P6.6 EVALUATION OF ARPS AND GDST 1-6 HOUR FORECASTS

Christopher W. Porter*[1], Eric M. Kemp[2], Janelle M. Janish[1], David E. Jahn[2]

(1) NSSL/Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma
(2) Center for Analysis and Prediction of Storms, Norman, Oklahoma

1. Introduction

Over the last few years MIT/Lincoln Labs (MIT/LL) has developed the Growth and Decay Storm Tracker (GDST), which is designed to track the storm envelope of existing convection, rather than individual cells, using radar data (Wolfson et al., 1999). The Advanced Regional Prediction System (ARPS) (Xue et al. 2000, Xue et al. 2001) has been developed by the Center for Analysis and Prediction of Storms and is used in operational tests to create real-time forecasts of convective events. The purpose of this study is to provide insight into the accuracy of ARPS and GDST 6-hr forecasts of position, coverage, and strength for five strongly-forced convective events. A 6-hr forecast period requires that storm system dynamical processes are considered. No dependence on dynamical processes is included within GDST. Hence, it is important to evaluate the accuracy of the GDST simulations beyond the first forecast hour and determine if ARPS can provide a more accurate forecast in this time period.

2. Methodology

Cases evaluated in this study are chosen from a suite of forecasts generated during the Spring Operations Period 1999 (SOP99). ARPS forecasts are made on a 3 km grid over a 960 x 852 km domain. The 3-km forecast grid is nested within a 9-km forecast grid centered over the central United States. This 9-km grid is, in turn, nested within a 30-km forecast grid that covers the contiguous United States. Using the 9-km forecast as the background field, WSR-88D Level-III reflectivity, rawinsonde sounding data, Oklahoma Mesonet data, satellite data and ASOS data are used to generate the initial field for the 3-km forecast. ARPS and GDST forecasts are analyzed on grids centered on various WSR-88D radar sites. A "fuzzy" validation technique (Hallowell et al. 1999), is used to verify all forecasts for composite reflectivity $\geq$ 41 dBZ. Rather than verify each grid point separately, a kernel is constructed around the grid point being scored, with the radius of the kernel (typically 5nm) dependent on the model resolution and the requirements of the user.

* *Corresponding Author Address*: Christopher W. Porter, NSSL, 1313 Halley Circle, Norman, OK, 73069; e-mail: Chris.Porter@nssl.noaa.gov

3. ARPS and GDST Results

Probability of Detection (POD), False Alarm Rate (FAR), and Critical Success Index (CSI) statistics for the ARPS and the GDST are averaged over all cases for each hour of the 6-hr forecast period (Figure 1). For the five cases studied, the ARPS forecasts tend to produce more skillful forecasts in the 2-6 hr forecast period. Generally, the GDST is able to produce a more accurate forecast over the initial 2-hr forecast period.

ARPS forecasts properly represent the observed mode of convection whether it be a broken line of isolated, intense convective cells or a solid, linear squall line. In addition, the ARPS forecasts capture the orientation of the convective systems well. However, ARPS also had some consistent forecast problems, primarily with spurious convection, areal coverage of reflectivities $\geq$ 41dBZ, and intensity of convection, all of which reduced ARPS forecasts' skill scores. The overall intensity of the model reflectivity is always too low at the initialization, due in part to the ice, snow, and hail fields not being included in the initial conditions (an inadvertent error affecting the 3-km SOP99 forecasts). Therefore, the skill scores at the initial time are not optimal. This can lead to an under prediction of the intensity or areal coverage of reflectivities $\geq$ 41 dBZ. But, in three of the five cases, the extent of the significant reflectivity generated by ARPS is actually greater than observed, especially near the end of the 6-hr forecast. This is due either to spurious convection (contributing to high FARs), or reflectivity values within portions of the convective systems that far exceed Level III observational data. Over predicting areal coverage and intensity may be an artifact of the model's 3-km grid resolution, which is still coarse relative to fine-scale microphysical processes.

In general, with the exception of one squall-line case, the GDST is able to discern the magnitude of propagation for convective systems as a whole.

