VERIFICATION OF MULTI-SENSOR, MULTI-RADAR HAIL DIAGNOSIS TECHNIQUES

Kiel L. Ortega[1,2], Travis M. Smith[1,2] and Gregory J. Stumpf[1,3]
[1]U. of Oklahoma/CIMMS; [2]NOAA/NSSL; [3]NOAA/NWS/MDL

1. Introduction

Stumpf et al. (2004) summarize many operational and experimental techniques for diagnosing the presence and size of large hail using a combination of radar reflectivity, reflectivity derivatives and environmental information. Ortega et al. (2005) examined a small set of cases as an initial effort to compare these techniques. This paper represents a continued effort to compare the performance of these techniques for assessing hail size by reporting on a large data sample from different areas of the continental United States consisting of various storm types.

There are many parameters that can affect the way storms are sampled by radar. Because most hail detection techniques require some amount of vertical integration of data, storm tilt and storm motion may adversely affect the calculations. If a storm is near-range to a radar, the upper portions of the storm are not sampled (in the so-called "cone of silence"). If radars have overlapping coverage, other radars may fill in the data missing from the first. If a storm is at long range but observed from multiple radars, the temporal sampling may be improved over a single radar.

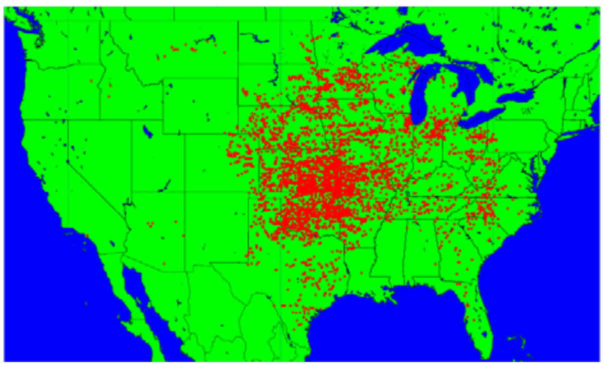

Figure 1: Distribution of the hail reports for the 106 cases analyzed

2. Hail and Storm Data

For 106 cases, we examined the following algorithm-produced products and compared them to hail reports in *Storm Data*:

- Single-radar, cell-based Maximum Expected Size of Hail (MESH) from the WSR-88D Hail Detection Algorithm (Witt et al. 1998a);
- Multi-radar, Cartesian grid data (1 km by 1 km) of MESH (Stumpf et al. 2004);
- Multi-radar, cell-based MESH (Stumpf et al 2002).

The 106 cases were spread throughout the United States, with most cases coming from the Central and Southern Plains, and Midwest regions. The cases occurred in all seasons and contain a variety of storm modes. The case dates range from 8 March 2002 to 11 December 2004. The 106 cases yielded 5671 hail reports. The distribution of hail reports is shown in Figure 1. The three-dimensional multi-radar reflectivity technique used to compute MESH for some of these comparisons is described by Lakshmanan et al (2005).

One must always maintain caution when using severe weather reports from *Storm Data* (Witt et al. 1998b). The major problem with comparing algorithm-predicted MESH to a hail report is that the report may not be the maximum hail size that fell. Incorrect report times and locations are also a problem.

Hail reports from *Storm Data* were quality controlled by hand. The first step was to compare the hail reports to a "hail swath" algorithm product (temporal maximum of multi-radar, maximum MESH grid; see Figure 2). The area within a 5 km radius from each report was searched with an automated

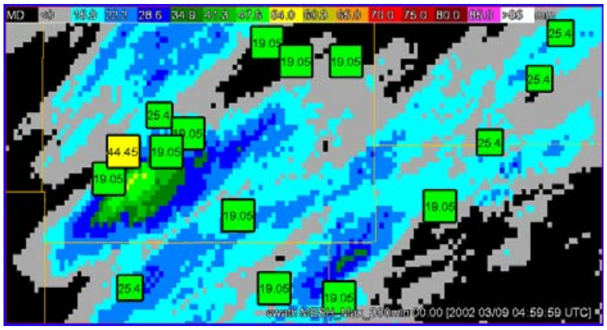

Figure 2: Example of a "hail swath" with hail reports overlaid

algorithm. Any reports receiving a missing value in the grid were further examined. Any report not associated with radar reflectivity greater than 40 dBZ in a time window from 15 minutes before to 5 minutes after the report time in the same 5 km search radius was considered a bad report and excluded from the report data set.

Sixty reports were excluded due to an error in the scoring program which incorrectly assigns missing values to reports near the merged domain edge. Some report locations were modified when the latitude/longitude of the report was found to be in error.

3. Hail Diagnosis Techniques

3.1 Cell-based versus Grid-based Techniques

Stumpf et al. (2004) provides an excellent overview of the procedures involved in cell-based and grid-based hail diagnosis. The major difference is that cell-based integrations occur with each storm, while grid-based integrations occur at each grid point in the merger domain. Cell-based integrations automatically include any tilting or twisting of reflectivity with height because they are based upon the maximum value of reflectivity at each vertical level in the storm. This will provide a different maximum value compared to the completely vertical integrations involved with the grid-based methods. Additionally, hail size estimates from grid-based vertical integration methods may vary with the resolution of the input 3D data field.

