Exploration of the NSSL Maximum Expected Size of Hail (MESH) Product for Verifying Experimental Hail Forecasts in the 2014 Spring Forecasting Experiment

Christopher J. Melick[1,2], Israel L. Jirak[1], James Correia Jr.[1,2], Andrew R. Dean[1], and Steven J. Weiss[1]

[1] NOAA/NWS/NCEP/Storm Prediction Center

[2]Cooperative Institute for Mesoscale Meteorological Studies , University of Oklahoma

Introduction and Motivation

- **Objective forecast verification of experimental severe weather forecasts was conducted in near-real time in the 2014 Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE), similar to what had been accomplished the prior two years (Melick et al. 2013).**
- **For the first time, however, individual hazard (tornado, wind, hail) probabilistic forecasts were produced by the Severe Desk led by the Storm Prediction Center (SPC) instead of one probabilistic forecast for total severe.** Preliminary local storm reports (LSR) served as the primary verification dataset when computing forecast verification metrics.
- **The goal of this work is to further explore radar-derived maximum expected size of hail (MESH) from the National Severe Storms Laboratory (NSSL) as a valuable surrogate to document the occurrence of hail and to verify the experimental hail forecasts.** Thus, gridded MESH fields retained from the 2014 HWT SFE time period were pursued as an alternative by using a procedure similar to that for severe storm reports. From this, a comparison of results between LSRs and MESH for objective forecast verification are presented.

2014 HWT SFE: Experimental Hail Forecast

- **Individual Hazard Product: SFE participants from the Severe Desk produced severe hail (≥ 1" in diameter) forecasts using the same probability contours in SPC Day 1 operational outlooks (5, 15, 30, 45, and 60%) with the option of adding contour lines (every 5%) for localized maxima.**
  - *Time Period:* 16-12Z outlook period was issued in the morning **by 15Z.**
  - *Philosophy:* The probabilistic forecasts were defined using radius of influence (**ROI**) within 25 miles[40-km] of a point.
  - *Grid:* GEMPAK graph-to-grid routine generated grid values (40-km grid; NCEP 212) of probabilities from drawn contours.
  - *Sample Size:* 17 days between May 5th – June 6th where forecast was issued and all verification data available.

2014 HWT SFE: Verification Datasets

- **LSRs: Hail reports (≥ 1" in diameter) received from National Weather Service forecast offices.**
  - *Raw Grid:* Counts over 20-hr valid period (through just after 12 Z) within 25 miles [40-km] of grid point on NCEP 212 grid
  - *Binary grid of LSRs:* One or more severe hail reports at each grid point constitutes an observed object
  - *PP LSR Grid:* Following Hitchens et al. (2013), "practically perfect"[PP] hindcast created by applying 2-D Gaussian smoother (sigma=120-km) to binary grid of LSRs.
- **MESH: High resolution grid (0.01 deg) of hourly maximum MESH obtained from NSSL**
  - *Raw grid:* Multi-hour (20-hr) maximum field is created to coincide with the 16-12Z forecast period
  - *Filtered grid:* Remove isolated pixels by using 2-D Gaussian smoother (sigma=0.01 deg).
  - *Neighborhood Max Grid:* Applied 40-km radius of influence [ROI]) to raw and filtered MESH 0.01 deg grid boxes.
  - *Interpolation:* 20-hr, neighborhood maximum MESH fields (raw and filtered) interpolated to NCEP 212.
  - *Final QC check:* NLDN flash count data used; MESH kept if 1 or more flashes within 40-km ROI of grid point
  - *Binary grid of MESH:* MESH ≥ 29mm (Cintineo et al. 2012) at each grid point constitutes an observed object.
  - *PP MESH Grid:* Similar to LSR process, PP hindcast produced from binary grids of raw and filtered MESH fields.

