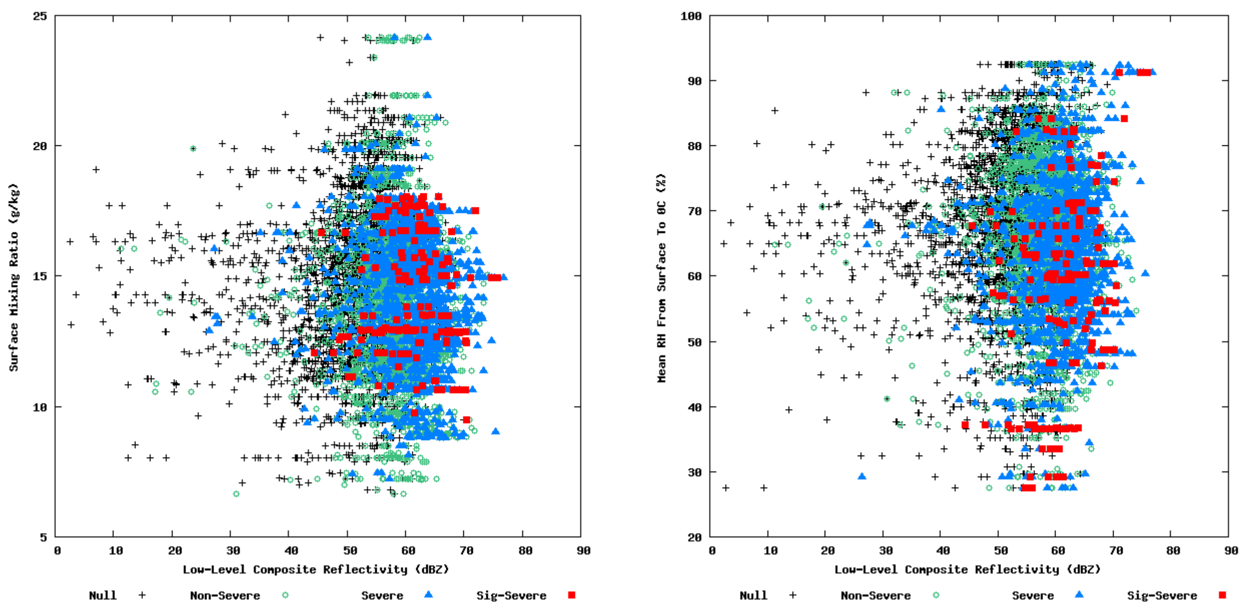

Figure 3: *Scatter plots of LLCR and NSE variables are shown. Surface mixing ratio is analyzed in the left plot and the mean RH from the surface to the height of 0°C is the right plot. Each point is color-coded according to recorded hail size. Red squares represent significant severe (>50.8mm), blue triangles represent severe (<50.8mm and >25.4mm), green circles represent non-severe (<25.4 mm), and black crosses represent null hail reports.*

- “Significant Severe”: Hail was greater than or equal to 2” (50.8 mm)

The goal of these plots was to determine if the NSE data improved the separation of the hail categories over the radar products alone. Overall, we found that NSE added little to the radar data.

Selected scatter plots are shown for each of the radar variables. Figure 3 shows that the addition of the NSE variables has little, if any, impact on the distribution of the hail categories when compared to LLCR. This is evident because the stratification of the hail sizes is essentially confined to the horizontal axis. These plots do show a relatively clear LLCR cutoff for severe hail near 53 dBZ. It is worth noting that the plot of mean RH below the melting level reveals significant hail events for both moist and dry environments. This emphasizes the complexity of the physical properties of hail. Ideally, large hail will shrink more quickly in a dry environment while small hail will be more prone to melting in a very moist environment. However, there is a cluster of several significant severe cases clustered in the driest environments in the study (RH < 40%). Our plots suggest that this idealized relationship is not a reliable indicator of surface hail size in actual storms.

Figure 4 contains the two plots of MESH that showed the best separation. While the separation of the categories on both plots is less than ideal, there is more than in the plots of LLCR. Somewhat surprisingly, significant severe hail shows the greatest response to the addition of the NSE parameters. The plot of surface mixing ratio shows this particularly well, with the significant reports beginning along a defined diagonal. In other words, for lower MESH values, large hail was associated with lower mixing ratios. Higher values of MESH (> 25 mm) show an even distribution of reports between 10-17 g/kg, suggesting that the addition of surface mixing ratio is not beneficial for these ranges. The H_{WBZ} has similar findings, with the largest impact for lower MESH cases (< 30 mm). These plots also reveal overall problems with MESH, especially for significant severe hail. Overall, MESH has an acceptable separation of the hail categories, with the majority of severe hail reports occurring over the severe MESH threshold (25.4 mm). However, significant hail is spread evenly from 25-100 mm, with many reports falling in the non-severe predicted range.

