P. 68 EVALUATING THE FUNDAMENTAL COMPONENTS OF A WARN-ON FORECAST SYSTEM IN A COLLABORATIVE REAL-TIME EXPERIMENT

Patrick T. Marsh[1,2], John S. Kain[2], Steven J. Weiss[3], Ryan A. Sobash[1], Fanyou Kong[4], Kevin W. Thomas[4], and Ming Xue[4]

[1]School of Meteorology, University of Oklahoma, Norman, OK
[2]NOAA/National Severe Storms Laboratory, Norman, OK
[3]NOAA/NWS/Storm Prediction Center, Norman, OK
[4]Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK

1. Introduction

During the 2010 Spring Forecasting Experiment (SE2010), conducted by NOAA Hazardous Weather Testbed Experimental Forecast Program, WRF-model guidance from a variety of convection-allowing model configurations was examined, including output from a 26-member high resolution ensemble produced by the Center for Analysis and Prediction of Storms (CAPS – see Xue et al. this workshop). The different model configurations and initialization procedures used in SE2010 included all the fundamental components of "Warn-on-Forecast" (WoF), a paradigm based on probabilistic prediction of severe convective phenomena using ensembles of storm-scale models (Stensrud et al. 2009). Specifically, different model configurations included a variety of models resolutions, dynamic cores, and physics options, applied in both deterministic and ensemble systems, while initialization procedures involved multiple assimilation approaches, both with and without radar data and other non-traditional data sources.

All of these factors are potentially important in bringing WoF to fruition. However, another of many scientific challenges facing WoF is how to construct reliable probabilistic information regarding severe convective phenomena when these phenomena will not be explicitly resolvable in larger domain model configurations for many years to come (e.g., explicit prediction of tornadoes will require grid spacing on the order of a few tens of meters). It may be possible to overcome this problem by identifying "extreme" model-generated features that have strong correlations with observed severe convective phenomena, and then using the former as surrogates for the severe phenomena in question. This "surrogate-severe" (SS) approach is fundamentally different from traditional applications of NWP for severe weather because it is phenomenon based. In particular, it relies on identification of explicit convective phenomena rather than environmental conditions that might support such phenomena.

Sobash et al. (2009) established the viability of this approach using several different SS diagnostic quantities. Among the quantities he examined, model-generated updraft helicity (UH) appeared to show the strongest correlation with observed reports of severe weather. UH is a measure of mid-level rotation in model-predicted updrafts and subjective assessments suggest that it is a useful surrogate for supercell thunderstorms (Kain et al. 2010), even when these storms are only crudely represented on the WRF model's native grid (Kain et al. 2008). Sobash et al. (2009) used a "neighborhood" approach based on the concepts in Theis et al. (2005) and Brooks et al. (1998) to produce severe-weather probability forecasts based on the locations of UH features in a deterministic model. During SE2010 these same concepts were applied to output from the CAPS high resolution ensemble and it was found that interpretation of derived probabilistic forecasts depends strongly on the parameters used for post-processing. Much of this has to do with the rarity of severe weather events. For example, when dealing with rare events in ensemble output, few grid points activate in any one member so that the likelihood of overlapping active points from multiple members can be quite low. Consequently, ensemble-based probabilities can be correspondingly low. However, application of various neighborhood approaches in post-processing can yield guidance products that are both consistent with the needs of severe weather forecasters (e.g., guidance for the probability of occurrence within a specified radius rather than at a point) and clearly convey the signal inherent within the model forecasts. This extended abstract presents examples of various derived products, their potential utility for current Outlook-scale severe weather forecasts, and their possible application within the focused scales of WoF for severe weather.

2. Dataset

SE2010 was held from 17 May through 18 June and included subjective evaluation of the CAPS ensemble initialized at 00 UTC with a 4-km grid spacing over the contiguous United States and run out 30 hours. This 26-member storm-scale ensemble consists of 19 Weather Research and Forecasting (WRF) Model – Advanced Research WRF (WRF-ARW) members, 5 WRF model Nonhydrostatic Mesoscale Model (WRF-NMM) members, and 2 Advanced Regional Prediction System (ARPS) members (Xue et al. this workshop). From this storm-scale ensemble, a subset of 15-members – the three control members, along with nine additional WRF-ARW members and three additional WRF-NMM members each with initial condition (IC) and/or lateral boundary condition (LBC) perturbations – were used to generate probability of exceedence plots of the hourly maximum UH at a grid point for further evaluation. This subset was selected because previous studies (e.g., Clark et al. 2009) indicated members containing IC/LBC perturbations contributed most strongly to ensemble spread.

