P2.3 DEVELOPMENT OF A COMPREHENSIVE SEVERE WEATHER FORECAST VERIFICATION SYSTEM AT THE STORM PREDICTION CENTER

Andrew R. Dean*[1,2], Russell S. Schneider[2], and Joseph T. Schaefer[2]
[1]Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK
[2]NOAA/NWS Storm Prediction Center, Norman, OK

1. INTRODUCTION

A truly valuable and informative verification system does not simply shine a light on the past through simple comparison between forecasts and verifying events. It should also provide prompt, valuable feedback to forecasters, including links to synoptic/mesoscale patterns and environments that are related to forecast performance, so that improvements can be made through a continuous learning process. This manuscript will describe the current efforts and future plans of the Storm Prediction Center (SPC), which is responsible for issuing forecasts of severe convection over the continental United States, to evaluate forecasts and advance the constructive feedback between the verification and the forecaster. The database itself will be briefly described, followed by several examples of unique verification products that are currently under development.

2. THE DATABASE

While forecast verification has long been a priority for the SPC (e.g. Weiss et al. 1980), recent advances in computing speed and technology have allowed a much more comprehensive verification database to be developed. This database was built using the open-source relational database system PostgreSQL. The database currently contains a record of severe storm reports collected by the National Weather Service (NWS) starting in 1950, as well as multi-decade digital archives of SPC's convective watch and outlook products. The presence of historical reports and forecasts in the database allows for new verification measures to be built back in time as they are developed.

In addition, this project is utilizing data from soundings and grid-based analyses to develop a climatology of severe storm environments, in order to provide a meaningful context for the forecasts and reports. Currently, data from SPC's hourly mesoscale objective analysis (Bothwell et al. 2002) is being stored in the database, which allows for an estimation of the convective environment in which a report occurred or a forecast issued. This will provide a new dimension to the verification data, offering a link between verification statistics and the synoptic and mesoscale backgrounds that drive the forecast process.

3. SOME EXAMPLES

With the aid of the new verification database, forecasts and events can be classified according to time of year, time of day, geographical location, convective environment, or any combination of these in order to provide a specific context for the computed forecast skill. Some examples of this are described below; these examples are only a small subset of new verification products that are under development.

3.1 *Using Environmental Data in Verification*

One important feature of the new verification system is the ability to combine verification data with information on storm environments, as shown in Figures 1 and 2. Environmental data in these examples was generated by SPC's hourly objective analysis routine.

Figure 1 is a map of reports for the period 2004-2005 that were associated with a low CAPE (100 mb mean layer (ML) CAPE < 1000 J/kg), low shear (0-6 km AGL shear < 10 m/s) environment. Probability of detection (POD) in any watch and in a tornado watch for each report type is listed in the lower-left corner of the plot. As would be expected, strong or violent (F2-F5) tornadoes and significant hail (≥ 2 inch diameter) were extremely rare in such environments, with none of the former and only one of the latter reported during the two-year period. Severe wind reports (≥ 50 knots [25.7 m/s] or damage consistent with such wind speeds) were the dominant report type in this environment; many of these reports were likely the result of microbursts from short-lived pulse convection. PODs were low, since widespread organized severe convection is not expected in such environments.

In contrast, strong or violent tornadoes and significant hail were much more prevalent in high CAPE (ML CAPE ≥ 2000 J/kg), high shear (0-6km AGL shear ≥ 20 m/s) environments, as shown in Figure 2. PODs for the various types of reports are very high, with 92% of such reports during the period being captured in a watch. Most reports that occurred in a high-CAPE, high-shear environment were confined to the Plains and Mississippi Valley, whereas the low-CAPE, low-shear reports were generally spread around the country. Most of the significant severe reports that were observed in the eastern third of the county were associated with low-CAPE, high-shear environments (not shown, see Schneider et al. 2006).