Another difference between the methods is the ability to track temporal/spatial trends in storm attributes. Cell-based techniques make this simple, as the attribute trend can be shown in graphical format. While this is not done with the grid-based techniques, the grids can be used to more easily examine the spatial extent of the attribute, as seen in Figure 2.

3.2 Storm Motion

Storm motion may have an adverse effect on the integrations, especially the grid-based techniques. This adverse effect occurs through a "false tilt." The "false tilt" occurs because of the time delay between the lowest radar scan and the highest radar scan. In this study, storm motions are taken into account when merging data together from multiple radars by floating the reflectivity along the storm motion vector using a time/space correction technique based on a storm segmentation and motion estimation method described by Lakshmanan et al. (2003).

3.3 Tilted Integration

method	RMSE	RMSE 95% CI	MAE	Bias	duration	type	integration	time-space correction?
1	27.90	27.44 – 28.37	20.95	13.95	20 min	SCIT		NA
2	22.95	22.20 – 23.57	17.37	9.24	20 min	MR SCIT		no
3	23.61	22.98 – 24.22	17.74	10.36	20 min	MR SCIT		yes
4	18.86	18.21 – 19.43	14.21	3.60	20 min	grid+MR SCIT	vertical	no
5	19.48	18.92 – 19.97	14.68	5.34	20 min	grid+MR SCIT	vertical	yes
6	17.47	16.82 – 18.05	13.04	-0.93	20 min	grid+MR SCIT	along storm tilt	no
7	17.37	16.82 – 17.83	12.98	0.23	20 min	grid+MR SCIT	along storm tilt	yes
8	17.84	17.35 – 18.31	13.46	2.73	entire case	grid	vertical	no
9	18.76	18.28 – 19.22	14.20	4.96	entire case	grid	vertical	yes
10	17.12	16.62 – 17.6	12.88	1.62	entire case	grid	along storm tilt	no
11	17.94	17.43 – 18.44	13.52	3.38	entire case	grid	along storm tilt	yes

Table 1: Root Mean Square Error (RMSE; mm), 95% RMSE confidence interval (RMSE 95% CI; mm), Mean Absolute Error (mm), Bias (mm), time window length around the reported hail time, computation type, integration strategy, and time-space correction information for each of the eleven techniques.

Actual storm tilt, which can be caused by high vertical wind shear, can also have an adverse effect on the integrations, especially with the grid-based techniques. Stumpf et al. (2004) explored using integration along the storm tilt because of these limitations with the grid-based technique. This study used automatically-determined storm tilts by fitting 2 least-squares lines – a west-to-east line and north-to-south line – to Multi-Radar Storm Cell Identification and Tracking algorithm (Stumpf et al 2002) 2-D features. A Barnes objective analysis was then computed on the west-to-east and north-to-south tilts and the result was used as input to the multi-radar reflectivity merging algorithm. The integration then took place along the tilt and the value is projected on the lowest elevation of the tilted column. Tilted integrations are essentially hybrid techniques, requiring both cell features and vertical grids.

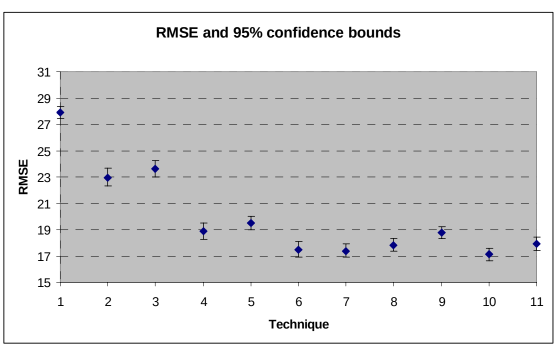

Figure 3: RMSE and 95% confidence intervals for the eleven techniques.

3.4 Integration Techniques

This study examined a combination of regular (cell and grid), tilted and time/space corrected techniques. For the multi-radar, cell-based techniques the maximum grid values of MESH (regular and tilted) were located within a 5 km radius around the cell and used as another diagnosis (noted as “grid-extracted” below). In all there are eleven techniques evaluated:

1. Single radar, cell-based MESH from the WSR-88D Hail Detection Algorithm
2. Multi-radar, cell-based MESH
3. Multi-radar, cell-based, time/space corrected MESH
4. Multi-radar, cell-based, grid-extracted MESH
5. Multi-radar, cell-based, time/space corrected, grid-extracted MESH
6. Multi-radar, cell-based, grid-extracted, tilted MESH
7. Multi-radar, cell-based, time/space

corrected, grid-extracted, tilted MESH
8. Multi-radar, grid-based MESH
9. Multi-radar, grid-based, time/space corrected MESH
10. Multi-radar, grid-based, tilted MESH
11. Multi-radar, grid-based, time/space corrected, tilted MESH

4. Results

The hail reports were compared to the above eleven hail diagnosis techniques. For the cell-based methods, the maximum cell-based MESH within a 10 km search radius from the hail report location and -15 to +5 minute time window around the report time was used for comparison. For grid-based methods, a “hail swath” for the time period of the entire case was produced and compared to the hail reports. The maximum value from the “hail swath” grid within a 5 km radius around the report was used for comparison. The root mean square error (RMSE), 95% bootstrap-estimated confidence interval for RMSE (Figure 3), mean absolute error (MAE), and bias were computed for each technique and are shown in Table 1.