2014 HWT SFE: Forecast Verification Metrics

- *Domain:* **All of the daily evaluations were restricted to a mesoscale "area of interest" for possible severe convection. This small domain was movable to locations in the eastern and central US.**

1. **Probability thresholds define binary event to construct 2x2 Contingency Table [Wilks 2006].**

Cont. Table	Observed *YES*	*Observed NO*
Forecast YES	"a" Hit	"b" False Alarm
Forecast NO	"c" Miss	"d" Correct Negative

Probability Threshold	*Forecast*	*Verification*
5%, 15%, 30%, 45%	Experimental Severe Hail	Binary grids of MESH or LSRs

2. **Direct comparison of probabilities between experimental hail forecast and PP:**

Fractions Skill Score **(Schwartz et al. 2010)**

$$\left(FSS = 1 - \left[\frac{FBS}{FBS_{worst}} \right], \text{ where } FBS = \frac{1}{N}\sum_{i=1}^{N}\left[NP_{F(i)} - NP_{O(i)}\right]^2 \text{ and } FBS_{worst} = \frac{1}{N}\left[\sum_{i=1}^{N} NP_{F(i)}^2 + \sum_{i=1}^{N} NP_{O(i)}^2\right] \right)$$

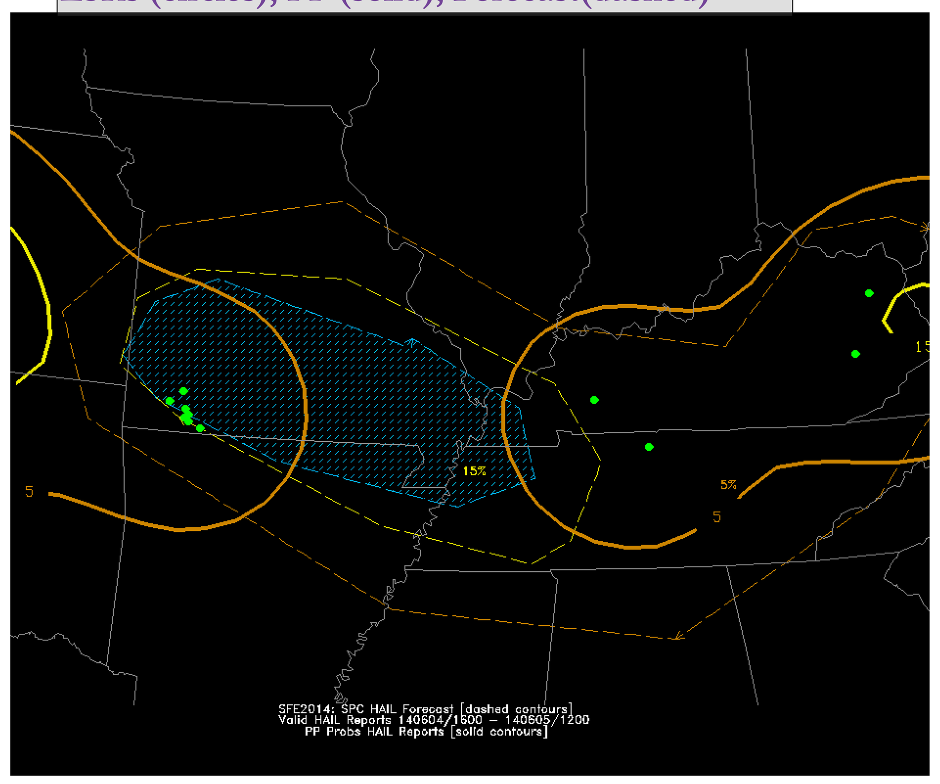
References:

