The Severe Hazards Analysis and Verification Experiment

Kiel L. Ortega
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and National Severe Storms Laboratory, Norman, Oklahoma

Travis M. Smith
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and National Severe Storms Laboratory, Norman, Oklahoma

Kevin L. Manross
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and National Severe Storms Laboratory, Norman, Oklahoma

Kevin A. Scharfenberg
National Weather Service Office of Climate, Water, and Weather Services, Norman, OK

Arthur Witt
National Severe Storms Laboratory, Norman, Oklahoma

Angelyn G. Kolodziej
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and National Severe Storms Laboratory, Norman, Oklahoma

Jonathan J. Gourley
National Severe Storms Laboratory, Norman, Oklahoma

Corresponding author address: Kiel L. Ortega, Cooperative Institute for Mesoscale Meteorological Studies, 120 David L. Boren Blvd., Norman, Oklahoma, 73072
Email: Kiel.Ortega@noaa.gov

ABSTRACT

During the springs and summers of 2006 through 2008, scientists from the National Severe Storms Laboratory and students from The University of Oklahoma have conducted an enhanced severe storm verification effort. The primary goal for the Severe Hazards Analysis and Verification Experiment (SHAVE) was the remote collection of high spatial and temporal resolution hail, wind (or wind damage) and flash flooding reports from severe thunderstorms. This dataset has a much higher temporal and spatial resolution than the traditional storm reports collected by the National Weather Service and published in *Storm Data* (10s of km^2 and 1 to 5 minutes versus 1000s of km^2 and 30 to 60 minutes), and also include non-severe reports that are not included in *Storm Data*. The high resolution of the dataset makes it useful for validating high resolution, gridded warning guidance applications.

SHAVE is unique not only for the type of data collected and the resolution of that data; but also for how the data is collected. The daily operations of the project are largely student-led and student-run. To complete the remote, high resolution verification the students use Google Earth™ to display experimental weather data and geographic information databases, such as digital phonebooks. Using these data, the students then make verification phone calls to residences and businesses, throughout the United States, thought to have been affected by a severe thunderstorm. The present article summarizes the data collection facilities and techniques, discusses applications of these data, and shows comparisons of SHAVE reports to reports currently available from *Storm Data*.

1. Introduction

In general, severe storm verification from the National Weather Service (NWS) have spatial and temporal scales similar to associated severe weather warnings (Hales and Kelly 1985), on orders of 1000 km^2 and tens of minutes. While this spacing may be useful in verifying warnings, the spacing is too coarse to verify new severe weather applications which have high temporal (order of 1-10 minutes) and spatial (order of 1 km) resolution. Reports considered non-severe by the NWS (e.g., hail less than 1.9 cm in diameter; NWS 2005) and confirmed reports of no severe weather near a severe thunderstorm are not widely available, yet are critical with respect to evaluating performance of high resolution applications and models. Furthermore, Witt et al. (1998b) and Trapp et al. (2006) have cited inconsistencies in the reported times, locations, and severe weather magnitudes (e.g., hail size) in the U.S. official climate record of severe weather events, the National Climate Data Center's *Storm Data* publication (NWS 2007).

The Warnings Decision Support System - Integrated Information (WDSSII; Lakshmanan et al. 2007; Hondl 2002) generates 3D grids of specialized experimental severe weather products by integrating data from multiple-radars and other sensors. With the onset of WSR-88D base-data distribution over the Internet (Droegemeier 2002) and inexpensive fast computing, it is possible to run WDSSII for the entire Continental United States (CONUS) at a resolution of approximately 1 km or better both horizontally and vertically, and 1 to 5 mins in time (Lakshmanan et al. 2006). Such a system is co-located at the National Severe Storms Laboratory (NSSL) and Storm Prediction Center (SPC), with experimental products shared on the web site http://wdssii.nssl.noaa.gov (Smith and Lakshmanan 2006). When these high-resolution data are coupled with

geographic information, it becomes possible to make a detailed assessment of when and where hazardous, severe weather events may have occurred. Evaluation of these research products requires an observational dataset more detailed than what is currently available in *Storm Data*.

The Severe Hazards Analysis and Verification Experiment (SHAVE), originally called the Severe Hail Verification Experiment (Smith et al. 2006), was conducted by NSSL each May-August during 2006-2008. SHAVE was designed to take advantage of the ability to blend high-resolution radar data with geographic information to accomplish the primary goal of collecting high temporal and spatial resolution hail reports in hail swaths from thunderstorms (Figures 1 and 2). These reports were collected via verification phone calls to locations along and near the storm's path immediately following the storm's passage. Meteorology undergraduate students enrolled at the University of Oklahoma were hired to make the phone calls. Scientists from NSSL and the University of Oklahoma's Cooperative Institute for Mesoscale Meteorological Studies provided guidance on regions to focus phone calls.

The initial scientific objectives of SHAVE were to:

- use the high-resolution verification data in the development of techniques for probabilistic hazard grids for severe thunderstorms (Ortega 2008),
- evaluate the performance of multi-sensor, multi-radar hail algorithms (Ortega et al. 2006),
- associate changes in hail size distribution with storm evolution, and
- enhance climatological information about hail in the United States.

Although the primary goal of SHAVE was to collect high resolution hail reports, each year the goals of the project were expanded. In 2007, wind and tornado damage were added to the verification efforts. In 2008, flash flooding was added to the verification efforts. These additional objectives often required different data collection strategies. Wind verification, for example, occurred the day following an event. Preliminary reports of wind damage, compiled by the SPC, were overlaid in a geographic information system (GIS) viewer and phone surveys were conducted in the areas surrounding the reports.

The objective for this paper is to explain the data collection facilities, techniques and surveys in detail, summarize the lessons learned from the techniques and surveys employed, discuss applications of these high resolution data and summarize the verification data collected. Section 2 will describe the data collection facilities and other project infrastructure. Section 3 describes the day-to-day operations of the project. Section 4 details the phone surveys employed by SHAVE to collect information on severe weather hazards. Section 5 illustrates uses of the SHAVE dataset and some initial results. Section 6 summarizes the project.

2. Data collection facilities

SHAVE was conducted within the National Oceanic and Atmospheric Administration's Hazardous Weather Testbed (HWT; Norman) as part of the Experimental Warning Program (Stumpf et al. 2008). SHAVE operations take place in NSSL's Development Laboratory located in the National Weather Center in Norman, OK. The major facilities required in SHAVE are listed below:

- *Computing cluster*: Described in detail by Lakshmanan et al. (2006). The computer cluster enabled the processing of radar data and derived products in a format suitable for displaying in a GIS viewer. Examples of this type of data are available at http://wdssii.nssl.noaa.gov.
- *Display machines*: The NSSL Development Laboratory (Figure 3) provides room for up to 15 computers with adequate display, processing and memory capabilities to run Google Earth™ Pro[1] and an Internet browser. The computers were used by the students during operations to connect to the computing cluster, display appropriate data and enter reports into a common database. The operations coordinator's (described in section 3) machine also had the capability to run utilities to help parse and send phone number information to the student callers.
- *Collaboration displays*: Four 1.27 m (50 inch) plasma screen monitors were used to display radar information relative to geographic features of the storm being investigated, a webpage displaying recent phone calls and an overview of the day's operations that could be viewed by visitors.
- *Telephone lines*: A telephone and headset was available for each student caller.
- *Geographic information*: The geographic information used in the project comes from several sources. Google Earth™ (Figure 4) provided street overlays and satellite views, as well as address and phone number information for locations such as businesses. Residential address and phone number data were obtained from two different sources: rural plat maps, which contain property ownership

[1] Google Earth™ Pro was adopted by SHAVE, however, any mapping service is suitable. SHAVE uses Google Earth™ because of its support of Keyhole Markup Language (KML). The use of Google Earth™ by NSSL is not an endorsement.

records and county directory information, and Delorme Street Atlas™ USA[2], which contains an address and phone number database for the CONUS. Plat maps provide non-digitized area street maps, the relative positions of residences on the streets and their phone numbers. The digital Delorme phone number database provided a more convenient method for searching and disseminating phone numbers to the student callers when compared to using the plat maps (Figure 4). Two major problems were found to occur with phone numbers: disconnected phone lines and inaccurate geolocations (latitude/longitude associated with an address).

3. SHAVE operations

There were three primary roles for SHAVE participants: project scientist, operations coordinator and student caller. The project scientists developed the survey scripts, operations plan, data collection infrastructure, and scientific objectives for the project. The operations coordinator was responsible for determining which storms to interrogate on a daily basis, collecting the phone numbers from plat maps and Delorme's database, and directing the student callers. The student callers were responsible for making the survey phone calls.