It is not readily apparent how to best synthesize and display ensemble guidance for extreme model-based phenomena. In an effort to address this issue, matrices of figures were constructed that display different techniques for computing the probability of the hourly maximum UH exceeding a threshold of 25 m^2s^{-2} at each grid point. Empirical evidence suggests that this threshold is useful for identifying mid-level mesocyclones in 4-kilometer WRF-ARW output (e.g., Sobash et al. 2009 and Kain et al. 2010), so the occurrence of UH greater than or equal to 25 m^2s^{-2} is considered an event. The matrices were designed to illustrate the sensitivity of the exceedence probability to: 1) search radius, and 2) spatial uncertainty on the computation, and to solicit feedback from forecasters and research scientists during the Spring Experiment on optimal ways to display the probability information.

a. Defining an Event

The infrequent nature of rare events makes it unlikely that two separate high resolution model forecasts would place extreme model-generated convective storm phenomena at the same grid point, even for generally similar mesoscale forecasts. Thus, when evaluating the probability that the hourly maximum UH would exceed 25 m^2s^{-2} from the 15-member subset, one would not expect to find very large probability values. This is consistent with the limited predictability on the convective scale and the associated low climatological frequency of rare events, which makes it difficult to convey statistically meaningful severe weather threats to the user community (e.g., Murphy 1991). Informal conversations with operational meteorologists suggest that both forecasters and users of hazardous weather information may not respond appropriately to the very low probability values that result from creating probability of exceedence guidance of rare events on a fine grid.

One potential remedy to this problem is to examine all grid points within a specified search radius for each occurrence of the designated event. As we would expect, a larger search radius tends to yield higher probabilities. For the SE2010, search radii of 4 kilometers (native model grid), 20 kilometers (5 grid points), and 40 kilometers (10 grid points) were examined. The latter radius is consistent with the methodology used in operational SPC Severe Weather Outlooks, which provide the probability of severe weather occurring within 40 km of any point.

One possible negative result to this approach is that by increasing the search radius, the specificity offered by high-resolution numerical models is diminished somewhat. No longer is a forecaster examining the probability of a given grid point exceeding a given threshold, instead the forecaster is examining the probability of *a* grid point exceeding a given threshold within the defined search radius. However, given the limited predictability on the convective grid scale and the comparatively low probability values at the grid point, use of a search radius is considered an appropriate method to identify a small region of enhanced threat.

b. Spatial Uncertainty

A basic concept of ensemble prediction systems (EPS) is that forecast errors typically result from: 1) initial condition uncertainty, and 2) model errors. Thus, use of different model configurations and IC/LBC perturbations should provide information about a possible range of forecast solutions. Ideally, an infinite member EPS would yield a probability density function (PDF) of potential forecasts that encompasses the atmospheric event. However, EPS do not have infinite members and therefore undersample the idealized PDF. Furthermore, most EPS are underdispersive owing to poor initial conditions, less than optimal IC perturbations, and poor physics. As such, the resulting spread of ensemble forecasts sometimes fail to capture the actual event. In the case of storm-scale ensembles, members of an underdispersive EPS might forecast

the occurrence of UH in similar locations and fail to capture all locations of where UH is possible.

One post processing approach to compensate for this underdispersion and undersampling is to apply a Gaussian kernel density estimator to the ensemble forecast (Brooks et al. 1998 and Silverman 1986). In this approach, every identified model-based event is represented using a 2-D Gaussian PDF, allowing us to incorporate a measure of uncertainty in predicting the location of an event. For each member of the ensemble, the event probability at a point is given by a linear combination of all amplitudes of all PDFs; for the entire ensemble, the probability is given by the average of the probabilities from all individual members. Using a similar approach, Schwartz et al. (2010) called this the neighborhood ensemble probability (NEP). Thus, by varying the standard deviation of the 2-D Gaussian distribution, the spatial uncertainty associated with each model-based event also varies. For SE2010, standard deviations of 0 (no uncertainty), 5 grid points, 10 grid points (small spatial uncertainty), 20 grid points (moderate spatial uncertainty), and 30 grid points (large spatial uncertainty) were examined.

