J4.3 SHORT-TERM RADAR NOWCASTING FOR HYDROLOGIC APPLICATIONS OVER THE ARKANSAS-RED RIVER BASIN

Matthew P. Van Horne*, Enrique R. Vivoni and Dara Entekhabi
Massachusetts Institute of Technology, Cambridge, MA

Ross N. Hoffman and Christopher Grassotti
Atmospheric and Environmental Research, Inc., Lexington, MA

1. INTRODUCTION

The demand for rainfall forecasts with high spatial and temporal resolution has increased recently from different sources including operational hydrologic forecasters, water managers and meteorologists. A general set of methods currently implemented for the generation of quantitative precipitation forecasts (QPFs) is radar nowcasting or short-term extrapolation (Browning and Collier 1989). Early short-term forecasting algorithms relied on pattern recognition of rainfall echoes to calculate cross correlation coefficients and use them to predict storm motion and features (Einfalt *et al.* 1990).

Nowcasting is the production of short-range (0-3 hour lead times) precipitation forecasts using extrapolation methods (Smith and Austin 2000). The widespread availability of high spatially and temporally resolved radar data in convenient digital form has given forecasters more accurate tools for the generation of forecasts for mesoscale phenomena (Browning and Collier 1989). Radar-based nowcasting is perceived to have benefits over numerical weather prediction (NWP) models for short forecast lead-times (Zipser 1990).

Nowcasting remains an active area of research, as there are still complications in its methodology and operational use. One drawback is the simplicity of pure advection models (see Brémaud and Pointin 1993; Johnson *et al.* 1998; Pereira Fo. *et al.* 1999; Wolfson *et al.* 1999; Handwerker 2002) that do not account for storm growth and decay (Smith and Austin 2000), considered main factors in determining the limit of predictability for nowcasting.

In this study, we present the application of a novel nowcasting model (Wolfson *et al.* 1999) for hydrometeorological forecasting in mid to large-scale basins. Improvements in the accuracy of short-term rainfall forecasts are achieved due to the high spatial and temporal resolution in radar rainfall data. The network of operational NEXRAD radars providing coverage over the Arkansas-Red River Basin (ARB) permits continuous radar-based nowcasting as a comprehensive short-term forecasting tool. Section 2 of this study introduces the nowcasting model used while section 3 discusses the data used. Evaluation criteria are described in section 4 and the results from case studies described in section 5 are available in section 6.

* Corresponding author address: Matthew P. Van Horne, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 48-320, Cambridge, MA 02319; e-mail: mvanhorne@alum.mit.edu

2.0 NOWCASTING MODEL

2.1 *Description*

The Growth and Decay Storm Tracker (GDST) developed by Lincoln Laboratory at the Massachusetts Institute of Technology (MIT LL) as part of the Integrated Terminal Weather System (ITWS; Evans and Ducot 1994) uses scale separation filtering to forecast mesoscale storm events (Wolfson *et al.* 1999). The GDST improves on traditional cross-correlation analysis by using an elliptical filter to separate the large-scale storm features from smaller-scale cells before correlation, thus enabling improved tracking of the storm envelope. Here we apply the GDST algorithm for large-scale hydrometeorological modeling over operational hydrologic basins using radar rainfall data at higher spatial/temporal resolutions than previously reported. The model is forced with radar rainfall measurements to produce forecasts for lead-times up to 120 minutes in 15-minute increments. Forecast verification uses radar rainfall images from the valid time for the forecast.

MIT LL originally designed the GDST to forecast line storm progression near airport areas to improve flight routing during severe weather (Forman *et al.* 1999). For example, preliminary testing at the DFW airport used NEXRAD reflectivity data at a temporal resolution of 6 minutes and spatial resolution of 1 km over a 440 km by 440 km area (Theriault *et al* 2000).

2.2 *Image Filtering*

Separation of the small-scale, short-lived cells from the large-scale, longer-lived features of the storm before the correlation analysis can result in more accurate envelope forecasting and thus better overall forecasting. The GDST accomplishes this by passing an elliptical filter over the input rainfall field (Wolfson *et al.* 1999). At each point in the field, the filter computes the average value of the filtered pixels, at 10º increments over 180º and assigns the maximum of the averages computed over the rotation to the base pixel in the filtered image. This filtering generates a smoothed large-scale image for use in the correlation analysis (Wolfson *et al.* 1999). The filter size and aspect ratio are two of the main parameters that affect model forecast skill. To determine the optimal filter size for each data set, filter size tests were performed to ensure optimal performance by the GDST. Optimal performance is determined by comparing the scores for each filter size, averaged over all lead-times.

