Application of the NCEP/EMC Short-Range Ensemble Forecast System (SREF) to Predicting Extreme Weather Events

M. Steven Tracton and Jun Du – NOAA/NWS/NCEP

1. Introduction and background

The National Centers for Environmental Prediction (NCEP) now runs operationally twice per day 11-member ensembles with its Global Medium Range Forecast model for medium-range (3-14 days) predictions (MREF). The fundamental principles and concepts of ensemble forecasting for MREF apply equally well to short-range, regional model based ensembles. However, medium and short-range forecast problems are different with regard to time and spatial scales of interest. In medium-range forecasting the phenomena of interest are the larger-scale, longer-lived circulation systems and associated weather, whereas at short ranges focus is on the mesoscale details of weather systems and associated sensible weather. Also, in Short Range Ensemble Forecasting (SREF) error is generally more important in forecast uncertainty (ensemble spread) than in MREF (e.g., with regard to physics parameterizations, land/surface interactions), and lateral boundary conditions are critical. Beyond these scientific issues, practical concerns relevant to an operational setting must be considered in development and display of MREF and SREF "user friendly" products and in their respective applications. Together, MREF and SREF are viewed as integral in a "seamless suite" of products that enable estimates in the forecast confidence of specific weather threats, first, in the context of the requisite larger-scale circulation pattern at longer ranges and, then, in the details of the relevant weather system and associated sensible weather in the short range.

To begin exploring the challenges and prospects of SREF, NCEP launched a pilot project in 1996 along the lines suggested by the first SREF Workshop (*Brooks et al.*, 1995). A 15-member ensemble was generated about once per week for an extended period using 80 km versions of the Eta model (*Black*, 1994) and the Regional Spectral Model (RSM) (*Juang and Kanamitsu*, 1994). Perturbations were provided by interpolation of " bred" initial conditions from the operational global ensemble system (*Toth, et al.*, 1997) and a variety of in-house analyses. The encouraging results from this study have been documented by *Stensrud et al.* (1999) and *Hamill and Colucci* (1997, 1998). Later, the pilot study was augmented by including 10 additional members, 5 runs each from the Eta and RSM initialized with regional "enhancement" of the global bred perturbations (*Tracton and Du*, 1998). Additionally, selected cases were rerun at higher resolution (40 and 48 km for RSM and Eta, respectively). Among the key findings, based on statistical verification and case study analysis, were that enhanced diversity of solutions (spread) was obtained with a multi-model ensemble, higher resolution, and the regional enhancement. More generally, the study illustrated the significance of uncertainties in short-range regional-model predictions, demonstrated the potential of SREF to provide operationally useful information, and provided a basis for a prototype operational system at NCEP.

That prototype consisted of the 10 Eta plus RSM members referred to above. That is, RSM and Eta control (unperturbed) forecasts plus runs of each model with two pairs of regionally enhanced bred modes from the global ensemble system. Lateral boundary conditions were provided by the control and respective perturbations from the global ensembles. This system was tested with 32 km resolution in the May, 1998 near real-time, multi-institutional experiment referred to as the Storm and Mesoscale Ensemble Experiment (SAMEX). The general aspects of the experiment and preliminary results have been documented by *Hou et al* (1999). A major conclusion, consistent with the pilot study referred to above, was that an "ensemble of multiple systems is close to optimal, probably because it represents most realistically the current uncertainties in the models and in the initial conditions". The second major result cited was that "perturbations in the physics, and lateral boundary conditions consistent with perturbations in initial conditions are both important for regional ensemble forecasting". From our own evaluation of the NCEP component of the complete SAMEX ensemble, a key finding was the importance of domain size in regard to the influence of lateral boundary conditions. In order to accommodate the experiment prescribed horizontal resolution (about 30 km), the domain was constrained to a relatively small area over the continental U.S. Only after evaluation and complementary experiments did it become apparent that the influence of boundaries was distinctly detrimental to the growth of spread in the interior of the domain (*Du and Tracton*, 1999).