3. Example

During the afternoon of 7 June 2010, several large supercell thunderstorms developed across the higher terrain of eastern Wyoming and western Nebraska. This convection ultimately grew upscale into a large Mesoscale Convective System (MCS) over western Nebraska and northeast Colorado and moved east-southeast through the overnight hours. For this paper, the ensemble forecast hour 29 (F29) is considered. For reference, the severe reports received by the National Weather Service for the hour preceding and following F29 are shown in Figure 1.

Ensemble probabilities generated on the native model grid (Figure 2, top-left panel) are quite detailed and contain numerous areas of zero/near-zero probability within the larger area of higher probabilities. When adding a small measure of spatial uncertainty to the location of probabilities (Figure 2, top-right panel) a contiguous region emerges across central Nebraska, but the maximum probability has decreased from 46.67% to 9.69%. The highest probabilities after applying a moderate measure of spatial uncertainty continue to focus on central Nebraska; however, the maximum probability is now only 6.44% (Figure 2, bottom-left panel). Applying large spatial uncertainty results in the loss of the 5% probability of exceedence contour (Figure 2, bottom-right panel).

When defining an event as an exceedence occurring within 20 kilometers of a grid point, the probabilities of exceedence increase as expected. Without accounting for spatial uncertainty, the largest probabilities are found in central Nebraska and have a maximum value of 53.33% (Figure 3, top-left panel). When accounting for small spatial uncertainty in the probabilities, the maximum probability decreases to 31.26%, however a smoother probability field is generated while still highlighting central and eastern Nebraska (Figure 3, top-right panel). After applying a moderate measure of spatial uncertainty the maximum probability is further decreased to 21.93%, but forecaster attention is still drawn to central and eastern Nebraska (Figure 3, bottom-left panel). When applying a large measure of spatial uncertainty, the resultant probabilities arrange in nearly concentric circles indicating that a grid point exists that is within or nearly within 20 kilometers of the largest grid point probabilities. Even so, the maximum probability is now 15.52% (Figure 3, bottom-right panel).

Expanding the definition of an event to exceedence within 40 kilometers of a given point increases the overall probabilities even further. The maximum probability of an event derived from the ensemble has increased to 66.67% (Figure 4, top-left panel). By introducing a small amount of spatial uncertainty the maximum probability decreases to 47.15%, however, a slight west-southwest to east-northeast elongation in the probability field becomes evident (Figure 4, top-right panel). Increasing the spatial uncertainty further (Figure 4, bottom-left panel) reduces the maximum probability to 36.78%. The highest probability values are arranged in nearly concentric circles, whereas the lower probability values capture hints of a southwest to northeast orientation to the probabilities. Increasing the spatial uncertainty still further reduces the maximum probability to 27.16% and becomes more circular in nature, centered in central Nebraska (Figure 4, bottom-right panel).

Comparing the probability of exceedence plots to the severe storm reports collected by the National Weather Service indicates that the ensemble's highest probabilities were observed slightly to the east of the severe weather reports. However, by redefining an event in terms of a search radius and introducing measure of spatial uncertainty, the ensemble probability fields still highlighted the region where severe weather occurred.

4. Concluding Thoughts

A key challenge for the WoF paradigm is to produce probabilistic guidance that has a high degree

of statistical reliability and resolution and is unambiguous for users to interpret. In this paper, it was shown that the character of event-based guidance products can vary substantially depending on the specific definition of an "event", the search radius for its occurrence, and the degree of spatial uncertainty associated with model predictions of it. It is anticipated that the WoF effort will require considerable research on each of these topics in order to optimize probabilistic forecasts of extreme and rare events such as tornadoes, large hail, and extreme wind gusts.

5. Acknowledgements

This research was supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Athena (a Cray XT4) at the National Institute for Computational Science (NICS; http://www.nics.tennessee.edu/)

6. References

Brooks, H. E., M. Kay, and J. A. Hart, 1998: Objective limits on forecasting skill of rare events. *Preprints, 19th Conference on Severe Local Storms*, Minneapolis, Minnesota, Amer. Meteor. Soc., 552-555.

Clark, A. J. and W. A. Gallus Jr., 2009: A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting,* **24**, 1121-1140.

Kain, J. S., S. J. Weiss, D. R. Bright, M. E. Baldwin, J. J. Levit, G. W. Carbin, C. S. Schwartz, M. L. Weisman, K. K. Droegemeier, D. B. Weber, K. W. Thomas, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931-952.