A typical day during the project began with the decision by project scientists whether or not to commence operations. On most days this coincided with the SPC's issuance of the 1630 UTC Day 1 Convective Outlook. Several other factors were also considered when deciding to commence operations, including:

[2] Any digital address and phone number database is suitable. The use of Delorme's software by NSSL is not an endorsement.

- whether hail producing storms will develop in areas that can be easily surveyed,
- whether widespread severe winds or wind damage had been reported the day before,
- whether flash flooding had been reported or if rainfall accumulations had exceeded flash flood guidance values,
- whether storms would form before 2130 LT (end of the experiment day),
- anticipated storm modes, and
- staff availability and fatigue.

Notification of operational status of the project was posted on the project website and distributed to participants via e-mail each morning. Participants reported to the operations area at the time specified in the notification message and operations began shortly thereafter.

The data collection effort proceeded with at least three participants, one serving as the operations coordinator, a project scientist, and student callers (ranging in number from one to five, depending on weather and availability). The project scientist was either on duty or on call each day to determine the status of operations and to help solve any technical problems. The operations coordinator made decisions on the type of storms to investigate, the exact storms which would be examined, and finding and disseminating geographically-referenced phone numbers to the student callers. The operations coordinator occasionally took part in making phone calls, if the workload allowed for it. The rest of the team made phone calls and collected data from the public about the particular threat under investigation.

The operations coordinator called an end to operations based on several conditions:

- storm activity had declined and was no longer anticipated to increase,
- storms were still active but it was past 2130 LT in the area of concern,
- storms were located in unpopulated areas or areas with inaccurate phone number geolocations,
- all possible locations for wind and flash flood impacts had been surveyed, or
- staff fatigue levels and anticipated level of operations for the following day(s).

At the close of operations, the operations coordinator ran a program that archived the data, and created and sent summary reports to the project participants and affected NWS offices.

4. Data collection

Phone surveys were developed for hail, wind, and flash flood events. Phone number collection was done under the philosophy of "more is better; all is best" (Figure 2). The original intent for hail surveys was to collect phone numbers in such a way such that cross sections could be completed at evenly spaced intervals along the storm's path (more specifically, the storm's MESH swath). However, it was quickly learned that this would not lead to evenly-spaced, high spatial resolution data for two primary reasons: 1) a low successful call rate[3] was achieved; 33.3% of the phone calls resulted in a successful

[3] A successful call was one in which the survey was marked as 'good' (meaning all information was acceptable) or 'questionable time' (meaning all information, except the reported times, was acceptable). Student callers subjectively determined whether information was acceptable or not.

data point (Table 1), and 2) calling all numbers along a cross-section was time consuming, leading to poor spatial resolution down the path of the storm.

Phone numbers of businesses and residences were collected using several sources. Rural county plat maps proved to be invaluable sources during SHAVE 2006. These maps essentially provided a county phone book with maps of area roads and residence's locations relative to the roads. These plat maps, combined with high resolution satellite photographs viewable in Google Earth, led to high confidence in the exact location of many SHAVE reports. During SHAVE 2006, phone numbers were also found using features within Google Earth (Figure 4). One feature was the "Layers" menu (bottom left in Figure 4). This menu allowed for the student callers to overlay many different features, including businesses. If a student caller clicked on these features, contact information appeared in a pop up window. Another feature was the "Search" menu (top left in Figure 4), which allowed for the student callers to find additional business listings not available through the "Layers" menu. These two features allowed for rapid, automatic placement of many phone numbers using very little time by the student callers, however, these numbers had a few concerns for their use. One was inaccurate geolocations associated with these points. The other is that nearly all of the search results and all of the features available are businesses. On many occasions, phone surveys with businesses required too much time because SHAVE student callers would need to either talk to the business' management (many of whom were usually not readily available) or the business would be reluctant to reveal the severity of the weather they received or any damage that may have occurred.

During SHAVE 2007 and 2008, most phone numbers were residential phone numbers found using Delorme Street Atlas™[4]. Delorme Street Atlas™[4] contains a phone number database, giving SHAVE participants access to addresses and phone numbers in the United States and Canada. The operations coordinator could search for phone numbers in Delorme, by either clicking on individual streets or using a search (e.g., a zip code). The operations coordinator could then export that data to Google Earth for georeferencing (Figure 4). The operations coordinator would save the data into a KML file and send it to the student callers. The digital phone number database and automatic georeferencing did not completely replace plat maps, as some rural locations still have poor addressing grids resulting in inaccurate georeferencing. However, the database did allow for many more numbers to become available to the student callers and provided a much easier method to procure numbers.

The Delorme database had some affects on the character of the database due to the large number of phone numbers made available. The first was a reduction in the number of 'questionable hail' reports (Table 1). The reason for this is that if a survey took place and the reported hail size was put in question by the student caller, then a "next-door" neighbor survey would be completed. If the questionable report was indeed questionable/inaccurate, the survey was eventually marked as not completed instead of 'questionable hail.' This was done to ensure that inaccurate data was not accidentally used in any post-processing of the reports. The second affect was an increase in 'questionable time' reports (Table 1). The increase was also due to the adoption of a "30-minute plus" rule during the 2007 experiment; if a survey occurred more than 30 minutes after hail fall, the survey was marked as 'questionable time.' The "30-minute plus" rule

[4] The use of Delorme's software by NSSL is not an endorsement.

was adopted after comments from student callers described a decreased quality of how people described the time period during which the hail was falling. For example, instead of getting time period descriptions like this: "the hail fell from 7:10 to 7:17 PM", student callers were more likely to get descriptions like this: "the hail started about 45 minutes ago and lasted about 5 minutes." Also, since many wind and flood surveys occurred the day following the event, most of those surveys were marked as 'questionable time.'

Shortly after arriving for daily operations, student callers would first survey regions that were potentially impacted by severe winds and flash flooding on the preceding day. We found the highest success rate collecting information about wind and flash floods the day following an event. It is surmised that this additional time enabled the public to adequately assess the damage that had been seen by the following day. Wind and flash flood calls would continue through the early afternoon until storms developed and posed a hail threat. Depending on the number of student callers available and the anticipated threats of the day, the calling team would split up and work multiple storms or different hazards. The calling areas were determined by the operations coordinator. The data collection strategies and survey questions for hail, wind and flash flooding are discussed below.

a. Hail

For hail storms, the desire was to sample a variety of storm environments (such as high instability and low wind shear or cold-core upper air systems) and storm types (such as supercells and "pulse" thunderstorms). For a typical day's operations, storms were selected by considering the following (in order of preference):

- the availability of accurate georeferenced phone numbers,
- temporal swath of maximum expected size of hail (MESH; Witt et al. 1998a) from the NSSL multi-radar, multi-sensor algorithm (Stumpf et al. 2004), and
- storm type and environment.

The phone surveys took place shortly (ideally within 60 minutes) following the storm's passage. Occasionally, SHAVE participants would be forced to delay surveys due to lightning in the area resulting in the public not wanting to be on the phone. NWS severe weather warnings also could cause delay in surveys as the public might also not want to be on the phone while a warning was still valid for their area.

The hail survey was developed in collaboration with NWS forecasters who regularly made verification phone calls. The survey was then reviewed to ensure that the questions would not lead those taking the survey to a specific answer. Student callers were trained not to ask leading questions in regard to hail size. For example, if a person said the hail was "between pea and softball", the student caller's follow up question would be "is it more like a pea or a softball?". The survey would then continue from there (in similar fashion), until a hail size was determined. The student callers were also trained to collaborate with each other in situations where the hail size might have been questionable. For instance, if a person surveyed seemed unsure about hail size, a "next-door" location would be surveyed to confirm or refute the reported hail size.