On the basis of the experiences above, NCEP/EMC actively engaged in constructing an operational SREF capability intended to run on its new IBM SP Class VIII supercomputer. The principal motivation has been and remains to provide guidance of a probabilistic nature on regional and relatively short-time scale weather events responsible for severe weather and heavy precipitation. It is widely agreed that ensemble based probabilities and/or measures of confidence hold the best potential for enhancing the ability to make user dependent informed decisions. The U.S. National Weather Service (NWS) now recognizes this fact by including in its "Vision 2005, the NWS Strategic Plan" committing the NWS to "provide weather, water, and climate forecasts in probabilistic terms by 2005".

2. Status

The current "operational" SREF configuration (beginning June 2001) consists of 10 members running from both 09 GMT and 21 GMT to 63 hours. In the context of the current NCEP IBM computer suite, this is the only approach that enables completion of runs about the same time as the operational Eta model runs (12 km) from 12 and 00 GMT. The SREF system consists of a control Eta model forecast plus 4 Eta perturbed runs and a control Regional Spectral Model (RSM) run plus 4 perturbed RSM forecasts. All forecasts are with 48 km resolution versions of the models over the operational Eta domain. Lateral boundary conditions are provided by the 9-hour forecasts from the respective members of the global model ensemble. Initial state perturbations are generated by independent "breeding" cycles, as for the NCEP global ensemble system, but in the context of the respective regional models.

SREF is currently run in what's referred to as a Real Time Test and Evaluation (RTT&E) mode leading (hopefully) to official NWS sanction as operational by spring, 2002. The RTT&E designation means SREF is running in the production side of the NCEP IBM with NCEP Central Operation production standards and support. However, products are not yet sanctioned nor distributed by the NWS as operational. They are available beyond NCEP on the SREF website (http://lnx48.wwb.noaa.gov/SREF/SREF.html) and in GRIB format from an NCEP server (Additional SREF products and displays are available at: http://eyewall.met.psu.edu/SREF/index.html.). SREF based products include maps from of ensemble mean and spread, "spaghetti" diagrams showing forecast diversity, and probability charts for precipitation, stability indices, etc. Recent additions include the ensemble based predominant precipitation type (rain, snow, freezing rain) and relative likelihood of each using the Baldwin et al. (1994) precipitation type algorithm.

Integral components of the RTT&E include education and training in the concepts and applications of SREF and assessing the performance and value of the SREF system. Education and training can be addressed, for example, by site visits and web based tutorials and a COMET computer-based training module (under construction). The system performance and value will be evaluated objectively (e.g., RMS errors, ranked probability scores, reliability diagrams) and subjectively on the basis of user feedback from NCEP Service Centers and selected ("Beta" site) NWS field offices. Unofficially, input is welcome from all quarters!!

Essentially the same SREF system as above had been running in a delayed mode (~ 8-10hr delay) from 12 and 00 GMT for about 14 months prior to June, 2001. Case studies (e.g., the January 2000 east coast snowstorm, Tracton and Du, 2001), preliminary objective diagnostics, and anecdotal evidence based on these runs unquestionably reinforced the proposition that SREF can be very valuable for operational forecasting applications. And this was supported unanimously by the Science and Operations Officers (SOOs) of NCEP Service Centers (HPC, SPC, AWC, TPC, MPC).

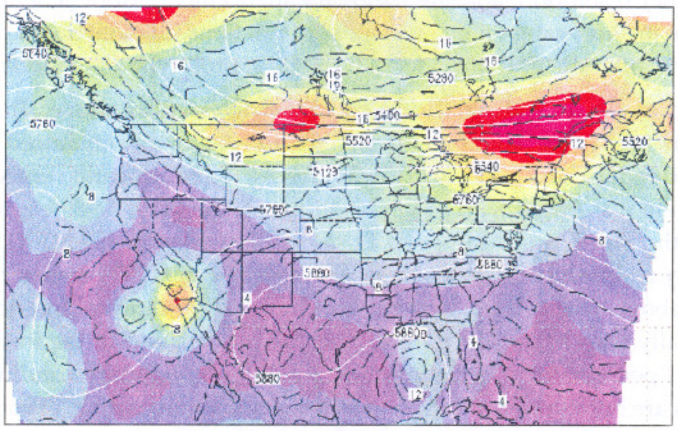

Fig. 1: Illustration of ensemble mean and spread,in this case 60 hour 500 mb height with mean abs abs vorticity superimposed.