Kain, J. S., M Xue, M.C. Coniglio, S. J. Weiss, F. Kong, T. L. Jensen, B. G. Brown, J. Gao, K. Brewster, K. W. Thomas, Y. Wang, C. S. Schwartz, and J. J. Levit, 2010: Assessing advances in the assimilation of radar data within a collaborative forecasting-research environment, *Wea. Forecasting*, in press.

Kain, J. S., S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010: Extracting unique information from high resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, in press.

Murphy, A. H., 1991: Probabilities, odds, and forecasts of rare events. *Wea. Forecasting,* **6**, 302-307

Silverman, B. W., 1986: *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 175 pp.

Sobash, R. A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, S. J. Weiss, and J. J. Levit, 2009: Forecast guidance for severe thunderstorms based on identification of extreme phenomena in convection-allowing model forecasts. *Preprints, 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction*, Amer. Meteor. Soc., Omaha, NE. CD-ROM 4B.6

Stensrud, D. J. , M. Xue, L. J. Wicker, K. E. Kelleher, M. P. Foster, J. T. Schaefer, R. S. Schneider, S. G. Benjamin, S. S. Weygandt, J. T. Ferree, and J. P. Tuell, 2009: Convective-Scale Warn-on-Forecast System. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. L. Levit, M. C. Coniglio, and M. Wandishin, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting,* **25**, 263-280

Theis, S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, **12**, 257–268.

