17.6 SEVERE WEATHER FORECASTS FROM AN ENSEMBLE OF HUMAN PERTURBED SIMULATIONS USING AN ADJOINT MODEL

Victor Homar[1], David J. Stensrud[1], Jason J. Levit[2]
[1]*NOAA National Severe Storms Laboratory, Norman, Oklahoma*
[2]*NOAA/NWS Storm Prediction Center, Norman, Oklahoma*

1. INTRODUCTION

The use of ensemble forecasting systems has become one of the most important recent advances in numerical weather prediction (Kalnay, 2003). Ensemble techniques provide an opportunity to explicitly incorporate the presence of uncertainties in observations and models, and to include them in the forecasting process. Identifying, quantifying and representing these uncertainties in the forecasting system has become a great challenge that is receiving primary attention in the weather research community (eg, Shapiro and Thorpe, 2004). Ensemble methodologies have notably benefited the skill of the medium-range forecasts in the synoptic scale, with the member generation strategies at this scale being basically focused upon the uncertainties in the initial conditions (ICs) and their evolution. Despite some debate about whether Monte Carlo perturbations or dynamically constrained methodologies are best for specifying a set of initial conditions (Hamill et al. 2000), the “breeding” of growing modes (Toth and Kalnay, 1993) and singular vector (Molteni et al. 1996) techniques are currently used at the major operational centers in United States and Europe. However, for applications in the mesoscale, methods to generate the perturbations needed for ensemble systems become more complex due to the larger and less known observational errors, more important model imperfections and users' sensitive dependence upon reliable forecasts. Xu et al. (2001, hereafter Xu01) propose a method to generate a short-range ensemble that would benefit from forecaster's guidance in identifying areas of forecast concern. They describe with detail the method to be used to generate such an ensemble, which involves running an adjoint model. They also present results of a test on a single case.

With the aim of assessing the value of short range numerical forecast ensembles to assist in the operational forecasting of threatening weather, the Storm Prediction Center and the National Severe Storms Laboratory conducted the 2003 Spring Program (SP03) experiment focused on the generation and interpretation of mesoscale short-range ensembles. Encouraged by the promising conclusions in Xu01, the SP03 included a sub-experiment aimed at testing their method for a longer period, using operational forecasters as drivers of the system. The underlying idea was to create a daily, customized ensemble to generate guidance for the severe weather to be forecast in the following 48 hours. Given the information obtained from an operational forecaster about the mesoscale-sized regions of potential severe weather, an adjoint model was used to compute the areas in the IC where adding perturbations would, in a linear sense, maximize the spread of the ensemble over those areas of concern and in directions deemed important by the forecaster. Essentially, ensemble dispersion was intended to be generated in specific areas and fields of interest as opposed to everywhere in the domain or following fast growing modes under global generic norms.

This study shows results from an evaluation of this ensemble during the experiment period, assessing the quality of the short-range probabilistic forecasts of severe weather events and precipitation. Comparison with the National Centers for Environmental Prediction (NCEP) operational deterministic Eta model and with the Short Range Ensemble Forecasting (SREF) system for the same period are provided.

Section 2 presents the model configuration and a description of the IC perturbation generation process. A description of the verification and comparison datasets used in the evaluation is presented in section 3. Section 4 shows the verification of severe weather events forecasts. Various skill scores of precipitation forecasts are shown and compared in section 5. Section 6 presents a summary and suggestions for future similar ensemble set-ups.

2. MODEL DESCRIPTION AND ENSEMBLE MEMBER GENERATION

The evaluated ensemble consists of 32 forecasts produced using the nonhydrostatic Pennsylvania State University-National Center for Atmospheric Research (PSU-NCAR) Fifth-generation Mesoscale Model (MM5V3) (Dudhia 1993, Grell et al. 1994). All simulations in the experiment are run with two domains interacting with a two way nesting strategy. The coarser

Corresponding author address: Victor Homar, NOAA National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069 ; e-mail: victor.homar@noaa.goc

Table 1. *Relative frequency of fields and forecast times selected by the forecasters and used to initialize the adjoint model integrations and IC perturbation generation.*

Field	**U**	**V**	**T**	**Q**	**Vort**	**Hgt**	**CAPE**	**SLP**	**W**
Freq. (%)	12.7	11.3	13.2	11.3	7.2	13.0	14.4	11.8	5.1