Figure 1 illustrates the ensemble mean and spread of a 60-hour forecast of 500 mb height. Spread is defined as the standard deviation of ensemble members from the ensemble mean. In theory and practice (as demonstrated by verification statistics), the ensemble mean on average is more skillful than any individual ensemble member. The ensemble mean will usually be "smoother" in appearance than any of the individual forecasts because the averaging filters the "unpredictable components", where unpredictable here means inconsistencies amongst ensemble members. Conceptually, considering anything with more detail than contained in the ensemble mean over specifies the inherent predictability; however, note that, if most of the ensemble members are similar in the amplitude and phase of even smaller-scale features, they will be retained in the ensemble mean. The ensemble mean is just the first order advantage of ensemble prediction. Its more significant use is in providing information on uncertainties and/or confidence. The most basic product addressing this is the ensemble spread. It reflects the overall degree of variability amongst the ensemble members - the larger values (towards red) indicating areas of substantial disagreement and hence less confidence in any individual prediction (or ensemble mean) and visa versa (more purple). The maps of spread thus provide an evolving measure of the relative confidence geographically and with respect to individual weather systems.

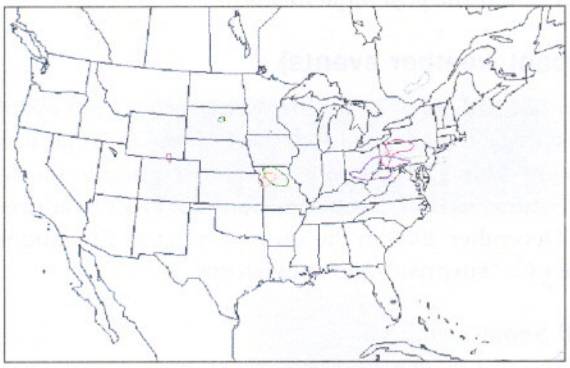

Fig.2: Example of spaghetti diagram, in this case of 12 hour accumulated precip .50" between 48-60hr

Figure 2 illustrates spaghetti diagrams with that for the 12-hour accumulated precipitation of .5". These diagrams are simply composite charts of a selected contour from each ensemble member plotted on the same chart. The obvious purpose is to convey the information content of each ensemble member in a sufficiently compact form to enable ready visualization and interpretation. When viewed over successive times into the forecasts, they provide information on the relative predictability as a function of forecast lead time and location (high where and when solutions are close and visa versa). Note too this is also a form of "graphical clustering" in that one can visually weigh the non-uniform distribution of solutions (if any) and thereby judge the relative likelihood of specific outcomes in terms of the number of forecasts pointing in that direction. Spaghetti diagrams of other parameters can provide specific information with regard to particular forecast issues. For example, the spaghetti diagrams for 1000-500mb thickness and for 850mb temperature are intended primarily for assessing the uncertainty in predicting the boundary between frozen and non-frozen precipitation; contour plots for MSLP relate to the position and (w.r.t. choice of contour value) the intensity of high and low pressure systems; isotach composite charts convey information about jet systems, and etc.

Figure 3 is an example of a probability chart derived from SREF output, in this case of the Lifted Index (LI) being less than 0° K. The LI is a measure of the vertical stability and hence an index for the potential for thunderstorms and/or severe weather. Probability estimates here are defined simply as the percentage of predictions out of the total (10) that satisfy the specified criterion. In principle probability information on any direct or derived model parameter can be output and calibrated statistically with regard to verifications over some training period. Additional examples are precipitation and low-level winds exceeding various threshold values. Note that the probability charts convey the net chance for a specific event, while the corresponding spaghetti diagrams describe the particulars of the alternative scenarios described by each member of the ensemble.