3.0 DATA

Forecast accuracy is largely influenced by the quality and spatial/temporal resolutions of the rainfall data used for short-term extrapolation. In this study, we compared two radar rainfall products from the Next Generation Radar (NEXRAD) system available over the entire United States: NEXRAD Stage III/P1 and Weather Services International (WSI) NOWrad product, with 2 km 15-minute native resolution (see Grassotti *et al.* 2002). WSI uses NEXRAD reflectivity data along with proprietary Z-R relationships to derive rain amounts without incorporating surface rain gauge data. The WSI product utilizes a "weather condition" approach to convert from reflectivity to rainfall rates with a precision of 1.27 mm hr^{-1}.

Grassotti *et al.* (2002) showed several characteristics of the two radar rainfall products through a comparison of the WSI data, Stage III/P1 and gauge data over the ABRFC. The main difference between the data sets was Stage III/P1 data tended to overestimate low rainfall rates (0-2 mm hr^{-1}) due to the mean bias correction performed when merging radar and gauge data. If systematic biases do not dominate WSI data, it should be superior to gauge data and roughly equivalent to Stage III/P1 data for the purposes of hydrologic forecasting (Grassotti *et al.* 2002).

The location used in this study is the Arkansas-Red River Basin River Forecasting Center (ABRFC) region. The ABRFC coverage area includes parts of Oklahoma, Texas, Arkansas, Missouri, Kansas, Colorado and New Mexico, and is over 500,000 km^2 in area. One sub-basin in the eastern portion of the ABRFC, the Illinois River basin (5851.8 km^2), is the focus of additional hydrologic analysis in this study.

4.0 MODEL PERFORMANCE CRITERIA

Full analysis of GDST forecasts requires a combination of meteorological and hydrological statistics. The combination of the two types of measures enables a better understanding of the space-time accuracy of the forecasted rainfall. A contingency table approach is the basis for meteorological scoring. The critical success index (CSI) is used for this purpose (Wilks 1995):

$$CSI = \frac{H}{H + M + FA} \qquad (1)$$

Hits (*H*) are pixels where both the forecast and observed values exceed a specified threshold, misses (*M*) occur when the observed pixel exceeds the threshold but the forecast pixel does not, and false alarms (*FA*) occur when only the forecast pixel exceeds the threshold. A high CSI score indicates a large degree of positional correspondence between the forecast and observed images.

Threshold-based and verification area CSI scoring are used in this study to discriminate model performance over varying rainfall intensity and varying degrees of spatial positioning. We define thresholds based on the cumulative rainfall distributions for each storm. The designations “low”, “median” and “extreme” correspond to the 0^{th}, 50^{th} and 90^{th} rain rate percentiles, respectively. A verification area CSI searches in a kernel surrounding the pixel in question in the observed image to verify the corresponding pixel in the forecast image.

These statistics provide a generalized description of the locational accuracy of the forecasts over a large area. The CSI only scores the percent of correctly placed non-zero pixels in relation to the observed image but does not address the geographical positioning of the rainfall forecasts. The nature of the CSI as a large-scale measure also does not indicate whether the forecast may be useful for hydrologic modeling at the basin scale.

A trio of hydrologic measures described by Smith *et al.* (2002) evaluates the accuracy of the intensity, extent and distribution of the forecasted rainfall within a catchment. The mean areal precipitation (*M(t)*), the fractional coverage with rainfall exceeding the low ($F_0(t)$), median ($F_{50}(t)$), and extreme ($F_{90}(t)$) thresholds, and the normalized distance to the basin outlet (*D(t)*) for each threshold accomplish this purpose.

5.0 CASE STUDIES

Three storm events were chosen for this study from the period of 1998-2000. Each of these storm events was chosen based on rainfall intensity and duration, since longer, more intense storms are hydrologically more significant.

5.1 *Event A*

This highly organized linear storm event impacted the ABRFC area for the duration of 5 October 1998. The storm initially developed as several individual cells over central Oklahoma and slowly developed into a linear band but did not move significantly during this organization. This storm was the result of a strong dynamic forcing event caused by a cutoff low in the upper atmosphere. The event impacted the Illinois River basin from 1200 UTC to 2000 UTC and then continued its decay while the remnants moved out of the ABRFC.

