P3.1 Development of MOS Thunderstorm and Severe Thunderstorm Forecast Equations with Multiple Data Sources

Kathryn K. Hughes *
Meteorological Development Laboratory
Office of Science and Technology
National Weather Service, NOAA
Silver Spring, MD

1. INTRODUCTION

Since September 2000, the Meteorological Development Laboratory (MDL) of the National Weather Service (NWS) has provided forecasts of thunderstorms and severe thunderstorms from 6 to 72 hours in advance. The Model Output Statistics (MOS) technique (Glahn and Lowry 1972) was used to generate forecast equations to predict the probability of thunderstorms and severe thunderstorms based on output from the 0000 and 1200 UTC cycles of the Aviation (AVN) run of the Global Spectral Model (GSM; Kanamitsu 1989). These thunderstorm probability forecasts are valid for 6-, 12-, and 24-h periods for the contiguous U.S. (CONUS).

Beginning May 2001, thunderstorm forecasts based on output from the medium-range forecast model (MRF) run of the GSM have been available for CONUS sites as part of the MRF MOS text message. The MRF MOS thunderstorm forecasts are generated out to 192 hours in advance for 12- and 24-h periods.

Cloud-to-ground lightning data from the National Lightning Detection Network (NLDN), provided by NASA's Global Hydrology Resource Center (GHRC) in Huntsville, Alabama, were used to define the presence of a thunderstorm in the MOS predictand data. Storm reports, consisting of reports of large hail, tornadoes, and high winds, were used to define the occurrence of a severe thunderstorm for the MOS development. The severe thunderstorm predictand was made conditional on the occurrence of a thunderstorm. Because all of the data in these predictand data sets were random in place and time, the observational data were placed on the same grid so both types of data could be analyzed and compared. Relative frequencies were developed from lightning data and storm reports, for every grid point, and for 6-, 12- and 24-h periods. These relative frequencies were then used as potential predictors in the equation development.

Details of the AVN MOS thunderstorm development and verifications of the skill of the new AVN MOS thunderstorm forecasts are presented in this paper. The AVN MOS thunderstorm forecasts are more skillful than the older MOS system which was based on output from the Nested Grid Model (NGM; Hoke et al. 1989). The MRF MOS thunderstorm development and the planned development of Eta-based thunderstorm and severe thunderstorm guidance are also discussed.

* *Corresponding author address:* Kathryn K. Hughes 1325 East-West Highway, Room 10380, Silver Spring, Md 20910-3283; e-mail: <Kathryn.Hughes@noaa.gov>

2. NGM-BASED THUNDERSTORM GUIDANCE

NGM-based thunderstorm and severe thunderstorm guidance were developed by using two different approaches. Six and 12-h probability forecasts (Bower 1993) are available in text messages for periods ending 12 to 60 hours after 0000 and 1200 UTC. In the equation development, multiple data sources were used to define the occurrence of a thunderstorm at a station. These data included the report of a thunderstorm in the station's hourly surface observations, a Manually Digitized Radar (MDR) report of Video Integrator and Processor level 3 (VIP3) or greater in a block approximately 115-135 km on a side (depending on latitude) and surrounding the station, or a report in the National Severe Storm Forecast Center's (NSSFC's) event logs that indicated a thunderstorm in the 115-135 km box. Similarly, the occurrence of a severe thunderstorm was indicated by the station's hourly observation or by a report of tornadoes, large hail, and/or damaging winds in the NSSFC logs for the station area.

Another NGM-based thunderstorm guidance product available to the user is the probabilistic and categorical (yes/no) forecast of thunderstorms and severe local storms for 18- or 24-h periods (Reap 1994). These forecasts are valid for the 12-36 h and 36-60 h projections after 0000 UTC and the 6-24 h and 24-48 h projections after 1200 UTC. These forecast products are available only in graphical form. The development of the forecast equations and the characteristics of the forecasts for the 24-h period differ significantly from the station-oriented guidance for the 6- and 12-h periods. First, the data used to define the occurrence of a thunderstorm were extracted from the NLDN. The hourly surface observations and radar data were not used. Secondly, a thunderstorm was defined by the occurrence of at least two cloud-to-ground lightning flashes in a grid block. Although the lightning data were random in space and time, each strike report contained information including the latitude, longitude, and time of the strike. This information was used to place the lightning data on a grid of 89x113 blocks, each approximately 48 km on a side, covering the CONUS and adjacent areas.