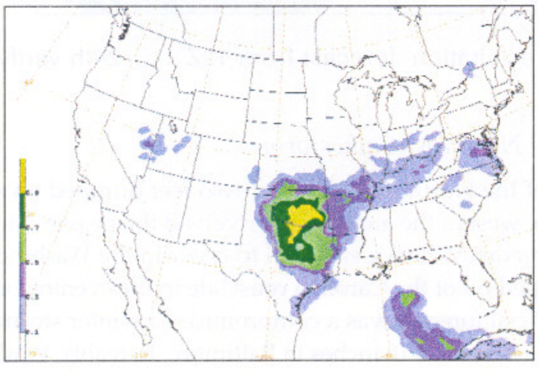

Fig. 3: Example of probability diagram, in this case the Lifted Index at 48 hour being less than 0°K.

Many additional products are or soon will become available. They include clustering, which here refers to grouping together ensemble members that are similar in some respect in regard to various fields and/or parameters. The cluster means effectively reduce the number of degrees of freedom relative to considering the complete set of individual forecasts, but not so much as the full ensemble mean (except if all the forecasts are alike). The clustering can be done at various levels over the model domain as a whole or for specific regions to focus on individual weather systems. Another set of products under development are meteograms, time traces at selected point locations. These can portray the time evolution of various quantities, such as surface temperature, in terms of possibilities or degree of uncertainty about the ensemble mean

5. Case studies (significant weather events)

Focus here is on the application of SREF, first, to an extreme precipitation event, the "surprise" heavy snowfall associated with the cyclogenesis along the east coast of the U.S. on 25-26 January 2000. This case was run retrospectively to assess whether it would have provided useful information in the context of the operational forecast problem and decision-making process. Second, the performance of SREF is addressed for the "millennium" snowstorm in the Northeast U.S., 30 December 2000 in the same context as the January storm except here, at least for Washington, DC, this was a case of a "surprise" NO snowstorm!

5.1 *The January 25-26, 2000 Snowstorm*

The storm of January 25-26, 2000 deposited a blanket of heavy snow from North Carolina northwards to New England and eastern New York State. As much as 20 inches were recorded in sections of North Carolina and up to 12-15 inches in the metropolitan Washington, D.C. area. The storms notoriety arises from the fact that it was unpredicted, certainly in regard to its severity in these regions, until about six hours before the heavy snow began accumulating. This largely, but not totally, reflects the fact that ALL (available) operational forecast models routinely available gave little, if any, clue to the imminence of this major weather event as late as the runs from 12Z January 24th. For the most part the models predicted the storm track and accompanying heavy precipitation too far to the east as, for example, in the operational Eta (32Km at that time) from 12 GMT 24 January shown in Fig. 4. As shown in Fig. 5, the SREF provided a clear "heads up" on the morning of the 24th for the possibility of a major snow event. This was especially true when the signal from the ensemble is considered in the context of independent information that morning from satellite imagery and radar observations, which suggested a storm track closer to the coast and precipitation further inland than the operational models were indicating. In actuality, the official forecast for Washington issued at 21Z January 24th called for only a 40% chance of light snow, and only some 6 hours later was the forecast changed drastically to reflect actual events.

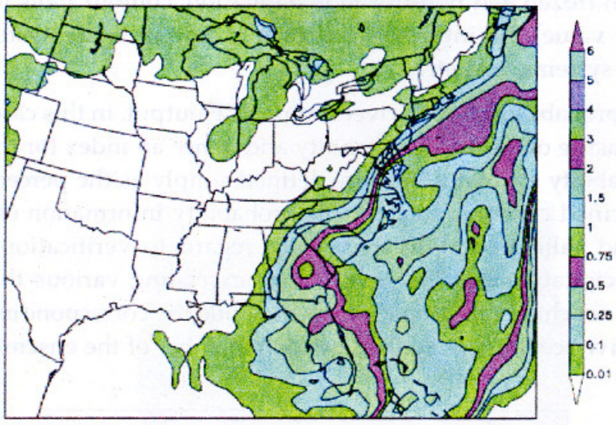

Fig. 4: Operational 24 hr Eta precipitation forecast from 12Z Jan 24th verifying 12Z Jan 25th .