The survey began by asking whether or not hail fell at the location. If hail did fall, the person was asked to estimate two hail diameters: the maximum and the average hail diameter. The average hail diameter was the person's estimate of the size of the

majority of the hail stones. If the hail size was large, usually greater than 44 mm, the survey would ask if the person had measured any hail stones or if they could measure one of the larger hail stones. Measurement of an average-diameter hail stone size was also requested, but generally unavailable[5]. The student callers then asked for the times the hail fall began and ended. While the student callers were trained to be specific and ask for the start and end of the *hail fall*, sometimes the start and end of the *storm* was reported by the person surveyed. The final question of the hail survey was how much of the ground was covered by the hail. Other information, such as damage from the hail or sightings of funnel clouds or tornadoes, was also recorded. This information was mostly for the benefit of the NWS offices which received SHAVE summary reports, but the damage information helped to confirm larger hail size reports whose size was not measured. Figures 1 and 2 show SHAVE and *Storm Data* verification from two storms, one in Lac qui Parle County, Minnesota on 27 July 2006 and the other from Yates County, New York, on 16 July 2008.

b. Wind

For wind surveys, there were two different modes of operation. In one mode, wind survey questions were asked following the hail survey. In this mode of operation, reports of no wind damage were not recorded as wind reports. This was due to the fact that the people being surveyed may not have been able to see any possible damage from wind yet. In the second mode, wind surveys were conducted a few hours after or the day following a wide spread wind event. In this mode of operation, local storm reports issued

[5] This was for several reasons, including the survey calling too long after a storm had passed and the hail had melted, but the person saved only the larger hail stones, or the maximum size was the average size.

by the NWS were used to find areas to survey and reports of no wind damage were recorded. During wind surveys, the degree of damage to structures or trees and the time the wind damage occurred was the primary information that the survey was intending to collect; if wind speed estimates or measurements were volunteered, the information was recorded. However, the survey only specifically asked for wind damage and the time the wind damage occurred[6]. Figure 5 shows SHAVE and *Storm Data* verification from a bow echo which produced widespread damaging winds over the Omaha, Nebraska, area on 27 June 2008.

c. Flash flooding

Flash flood surveys were conducted if any of the following three criteria were met: 1) 6-hour precipitation accumulation exceeded flash flood guidance (FFG) values less than 24 hours ago, 2) a flash flood warning or urban/small stream advisory was issued by an NWS forecast office during the past 24 hours, or 3) a survey for a different severe weather threat indicated flash flooding was a problem. The ratio product that compared precipitation to FFG values was generated locally at NSSL by acquiring the latest CONUS mosaic of FFG values from the NWS's National Precipitation Verification Unit and comparing it to rainfall estimates from NSSL's National Mosaic and Quantitative precipitation estimation system (NMQ; http://www.nmq.noaa.gov/). It was noted that the stitched CONUS grid of FFG values exhibited discontinuities at River Forecast Center (RFC) boundaries. These boundaries were discovered to be a result of the different methods used in the NWS to derive FFG and now gridded FFG (GFFG)

[6] The one exception was if the survey called an airport. The survey would then ask for anemometer measurements.

values. The 6-hr ratio product was the primary tool used to guide the location of the phone calls; however, there were occasions when NWS-issued warnings, advisories, and ratio products at 1- and 3-hr duration were relied upon. The objective of the survey questions was to find information about flooding that resulted from rain that fell less than six hours prior to the onset of flooding. In order to avoid flooding reports resulting from levee or dam breaches, or from flooding of main river stems, student callers cross-referenced a static flow accumulation map derived from a digital elevation model to check the size of the contributing basin areas. The flow accumulation map was color-coded to highlight regions with flow accumulations less than 260 km^2, which is generally accepted as the threshold basin area for rainfall-induced flash floods (see e.g., Davis 1998; Reed et al. 2007). Satellite views of the river channels available in Google Earth™ were also used to avoid large-scale fluvial flooding events.

The survey questions about flash flooding initially came from the *Storm Data* Directive, but were modified and supplemented based on conversations with forecasters at the NWS's Arkansas-Red Basin RFC and researchers at the Office of Hydrology. The questions were focused on the location, nature, magnitude, frequency, and times flooding began and ended. In addition, any evacuations or rescues were included in the reports. The caller first determined if the resident had experienced any type of flooding problems. If the answer was no, then the report was recorded as a null event and the call was terminated. If the response was positive, then the caller determined if the impacts of flooding were inundation or damage to property including croplands, road or bridge closures, rivers or creeks out of banks, or backups due to insufficient drainage. Any or all of the previous impacts could have been recorded in the report by the caller. The

webpage also enabled the callers to type in comments as needed. The next question attempted to determine the magnitude of flash flooding through the depth and/or horizontal extent of the floodwaters. Caution was exercised in the survey to insure residents didn't approach moving or standing waters to estimate their depth. Several reference values of reported water depth were listed relative to car tires, mailboxes, or windows and then later converted to heights in metric units. In the case of flooded creeks or streams, questions focused on the lateral extent water was out of its banks. Questions about the frequency of flooding were a supplement to the *Storm Data* Directive and were found to be particularly informative. The caller would ask if the type of reported flooding occurs every time it rains, only when it rains hard, about once a year, about once every 5, 10, 50, or 100 years, or "never seen it like this before". Finally, the caller attempted to record the beginning and end times when flash flooding occurred or was still ongoing. Figure 6 shows an example of flash flood verification done by SHAVE and compared to reports available in *Storm Data* for an event in Central Indiana on 7 June 2008.

5. Dataset Uses and Results

As shown in Table 2, SHAVE provides a unique database mostly populated with reports (non-severe and confirmations of 'null' events) not readily available from *Storm Data*. These data are important for algorithm evaluation work and provide a fuller account of an event than just using 'severe' reports. The data collected during SHAVE have several possible uses due to the variety of questions asked during the surveys. While maximum hail size and flooding depth have obvious applications in algorithm

evaluation, questions such as how much of the ground was covered by hail and how often the observed flooding occurs could have very unique applications. Work continues into applications of the wind damage verification data. Two difficulties arise in the current method SHAVE student callers use to collect wind damage verification, one being the difficulty in pinpointing areas in which to make verification phone calls for wind damage; further, a member of the public being surveyed may not know the full extent of wind damage in their area. Another is the wind damage data itself. Currently, it is more qualitative (description of damage) than quantitative (wind speed). One possible solution to this problem is using the Enhanced Fujita scale (LaDue and Mahoney 2006) and is currently being evaluated by SHAVE project scientists.

Hail data collected by SHAVE can be applied to a number of different research goals. One is evaluating gridded, multi-sensor hail diagnosis techniques (Stumpf et al. 2004). Wilson et al. (2009) used SHAVE hail data to investigate the skill in using gridded MESH as a synthetic verification tool and evaluated a number of different techniques for their skill in severe hail prediction. Preliminary results found limited skill in both evaluations, which may be a result of the wide range of hail sizes for a given MESH value as revealed by SHAVE hail data (Figure 7). Another, more unique, use of SHAVE hail data is to evaluate how well the public reports on hail sizes. Figure 8 shows comparison between the estimated hail size and the measured hail size. The overall bias was -2 mm with a root mean square error of 12 mm, suggesting those who provided both estimates and measurements of maximum hail size slightly underestimated the measured hail size, but their estimates were within ±12 mm of the measured hail size. Further, as Figure 9 shows, an evaluation of how frequently the public estimates hail sizes to well

known objects (such as coins or golf balls) instead of the true hail size (as revealed through measurements) could also be performed using SHAVE data.

Flash flood data is currently being applied to evaluate differences between legacy, county-based FFG and new gridded FFG (GFFG; Erlingis et al. 2009). An envisioned outcome from this study will be a simple table reporting the ratios of quantitative precipitation estimate (QPE)/FFG and QPE/GFFG for 1-, 3-, and 6-hour duration that result in the greatest skill in predicting flash floods. The flash flood data also has applications beyond algorithm evaluation. The flash flood survey asks for both the impact of the flooding (Figure 10) and the frequency of the observed flooding (Figure 11)[7]; these data could be combined to identify locations that are prone to flooding. For example, if a local SHAVE-type experiment occurred and found that a pond or stream was prone to severe flooding every time it rained, permanent warning signs could be placed near the flood prone area.

6. Summary

SHAVE presents a framework for conducting high resolution, remote verification of severe storms. The data collected during SHAVE provide a unique dataset that contribute to current research to more accurately and confidently (due to the high resolution) relate radar data to ground truth. The dataset is unique not only for its spatial accuracy and resolution, but also the types of reports contained within the dataset as a majority of the reports are 'no hail/wind/flood' (Table 2). Also some of the data obtained

[7] Severe flooding was considered standing water greater than 0.91 m in depth or moving water greater than 0.15 m in depth.

through SHAVE surveys has uses beyond meteorological uses, such as the economical impacts of flash flooding due to the frequency and impacts of the flooding.

The project is unique in the fact that, except for the building of the project's infrastructure and initial training, the project is largely student-led and student-run. For a sponsoring organization, this provides for an inexpensive project that yields a large amount of data. Opportunities to learn about severe storm forecasting, radar interrogation of storms and severe storm hazards provide the students with an extraordinary learning experience and potential research topics for course work. Further, the students must apply critical thinking and teamwork during the verification process; combined with invaluable face time with project scientists, the project also provides participating students a great professional development experience.

The project website, which includes the project plan, facilities summary, example cases and a data archive, is available at http://ewp.nssl.noaa.gov/projects/shave/. SHAVE is expected to continue in 2009 and beyond, with focus on dual-polarization radar data and possible expansion into winter weather verification.