Fcst Time (h)	**+12**	**+18**	**+24**	**+30**	**+36**	**+42**	**+48**
Freq. (%)	0.5	0.9	11.3	26.4	47.0	8.3	5.6

domain has 66 × 46 grid points, 90 km grid spacing and covers the continental United States (CONUS), south Canada, the Gulf of Mexico, easter North Pacific and western North Atlantic. The inner domain has 157 × 97 grid points, 30 km grid spacing and covers basically the CONUS. All simulations are performed on 24 sigma levels, with higher concentration at lower altitudes to better resolve near-ground processes. Subgrid moist convection is parameterized using the Kain and Fritsch (1990) scheme. A simple microphysics scheme that allows for ice concentration at temperatures below freezing is used. Boundary layer processes are parameterized using the Eta PBL (Janjic, 1994) scheme together with the Dudhia (1996), five-layer simple soil model. Cloud and clear air radiative effects, as well as water vapor, carbon dioxide and ozone concentrations are considered in the radiative scheme. Both coarser and inner domains use the same parameterizations for all simulations.

This experimental ensemble of the SP03 experiment (MM5ADJ) ran from April 28 to June 6, but only from Monday-Friday of each week. The Eta model forecasts at 1200 UTC are used to provide the ICs for the MM5ADJ. Lateral boundary conditions for the MM5ADJ also come from the 1200 UTC Eta model. To generate the set of different IC, the method detailed in Xu01 was followed: each day a human forecaster was asked to identify 16 features of interest in the control run such that severe weather for the following day (day 2) was, in the forecaster's opinion, most sensitive to the feature. The forecaster was able to select structures at any time (in 6 h intervals) from the Eta forecast, on the following fields: horizontal and vertical wind components, temperature, specific humidity, geopotential height, sea level pressure, vorticity and pseudo-Convective Available Potential Energy (pseudo-CAPE, defined as temperature difference between middle an low levels). The fields in this list were predetermined to allow easy operational implementation but there is no other limitation on the feature of interest but to be differentiable with respect to the model fields. Table 1 shows the distribution of fields used by the forecasters during the whole experiment. There is not a clearly preferred subset of fields by the forecasters, with perhaps less tendency to use the vertical velocity or vorticity fields. This may be related to the traditional fields involved in mesoscale conceptual models used to identify the areas of concern and also the confidence of the forecaster on the numerical forecast of vertical velocity. The distribution in forecast hours (Table 1) shows that almost 3 out of 4 times the forecaster defined the feature of interest at T+30h or T+36h, the time of the afternoon convection for day 2 forecast. A total of 27 cases are finally available for this evaluating the ensemble guidance potential.

For each of the 16 selected features of interest, an adjoint model integration on the coarser domain was correspondingly initialized and the sensitive areas of each forecaster-specified feature to the IC were derived. The adjoint model used is the MM5 Adjoint Modeling System (Zou et al., 1997, 1998) developed by NCAR. The code is derived from a simplified version of the forward MM5. The adjoint runs have no parameterized convection but include explicit microphysics, radiation scheme and surface processes. Once the sensitivity fields were obtained from the adjoint, the horizontal wind fields and temperature sensitivities were rescaled to an amplitude within an estimation of the typical analysis error on the mesoscale. Table 2 shows the maximum perturbation amplitude used for each field. Then, two MM5 simulations were run for each highlighted feature, each one perturbed in opposite direction (positive and negative). Since the forecaster was requested to highlight 16 features each day, 32 perturbed simulations were produced to form the MM5ADJ ensemble.

Table 2. *Maximum and typical amplitude of the IC perturbations.*

Field	**U & V**	**T**
Max.	8.0 m s^{-1}	4.0 K
Typical	2 m s^{-1}	1.5 K

3. VERIFICATION AND COMPARISON DATA

Ensemble forecasts are expected to provide information on the uncertainty of the forecast aspect of interest. We evaluate here the ability of the MM5ADJ ensemble to forecast probabilities of severe weather and precipitation. All verification and comparison fields used in this study are remapped to the MM5 30 km domain, in order to facilitate and ensure a fair comparison among them. Two observational datasets are used in this verification:

a) *SPC severe weather reports*: The severe weather probabilistic forecasts are verified using the SPC severe weather reports database. This database contains a quality controlled list of tornado, large hail and wind damage reports in the United States with information about the intensity of the event and its location in space and time. Figure 1b shows an example of the reports in the SPC database for the same period shown in Fig. 1a. A gridded field on the MM5 domain is created by setting the grid points with at least one report in its grid box to the value of 1. This field does not contain information about the type, intensity or density of reports within the grid box but it is consistent with the probabilistic forecast that it verifies. Here, the models' forecasts strictly refer to the occurrence of severe weather within a gridbox rather than to the type, intensity or density of events (see next section).

b) *NCEP/CPC Stage IV precipitation*: In order to verify the precipitation forecasts we use the NCEP/CPC 4km Stage IV data every 6 h. This dataset is based on a multi-sensor hourly analysis and it is quality controlled manually. The precipitation remapping is performed maintaining the original amount of precipitation in the MM5 grid as done by NCEP for grid interpolation and QPF verification (Mesinger, 1996).