5.2 *The December 29-30 2000 Northeast Snowstorm*

Very heavy snow fell over much of the Northeast with over two feet dumped on northern New Jersey. But amounts dropped off dramatically just to the west of the areas which received the heaviest accumulations. The Eta as late as the run from 12 GMT 29 Dec (Fig. 6) predicted heavy snowfall to occur in the Washington/Baltimore corridor associated with the surface low which developed off of the Carolina coast late in the evening on 29 December. The MRF predicted a lighter accumulation, and the official forecast was a compromise - a winter storm warning across the DC area, with predictions of 3-6 inches in Washington and 5-10 inches in Baltimore. In reality, Washington and Baltimore woke up on the morning of the 30th with sunny skies and, very surprisingly, no snow. As shown in Fig. 7, the ensemble indicated a 30-40% chance of significant snow; thus, in this case SREF gave a "heads up" (60%) for the chance of no snowstorm.

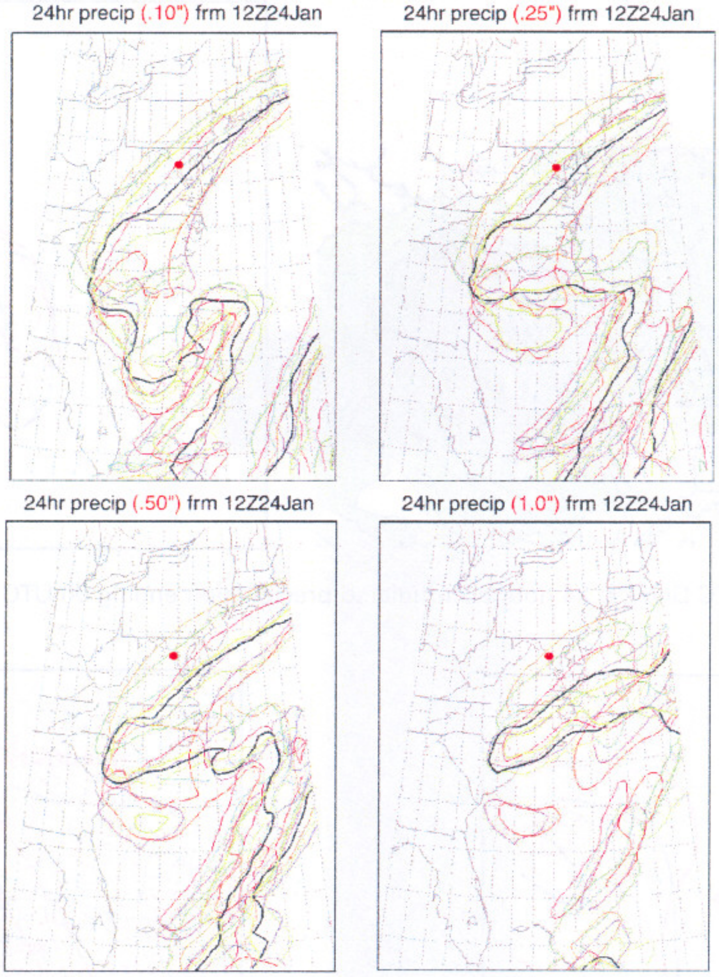

Fig. 5: 24- Hour ETA ensemble from 12Z Jan. 24TH spaghetti diagrams of 12 hr accumulated precip for period ending Jan 25th ; DCA =>