3. DEVELOPMENT OF AVN- AND MRF-BASED THUNDERSTORM GUIDANCE

AVN MOS thunderstorm and severe thunderstorm forecasts have been generated since September 2000. Forecasts are produced for 6-, 12-, and 24-h periods for both the 0000 and 1200 UTC cycles, out to 84 hours in advance. The 6- and 12-h guidance out to 72 hours in advance is available in AVN MOS messages (Dallavalle and Erickson 2000). The 24-h guidance out to 84 hours, as well as the 6- and 12-h guidance out to 72 hours in advance, are also available in binary form as BUFR (Binary Universal Form for the Representation of meteorological data) messages.

MRF MOS thunderstorm forecasts have been generated since May 2001. These forecasts are produced for 12- and 24-h periods out to 192 hours in advance from 0000 UTC and are available in the new MRF MOS messages (Erickson and Dallavalle 2000). The 12- and 24-h thunderstorm forecasts are also available in binary format as BUFR messages. No severe thunderstorm guidance is available in the MRF MOS.

3.1 Data Collection

The cloud-to-ground lightning data were obtained from NASA's GHRC in Huntsville, Alabama. These lightning data were measured by the NLDN, maintained by Global Atmospherics, Inc., Tucson, Arizona. The cloud-to-ground lightning data were the only source used to define the presence of a thunderstorm for 6-, 12-, and 24-h periods. An upgrade was made to the NLDN in 1994 (Cummins et al. 1998), so data before 1994 were not included in the development. Because the data are random in time and place, the lightning data were placed on a grid (Fig. 1). The small markers show the center of each grid box and the coverage of the gridded thunderstorm forecasts. The larger markers show the CONUS MOS stations available in the AVN/MRF MOS text messages. A nearest neighbor approach was used to assign the closest grid point to a MOS station to generate the thunderstorm and severe thunderstorm text messages.

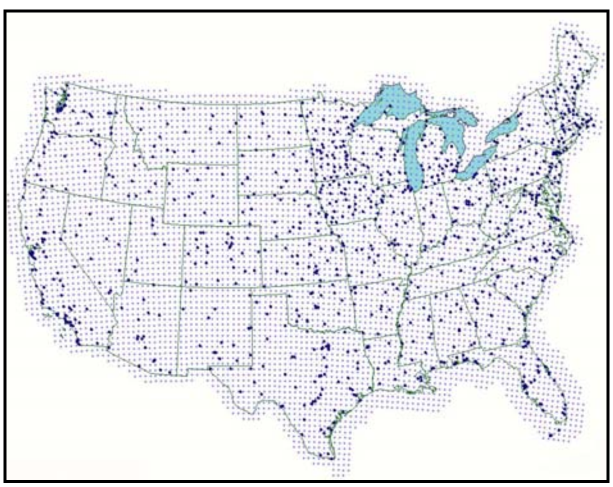

Figure 1. Grid used for the thunderstorm development. The small markers show the center of each grid box. The large markers show the CONUS MOS sites.

The predictand data for the severe thunderstorm consisted of individual reports of tornadoes, large hail, and thunderstorm wind gusts or damage, compiled and quality controlled by the Office of Climate, Water, and Weather Services (OCWWS) from Storm Data reports. These storm reports were obtained by NWS Forecast Offices and were stored in the NWS severe weather data base. For consistency with the lightning data, only storm reports from 1994 and after were included in the development. A subset of the grid shown in Fig. 1. was used for the severe thunderstorm reports. The borders of the severe thunderstorm grid did not extend beyond the land areas of the CONUS. Because of the sparsity of the storm data, a Geographical Information System was used to study the relationship of the storm reports to the population density of the area covered by the grid. Data from areas with very low population density were not included in the developmental sample. This modification seemed to improve slightly the forecast equations. Note that the final severe thunderstorm forecasts are generated for the full CONUS. In future studies, it may be beneficial to include only areas in the sample where any storm reports occurred rather than relying strictly on population density.

3.2 Predictand Definition

A thunderstorm was defined by one cloud-to-ground lightning strike during the defined period. This was a change from the requirement of two strikes used in the earlier NGM 24-h thunderstorm development. Recent studies have shown the requirement of two strikes is no longer necessary (Huffines and Orville 1999), so for this development one or more cloud-to-ground lightning strikes defined the thunderstorm events. All strikes for a given hour within a given box were added up and assigned to the grid box labeled by the center point. Hours with no reports of lighting were simply considered non-thunderstorm events. It was assumed there were no missing cases of thunderstorms. The hourly thunderstorm reports were then summed for the appropriate 6-, 12-, or 24-h period.