Acknowledgements

The authors would like to acknowledge the student callers: Jenifer Bowen, Chad Echols, Jessica Erlingis, Margaret Frey, Steve Irwin, Carey Legett, Tiffany Meyer, James Miller, Kelsey Mulder, Christa Riley, Brett Roberts and Rachael Sigler, who made possible the collection of this unique dataset. Thanks to Jennifer Palucki, Rick Smith, Alan Stewart and many others for their comments on the hail portion of the survey script. John Schmidt, Ernie Wells, and Seann Reed supplied useful comments on the flash flooding script. Robert Coggins, Karen Cooper, Vicki Farmer, Zac Flamig, Doug Kennedy and Joe Young provided essential computer support for the project. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA17RJ1227, U.S. Department of Commerce and the National Severe Storms Laboratory's Director's Discretionary Research Funds. Mentions of specific software in this manuscript are not endorsements of said software by the funding sources.

References

Davis, R. S., 1998: Detecting time duration of rainfall: A controlling factor of flash flood intensity. Preprints, *Special Symposium on Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., 258-263.

Droegemeier, K.K., and Coauthors, 2002: Project CRAFT: A test bed for demonstrating the real time acquisition and archival of WSR-88D Level II data. Preprints, *18th Int. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, Amer. Meteor. Soc., Orlando, FL, 5.20.

Erlingis, J. M., J. J. Gourley, T. M. Smith, and K. L. Ortega, 2009: Development of detailed database of flash flood observations. Preprints, *23rd Conf. on Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., JP2.8.

Hales, J. E., Jr., and D. L. Kelly, 1985: The relationship between the collection of severe thunderstorm reports and warning verification. Preprints, *14th Conf. on Severe Local Storms*, Indianapolis, IN, Amer. Meteor. Soc., 13–16.

Hondl, K., 2002: Current and planned activities for the warning decision support system-integrated information (WDSS-II). Preprints, *21st Conf. on Severe Local Storms*, Amer. Meteor. Soc., San Antonio, TX, 4.8.

LaDue, J. and E. A. Mahoney, 2006: Implementing the new Enhanced Fujita scale within

the NWS. Preprints, *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 5.5.

Lakshmanan, V., T. Smith, K. Cooper, J. Levit, K. Hondl, G. Stumpf, and D. Bright, 2006: High-resolution radar data and products over the continental united states. Preprints, *22nd Intl. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 9.7.

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

NWS, cited 2005: WFO severe weather products specification. National Weather Service Instruction 10-511. [Available at http://www.weather.gov/directives/].

NWS, cited 2007: *Storm Data* preparations. National Weather Service Instruction 10-1605. [Available at http://www.weather.gov/directives/].

Ortega, K. L., T. M. Smith, and K. A. Scharfenberg, 2006: An analysis of thunderstorm hail fall patterns in the Severe Hail Verification Experiment. Preprints, *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., P2.4.

----------, 2008: Severe weather warnings and warning verification using threat-

areas. M.S. Thesis. School of Meteorology, The University of Oklahoma, 48 pp.

Reed, S. M., J. Schaake, and Z. Zhang, 2007: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *J. Hydrology,* **337,** 402-420.

Smith, T. M., and V. Lakshmanan, 2006: Utilizing Google Earth as a GIS platform for weather applications. Preprints, *22nd Intl. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 8.2.

----------, K. L. Ortega, K. A. Scharfenberg, K. Manross, and A. Witt, 2006: The Severe Hail Verification Experiment. Preprints, *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 5.3.

Stumpf, G. J., T. M. Smith and J. Hocker, 2004: New hail diagnostic parameters dervied by integrating multiple radars and multiple sensors. Preprints, *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., P7.8.

----------, and Coauthors, 2008: Experimental warning guidance for convective weather

threats. Preprints, *24th Intl. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., P1.3.

Trapp, R.J., D.M. Wheatley, N.T. Atkins, R.W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21**, 408–415.

Witt, A., 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.

----------, ----------, ----------, E.D. Mitchell, J.T. Johnson, and ----------, 1998b: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13**, 513–518.

Wilson, C. J., K. L. Ortega, and V. Lakshmanan, 2009: Evaluating multi-radar, multi-sensor hail diagnosis with high-resolution hail reports. Preprints, *25th Intl. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography and Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., P2.9.

List of tables

Table 1: Summary of phone surveys complete during SHAVE.

Table 2: Summary of data points collected during SHAVE.

List of figures

Figure 1: SHAVE and *Storm Data* report data overlaid in Google Earth™ on a NSSL radar-based hail swath from a storm on 27 July 2006 in Lac qui Parle County, Minnesota. Where NWS reports are attached by a line, it is one report in *Storm Data* with a start and end location.

Figure 2: SHAVE and *Storm Data* overlaid in Google Earth™ on a NSSL radar-based hail swath for a storm on 16 July 2008 in Yates County, New York.

Figure 3: SHAVE student callers working in NSSL's Development Laboratory.

Figure 4: Image showing phone number information overlaid in Google Earth™. At left are the two menus for searching and overlaying different layers (such as businesses). The red telephone icons on the map indicate where SHAVE has found phone numbers either through plat maps or Delorme's phone number database. For privacy reasons, contact information has been changed to generic values, such as phone numbers to 555-555-5555. NSSL's use of Delorme's database is not an endorsement.

Figure 5: SHAVE and *Storm Data* reports overlaid in Google Earth™ on a NSSL radar-based hail swath for a storm on 27 June 2008 near Omaha, Nebraska. Where NWS reports are attached by a line, it is one report in Storm Data with a start and end location.

Figure 6: SHAVE and *Storm Data* verification in Google Earth™ from a flash flood event in central Indiana on 7 June 2008. *Storm Data* flash flood reports are represented by polygons.

Figure 7: Distribution of reported maximum hail sizes for different MESH values. 3921 SHAVE hail reports were used, of which 1539 were severe (dia. ≥ 19 mm).

Figure 8: Estimated hail stone diameter compared to measured hail sizes. The diagonal line is estimated hail size equals measured hail size.

Figure 9: *Storm Data* and SHAVE maximum hail size distributions. Highlighted are common size classifications. The bottom graph is a zoom of the 50.8 to 177.8 mm size range of the top graph. *Storm Data* reports from October 2006 to June 2008 were used.

Figure 10: Impacts from surveyed flash flood events during SHAVE.

Figure 11: Relative frequency histograms of flash flood frequency for non-severe and severe reports collected during SHAVE.

Table 1: Summary of phone surveys complete during SHAVE.

		2006	2007	2008	Total
Survey Completed Information Accepted	Good	4880 (35.2%)	2180 (15.4%)	448 (1.3%)	7508 (12.2%)
	Questionable Time (did not know or survey occurred 30+ min after storm passage)	658 (4.7%)	2705 (19.1%)	9604 (28.8%)	12967 (21.1%)
Survey Completed Information Rejected	Questionable Hail (reported hail size did not correlate to nearby reports)	371 (2.7%)	127 (0.9%)	63 (0.2%)	561 (0.9%)
	Questionable Location (inaccurate geolocation of address)	42 (0.3%)	34 (0.2%)	140 (0.4%)	216 (0.5%)
	Wrong Location (phone number at different location)	47 (0.4%)	30 (0.2%)	33 (0.1%)	110 (0.2%)
Survey Not Completed	Busy/Intercept Operator	777 (5.6%)	638 (4.5%)	1297 (3.9%)	2712 (4.4%)
	No Answer	5485 (39.6%)	6985 (49.4%)	17662 (52.9%)	30132 (49.1%)
	Do Not Call/Disconnected	1286 (9.3%)	1413 (10.0%)	4091 (12.3%)	6790 (11.1%)
	Other	307 (2.2%)	41 (0.3%)	10 (0.1%)	358 (0.5%)

Table 2: Summary of data points collected during SHAVE.