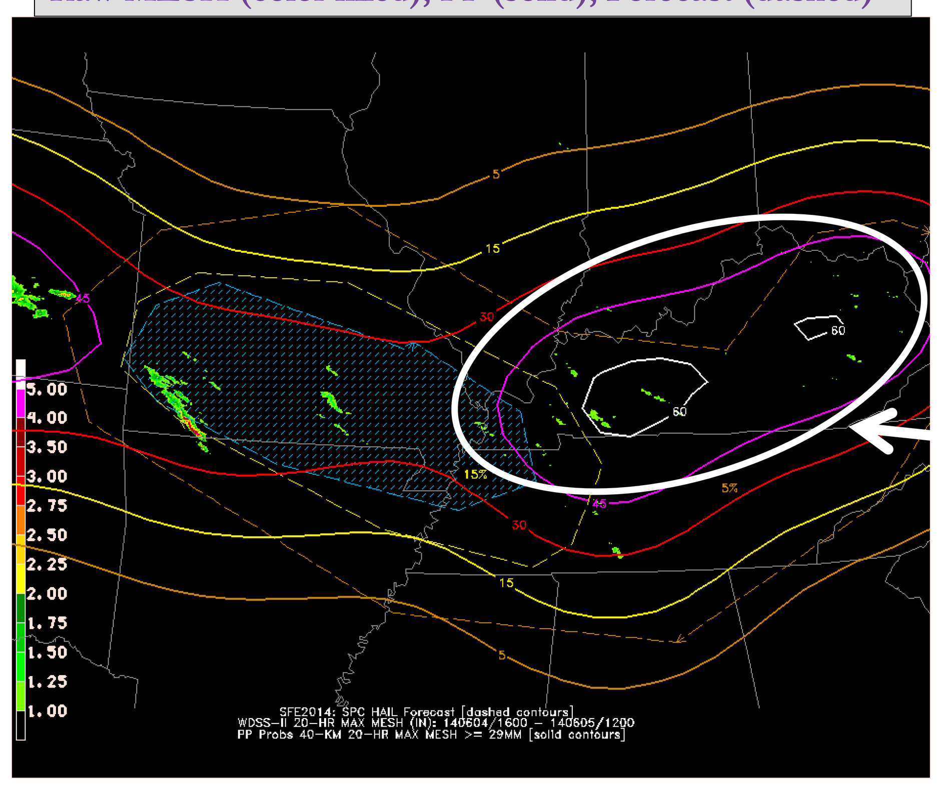
- Cintineo, J.L., T.M. Smith, V. Lakshmanan, H.E. Brooks, and K.L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. Wea. Forecasting, **27**, 1235–1248.
- Hitchens, N.M., H.E. Brooks, and M.P. Kay, 2013: Objective limits on forecasting skill of rare events. Wea. Forecasting, **28**, 525–534.
- Melick,C.J, I.L. Jirak, J. Correia Jr, A.R. Dean, and S.J. Weiss, 2013: Utility of Objective Verification Metrics during the 2013 HWT Spring Forecasting Experiment. Preprints, 38th Natl. Wea. Assoc. Annual Meeting, Charleston, SC, Natl. Wea. Assoc., P1.27.
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. Wea. Forecasting, **25**, 263–280.

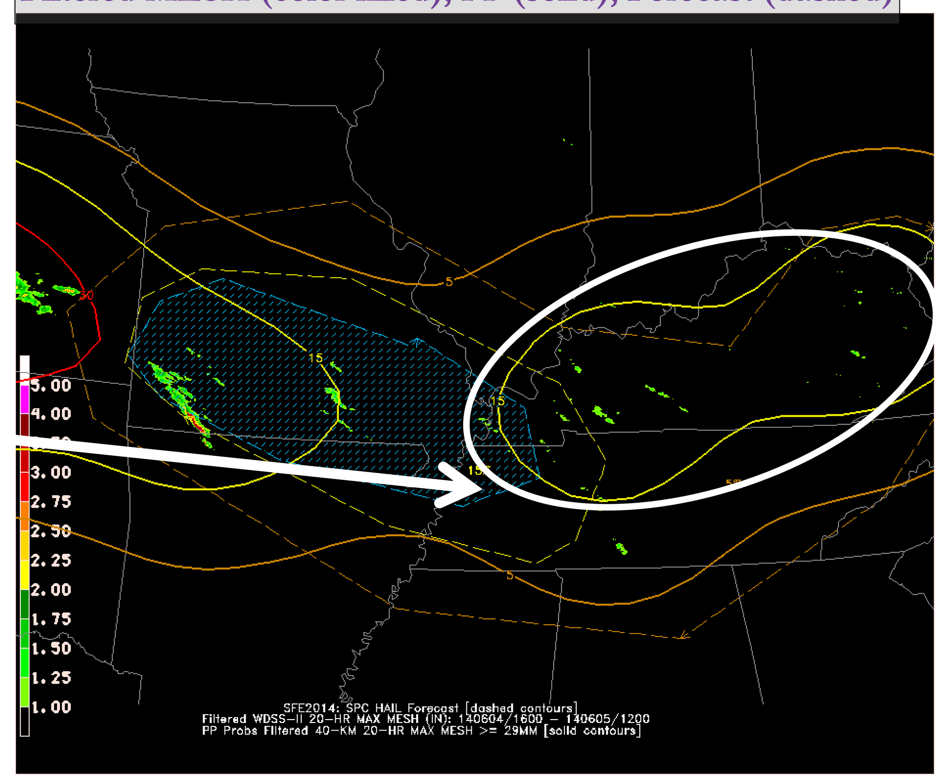
Acknowledgments:
This poster was prepared with funding provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