3.2 *Report Clustering*

Clustering reports is another way to examine the spatial aspect of forecast verification. Traditionally, watch POD has been calculated using all reports, but reports that are isolated are not necessarily expected to be in a watch, since issuing watches when there is a threat for only very isolated severe convection will likely result in numerous false alarms. For example, according

---

[*] Corresponding author address: Andrew R. Dean, CIMMS, Univ. of Oklahoma, National Weather Center, Suite 2300, Norman, OK 73072-7268; e-mail: andy.dean@noaa.gov

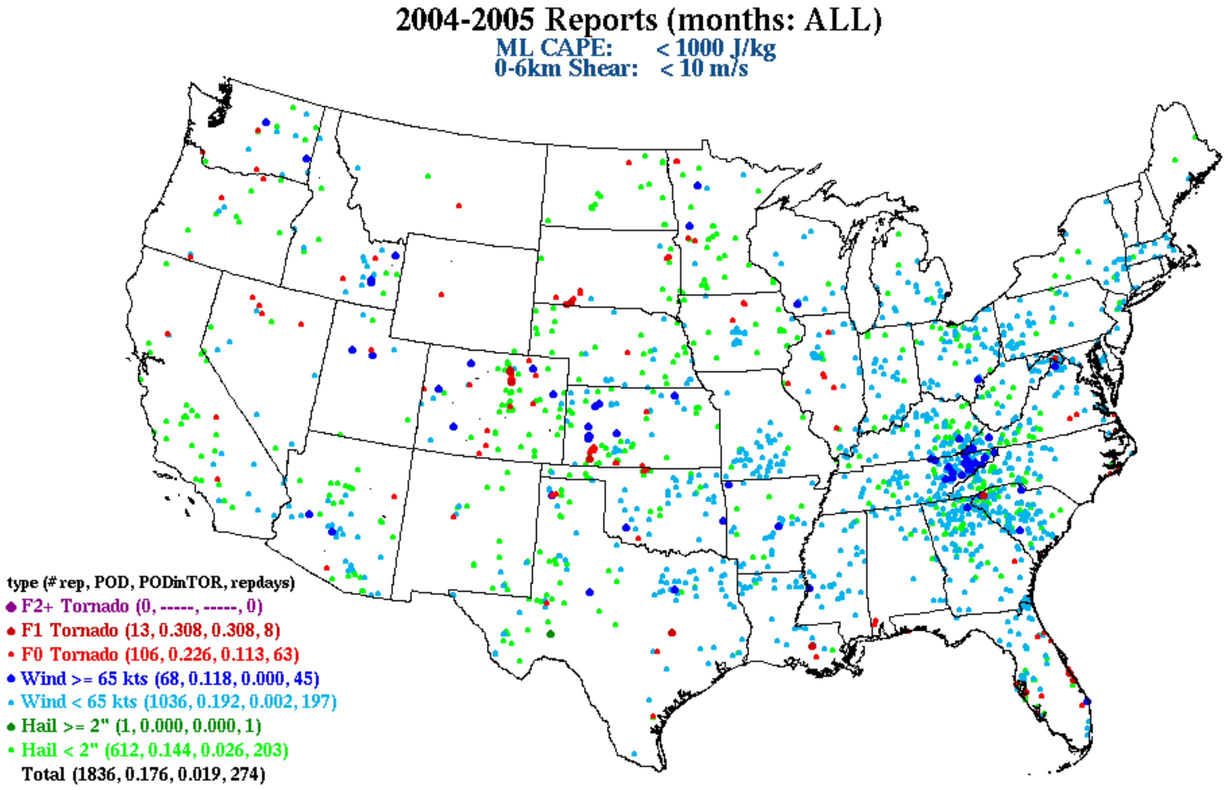

Figure 1: A map of all reports that occurred in low-CAPE, low-shear environments for the period 2004-2005. The legend includes data for the number of reports, POD in any watch, POD in tornado watches, and the number of report days for each report type. Environmental data is provided from SPC's surface objective analysis routine.