5.2 *Event B*

This storm developed over western Oklahoma as an unorganized, seemingly chaotic rainfall event, early on 4 January 1998. The degree of organization varies over the course of the day, ranging from well-defined individual thunderstorms to large areas of stratiform rain to linear groups of convective cells. This event was not as well organized as Event A but still impacted the Illinois River basin for a period of approximately 14 hours attaining similar areal averaged precipitation levels.

5.3 *Event C*

This storm event begins as a mix of linear and chaotic elements, loses its organization for a period then recovers some of its linear features. Throughout the period of 1200 UTC 13 April 1999 through 1200 UTC 14

April 1999, there is a constant presence of linear characteristics but not as strongly organized as Event A. Several different groups of individual convective cells emerge at various times but were never able to maintain a linear formation for significant periods. This storm did not greatly affect the Illinois River basin, only providing rainfall for a few hours on the morning of 14 April 1999.

6.0 RESULTS

Assessing the GDST forecast value in a hydrometeorological forecast is the focus of this section. The performance measures were applied to forecasts generated from data of different spatial and temporal resolution to explore the impact on forecast accuracy.

6.1 *Threshold Dependence*

A comparison between NEXRAD Stage III/P1 and WSI data (4 km and 1-hour resolution) was performed to determine the relative accuracy between the data sources. The finer resolution (2 km, 15-minute) WSI data allows a wider range of spatial and temporal resolutions to be used in evaluating GDST forecast skill. The cumulative rainfall rate distributions for each of the two data sets underlie the threshold-dependent curves shown in Figure 1.

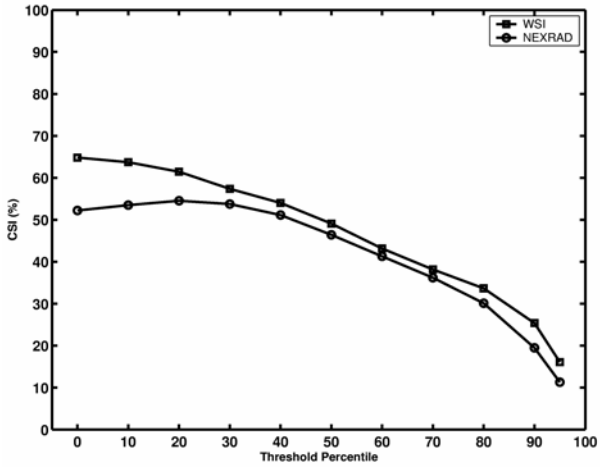

Figure 1. CSI as a function of threshold percentile for Event A using NEXRAD and WSI 4 km, 60 minute resolution data. The data points are the mean 60-minute forecast scores over the duration of the event.

Difference between WSI and NEXRAD CSI scores in Figure 1 for low threshold values is attributed to the mean bias correction algorithm used for incorporating radar data with sparse rain gauges. This results in the overestimation of low rain rates as compared to the WSI product (Grassotti *et al.* 2002). The two data sets are similar over the majority of their range and are especially close at high rainfall rates. That similarity implies that the two data products are interchangeable for the purposes of this study. For the remainder of this paper, WSI data will be used in GDST analysis due to the data availability with increased temporal and spatial resolution and the higher forecast accuracy over the entire range. Decreases in CSI with increasing threshold are expected and may be important for hydrometeorological forecasting.

6.2 *Lead-time Dependence*

All three events were evaluated to examine the decrease in forecast skill as a function of lead-time and storm type. Figure 2 shows forecast skill (CSI) as a function of lead-time for Events A, B, and C for low, median and extreme threshold rainfall rates.

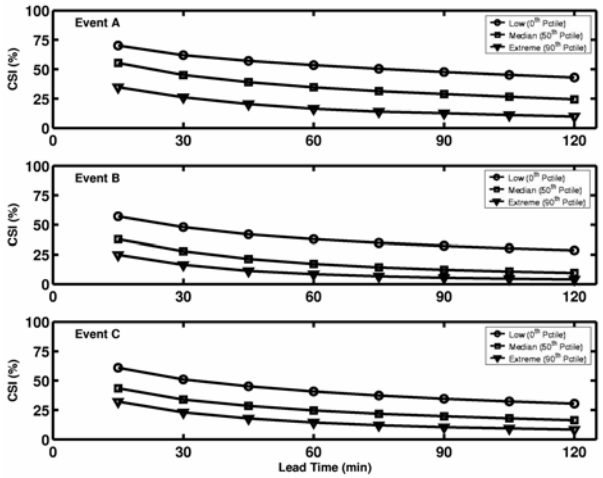

Figure 2. CSI scores versus lead-time for the three storm events using WSI 4 km, 15-minute data. The data points shown are the average values over the duration of the storm.