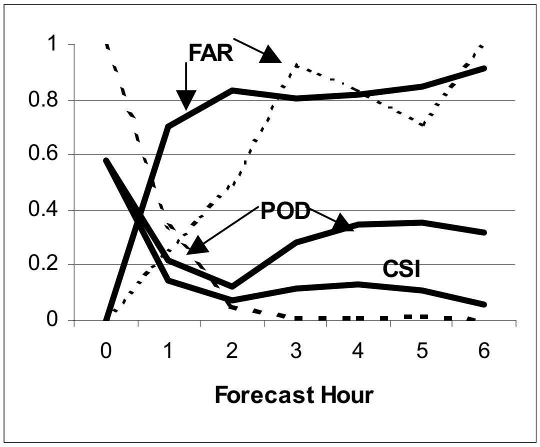

Figure 1 Composite ARPS (solid lines) and GDST (dashed lines) POD, FAR, and CSI values averaged over all cases for each forecast hour. GDST CSI values (not shown) are nearly identical to its POD values.

This allows forecasts to contain some measurable skill up through one hour into the forecast period. However, in all cases, the GDST forecast skill degrades quickly, typically containing very little skill by the second hour.

GDST POD and CSI values are higher than averaged ARPS values through the first forecast hour. Beyond the second hour, the ARPS's POD and CSI values rise, with average POD values near 0.35 and average CSI values near 0.13. The GDST produces lower FAR values throughout the first two hours, but both methods generate high FAR values from 3 to 6 hrs into the forecast period.

Unfortunately, the size of the verification domains (either 400 km x 400 km or 255 km x 255 km) and length of forecast period (6 hrs) allows convective lines within the GDST forecasts to propagate beyond the domain boundary within the forecast period. This impacts the statistical results in an artificial manner by decreasing POD and CSI scores. This can not be prevented because reflectivity observations are only available to a limited distance from the radar. Therefore, simply increasing the size of the verification domain will not necessarily alleviate this problem. However, the use of composited radar observations from multiple radar sites is a potential remedy.

4. Forecast Value vs. Forecast Skill

To date, GDST verification studies have focused on the value of the forecast, using methods to increase statistical scores to agree with subjective feedback (Hallowell et al. 1999). The value of a forecast relates to user-specific costs or utilities that are expected to arise from using the forecasts. Most of the statistics generated in this report are also attempts to assess value by using the tolerance kernel.

However, studying the skill of forecasts can be very profitable as well. Forecast skill measures the general association between the forecasts and the observations, usually relative to a set benchmark (e.g., random forecasts, climatology, persistence, etc.). This report has attempted to address forecast skill by subjectively evaluating each case hour by hour. In some cases, these evaluations have identified genuinely good features (e.g., convective initiation and maintenance, storm type, storm orientation, etc.) in forecasts that have poor verification statistics. Conversely, these evaluations have also identified undesirable features in highly-rated forecasts (e.g., storms too intense, spurious convection, over forecasting convection, etc.).

Further discussion on forecast value and forecast skill in addition to a comprehensive analysis of each convective case from this study can be found at http://www.nssl.noaa.gov/~porter/cwpdt00/finalreport.pdf.

5. References

Hallowell, R.G., M.M. Wolfson, B.E. Forma n, M.P. Moore, B.A. Crowe, T.M. Rotz, D.W. Miller, T.C. Carty, S.F. McGettigan, 1999: The Terminal Convective Weather Forecast Demonstration at the DFW International Airport. *Preprints*, *8th Conf. on Aviation, Range, and Aerospace Meteorology* , Dallas, TX, AMS, Boston, MA.

Wolfson, M. M., B. E. Forman, R. G. Hallowell, and M. P. Moore, 1999: The Growth and Decay Storm Tracker. *Preprints, 8th Conf. on Aviation, Range, and Aerospace Meteorology*, Dallas, TX, Amer. Meteor. Soc., 58-62.

Xue, M., K. K. Droegem eier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics*, **75**, 161-193.

Xue, M., K. K. Droegemeier, V . Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D. -H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Physics*, in press.