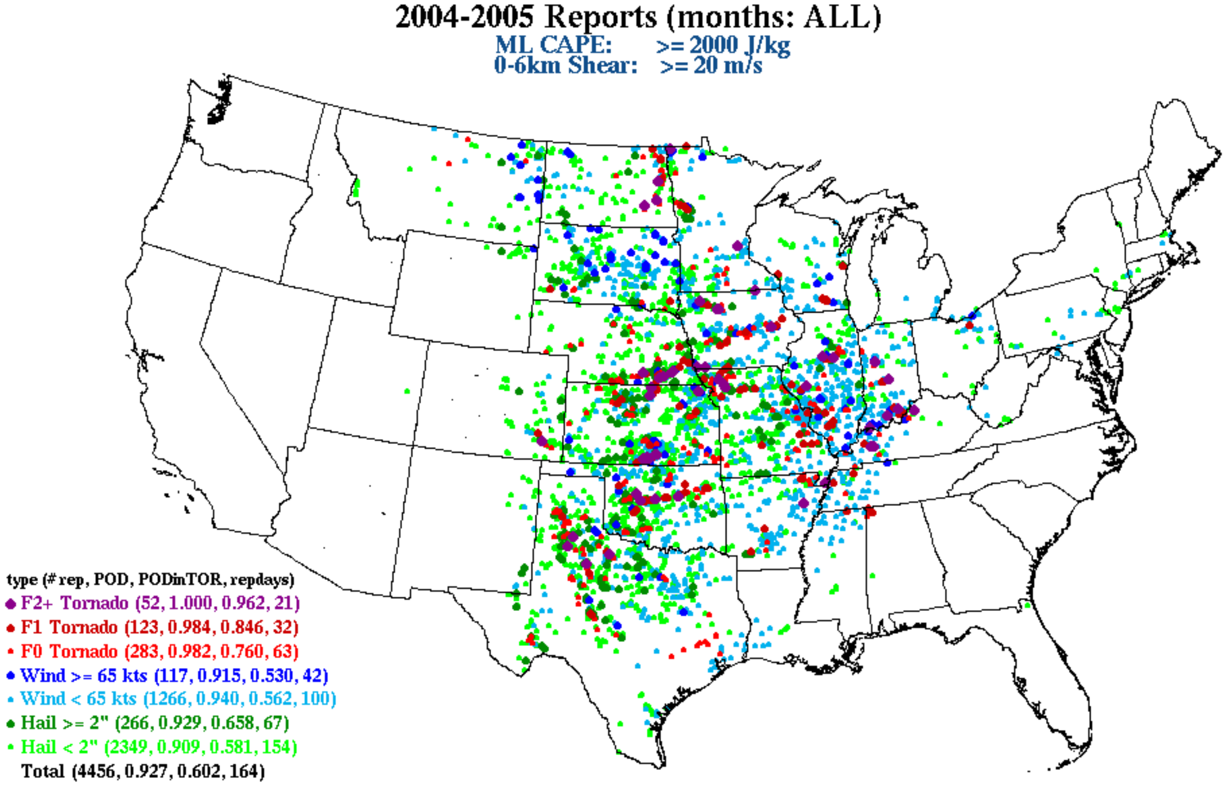

Figure 2: The same as figure 1, only for reports in high-CAPE, high-shear environments.