6. Discussion

The bottom line from the case studies (and other considerations) is that in the face of the respective deterministic operational model forecasts to the contrary, the ensemble runs sent a distinct signal ("heads-up") for the possibility of heavy snow in the Washington, DC region in the January 25-26th case and for the possibility of no snowstorm in the December 29-30th case. An issue is whether these signals would have made a difference in the NWS operational (busted) forecasts. The question relates to the more general issue of whether and how to convey uncertainty in forecast products. The answer clearly depends upon the specific needs and requirements of users, which varies from the general interest of the "person on the street" to more sophisticated applications that can benefit from cost versus loss considerations as a function of user specific critical thresholds ("threshold of pain"), e.g., diverting aircraft from potentially affected airports. To fully exploit SREF, therefore, it is incumbent to educate both forecasters and users on the fundamental concepts and applications of ensemble prediction. In addition to the continued general use of SREF, an important test of this is that SREF will be a critical and necessary component in HPC's upcoming Winter Weather Experiment that began November 1, 2001 and extends through early spring, 2002. The intent is to enhance the suite of products and services available from HPC to assist NWS field offices (WSOs) in delivering improved winter weather services to the public - especially to improve lead time and probability of detection in the prospects of winter storms and related temporal and spatial distributions of frozen versus non-frozen precipitation. To a large extent, the experiment was motivated by the potential of SREF referred to above and illustrated by the SREF case studies above. Products derived from SREF will include PQPF for individual county dependent "critical" thresholds, coupled with the probabilities of precipitation type (rain, snow, and freezing rain) derived from the Baldwin/Schichtel algorithm used in the operational Eta and MRF models. An example of this product is shown in Fig. 8. Note that the envelope of possible storm tracks has been added for subjective considerations. The combination (conditional probability) of the PQPF and precipitation type probabilities (together with all other available model guidance, satellite and radar data) will then serve as the basis in "chat room" discussions between HPC and field offices in issuance of official public Winter Weather Advisories and Warnings.

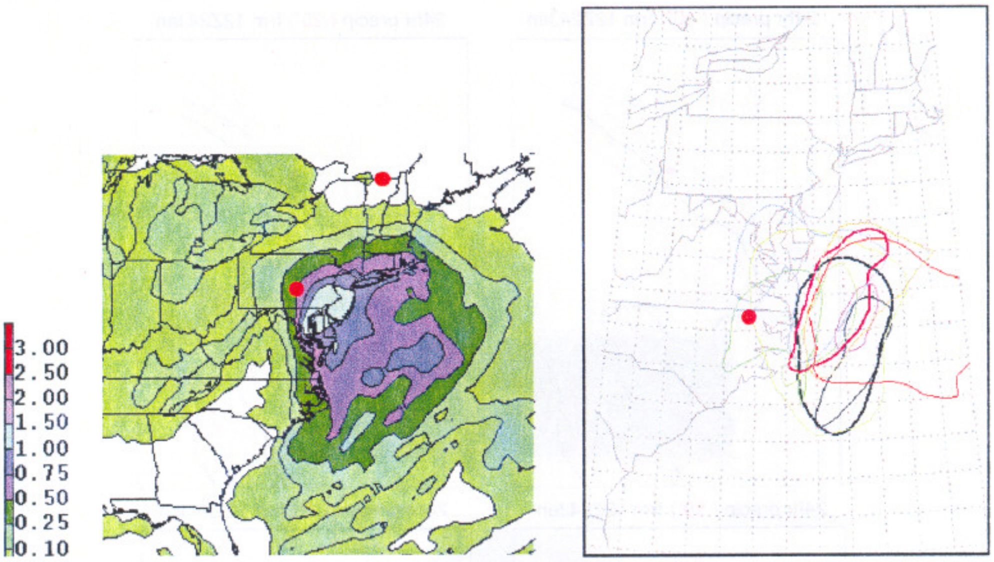

Fig. 6: ETA from 12 UTC Dec 29: 24 hour accumulated precipitation ending 00 UTC 31 Dec.