Case Studies: Verification of Probabilistic Hail Forecasts using MESH versus LSRs

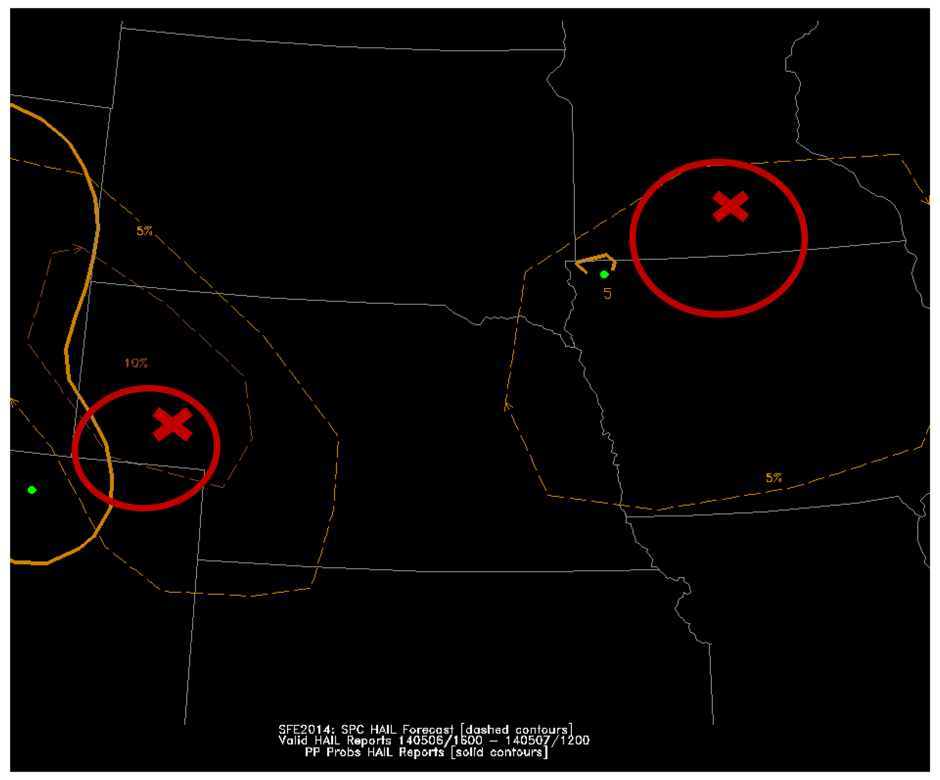
LSRs (circles), PP (solid), Forecast(dashed) | Raw MESH (color filled), PP (solid), Forecast (dashed) | Filtered MESH (color filled), PP (solid), Forecast (dashed)