The severe thunderstorm reports were also assigned to specific grid boxes and summed over the appropriate periods. The severe thunderstorm predictand was made conditional on a thunderstorm. Thus, only thunderstorm events were considered. If a thunderstorm occurred and no severe thunderstorm reports were received, then the event was non-severe. If a thunderstorm occurred and severe thunderstorms were reported, then the event was severe.

3.3 Relative Frequencies

Five years of the cloud-to-ground lighting and storm data (April 1994 - March 1999) were used to develop monthly lightning and severe thunderstorm climatologies for the CONUS for 6-, 12-, and 24-h periods. Over 128 million cloud-to-ground lightning strikes, and 100,000 storm reports were processed to generate monthly relative frequencies for each period at every grid point used in the development. These relative frequencies were then offered as potential predictors and were often selected by the linear regression program as part of the forecast equations. Because the severe thunderstorm reports were relatively rare, the severe thunderstorm relative frequencies tended to be discontinuous in space. Unusually high forecasts of conditional severe thunderstorms at isolated stations sometimes resulted from this predictor. This problem was reduced after a 9-point spatial smoothing was applied to the severe thunderstorm relative frequencies.

3.4 Predictors

Forecast equations were developed by applying linear multiple regression techniques to relate the observations of thunderstorms or severe thunderstorms (the predictand) to forecast variables (the predictors) from the AVN or MRF model, and the relative frequencies. A regionalized approach was used for the equation development. Data for a group of stations, or a group of grid boxes, were combined to increase the sample size and the stability of the equations. For the AVN and MRF MOS thunderstorm and/or severe thunderstorm development, all of the grid boxes were combined into one region. Separate forecast equations were developed for each projection and forecast period.

The potential predictors offered to the regression program included AVN or MRF model output fields interpolated to the grid points, variables derived from the model output fields, and climatic variables for each grid point. The most important predictors were the mean relative humidity, the vertical velocity at 700 and 500 hPa, the surface-based CAPE, moisture divergence at 925 and 850 hPa, model forecasts of convective precipitation amount offered in binary form, various stability indices, the product of the K-index times the relative frequency (Reap and Foster 1979), u- and v-wind components at 700 and 500 hPa, and the difference between the equivalent potential temperature at 850 and 700 hPa.

3.5 Seasonal Stratification

The developmental data were stratified into three sea sons. The AVN and MRF MOS equations were developed for spring (March 16-June 30), summer (July 1 - October 15), and cool (October 16 - March 15) seasons. Three seasons of AVN model and observational data, from July 1997 through March 2000, were available to develop the summer and cool season AVN MOS thunderstorm and severe thunderstorm equations. Four seasons of AVN and MRF model and observational data, from April 1997 through March 2001, were available for the AVN MOS spring thunderstorm and severe thunderstorm equations, and the MRF MOS spring, summer, and cool season thunderstorm equations.

3.6 Projections and Products

Equations were developed to predict the probability of a thunderstorm and the conditional probability of a severe thunderstorm for 6-h, 12-h, and 24-h periods from both the 0000 and 1200 UTC cycles of the AVN model. The equations generate guidance for 6-h periods valid 6-12, 12-18,..., 60-66, and 66-72 hours in advance. The equations for the 12-h periods generate forecasts valid 6-18, 18-30,..., 54-66, and 60-72 hours in advance. The equations for the 24-h periods generate forecasts for periods of 12-36, 24-48,..., 48-72, and 60-84 hours in advance. The 6- and 12-h forecasts are available in the AVN MOS text messages. All the 6-, 12- and 24-h forecasts are available in the AVN MOS BUFR messages.

Equations were also developed to predict the probability of a thunderstorm for 12- and 24-h periods from the MRF model. The equations generate forecasts for 12-h periods valid 12-24, 24-36,..., 168-180, and 180-192 hours in advance. The equations for the 24-h periods generate forecasts valid 12-36, 24-48, ..., 156-180, and 168-192 hours after 0000 UTC. All the 12-h, and the 24-h forecasts spanning the 1200-1200 UTC period are available in the MRF MOS text messages. All the 12- and 24-h forecasts are available in the MRF MOS BUFR messages.