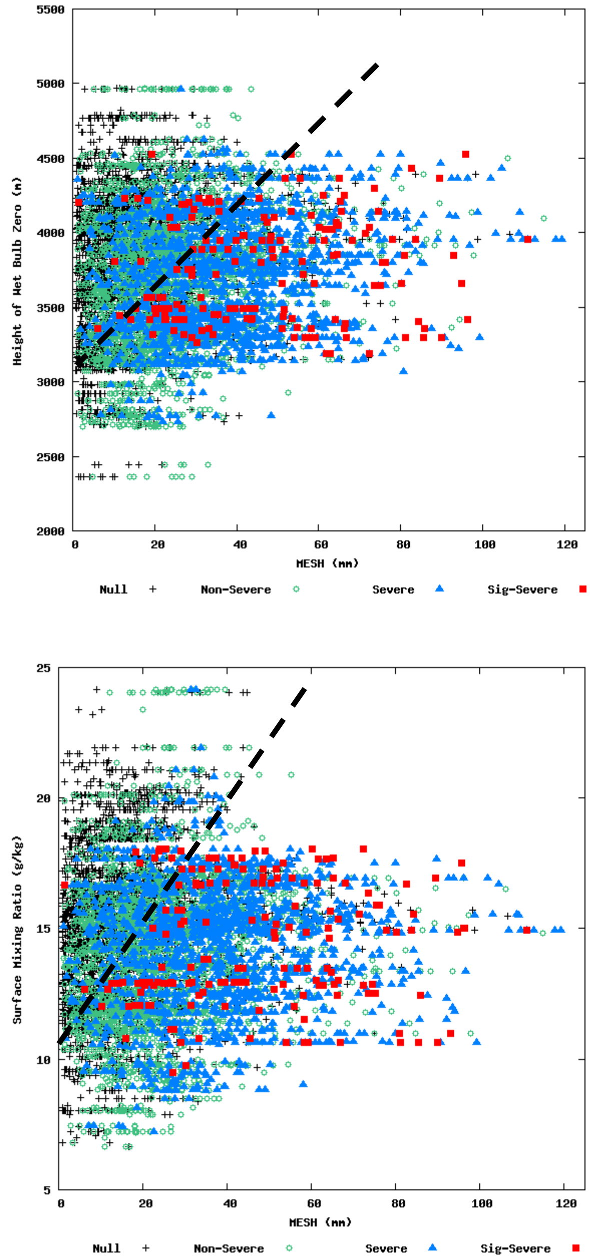

***Figure 4**: Same as Figure 3, except with MESH plotted on the x-axis against the H_{WBZ} in the upper plot and surface mixing ratio in the lower plot.*

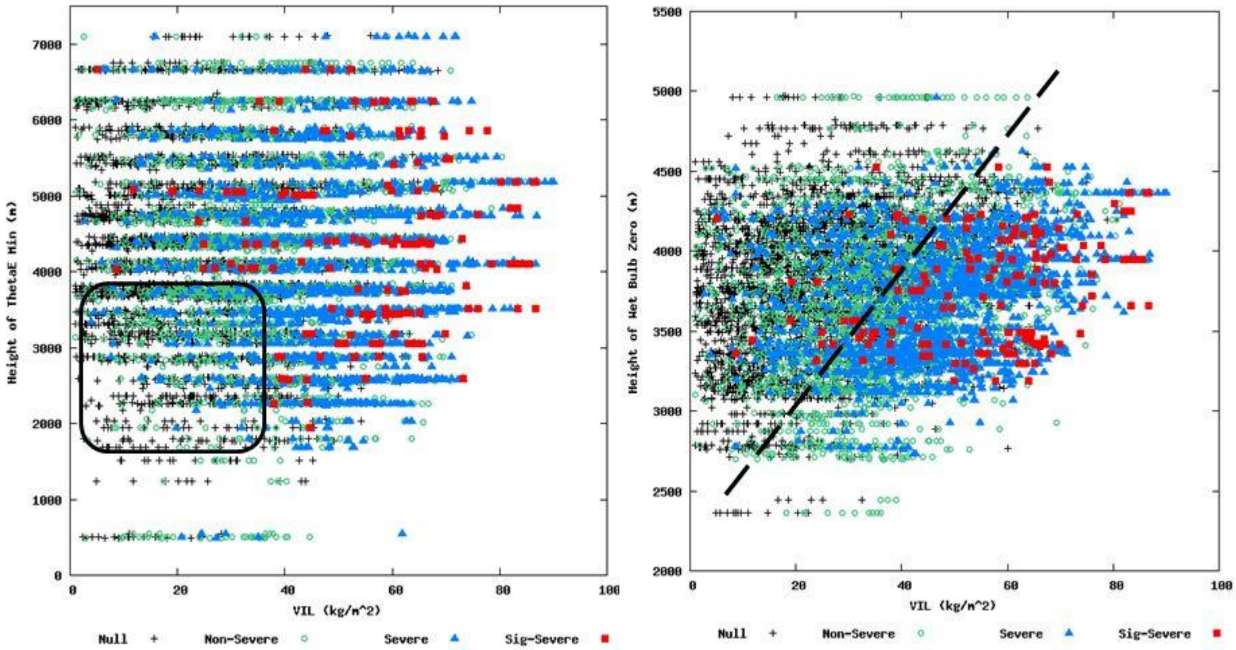

Figure 5*: Same as Figure 3 with VIL being compared to H_{TEM} on the left and H_{WBZ} on the right. The box on the left plot highlights a region with markedly reduced severe and significant severe reports. The dashed line on the right denotes the approximate division between severe/significant severe hail and non-severe reports.*