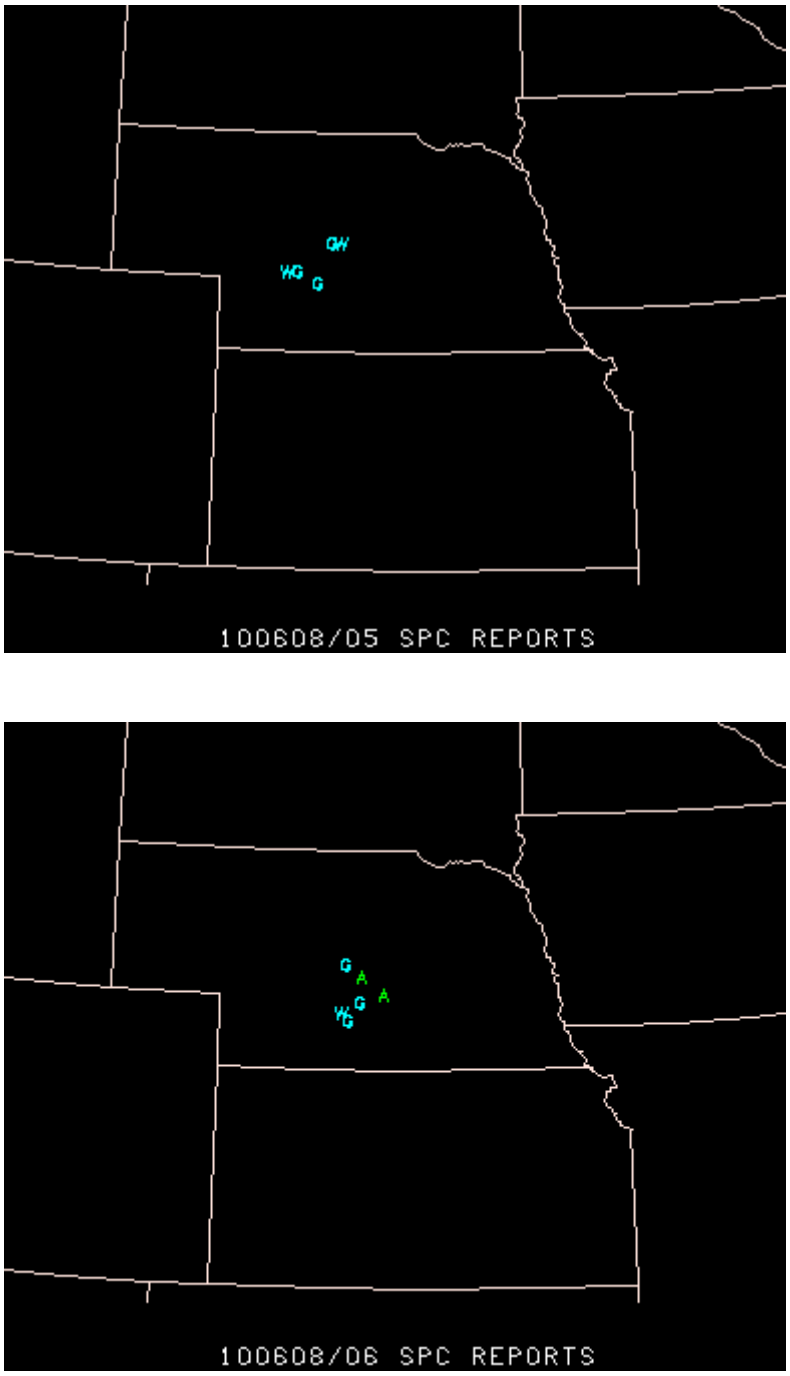


Figure 1. Local storm reports over the central plains for (top) 04-05 UTC 8 June 2010 and (bottom) 05-06 UTC 8 June 2010. Symbol meanings are **W** for severe thunderstorm wind damage, **G** for severe thunderstorm wind gusts, and **A** for severe hail report.