In addition to the objective verification against the observational datasets, we assess the relative value of the MM5ADJ by comparing it to other available short-range forecasts for the same period:

a) *Subjective Day2 Outlooks*: During the SP03 a forecaster issued a severe weather outlook for Day2, following the same guidelines used for the daily operational SPC outlooks (Kay and Brooks, 2000). Note that the operational SPC outlooks are theoretically designed to forecast the probability of severe weather within 20 miles of a point, which is equivalent to an square area with 57 km on each side. Admittedly, remapping the probabilities to a 30 km grid produces a shift towards overforecasting (increasing the false alarms). However, it is uncertain how sensitive the forecaster is to this definition when defining the outlook. On the other hand, the SPC outlooks are issued using 5 discrete probability categories: 0.00, 0.05, 0.15, 0.25 and 0.35. Again, in order to ensure fair comparison among the forecasts, we will project the models' forecasts to the SPC values. For instance, all probabilities above 0.35 from the models are considered 0.35 in the verification.

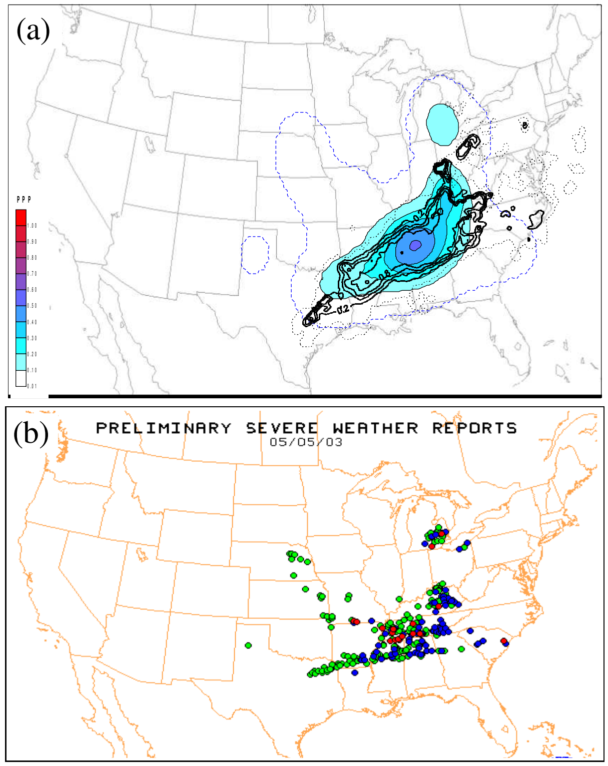

Fig. 1 *a) Probability of severe weather (black lines in 20 % interval) using the SCP parameter from the 32 member MM5 Adjoint ensemble for the 24 h period from 1200 UTC May 5 to 1200 UTC May 6 2003. Shaded field shows the Practically Perfect Prog. Dashed lines depict the 1% probabilities. b) Storm reports from the same period. Red circles represent tornadoes, green circles represent large hail and blue circles represent damaging winds.*

b) *Operational Eta*: We include the 1200 UTC run from the NCEP Eta to add a deterministic model in the comparison. Probabilistic forecasts from this model are trivially created by setting the field to 1 (actually, 0.35 for the severe weather forecasts) when the condition for severe weather is satisfied and 0 otherwise.

c) *Short Range Ensemble Forecasting (SREF)*: The NCEP ensemble for short range forecasting (e.g. Hamill and Colucci 1997) at that time consisted of 10 members, five Eta and five RSM (Regional Spectral Model). The SREF forecasts would have provided a good opportunity to compare the experimental MM5ADJ ensemble, that uses human perturbed IC, against the dynamical method of breeding of growing modes used in the SREF. Unfortunately, only 11 days are available to perform the comparison for the period that the SP03 lasted. This hampers the statistical significance of such comparison. As a solution of compromise we will perform the verification process twice, one using the 27 cases where the MM5ADJ is available and another with the 11 days available from the SREF forecasts.