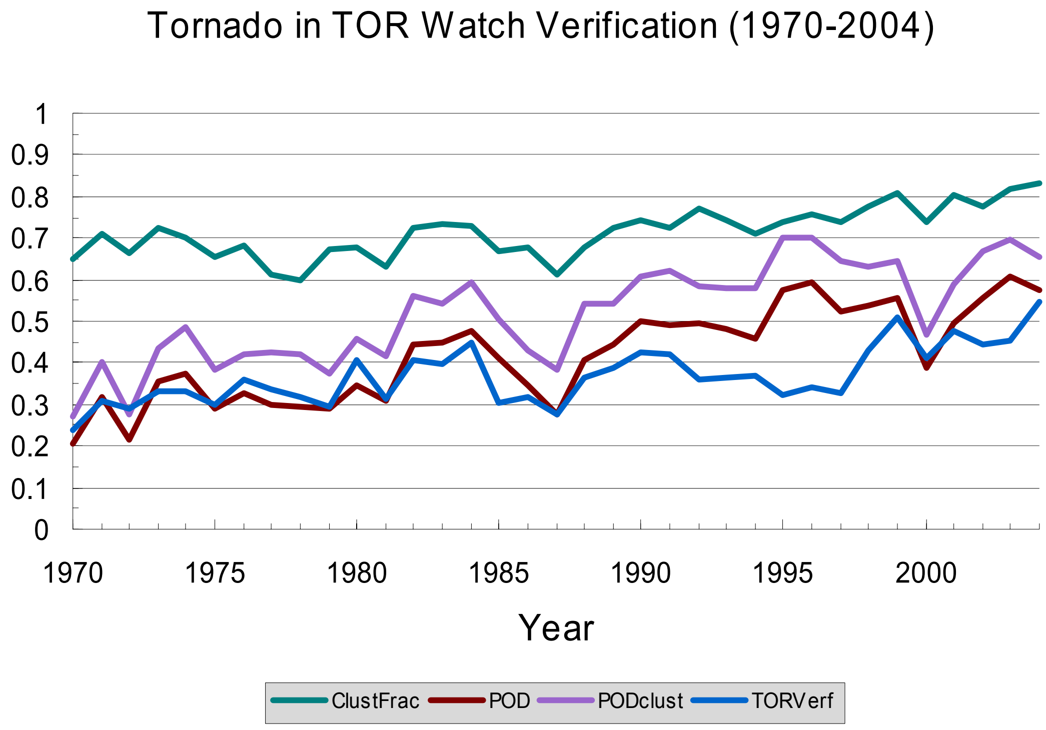

Figure 3: A time series of POD (tornado reports in tornado watches), watch verification, and clustering data for the period 1970-2004. 'ClustFrac' = fraction of tornado reports that clustered in groups of 2 or more, 'POD' = POD of all tornado reports, 'PODclust' = POD of only clustered tornado reports, TORverf = fraction of tornado watches that verified with at least 2 tornado reports.

to current National Weather Service directive, tornado watches should be issued when two or more tornadoes (or at least one F2-F5 tornado) are expected in the watch area. In order to determine which reports should ideally be in a watch, reports can be clustered using the minimum area (8,000 mi^2 [around 20720 km^2]) and duration (2 hours) of a watch as the clustering criteria (Dean and Schaefer 2006). A group of 2 or more tornadoes within an 8,000 mi^2 area that occur within 2 hours of each other are considered to be "clustered" according to the watch criteria.

Figure 3 shows a time series for the period 1970-2004 of overall POD for tornado reports in tornado watches, POD of clustered tornado reports, the fraction of tornado watches that verified with at least two tornado reports, and the fraction of tornado reports that clustered according to the criteria listed above. The POD of clustered tornado reports is consistently around 0.1 higher than the overall POD; in the last decade, between 60-70% of clustered tornado reports have occurred in tornado watches. Figure 3 also shows an increase in the fraction of clustered reports with time, likely reflecting the increase with time in the total number of tornado reports which has been observed (Weiss and Vescio 1998). Several notable yearly minima present in all of the verification measures are associated with minima in the fraction of clustered reports (particularly in 1981, 1987, and 2000), suggesting a link between report clustering and apparent predictability.

3.3 *Temporal Aspects of Verification*

Severe convection is not equally likely to occur at any time of day or year; consequently, it is essential to examine the relationship of forecast performance to the time and/or date of the event. For example, Figure 4 shows the distribution of tornado POD in watches by day of year for the period 1996-2005. The data are smoothed using a Gaussian kernel density estimator, following the method of Brooks et al. (2003).

PODs are highest in the cool season and spring, when dynamic, synoptically-evident events tend to dominate. During the transition from spring to summer, PODs decrease and the difference between overall POD and tornado watch POD increases, which reflects a higher incidence of tornadoes in severe thunderstorm watches. PODs are lowest in the summer months, when events tend to be weaker and less widespread, as indicated by the increase in the fraction of F0 tornadoes during this time.