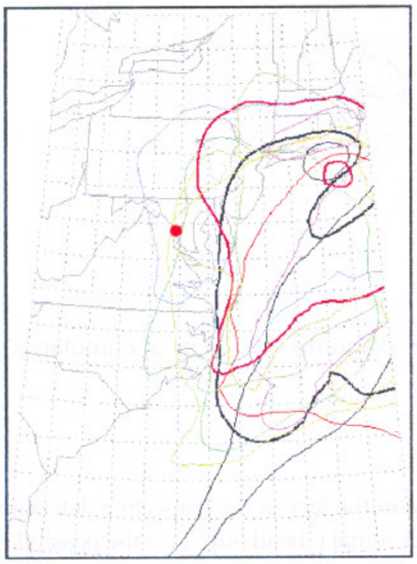

Fig. 7: SREF 24hr (upper left) and 36 hr (middle left) spaghetti of for .50" 12 hr precip ending 12 UTC 30 Dec and 00 UTC 31 Dec, respectively

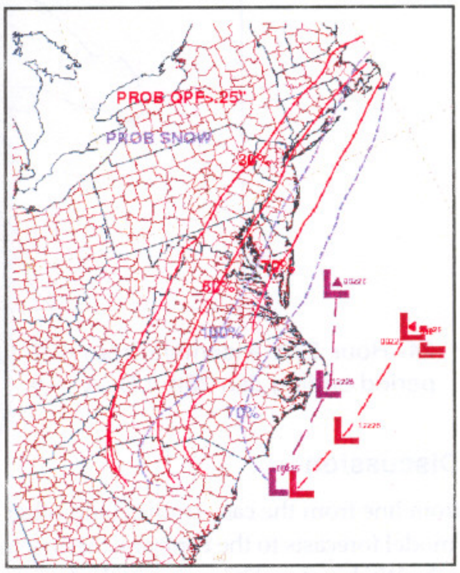

Fig. 8: Example of probabilistic snow product generated from the SREF

References

Brooks, H. E., M.S. Tracton, D.J. Stensrud, G. DiMego, and Z. Toth, 1995: Short range ensemble forecasting: Report from a workshop, 25-27 July 1994, *Bull.Amer.Meteor.Soc.*,**76**,1617-1624.

Baldwin, M., R. Treadon, and S. Contorno, 1994: Precipitation type prediction using a decision Tree Approach with NMC's Mesoscale eta Model. Preprints, 10th Conf. On Numerical Wea. Prediction, Portland, OR, *Amer. Met.Soc.*, 30-31.

Black, T., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting*, **9**, 265

Du, J.,and M.S. Tracton, 1999: Impact of lateral boundary conditions on regional-model ensemble prediction. Research Activities in Atmospheric and Oceanic Modeling (edited by H.Ritchie), Report 28, CAS/JSC Working Group Numerical Experimentation (WGNE),*WMO/TD*-**No. 942**, 6.7-6.8.

Hamill, T.M and S.J Colucci,1997: Verification of Eta-RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, **125**, 1322-1327.

Hou,D., E.Kalnay, and K. Drogemeir, 1999:Objective verification of the SAMEX '98 ensemble forecasts. *Mon.Wea.Rev.*, **129**, 73-91.

Juang, H.M., and M. Kanamitsu, 1994: The NMC nested regional spectral model. *Mon.Wea.Rev.*, **122**, 3-26.

Stensrud, D.J., H.E. Brooks, J.Du, M.S.Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon.Wea.Rev.*, **127**, 433-446.

Toth, Z., and E. Kalnay, 1997: Ensemble Forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.

Tracton, M.S., J. Du,. Z. Toth, and H. Juang,1998: Short-range ensemble forecasting (SREF) at NCEP/EMC. Preprints, 12th Conf. on Numerical Weather Prediction. Phoenix, AZ, *Amer. Meteor.Soc.*, 269-272.

Tracton, M.S. and J. Du, 2001: Application of the NCEP/EMC short-range ensemble forecast system (SREF) to predicting extreme precipitation events. Preprints, Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, Albuquerque, NM, *Amer. Meteor. Soc.*, 64-65.