d) *Practically Perfect Prog*: Although this field is not a forecast, we will use it as a measure of the upper limit of a probabilistic forecast provided some limits in generating the forecast (e.g. smoothness, size and spatial continuity of significant probability areas). Brooks et al. (1997) discuss the concept of the practically perfect progs (PPP) and present the main characteristics. Essentially, the PPP field is constructed by using a nonparametric density estimation with a two-dimensional Gaussian kernel for each gridpoint with a report in the observational dataset. The parameters that define the kernel are calculated by Brooks et al. (1997) from the statistical properties of the climatology of SPC operational outlooks. Figure 1a depicts the PPP field obtained from the reports shown in Fig 1b and provides an example showing the main characteristic of the PPP - the hypothetical forecast is as accurate as could be expected for a forecaster already aware of the reports.

4. VERIFICATION OF SEVERE WEATHER FORECASTS

To diagnose "severe weather" from model output fields, we use the Supercell Composite Parameter (SCP, Thompson et al. 2002). The SCP is calculated as:

$$SCP=\frac{muCAPE}{1000\,J\,kg^{-1}}\times\frac{0-3\,km\,SRH}{100\,m^{2}\,s^{-2}}\times\frac{BRN\ shear}{40\,m^{2}\,s^{-2}}$$

The probability of occurrence of severe weather during a 24 h period at every grid point is defined here as the number of ensemble members having a SCP larger than 1 and some convective rain at that grid point anytime during that 24 h period, divided by the number of ensemble members. We use the threshold of SCP larger than 1 as it is reported by Thompson et al. (2003) in order to differentiate supercell storms in both observed and Rapid Update Cycle-2 analysis/forecast model proximity soundings. Each day, two probabilities for the forecast of occurrence of severe weather are produced for each of the four available forecasts sources in the comparison: one for the 0 – 24 h and another for the 24 – 48 h period. Figure 1a shows an example of the distribution of this probabilistic forecast from the MM5ADJ, showing for this case values up to 80% from the lower Mississippi Valley to the Ohio and Tennessee Valley.

A global verification of the probabilistic forecasts is done by using the attributes diagram (Wilks, 1995). This diagram shows the observed frequency of an event as a function of the forecast and allows an interpretation of skill for each forecast category separately. The attributes diagram also allows one to interpret the reliability, resolution and uncertainty of each forecast interval (Wilks, 1995). Figure 2 shows the attributes diagram for all the forecasts compared in this study. The observed climatological frequency is 0.016 for the 27 cases and 0.024 for the 11 "SREF cases" (the days were SREF data are available are mostly during the first two weeks of May 2003, which was the most active period of the SP03 experiment). Not surprisingly for the prediction of rare events, all forecasts in the comparison show good skill at predicting no occurrence of severe events (0.00 probs), with the human outlooks showing the higher reliability at this category. For *low* (0.05) and *moderate* (0.15) probabilities, the MM5ADJ is the only forecast showing