		2006	2007	2008	Total
Hail	'No hail' report	2349	1469	2859	6677
	Non-severe report (dia. < 19 mm)	1629	1292	1847	4768
	Severe report (dia. ≥ 19 mm)	1353	1727	2311	5391
	Significant-severe report (dia. ≥ 51 mm)	47	153	186	386
	Measured hail reports (maximum size only)	42	56	150	248
Wind	'No wind' report	N/A	218	297	515
	Wind damage or speed report	N/A	224	561	785
Flooding	'No flood' report	N/A	N/A	1843	1843
	Flooding report	N/A	N/A	892	892

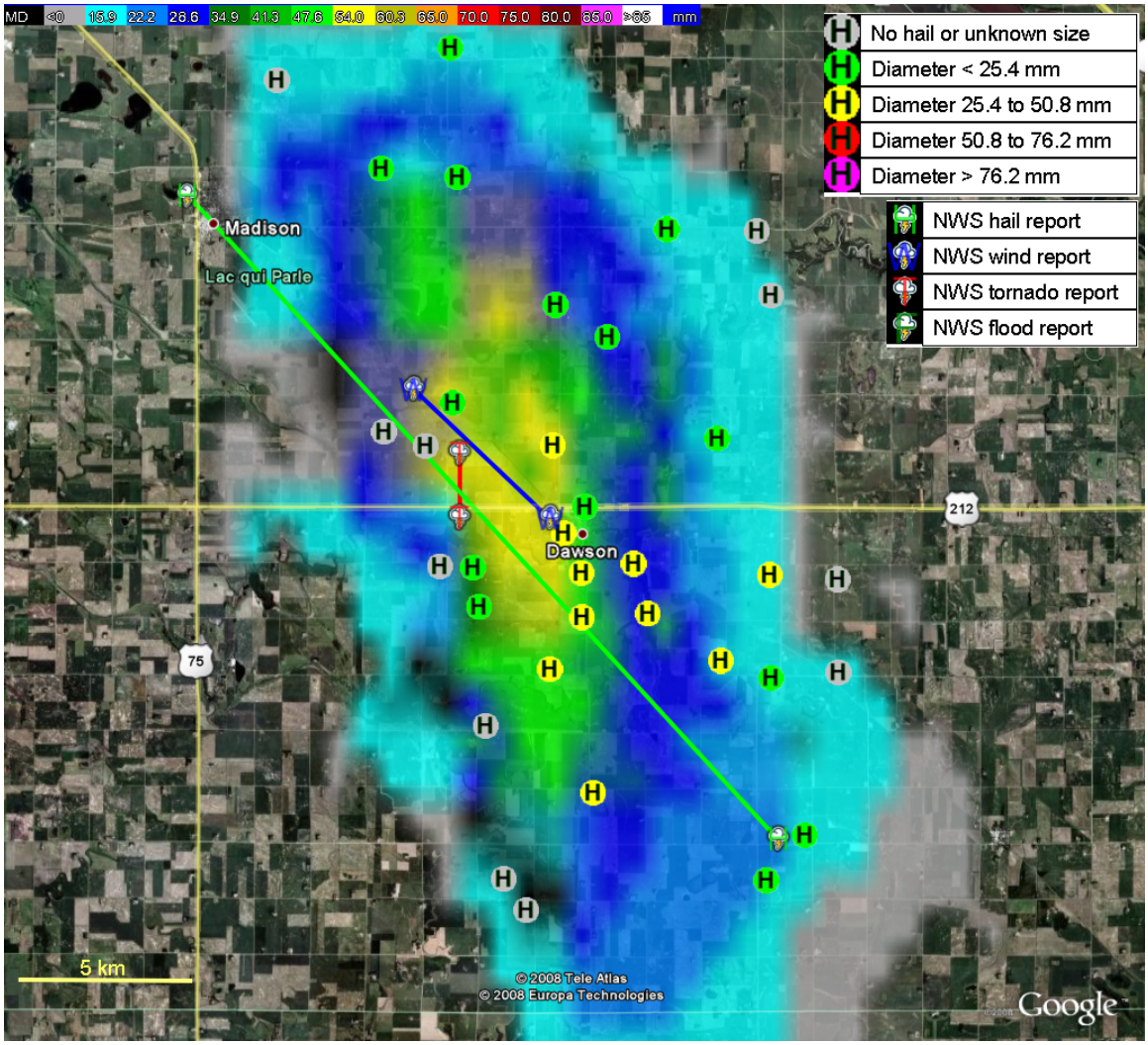

Figure 1: SHAVE and *Storm Data* report data overlaid in Google Earth™ on a NSSL radar-based hail swath from a storm on 27 July 2006 in Lac qui Parle County, Minnesota. Where NWS reports are attached by a line, it is one report in *Storm Data* with a start and end location.

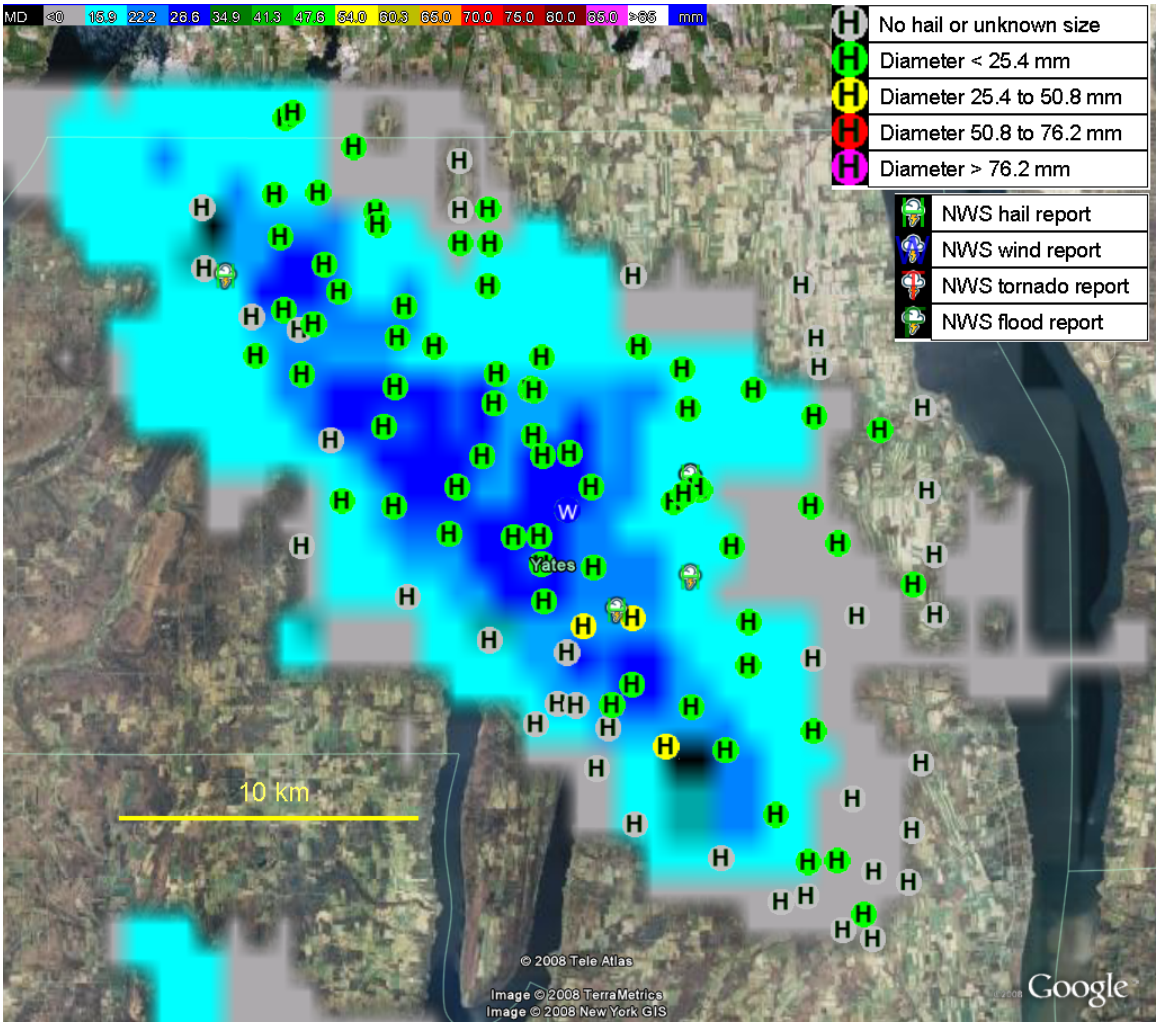

Figure 2: SHAVE and *Storm Data* overlaid in Google Earth™ on a NSSL radar-based hail swath for a storm on 16 July 2008 in Yates County, New York.

Figure 3: SHAVE student callers working in NSSL's Development Laboratory.