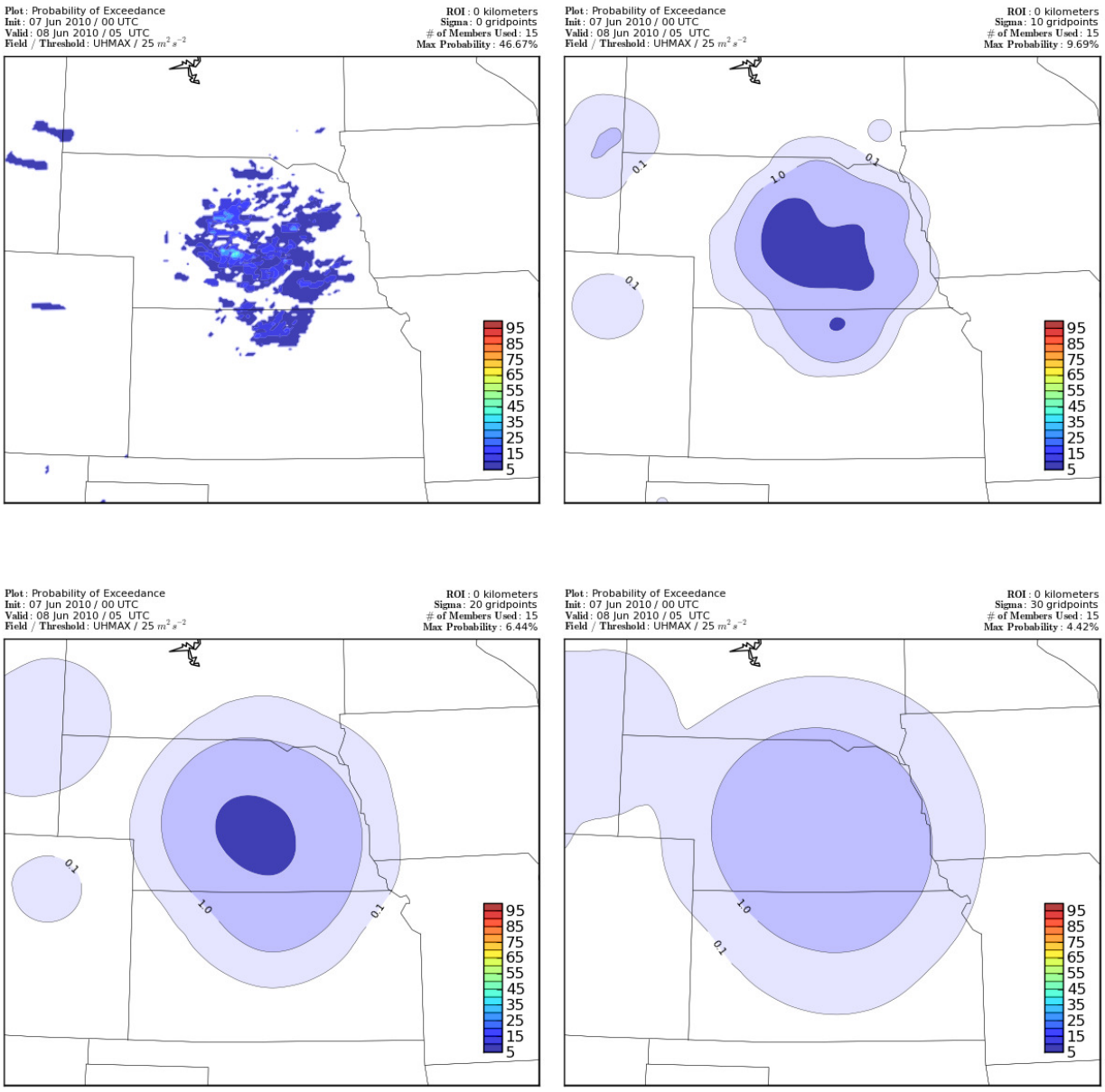


Figure 2. Probability of Exceedence (PoE) plots of hourly maximum updraft helicity using an exceedence threshold of 25 m^2s^{-2}. PoE were computed on the native grid with no smoothing in the upper-left, a sigma value of 10 grid points in the upper-right, a sigma value of 20 grid points in the bottom-left, and a sigma value of 30 grid points in the bottom-right.

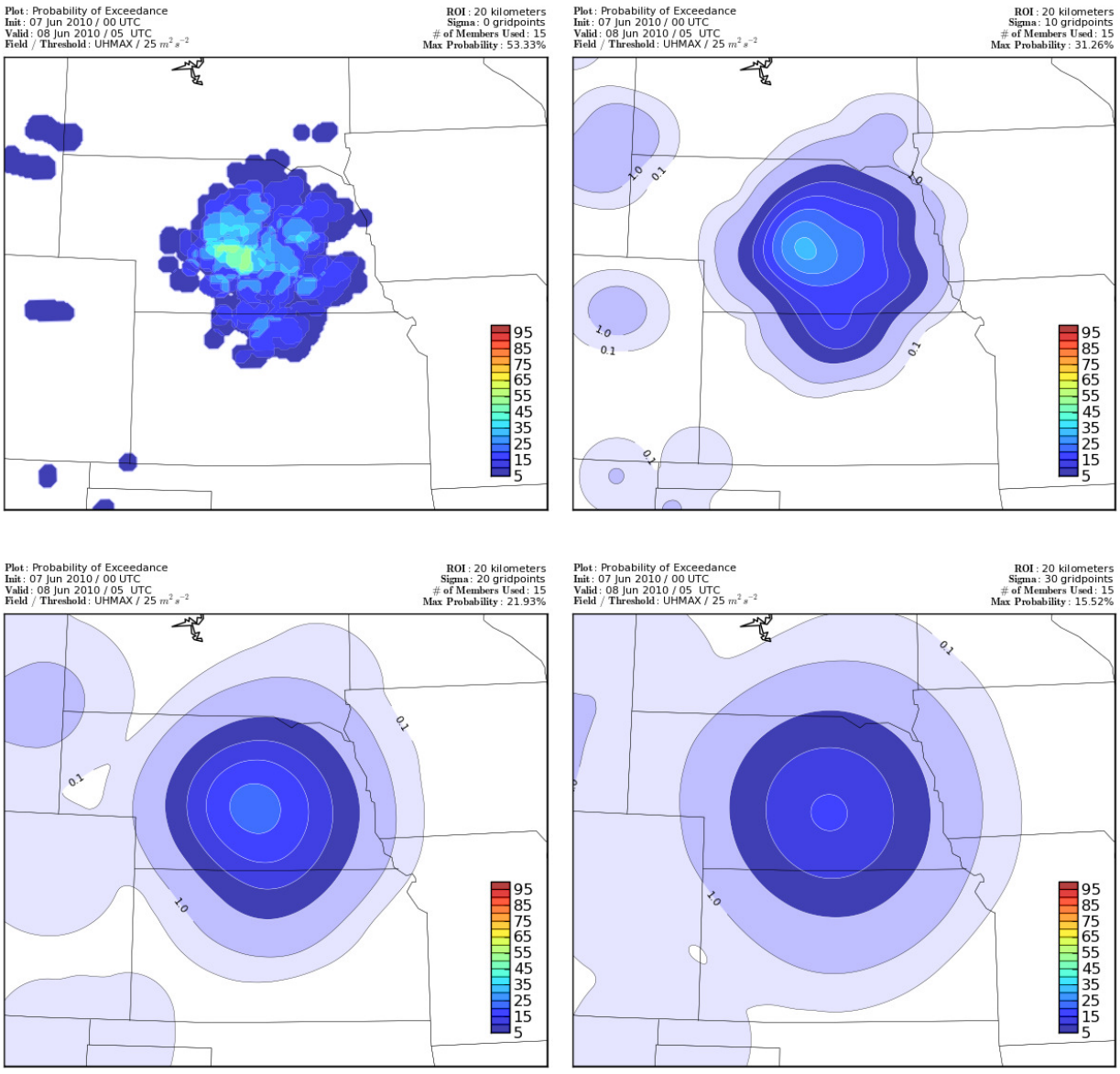


Figure 3. Same as in figure 2, however using a radius of influence of 20 kilometers (5 grid points).