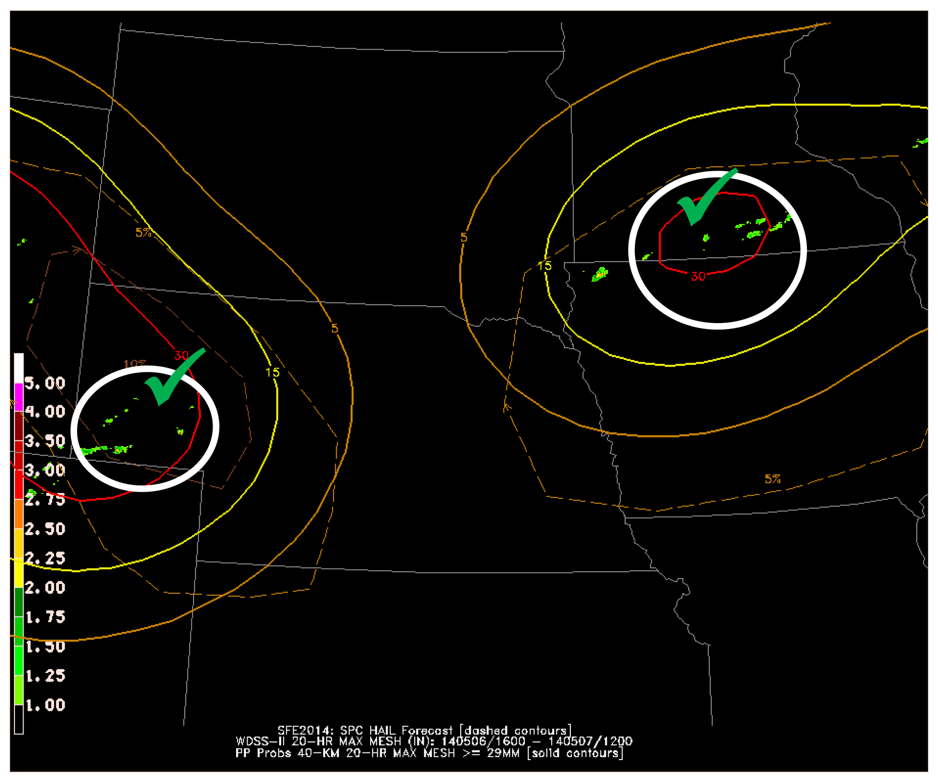


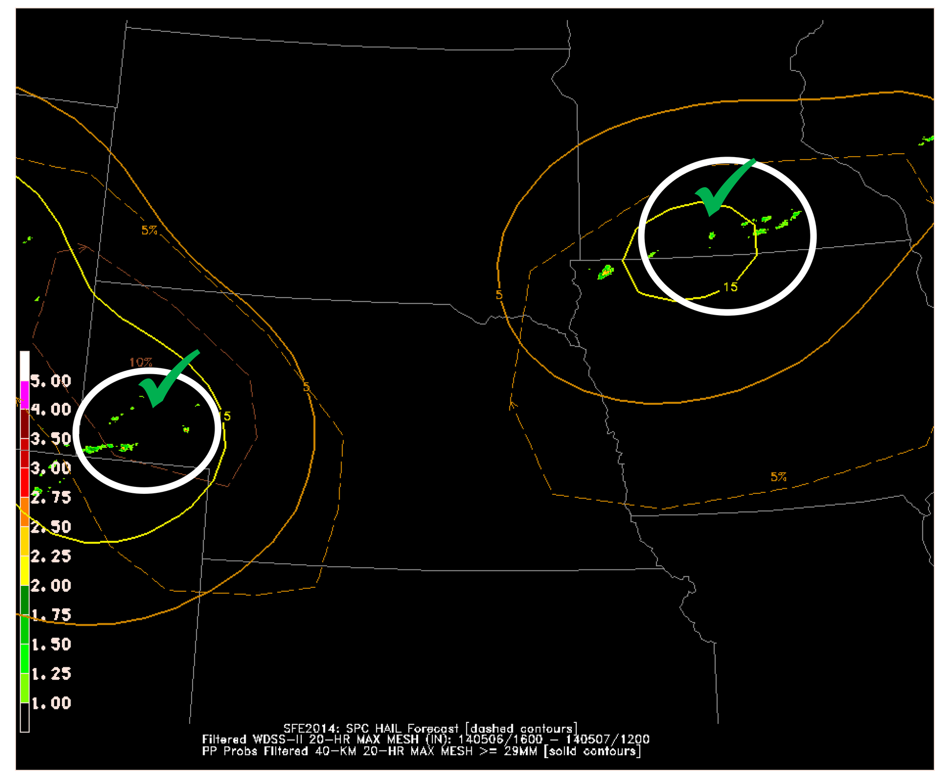




May 20th 2014: →All hail reports in 5% forecast contour but disjointed PP →Connected probability area from MESH PP is more useful →After filtering MESH, 15% PP more appropriately matches sparse coverage