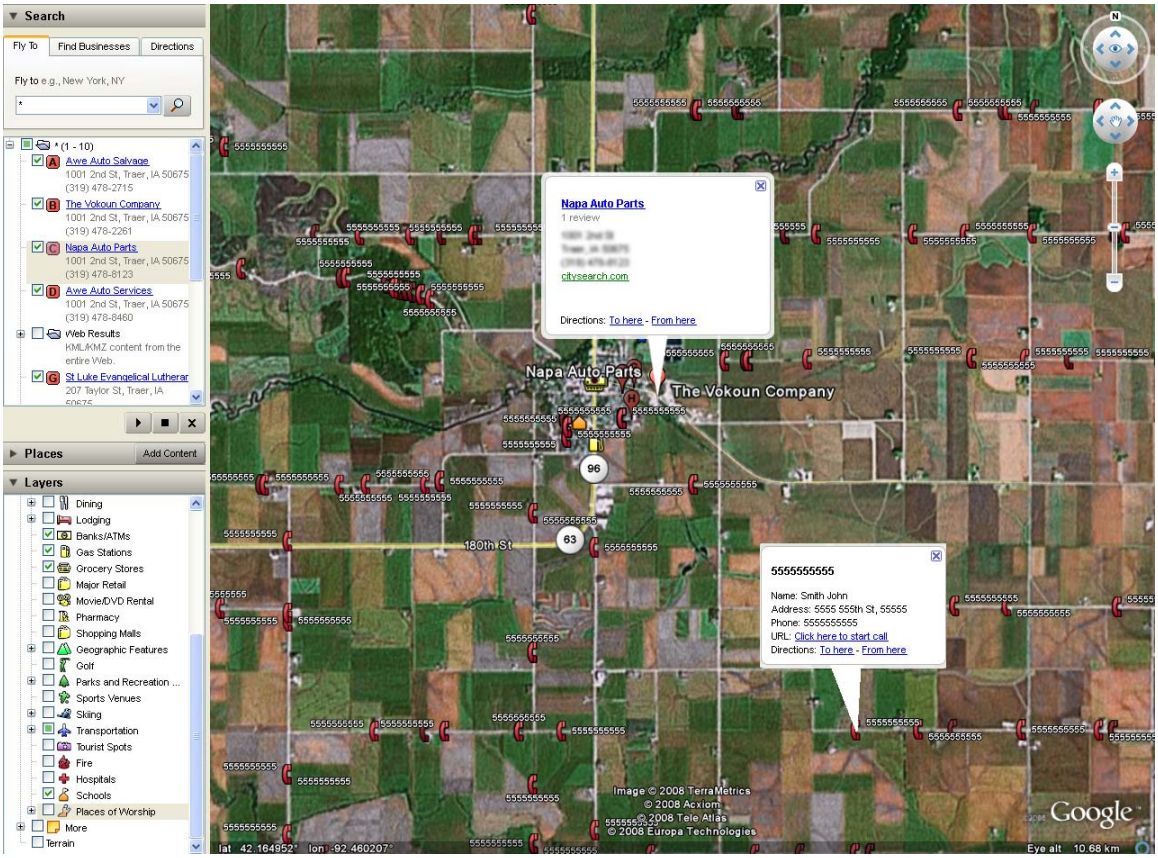

Figure 4: Image showing phone number information overlaid in Google Earth™. At left are the two menus for searching and overlaying different layers (such as businesses). The red telephone icons on the map indicate where SHAVE has found phone numbers either through plat maps or Delorme's phone number database. For privacy reasons, contact information has been changed to generic values, such as phone numbers to 555-555-5555. NSSL's use of Delorme's database is not an endorsement.

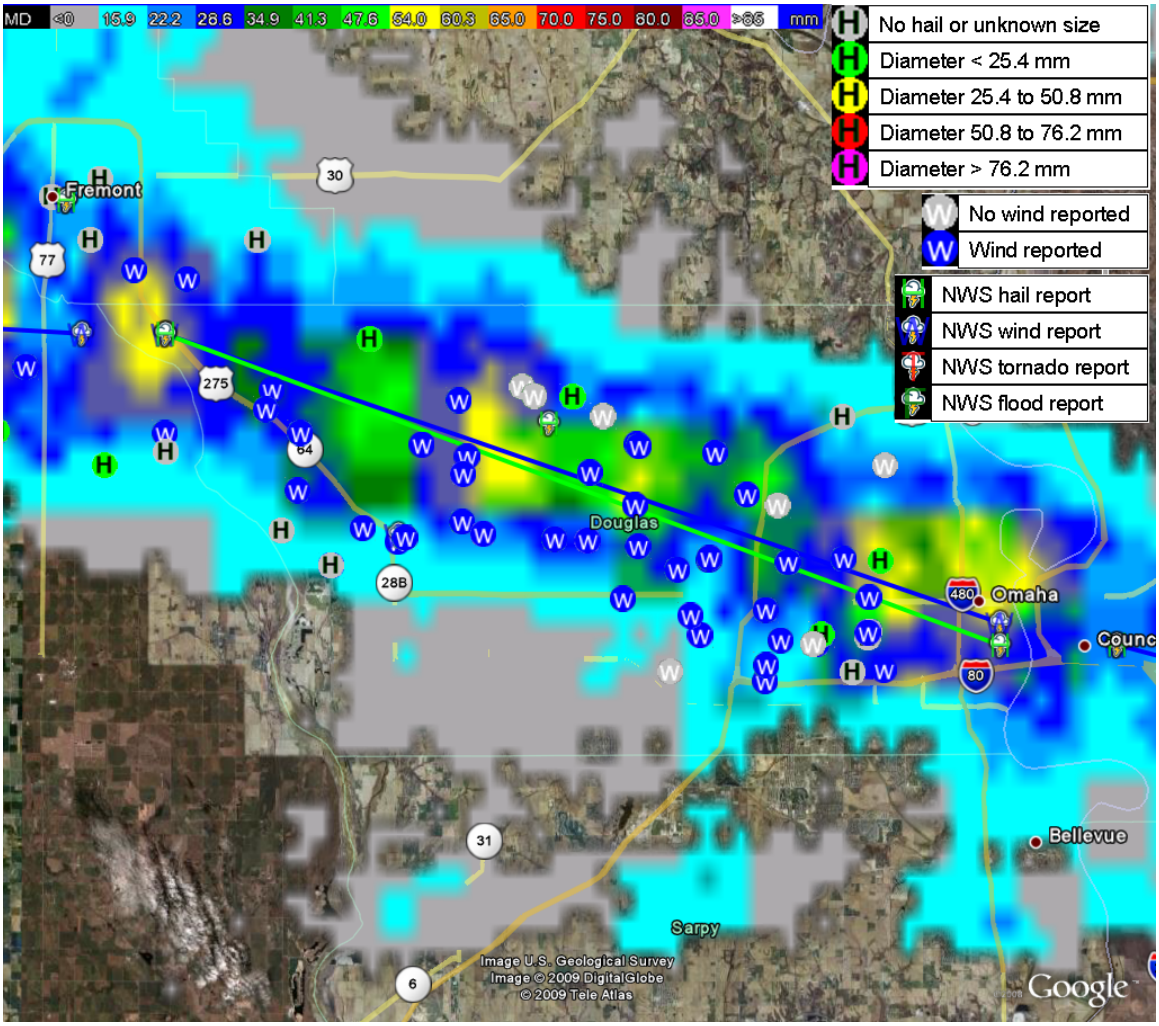

Figure 5: SHAVE and *Storm Data* reports overlaid in Google Earth™ on a NSSL radar-based hail swath for a storm on 27 June 2008 near Omaha, Nebraska. Where NWS reports are attached by a line, it is one report in Storm Data with a start and end location.