Even though fall and winter events are not nearly as frequent as spring events (Figure 5), such events are equally likely to be detected in any watch and more likely to be detected in a tornado watch when they do occur. The cool season events are also more likely to be stronger tornadoes, as the fraction of F0 events is lowest during these months; this likely is a contributing factor to the high POD of these events.

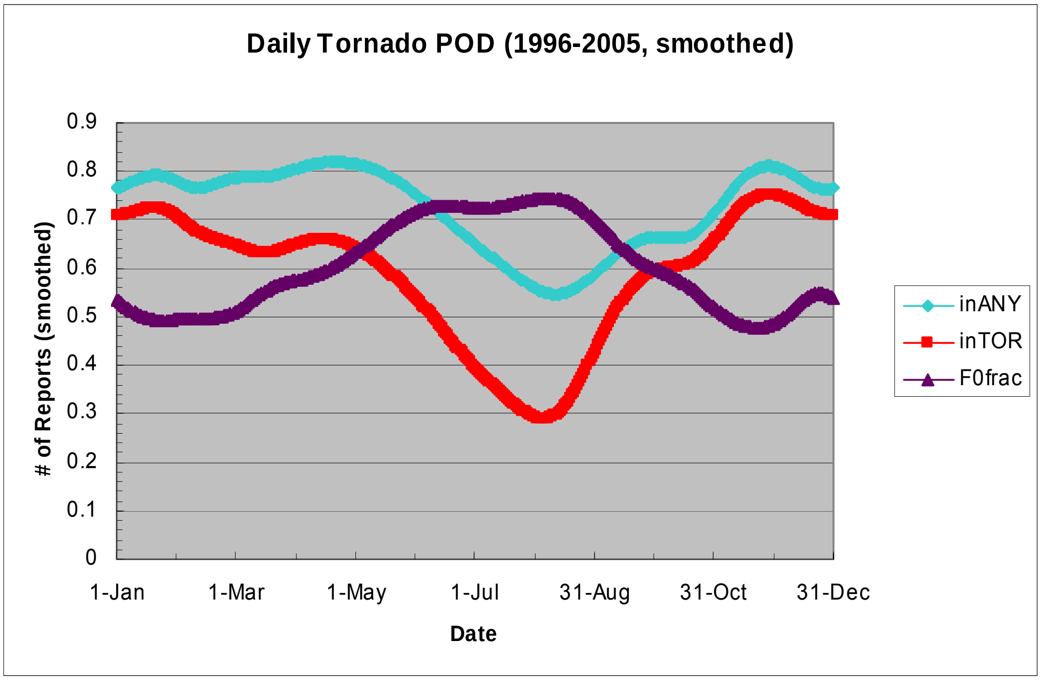

Figure 4: A time series of POD by day of year, for the period 1996-2005. Data are smoothed in time using a Gaussian kernel density estimator with smoothing parameter h = 15 days (Brooks et al., 2003). 'inANY' = POD of a tornado report in any watch, 'inTOR' = POD of a tornado report in a tornado watch, 'F0frac' = fraction of tornado reports that were rated F0.

4. CONCLUSION

As the verification database continues to evolve, forecasters will be able to slice through the data described above to determine the specific types of environments and cases where not only they do well, but also cases where they do not do as well. In the future, it is anticipated that radar and lightning data will be added to the verification database, providing even more context for the verification data. Through the development of a comprehensive integrated database containing severe weather forecasts and reports, remote sensing information about convective storm occurrence, and environmental information associated with each report, forecasters will gain new tools to better understand attributes of their decision-making processes, providing them with unique capabilities to improve severe weather forecasts.

Acknowledgements: We would like to thank Steve Weiss for his thoughtful review of this paper. Numerous discussions with SPC staff helped to clarify the ideas presented. Andrew R. Dean was funded under NOAA-University of Oklahoma Cooperative Agreement #NA17RJ1227, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of NOAA or the Department of Commerce.