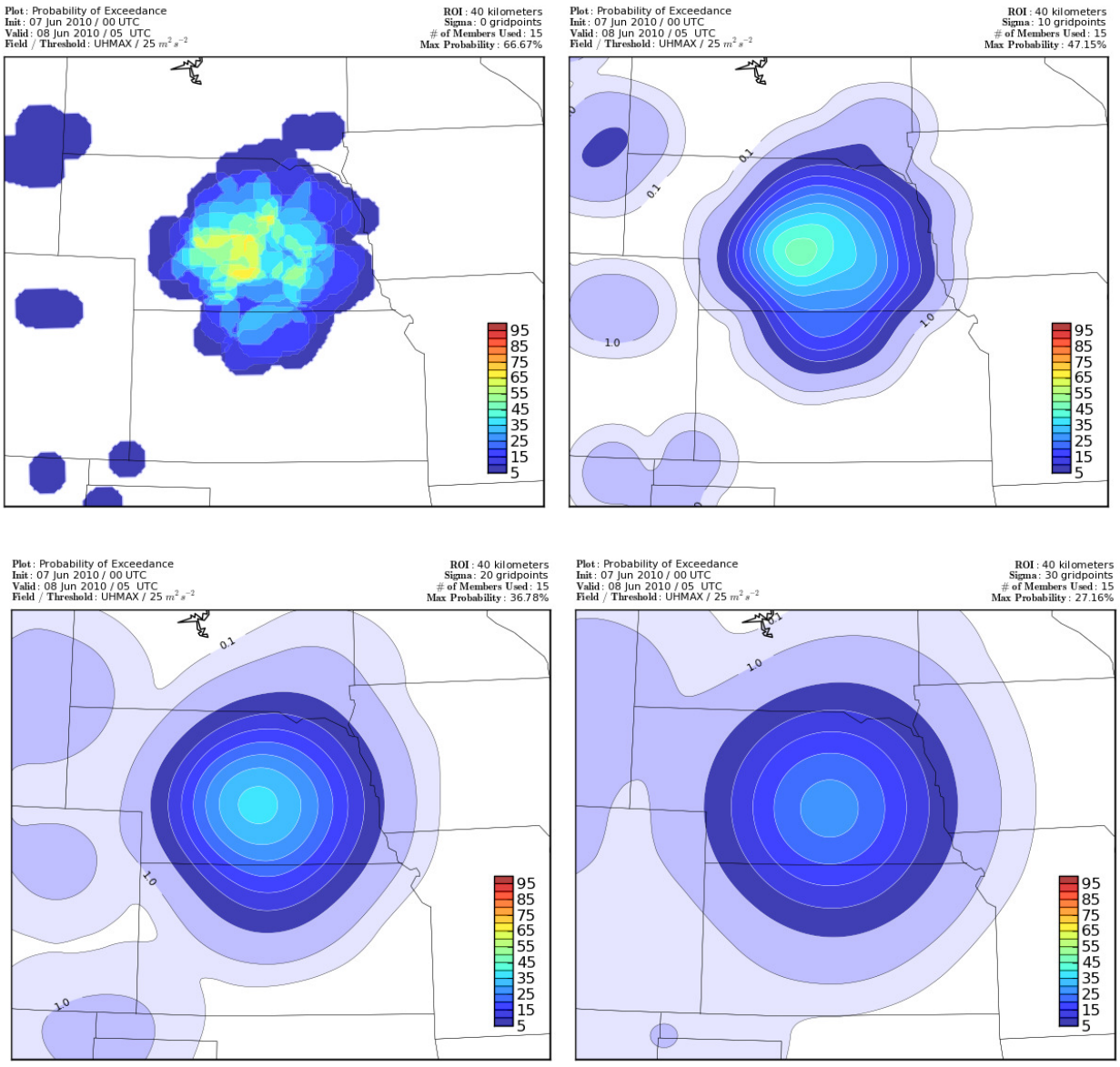


Figure 4. Sames as in figures 2 and 3, except the raidus of influence is changed to 40 kilometers, or 10 grid points.