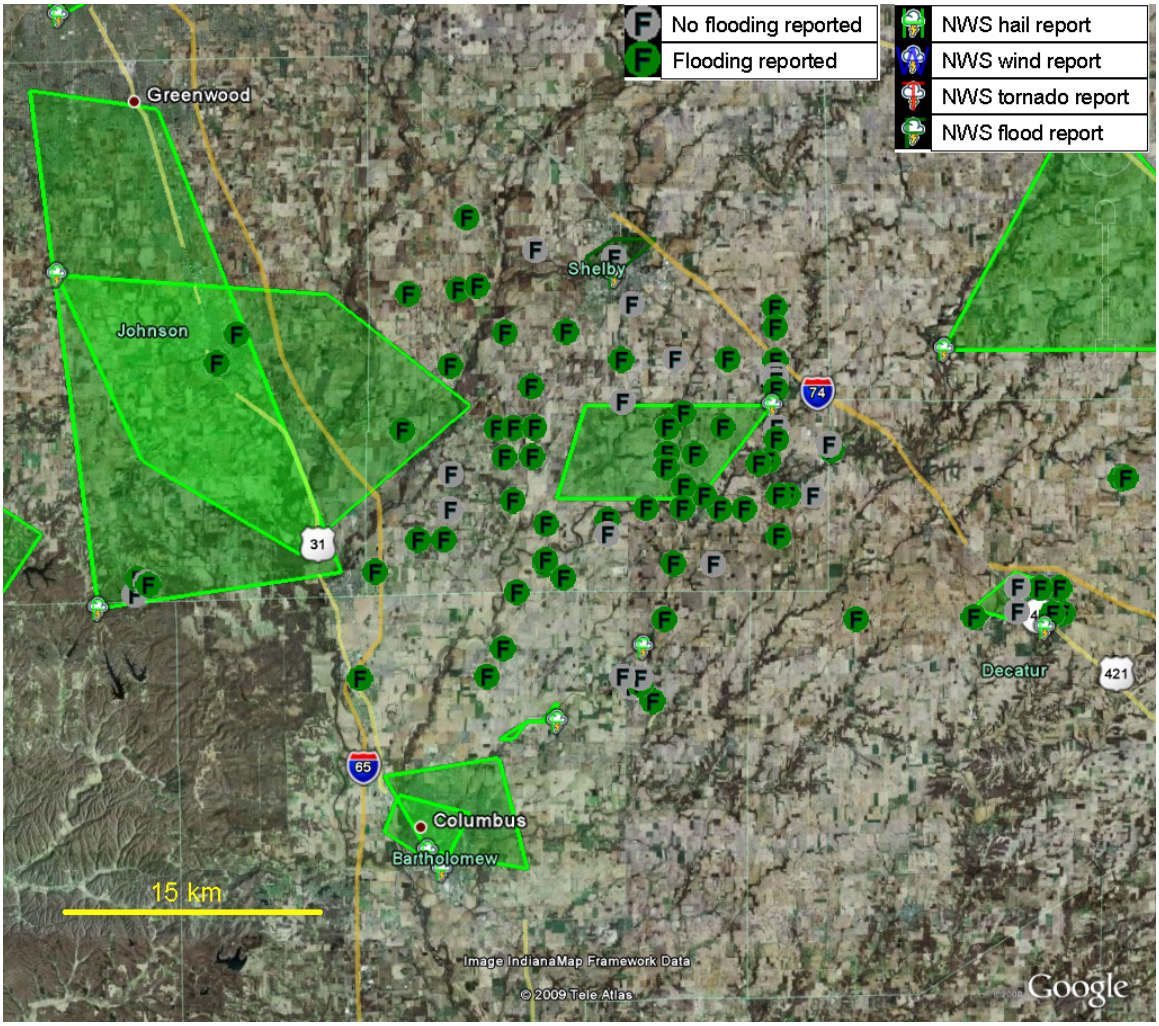

Figure 6: SHAVE and *Storm Data* verification overlaid in Google Earth™ from a flash flood event in central Indiana on 7 June 2008. *Storm Data* flash flood reports are represented by polygons.

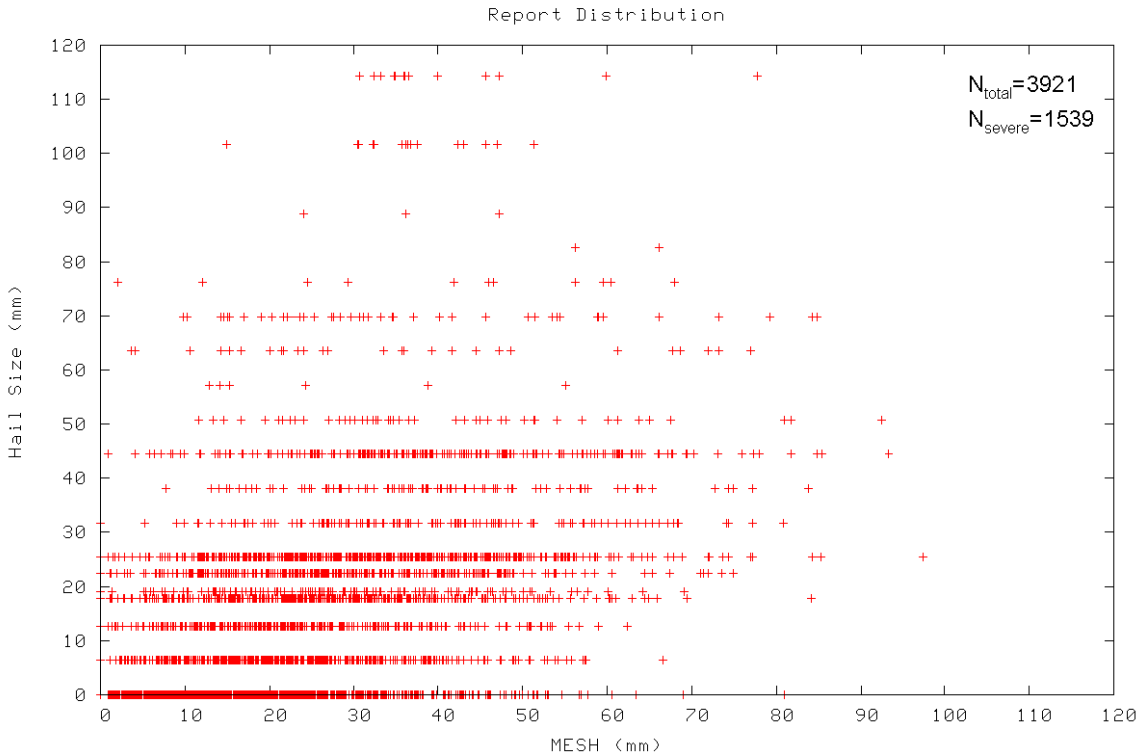

Figure 7: Distribution of reported maximum hail sizes for different MESH values. 3921 SHAVE hail reports were used, of which 1539 were severe (dia. ≥ 19 mm).

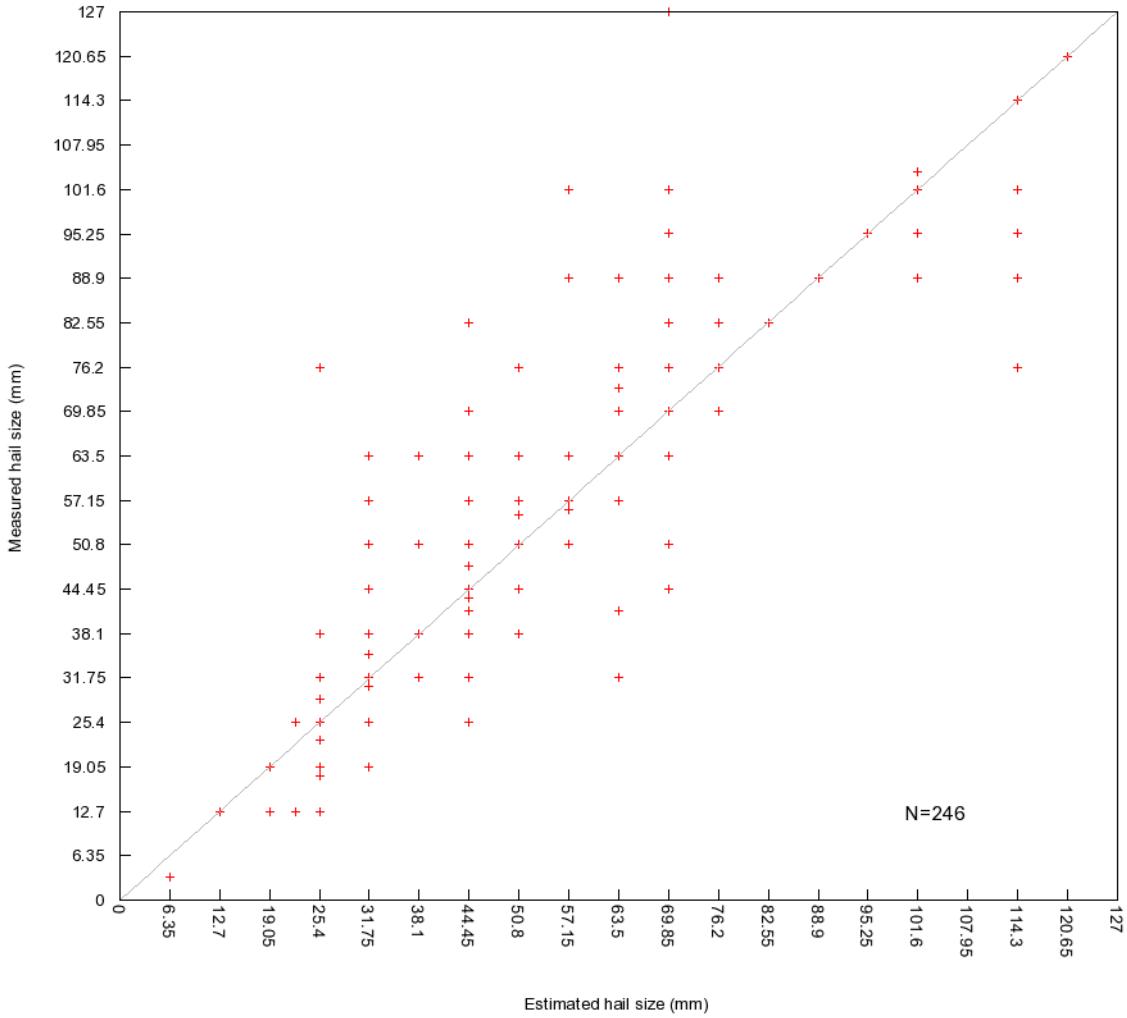

Figure 8: Estimated hail stone diameter compared to measured hail diameters. The diagonal line is estimated hail size equals measured hail size.