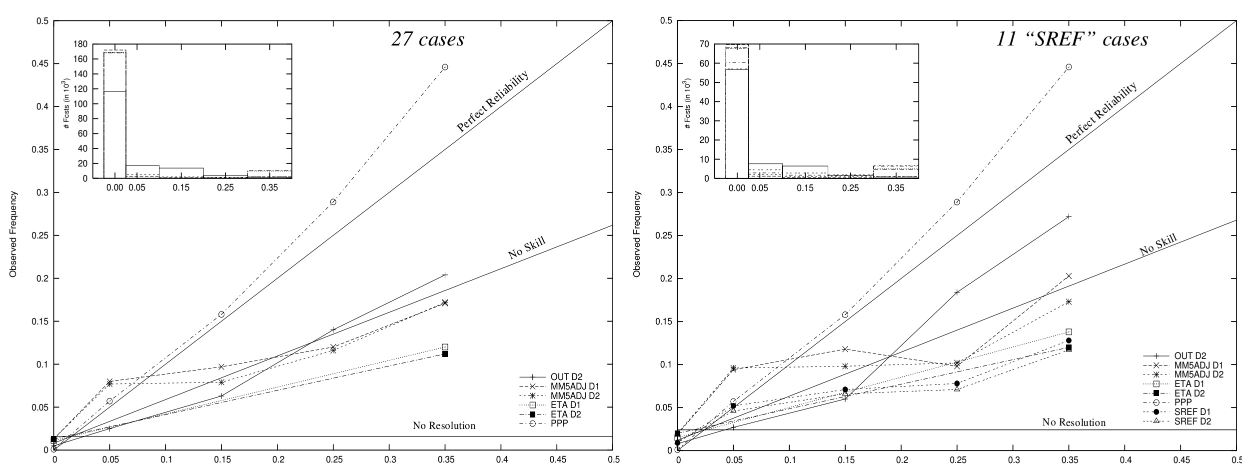

Fig 2. *Attributes diagrams for the probability of severe weather as obtained from: SP03 preliminary Day2 outlooks, T+24 h and T+48 h SCP forecasts from the MM5ADJ, Eta and SREF. Left panel was calculated using all 27 days available for the MM5ADJ experimental period and right panel contains results only from the days the SREF output were available during that test period.*

some skill, with especially good reliability at the *low* category. The fact that the MM5ADJ ensemble is underforecasting for both day1 and day2 in the *low* category may indicate that a more adequate threshold for the SCP parameter may be needed for this system rather than the one used (SCP>1). For high probabilities (when a majority of the ensembles agree), the MM5ADJ shows no skill in predicting severe weather. This is an indirect indication of underdispersion in the ensemble since it is a result of high ensemble member agreement yet not reproducing the observations. The human outlooks, however, become skillful at the high probability categories, revealing the forecasters skill when they show high confidence on the intensity of the situation of the day and decide to use high probabilities in the outlook. The Eta forecasts are clearly hampered in this type of probabilistic verification but it is still significant that the model overforecasts the severe weather. In only 11 to 12 % of the times in which the model forecasts SCP > 1 and convective rain in a gridpoint is severe weather actually reported in the gridbox during the 24 h period.

Regarding the results from the 11 “SREF cases”, similar scores to those from the 27 days sample are obtained for the outlooks, MM5ADJ and Eta model, with some minor but notable differences. The outlooks show a remarkable increase in skill and reliability for the high probability categories during this convectively active period. Also, the MM5ADJ has generally improved skill at all categories. Focusing on the SREF results, it shows almost perfect reliability (better than the PPP for day1) for the *low* category but no skill for higher probabilities. This result clearly shows the advantage of the MM5ADJ over the SREF in forecasting highly probable episodes of severe weather (usually the most intense and damaging). This is most likely a consequence of the customized design of the MM5ADJ to focus on the areas of severe weather threat, whereas the SREF system is designed to cover a wide variety of mesoscale forecast aspects and shows its strength at the lower probability range.

As expected, and although the SREF is forecasting the *low* category with great success for this small sample size, all forecasts considered in this comparison are far from the potentially attainable limit set by the PPP field.

5. VERIFICATION OF PRECIPITATION FORECASTS

The 6-hourly accumulated precipitation forecasts from the MM5ADJ, Eta model and SREF are verified using the NCEP/CPC Stage IV dataset. Although probabilistic fields can be derived from an ensemble of precipitation forecasts, we will show here only results from the

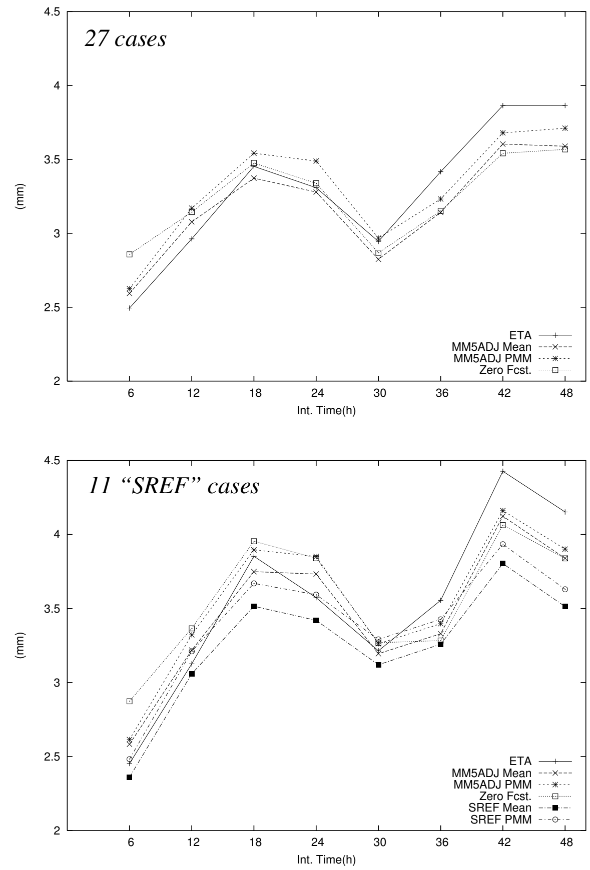

Fig. 3 *Root mean square error of the 6-hourly accumulated precipitation. Top panel include the 27 cases of the SP03 experiment and the bottom panel includes the 11 available cases from the SREF during that period.*

