Verification Development:
An events-oriented approach to verifying high resolution forecasts of wintertime precipitation

August 4, 2006

Submitted by:

Professor Paul J. Roebber
Atmospheric Science Group
Department of Mathematical Sciences
University of Wisconsin at Milwaukee
Milwaukee, WI 53201

Project Summary

An events-oriented verification procedure is designed and tested using data obtained from the DTC Winter Forecast Experiment (DWFE). The proof-of-concept is applied to mesoscale forecasts from the ARW and NMM dynamic cores of the WRF for synoptic precipitation systems, subjectively and objectively categorized as warm air advection, differential cyclonic vorticity advection, frontal circulation and upper jet streak. The method is designed to generalize readily to a wide range of precipitation phenomena. The procedure verifies several performance aspects as a function of precipitation feature: precipitation location, area, orientation and intensity. This report covers research results to date, documented below: final results are pending.

1. Project Description

With the advent of high resolution numerical weather predictions, a new challenge is to develop verification methods that properly assess the information content of these data. Traditional verification measures, when applied to such data, can lead to improper conclusions concerning model forecast performance (e.g., Baldwin et al. 2001; Roebber et al. 2004). New methods must be developed which account for the high degree of spatial variability in meteorological fields such as precipitation, and that can be used to assess the relative utility of forecasts, based on the needs of the end user. For example, a forecast that contains localized, intense precipitation is likely to contain large errors at particular points relative to a forecast that has a more homogenous distribution (such as with a coarse resolution numerical model), yet the former forecast may provide a more realistic conceptual basis for the observed evolution owing to its ability to capture the scales and amplitudes of embedded structures, and may be more useful to mesoscale forecasters.

Methods that attempt to assess the realism of the forecast have been denoted "events-oriented" verifications (e.g., Baldwin et al. 2001). A number of ideas have been proposed (e.g., Anthes 1983). Recently, Fowle and Roebber (2003) applied an events-oriented verification approach to NWP forecasts of convection and found that a 6 km grid space model was capable of providing skillful information concerning the occurrence and timing of organized modes of convection, but that the model had difficulty with weakly forced convective modes and was still subject to non-negligible errors in forecast location. Earlier, Roebber and Gehring (2000) applied an events-oriented technique to the occurrence and location of lake-breeze events on the western shores of Lake Michigan, and showed that a high resolution model had considerable skill in predicting the occurrence and relative westward extent of such events. These studies provide results that are directly accessible to forecasters, while simultaneously reveal useful information to model developers concerning continuing problems (e.g., in the case of lake breeze development, the lack of a diurnal cycle in lake surface temperatures contributed significantly to a bias in the westward extent of the lake breeze).

In the present study, we develop and test an events-oriented verification procedure for a specific type of pheneomena that will serve as a prototype for more general model verifications. We apply the procedure to synoptic precipitation associated with warm air advection, frontal circulations, differential cyclonic vorticity advection and upper-level jet streaks (note, however, that the procedure is not specific to this type of event, but rather could be generalized to other phenomena, such as lake effect snow, coastal fronts, upslope precipitation and warm sector convection).

The verification focuses on several aspects as a function of precipitation feature: precipitation location, area, orientation and intensity. These aspects are characterized and verified as follows. First, all contiguous precipitation areas are identified in the gridded observational [NCEP/Office of Hydrology hourly multisensor National Precipitation Analysis (NPA)] and DWFE model data (all datasets are analyzed on the DWFE grid). Twenty-four hour accumulated precipitation for the period beginning at 1800 UTC is subjected to agglomerative hierarchical Cluster Analysis (Everitt 1980; see Marzban and Sandgathe 2006 for a review) based on spatial coordinates. Inter-cluster distance (similarity) is measured using Euclidean Distance and the SLINK (Single Link or shortest distance) criteria. Manual inspection of a variety of dates revealed that the SLINK variant was optimal for the 24-h precipitation data under consideration, with three or fewer clusters for each date.

A subjective, synoptic classification of each of the identified precipitation clusters for the period 15 January – 31 March for 2001-05 is produced. Precipitation classes are categorized as above (warm air advection, frontal circulation, differential cyclonic vorticity advection and upper-level jet streak). These classes are identified using analyses at the surface and upper levels, as well as individual station reports and skew-T plots.

Next, the precipitation clusters are processed following the procedure of Davis et al. (2006). The resulting fit of a bivariate normal ellipsoid produces the following measures for each cluster: precipitation centroid (latitude and longitude), the length of the major and minor axes and the major axis orientation (where a positive angle denotes a northeast-to-southwest direction).