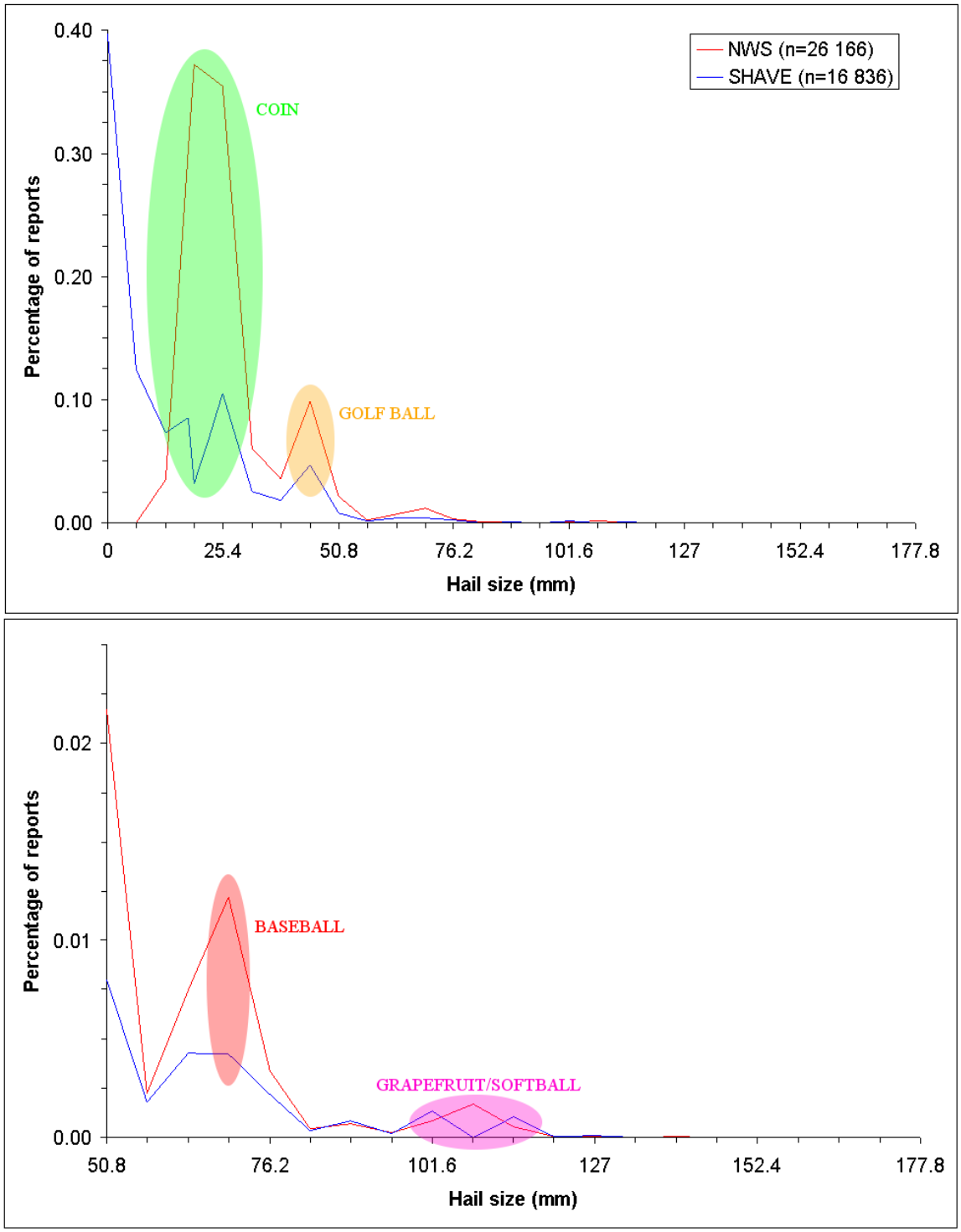

Figure 9**:** *Storm Data* and SHAVE maximum hail size distributions. Highlighted are common size classifications. The bottom graph is a zoom of the 50.8 to 177.8 mm size range of the top graph. *Storm Data* reports from October 2006 to June 2008 were used.

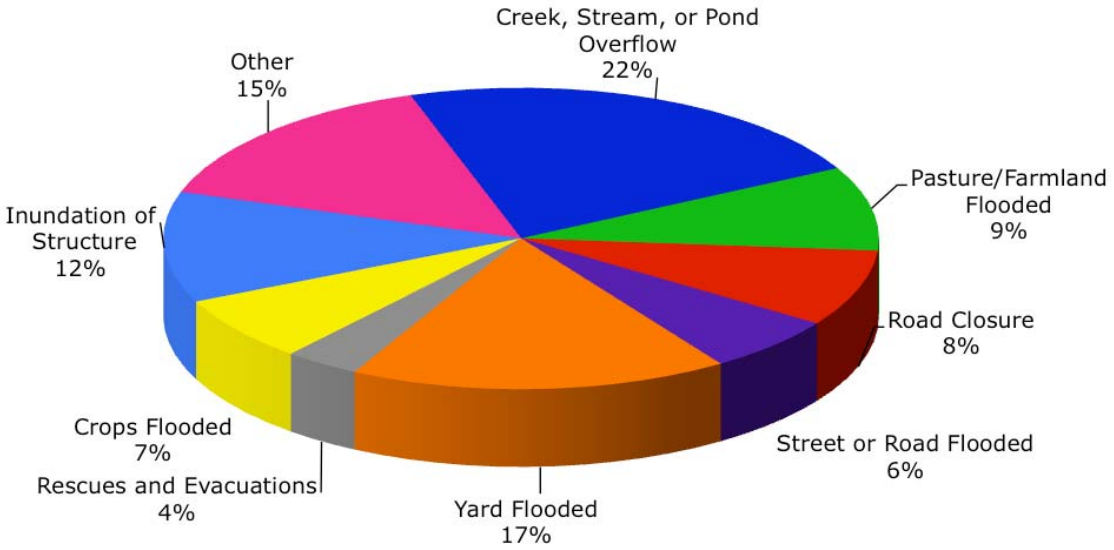

Figure 10: Impacts from surveyed flash flood events during SHAVE.

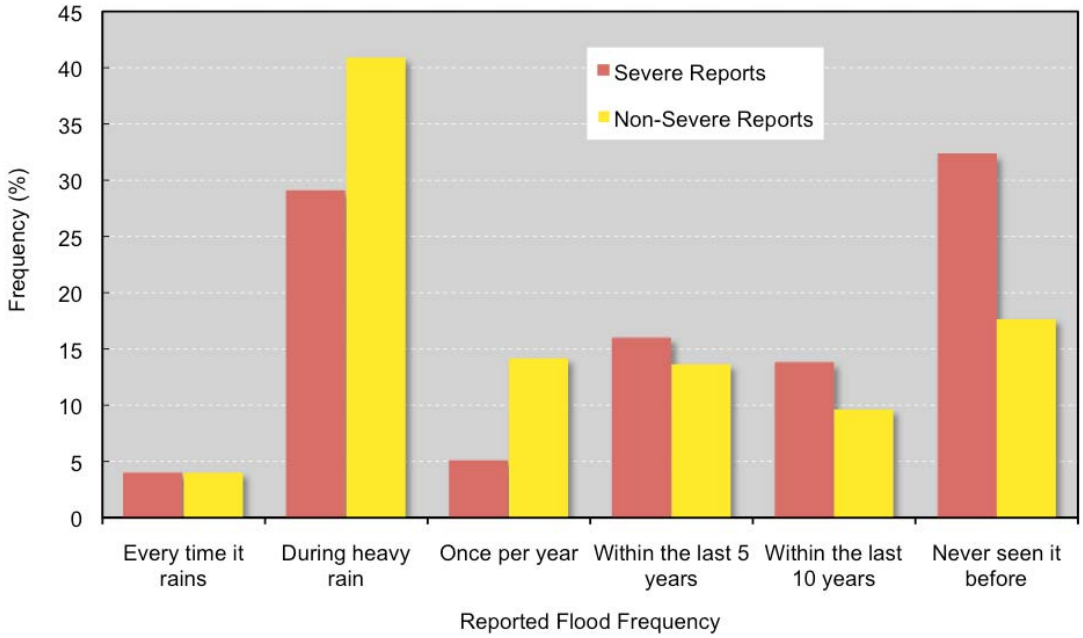

Figure 11: Relative frequency histograms of flash flood frequency for non-severe and severe reports collected during SHAVE.