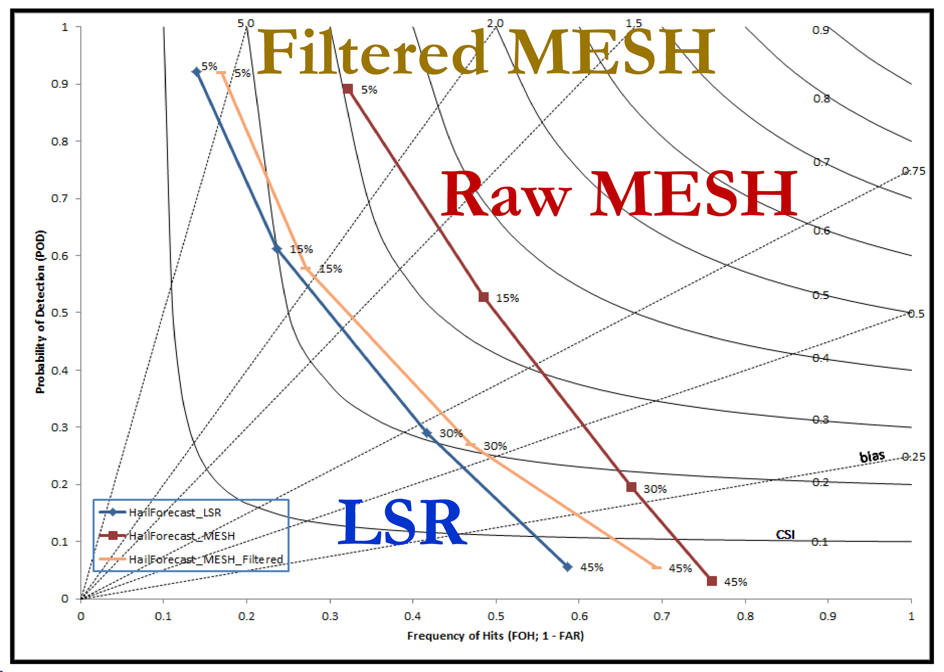




May 6th 2014: →Marginal verification at 5% with isolated hail LSRs → MESH tracks over MN, NE, and WY show greater spatial coverage →The excessive 30% contour is reduced to 15% on Filtered MESH

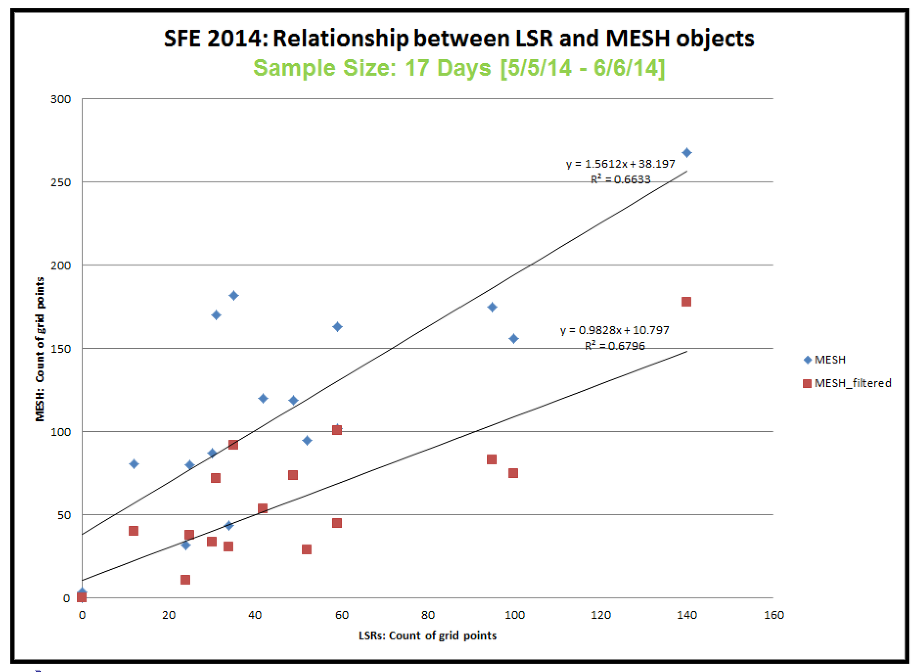
Statistical Forecast Verification Results: LSRs, Raw MESH, and Filtered MESH