A main feature shown in Figure 2 is higher Event A scores were expected based on the more organized linear squall line relative to the other events. As expected Figure 2 showed the dependence of GDST forecast skill on storm organization. The true value of these forecasts may not be entirely represented by Figure 2 since no extended kernel in the observed image was used for verification.

6.3 *Verification Area Dependence*

The size of the kernel used for forecast verification is an important parameter affecting the resultant CSI score. Scores calculated using a large kernel size might not provide relevant information about the forecast quality or accuracy. However, small kernels may not capture the full forecast value in cases where exact prediction is not a necessity. Operationally, the chosen kernel size reflects the importance of a shift or lag in the forecasted image, with kernel size varying inversely with necessary precision. Figure 3 shows the effect of varying the size of the verification kernel on the one-hour CSI scores, using WSI 2 km, 15-minute data for Events A and C.

Over a comparable kernel area, 484 km^2, Events A and C scored within the 50-70% range obtained by Cartwright *et al.* (1999). This method of scoring may be beneficial for hydrological analysis as rain forecasted within a few kilometers of its actual location may still be useful for hydrologic modeling.

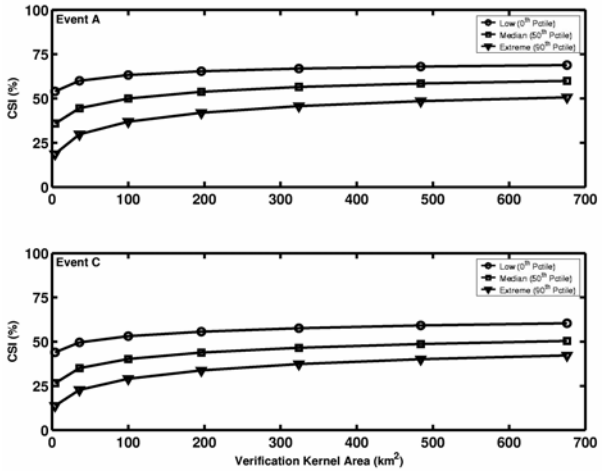

Figure 3. CSI versus verification area size for Events A (top) and C (bottom) using WSI 2 km, 15-minute data. The data points are the average values over the duration of the event at each verification area size.

6.4 *Comparison to Persistence Forecast*

Persistence is used to evaluate if the GDST forecasting method provides an improvement over a simpler method. Figure 4 shows the areas where the forecasting methods (GDST or persistence) either over- or underestimate rainfall for the 60-minute lead-time forecast in Event A. In this specific case, the GDST generally overestimated the leading edge of the storm, leading to more risk-averse hydrologic forecasting. Another benefit of the GDST over persistence is the distribution of the over- and underestimated pixels throughout the storm area where, predictably, persistence resulted in large contiguous areas of over- and underestimation. Overall, the use of filtering and extrapolation within the GDST result in a more accurate forecast as compared to persistence.

6.5 *Hydrologic Analysis*

The meteorological statistics analyzed to this point provide only one insight into the quality of rainfall forecasts while basin-based hydrologic statistics show different features and further explore the applicability of the forecasts. Figure 5a shows the location of the Illinois River basin within the ABRFC and a 4 km resolution digital elevation model (DEM) of the basin (5,851.8 km^2). In Figures 5b-d, the solid lines represent the observed images while the dashed lines represent the 60-minute forecast images for Event A. Figure 5b shows the values of the 60-minute M(t), mean areal precipitation, over the Illinois River basin. The GDST predicts magnitudes and timings of the peaks moderately well but the magnitude overestimation will result in more risk-averse flood forecasting.