5. REFERENCES

Bothwell, P. D., J. A. Hart, and R. L. Thompson, 2002: An integrated three-dimensional objective analysis scheme in use at the Storm Prediction Center. Preprints, *21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., J117-J120.

Brooks, H. E., C. A. Doswell III, and M. P. Kay, 2003: Climatological estimates of local daily tornado probability for the United States. *Wea. Forecasting*, **18**, 626-640.

Dean, A. R. and J. T. Schaefer, 2006: The skill of forecasting relatively isolated severe thunderstorm events. Preprints, *18th Conf. on Probability and Statistics in the Atmospheric Sciences*, Atlanta, GA, Amer. Meteor. Soc., CD-ROM (1.1).

Schneider, R. S., A. R. Dean, S. J. Weiss, and P. D. Bothwell, 2006: Analysis of estimated environments for 2004 and 2005 severe convective storm reports. Preprints, *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., CD-ROM (3.5).

Weiss, S. J., D. L. Kelly, and J. T. Schaefer, 1980: New objective verification techniques at the National Severe Storms Forecast Center. Preprints, *8th Conf. on Weather Analysis and Forecasting*, Denver, CO, Amer. Meteor. Soc., 412-419.

Weiss, S.J., and M. D. Vescio, 1998: Severe local storm climatology 1955-1996: Analysis of reporting trends and implications for NWS operations. Preprints, *19th Conf. on Severe Local Storms*, Minneapolis, MN, Amer. Meteor. Soc., 536-539.

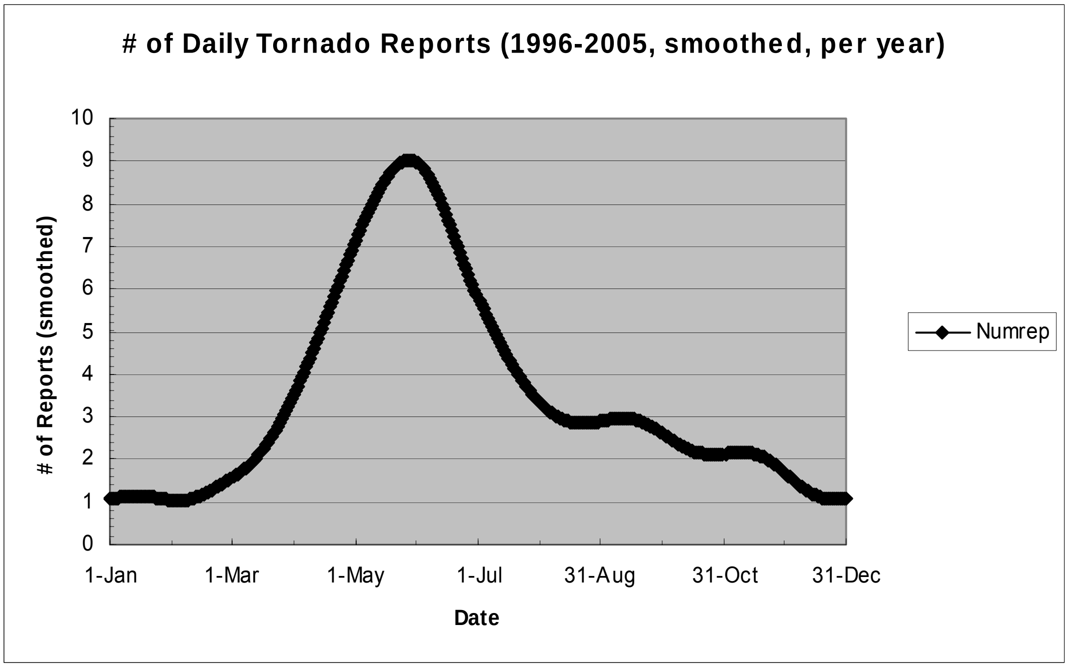

Figure 5: A time series of daily tornado counts for the period 1996-2005. Data are smoothed in time as in figure 4.