The last set of plots is shown in Figure 5. Out of the three radar products analyzed, the NSE variables improved the performance of VIL the most. This is somewhat surprising, considering the prior research on the problems with using VIL as a proxy for hail size (Edwards and Thompson, 1998). This lack of skill is evident in the plot of H_{TEM}, with no clear horizontal separation of hail size categories. While the height of theta-e min plot does not provide a clear overall pattern, it contains one particularly interesting feature. For the VIL values of less than 40 kg/m^2 and heights of the theta-e minimum below 4,000 m, there is a noticeable drop in the concentration of severe hail reports. For low VIL storms with H_{TEM} greater than 4,000 m, there are more severe cases. Investigation of the data suggests that MESH had a particularly difficult time diagnosing low VIL storms that produced significant hail. This plot provides a potentially useful parameter to partially explain the presence of large hail within low VIL storms. The H_{WBZ} plot displays the best separation out of the 15 combinations of radar and NSE data.

The ineffectiveness of NSE variables as hail size estimators may be due to the coarse temporal and spatial resolution of the RUC analysis. Since the analysis takes place on a grid size of 20 km, only a few pixels comprised the entire storm track.

4. CONCLUSIONS

This study used high resolution hail data to analyze existing radar products and the impact of combining NSE data with these products. The purpose of this fine scale analysis was to investigate possible components for the development of a future gridded surface hail product. Our initial conclusions are:

1. The addition of NSE variables adds very little to radar products when determining expected hail size.

2. MESH shows some separation between the different hail size categories, but needs further tuning, especially when considering significant severe hail (> 50.8 mm).

3. Significant severe hail events occur in both environments of low and high values of mean RH below 0°C.

4. Significant hail events in low VIL scenarios (< 40 kg/m^2) could be related to the height of the theta-e minimum.

Possibilities for future work include applying different techniques for matching the SHAVE reports to the surrounding environments, such as averaging an area around the report instead of using a strict point match technique. However, such an analysis would still be restricted by the resolution and accuracy of the source of the NSE data. Until high resolution and accurate NSE data is available, it will be difficult to apply the real-time thermodynamic environment in a surface hail fall product.

5. ACKNOWLEDGMENTS

This work was prepared by the authors with funding provided by National Science Foundation Grant No. AGS-1062932, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA08OAR4320904, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

6. REFERENCES

Brimelow, Julian C., Gerhard W. Reuter, Eugene R. Poolman, 2002: Modeling Maximum Hail Size in Alberta Thunderstorms. *Wea. Forecasting*, **17**, 1048–1062.

Edwards, Roger, Richard L. Thompson, 1998: Nationwide Comparisons of Hail Size with WSR-88D Vertically Integrated Liquid Water and Derived Thermodynamic Sounding Data. *Wea. Forecasting*, **13**, 277–285.

Jewell, Ryan, Julian Brimelow, 2009: Evaluation of Alberta Hail Growth Model Using Severe Hail Proximity Soundings from the United States. *Wea. Forecasting*, **24**, 1592–1609.

Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The Warning Decision Support System - Integrated Information. *Wea. and Forecasting,* **22,** 596-612.

Marzban, Caren, Arthur Witt, 2001: A Bayesian Neural Network for Severe-Hail Size Prediction. *Wea. Forecasting*, **16**, 600–610.

Ortega, K. L., T.M Smith, and G. J. Stumpf, 2006: Verification of Multi-Sensor, Multi-Radar Hail Diagnosis Techniques. *Preprints, 22nd Conf. on IIPS,* Atlanta, GA.

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.

Stumpf, G.J., T.M. Smith, and A.E. Gerard, 2002: The Multiple-Radar Severe Storm Analysis Program (MR-SSAP) For WDSS-II. *Preprints, 21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., 138-141.

Stumpf, G.J., T.M. Smith, and J. Hocker, 2004: New Hail Diagnostic Parameters Derived By Integrating Multiple Radars and Multiple Sensors. *Preprints, 22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., CD-ROM, P7.8.

Witt, Arthur, Michael D. Eilts, Gregory J. Stumpf, J. T. Johnson, E. De Wayne Mitchell, Kevin W. Thomas, 1998: An Enhanced Hail Detection Algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286–303.