An important step is to determine if an artificial neural network (ANN) can be trained to recognize the subjectively defined precipitation classes. Neural networks are particularly adept at pattern recognition (e.g., Roebber et al. 2003 and references therein). Accordingly, the 2001-04 data are used to train multilayer perceptrons (MLPs) for each of the 4 precipitation classes, while the 2005 data are used to test network performance. Note that while the use of an ANN is not strictly necessary, since we have already manually categorized these cases, it is important to test this methodology for feasibility and potential generalization. Inputs vary for each network, but are based on minimalistic analyses derived from daily average data from the NCEP/NCAR Reanalysis dataset, such as the coordinates of surface low pressure systems, 500 hPa vorticity and 300 hPa wind speed maxima, respectively. Once a suitable network is obtained, it can then be applied to the DWFE forecasts. ANN performance is assessed using 2x2 contingency verification classes (forecast and observed; not forecast and not observed; forecast but not observed – if no observed event of that type or no event; not forecast but observed – if no forecast event of that type or no forecast event) and associated statistics (probability of detection, false-alarm ratio, true-skill statistic).

The verification is conducted as a function of precipitation category. The offset of the forecast precipitation centroid relative to the observed data, is used to determine translation errors. Even when a forecast centroid is correct, however, the precipitation overlap may not be exact, owing to rotational errors (consider a northeast-southwest precipitation band compared to a northwest-southeast band, centered on the same location). Rotational errors is assessed through comparison of the analyzed and forecast angle of the major axis. These two errors are summarized through percentile distributions. Areal coverage errors can be summarized over a set of events through percentile distributions for each class, based on the derived ellipsoids. Given that an event is both forecast and observed, we will verify the model performance in prediction of precipitation intensity.

Precipitation intensity class frequencies (low, moderate, high) are used to characterize precipitation intensity and a probability score (Brier 1950) is used to verify the model forecasts.

2. Project Results and Deliverables

At this stage, most of the work has been completed. What remains is to finalize the training of the ANNs and produce the final verification statistics. To date, a network has been trained for each of the four precipitation categories based upon the schematic input data. An independent analyst is conducting a second, subjective assessment of the precipitation categories to compare against the P.I.'s categorization for the test period. This consistency comparison between analysts will reveal an expected upper bound for ANN accuracy. Currently, the networks are within 30% of the P.I.'s categorizations (e.g., for differential cyclonic vorticity advection, the ANN agrees with the analyst that CVA is occurring in a given case 68% of the time, and that CVA is not occurring 67% of the time). Some additional experimentation with inputs, for example, those that quantify field shapes as well as locations, may be needed. Further results will be presented during the presentation on August 18.

Components of the project that could be offered as part of the DTC software suite include: C++ code that performs a cluster analysis of gridded precipitation data; FORTRAN code that derives synoptic elements (location and intensity of surface cyclones and anticyclones; location and intensity of 500 hPa vorticies; location and magnitude of 300 hPa jet streaks) from NCEP/NCAR Reanalysis and DWFE model data; C++ code that determines whether observed and model precipitation is the result of warm air advection, frontal circulations, differential cyclonic vorticity advection and/or an upper-level jet streak; and FORTRAN code that processes defined clusters to produce precipitation objects (location, orientation of major axis, and length of the major and minor axes).

One outcome of this study is verification of precipitation events as defined by the synoptic forcing. How model performance with respect to precipitation forecasts varies with synoptic forcing has not yet been quantified. With this approach, we will be able to address several aspects of performance: location errors (translation and rotation), areal bias, and misrepresentation of intensity. These measurements should provide new insights into model physics that might impact improvements in operational forecasting.

A multi-part paper is anticipated from this research. The first part will detail the methods used to analyze the precipitation and to categorize the synoptic forcing, and will provide a summary of "climatology" of the models and observations. The second part will document the forecast verification results and provide suggestions as to causes for the forecast errors.

References

Anthes, R.A., 1983: Regional models of the atmosphere in middle latitudes. *Mon. Wea. Rev*., **111**, 1306-1335.

Baldwin, M.E., S. Lakshmivarahan, and J.S. Kain, 2001: Verification of mesoscale features in NWP models. Preprints, *Ninth Conf. On Mesoscale Processes*, Fort Lauderdale, FL, Amer. Meteor. Soc., 255-258.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev*., **78**, 1-3.

Davis, C., B. Brown and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev*., **134**, 1772–1784.

Everitt, B. S. 1980. *Cluster Analysis*. Second Edition, Heinemann Educational.

Fowle, M.A. and P.J. Roebber, 2003: Short-range (0-48 h) numerical prediction of convective occurrence, mode and location. *Wea. Forecasting*, **18**, 782-794.

Marzban, C., and S. Sandgathe, 2006: Cluster analysis for verification of precipitation fields. In press, *Wea. Forecasting.*

Roebber, P.J. and M.G. Gehring, 2000: Realtime prediction of the lake breeze on the western shore of Lake Michigan. *Wea. Forecasting*, **15**, 298-312.

Roebber, P.J., S.L. Bruening, D.M. Schultz and J.V. Cortinas Jr., 2003: Improving snowfall forecasting by diagnosing snow density. *Wea. Forecasting*, **18**, 264-287.

Roebber, P.J., D.M. Schultz, B.A. Colle and D.J. Stensrud, 2004: The risks and rewards of high resolution and ensemble numerical weather prediction. *Wea. Forecasting*, **19**, 936-949.