verification of the skill of single precipitation fields from the ensembles such as the mean and the probability mean matched (PMM) precipitation. The PMM (Ebert, 2001) is calculated as the ensemble mean, rescaled locally using the global (all ensemble members) distribution of precipitation. This field is intended to possess the representativity and properties (in terms of smoothing out uncertain features from the individual members) of the ensemble mean but a similar precipitation distribution to the precipitation field of an individual deterministic member.

The root mean square error (RMSE) of the 6-h forecast precipitation is shown in Fig. 3 as a function of the integration time. A measure of the mean total precipitation is given through the labeled “Zero Fcst”. The diurnal cycle in the precipitation and the RMSE clearly emerges in the plots. Maximum precipitation is observed at T+18 and T+24 h, which corresponds to the period from 00 – 12 UTC. Using the sample of 27 cases (Fig. 3, top panel), the Eta model has the lowest RMSE for the first 12 hours but it becomes the forecast

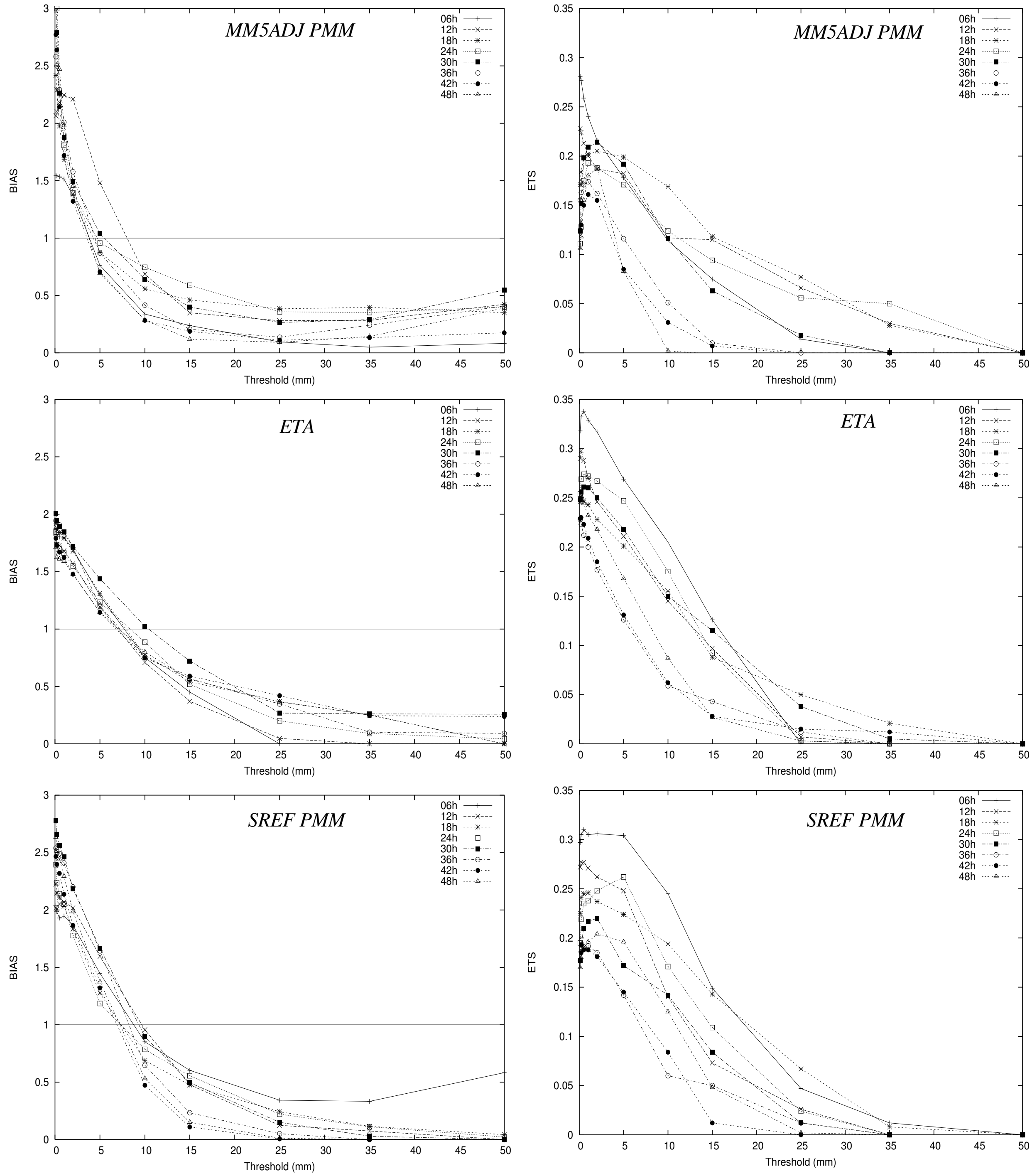

Fig. 4 *Six hours accumulated precipitation BIAS and ETS as a function of precipitation threshold for the 11 "SREF cases". BIAS plots show the reference line of BIAS=1.*

with highest error during the second 24 h of the period. The ensemble mean of the MM5ADJ consistently produces relatively low RMSE during all 48 h. However, results for the MM5AJD PMM show higher RMSE that the ensemble mean for all time spans. This is not an unexpected result since the RMSE score strongly penalizes small-scale structures in the field, and the PMM is designed to have enhanced small scale features compared with the ensemble mean. Note how the lowest RMSE for the last hours of the forecast period is obtained by a constant precipitation of zero ("Zero Fcst"), indicating almost no value for any of the forecasts at those time spans.

Restricting the sample size to the 11 "SREF cases" (Fig. 3, bottom panel) the RMSE increase generally due to the higher average convective activity compared to the SP03 27 cases. Results for the SREF forecasts show that the SREF mean has the lowest RMSE among the compared forecasts for all time spans.

To provide a global skill score for the precipitation forecast not so sensitive to errors on a single day, we compute the BIAS and Equitable Threat Score (ETS), computed as:

$$BIAS = \frac{F}{O} \qquad ETS = \frac{C-E}{F+O-C-E}$$