The scores show that pure cell-based techniques have a high bias in hail size estimation that contributes, in part, to larger RMSE and MAE scores than the grid-based techniques. The difference in MSE between the best performing, purely cell-based technique (technique 2; RMSE = 22.96) and the poorest performing, purely grid-based technique (technique 9; RMSE = 18.76) is quite glaring.

The three worst techniques in terms of both MSE and bias were the three cell-based methods. In fact the worse performance is from the current WSR-88D single radar, cell-based method! This agrees with the study by Ortega et al. (2005), which showed one case quantitatively and several cases qualitatively, where the multiple radar techniques outperformed the single radar techniques. The three best techniques in both MSE and bias were techniques using tilted integration.

5. Conclusions

Two-dimensional gridded fields of hail size based on integration of a three-dimensional reflectivity field have multiple advantages over the WSR-88D Hail Detection Algorithm. Because the MESH field is produced from a multiple-radar reflectivity field, radar sampling issues such as the “cone of silence” are reduced or eliminated. Two-dimensional spatial fields of MESH show the horizontal extent of the hail swath, while the WSR-88D HDA only provides one value of MESH for a storm cell and does not summarize past hail activity. The skill scores for the grid-based methods are much improved over the cell-based MESH.

While the procedure for calculating verification statistics (RMSE, etc) for the grid-base and cell-based techniques are not exactly the same because of the differences in durations of the time windows used, we believe that the results may be compared because the time errors associated with the hail reports were eliminated from both types of comparisons. While the comparisons were done quite precisely with respect to spatial relation between the report and algorithm output, the temporal relation was not as firmly established (because of the use of “hail swaths” for the gridded techniques). An extension for a better one-to-one comparison would be to make a “hail swath” using a time window similar to the search time window for the cell-based method comparison.

One method of hail diagnosis not explored was the use of dilation of the reflectivity field. Dilation should help deal with storm tilt caused by high vertical wind shear and storm movement, but is more resource-intensive than the tilted integrations.

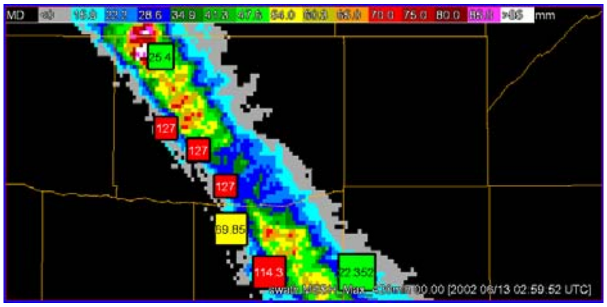

Figure 4: "Hail swath" compared to hail reports showing difference in location between the swath centerline and maxima, and hail report locations

The location of hail fall compared to the "hail swath" needs to be further explored. Figure 4 illustrates how the hail reports were often downwind and to the right of the storm motion when compared to the "hail swath." A hail size verification project is currently being planned to make hail report cross-sections through "hail swaths" to create a high resolution data set for future use.

Further investigation is needed involving the additional parameters for hail diagnosis described in Stumpf et al. (2004). A detailed look into storms that are currently problematic: storms in the "cone of silence," highly sheared storms, fast moving storms, storms at long ranges, and storms in rapidly evolving near-storm environments will expand this work in the future. Future work will also expand the hail detection algorithms to incorporate polarimetric variables.

6. Acknowledgements

Funding for this research was provided under NOAA-OU Cooperative Agreement NA17RJ1227. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Severe Storms Laboratory, the National Weather Service Storm Prediction Center or the U.S. Department of Commerce. Thanks to Arthur Witt, Valliappa Lakshmanan, and Kim Elmore for their comments on the manuscript.

References

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

Lakshmanan, V., R. Rabin, and V. DeBrunner, 2003: Multiscale storm identification and forecast. *J. Atm. Res.*, 367–380.

Lakshmanan, V., T. Smith, K. Hondl, G. J. Stumpf, and A. Witt, 2005: A real-time, three dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity and derived products. *Weather and Forecasting*, Accepted.

Ortega, K. L., T. M. Smith, G. J. Stumpf, J. Hocker and L. Lopez, 2005: A comparison of multi-sensor hail diagnosis techniques. *Preprints, 21st Conf. on IIPS*, San Diego, CA, Amer. Meteor. Soc., CD-ROM, P1.11.

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, 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.

Witt, A., M. D. Eilts, G. J. Stumpf, E. D. Mitchell, J.T. Johnson, and K. W. Thomas, 1998b: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13**, 513-518.