Performance Diagram at Standard Thresholds



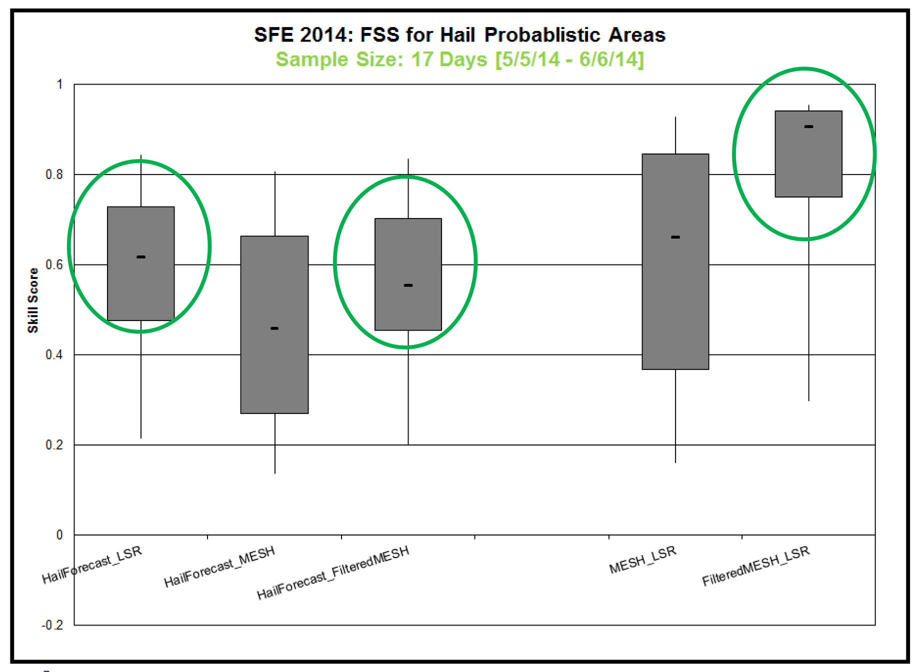
→ Using MESH resulted in similar POD but higher FOH compared to LSRs

Scatter Plot Correlation for Observed Objects



→Many more MESH objects than LSRs, but filtering reduces number

FSS: Direct Comparison between Probabilistic Areas



→ Higher FSS with greater similarity for LSRs and Filtered MESH

Summary and Conclusions

- Objective forecast verification expanded in the 2014 HWT SFE to evaluate experimental probabilistic forecasts for individual severe hazards (tornado, wind, hail) with primary observations coming from LSRs. However, MESH plots were also created to test alternative verifying data.
- A more formal comparison of skill using MESH occurred post SFE where a similar procedure to LSRs was followed. The results show MESH to have greater number of observed objects in contrast to severe reports resulting in lower false alarms and higher CSI at all probability thresholds. FSS increased with Filtered MESH as it better agreed with the forecast and LSRs.
- MESH tracks proved to be valuable in identifying events in low-density population areas but also as an independent dataset to supplement hail LSRs. Future efforts will incorporate forecast verification metrics from both LSRs and MESH in a side-by-side diagnosis in subsequent SFE years. In addition, methods to realistically combine both verification datasets will be investigated.