where, F is the number of forecast points above a threshold, O the number of observed points above a threshold, C the number of points with both forecast and observations above a threshold, T the total number of gridpoints in the forecast and E=F·O/T. For the sake of brevity, we show BIAS and ETS results for the 11 "SREF cases" and only the PMM is used for the MM5ADJ and SREF ensembles. This field shows better BIAS and ETS scores than the ensemble mean (not shown).

All forecasts overpredict (BIAS > 1) precipitation amounts lower than 5 mm, though the Eta model has the lowest bias. As expected, for larger precipitation amounts, all models underpredict (BIAS < 1) precipitation, decaying to small bias with increasing threshold. However, the Eta model has smaller bias for the 10 and 15 mm thresholds. It is noteworthy that the MM5ADJ has larger biases for small precipitation amounts but does not decay to 0 as the other two systems do for very high thresholds (> 25 mm /6h).

Regarding the ETS, similar behavior is detected in the results. The Eta and especially the SREF obtain better scores for low precipitation thresholds, whereas the MM5ADJ forecasts are better for the highest precipitation thresholds. This consistent result in the precipitation verification may be a consequence of the particular design of the MM5ADJ ensemble, focused on severe weather and active situations where higher precipitation rates are more likely, but it may also have influence from the fact that the MM5ADJ uses the Kain and Fritsch (1990) convective scheme whereas the Eta and SREF use the Betts-Miller (1986) convective adjustment parameterization.

6. CONCLUSIONS

This paper presents some verification results of an ensemble created using human-generated input and an adjoint model. Following the ensemble generation process described in Xu01, an experiment was conducted during the 2003 SPC/NSSL Spring Program. Results from that experiment are verified against severe weather reports and 6-hourly accumulated precipitation. In order to assess the value of the experimental ensemble we compare the verification scores to other available short-range operational forecasts. The SPC Day2 convective outlooks, the Eta model and the SREF system are included in the comparison. Unfortunately, the SREF dataset is incomplete for the period of the experiment and no satisfactory comparison can be carried out. However, from the small sample of 11 cases some differential characteristics emerge between the SREF and the experimental MM5ADJ system.