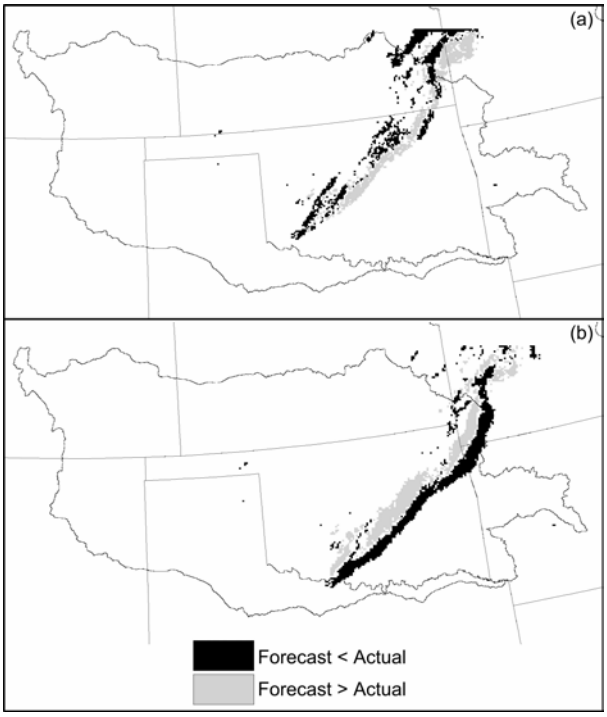

Figure 4. The areas where the GDST (a) and the persistence (b) 60-minute forecasts erred in their rain rate magnitude estimates. Black areas indicate underestimation of rain rate magnitude while gray areas indicate overestimation.

Figure 5c shows the values for $F_0(t)$, the fractional rainfall coverage over the Illinois River basin. Figure 5d shows normalized distance to outlet, $D(t)$, values for the observed storm event and the 60-minute forecasted storm event. The horizontal line at $D(t)$ = 0.58 represents the normalized distance for uniform rain over the basin. The GDST 60-minute forecast curve shows a slight lag in time behind the observed curve for both $F(t)$ and $D(t)$.

7.0 Conclusions

High spatial/temporal resolution radar precipitation estimates greatly improve nowcasting for hydrologic modeling. The GDST nowcasting model produces forecasts that have potential for hydrologic use as illustrated by a combination of meteorological and hydrologic measures. Rainfall threshold evaluation of the GDST provided valuable information about the model capability to correctly locate different intensities of rainfall. The combination of the extreme threshold with an increased verification kernel shows that the GDST forecasts high rainfall rates with moderate accuracy. Hydrologic usefulness of GDST forecasts is further shown via a trio of basin-based measures and the GDST performs reasonably well in its prediction of rainfall intensity, extent and distribution over the Illinois River basin. Active research continues in the use of GDST forecasts for hydrometeorologic modeling.