The SREF shows almost perfect probabilistic severe weather forecasts for low probability events but no skill for high probability categories (> 0.06). The MM5ADJ, however, shows some skill at most of the probability categories, being only overcome by the human outlook at the high probability end. This suggests that human-generated ensembles could be a useful addition to the model guidance provided for severe weather situations.

The comparison of precipitation skill scores confirms the low biases and high ETS for small precipitation amounts of the Eta model, and highlights the better skill of the SREF for low intensity events (precipitations lower than 25 mm/6h) and the relatively good results of the MM5ADJ forecasts for high precipitation thresholds.

Future work will involve the verification of probabilistic forecasts of precipitation and a more detailed analysis of the suggestion that emerges from this study indicating the encouraging skill of the human-generated ensemble system on very active events (high probability of severe weather and heavy amounts of precipitation). Limiting the verification area to areas of concern in the forecast could help to further understand these results.

Acknowledgments: The authors would like to thank Dr. David Bright for the helpful discussions and providing the available SREF data, Dr. Mike Baldwin for his assistance with the GRIB conversion and remapping process, Dr. Harold Brooks for the discussions on the Practically perfect prog field. The adjoint calculations were performed on the Linux supercomputer located at the Oklahoma Supercomputer Center for Education and Research (OSCER) at the University of Oklahoma. This work was performed while the leading author held a National Research Council Research Associateship Award at the NOAA/National Severe Storms Laboratory.

REFERENCES

Betts, A. K., and M. T. Miller, 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.,* **112**, 677–691.

Brooks, H. E., M. P. Kay, and J. A. Hart, 1998:

Objective limits on forecasting skill of rare events. *Preprints*, 19th Conf. on Severe Local Storms. Minneapolis, MN, Amer. Meteor. Soc., 552-555.

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

Dudhia, J. 1993: A nonhydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and the simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.,* **121**, 1493–1513.

Jimy Dudhia, 1996: A multi-layer soil temperature model for MM5, Preprint from the Sixth PSU/NCAR Mesoscale Model Users' Workshop.

Ebert, E. E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation *Mon. Wea. Rev.*, **129,** 2461-2480

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Tech. Note NCAR/TN-398+STR, National Center for Atmospheric Research, Boulder, CO, 122 pp.

Hamill, T. M., Snyder, C., Morss, R. E. 2000: A Comparison of Probabilistic Forecasts from Bred, Singular-Vector, and Perturbed Observation Ensembles
Mon. Wea. Rev. **128,** 1835-1851.

Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta–RSM short-range ensemble forecasts. *Mon. Wea. Rev.,* **125**, 1312–1327.

Janjic, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.,* **122**, 927–945.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47**, 2784–2802.

Kalnay, E., 2003: Atmospheric Data Assimilation and Predictability. *Cambridge University Press*. 342 pp.

Kay, M. P., and H. E. Brooks, 2000: Verification of probabilistic severe storm forecasts at the SPC. *Preprints*, 20th Conference on Severe Local Storms, Orlando, Florida, American Meteorological Society, 285-288.

Mesinger, F., 1996: Improvements in quantitative precipitation forecasts with the Eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.,* **77**, 2637–2649.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Royal Meteor. Soc.,* **122,** 73–119.

Shapiro, M. A. and Thorpe, A. J. 2004: THORPEX International Science Plan. World Meteorological Organization. 38 pp.

Thompson, R. L., Edwards,R., Hart, J. A., Elmore,K. L., Markowski, P. 2003: Close Proximity Soundings within Supercell Environments Obtained from the Rapid Update Cycle. Wea. Forecasting, **18**, 1243-1261

Thompson, R. L., R. Edwards, and J. A. Hart, 2002: Evaluation and interpretation of the supercell composite and significant tornado parameters at the Storm Prediction Center. Preprints, *21st Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., J11–J14.

Toth Z. and Kalnay E., 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Xu, M., Stensrud, D. J., Bao, J-W, Warner,T. T. 2001: Applications of the Adjoint Technique to Short-Range Ensemble Forecasting of Mesoscale Convective Systems *Mon. Wea. Rev.,* **129**, 1395-1418.

Zou, X., Vandenberghe, F. Pondeca, M., and Kuo, Y-H., 1997: Introduction to adjoint techniques and the MM5 adjoint modeling system, NCAR Tech Note NCAR/TN-435+IA.

Zou, X., Huang, W., and Xiao, Q., 1998: A User's Guide to the MM5 Adjoint Modeling System, NCAR Tech. Note NCAR/TN-437+IA.