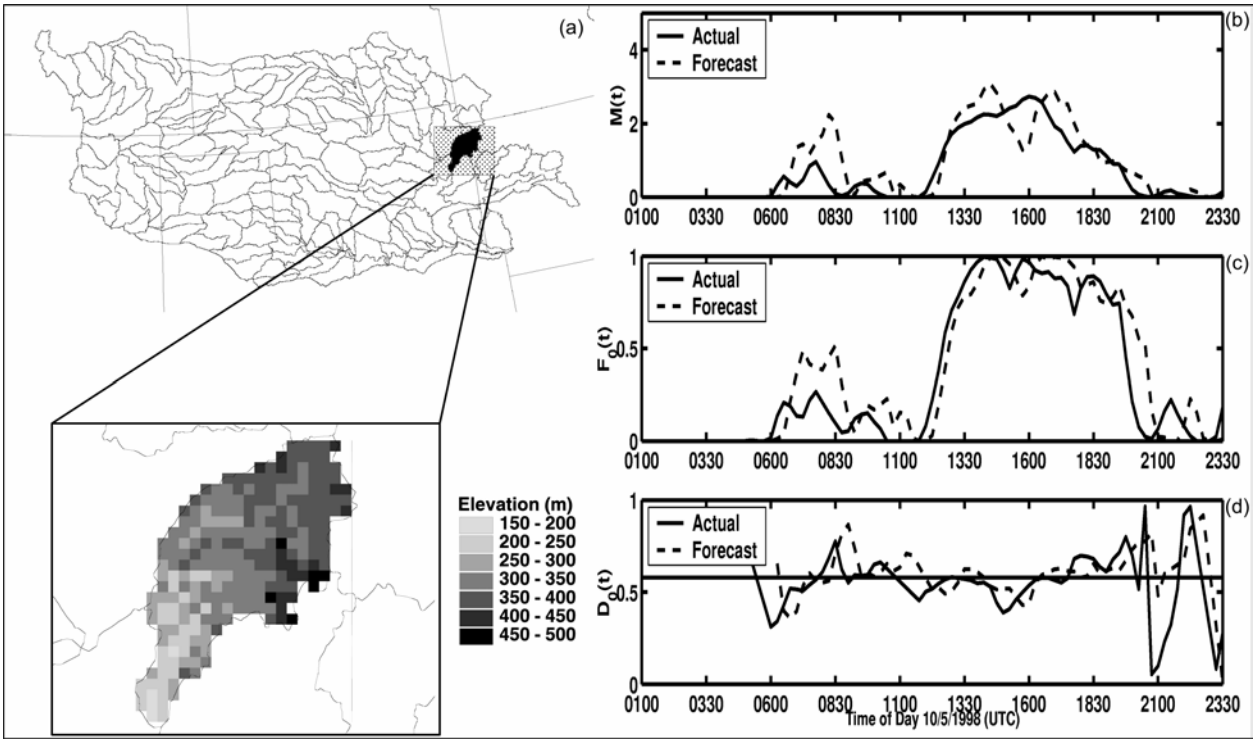

Figure 5. a) Location of the Illinois River basin. b-d) Hydrologic analysis of Event A over the Illinois River basin. Forecast values are dotted lines and observed values are the solid lines.

8.0 Acknowledgements

We thank MIT Lincoln Lab for the use of the GDST nowcasting model and particularly Barbara Forman and Bob Hallowell for help with specific problems. Weather Services International, Inc and the NWS Hydrology Laboratory provided the radar rainfall data used in this study. This work was partially funded by the US Army Research Office (contract DAAD19-00-C-0114) and the American Meteorological Society Dr. Pedro Grau scholarship (2001-2002).

9.0 References

Brémaud, P.J. and Y.B. Pointin, 1993: Forecasting heavy rainfall from rain cell motion using radar data. *J. Hydrol.*, **142**, 373-389.

Browning, K.A. and C.G. Collier, 1989: Nowcasting of Precipitation Systems. *Rev. Geophys.*, **v.27, n.3**, 345-370.

Cartwright, T.J., M.M. Wolfson, B.E. Forman, R.G. Hallowell, M.P. Moore and K.E. Theriault, 1999: The FAA Terminal Convective Weather Forecast Product: Scale Separation Filter Optimization. *29th International Conf. on Radar Meteorology,* Montreal, Quebec, 852-855.

Einfalt, Thomas, Thierry Denoeux and Guy Jacquet, 1990: A Radar Rainfall Forecasting Method Designed for Hydrological Purposes. *J. Hydrol.*, **114**, 229-244.

Evans, James E. and Elizabeth R. Ducot, 1994: The Integrated Terminal Weather System (ITWS). *The Lincoln Laboratory Journal*, **v.7, n.2**, 449-473.

Forman, B.E., M.M. Wolfson, R.G. Hallowell, and M.P. Moore, 1999: Aviation User Needs for Convective Weather Forecasts. *Presented at the American Meteorological Society 79th Annual Conf.*

Grassotti, Christopher, R. N. Hoffman, E. R. Vivoni and D. Entekhabi, 2002: Intercomparison of Radar and Rain Gauge Observations over the Arkansas-Red River Basin. *Submitted to Wea. Forecasting*.

Handwerker, Jan, 2002: Cell tracking with TRACE3D – a new algorithm. *Atmospheric Research*, **61**, 15-34.

Johnson, J.T., Pamela L. MacKeen, Arthur Witt, E. DeWayne Mitchell, Gregory J. Stump, Michael D. Eilts and Kevin W. Thomas, 1998: The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm. *Wea. Forecasting*, **13**, 263-276.

Pereira Fo., Augusto J., Kenneth C. Crawford and David J. Stensrud, 1999: Mesoscale Precipitation Fields. Part II: Hydrometeorological Modeling. *J. Appl. Meteor.*, **38**, 102-125.

Smith, James A., Mary Lynn Baeck, Yu Zhang, and Charles A. Doswell III, 2001: Extreme Rainfall and Flooding from Supercell Thunderstorms. *J. Hydromet.*, **2**, 469-489.

Smith, K.T. and G.L. Austin, 2000: Nowcasting precipitation - a proposal for a way forward. *J. Hydrol.*, **239**, 34-45.

Theriault, K.E., M.M. Wolfson, B.E. Forman, R.G. Halowell, M.P. Moore and R.J. Johnson Jr., 2000: FAA Terminal Convective Weather Forecast Algorithm Assessment. Preprints, *Ninth Conf. on Aviation, Range, and Aerospace Meteorology*.

Wilks, D.S., 1995: *Statistical Methods in Atmospheric Sciences*. Academic Press, New York, 464 pp.

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

Zipser, E., 1990: Rainfall predictability: When will extrapolation-based algorithms fail? In *Eighth Conf. on Hydrometeorology*, 138-142.