4. SEPTEMBER 20, 2000 CASE STUDY

On September 20, 2000, a cold front moved from west to east across much of the eastern United States, resulting in severe weather, including seven tornadoes, across Indiana, Ohio, Kentucky, and New York. The observed cloud-to-ground lightning strikes (light markers) and the severe weather reports (dark markers) were mapped on the thunderstorm grid (Fig. 2) for the 24-h period ending 1200 UTC, September 21, 2000. The AVN MOS 24-48 h probability of thunderstorm guidance generated on 1200 UTC, September 19, 2000, valid for the period ending 1200 UTC, September 21, 2000, is shown in Fig. 3. Figure 4 shows the forecasts of the unconditional probability of severe thunderstorms (the product of the probability of thunderstorms multiplied by the probability of conditional severe thunderstorms) for the same period. The AVN MOS thunderstorm and severe thunderstorm forecasts indicated a high likelihood of thunderstorms in an area from Ohio to Louisiana and a relatively high probability of severe thunderstorms in southern Michigan, Ohio, and Kentucky. Nearly all the thunderstorms were contained in the 20% probability contour. Many of the severe weather events occurred within the 5% probability contour.

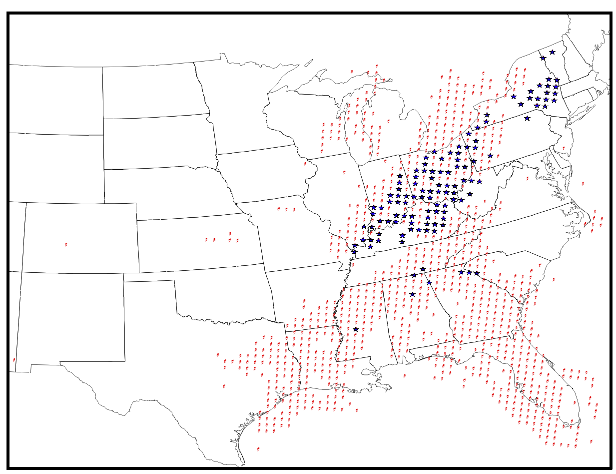

Figure 2. Cloud-to-ground lightning strikes (light markers) and severe storm reports (dark markers) for the 24-h period ending 1200 UTC September 21, 2000.

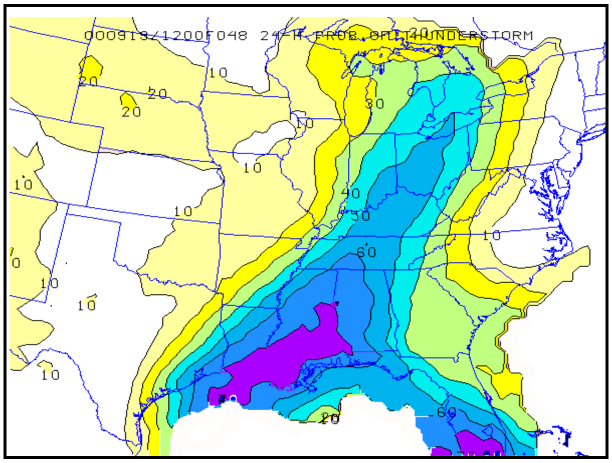

Figure 3. Probability of thunderstorm forecasts valid for the 24-48 h period ending 1200 UTC, September 21, 2000.

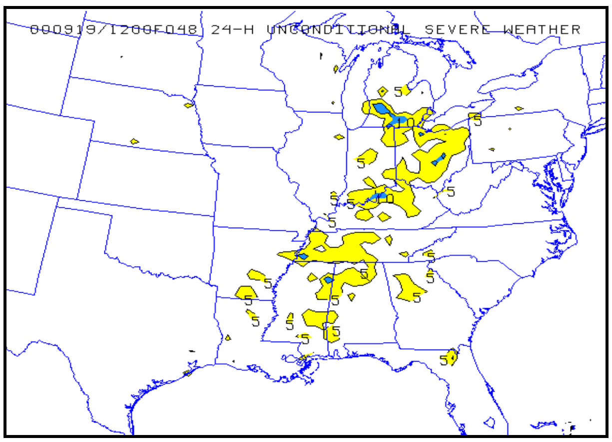

Figure 4. Same as Fig. 3, except for the unconditional probability of severe thunderstorm forecasts.

5. VERIFICATION

Climate forecasts derived from the relative frequencies were used to assess the skill of the thunderstorm and severe thunderstorm probability forecasts. All of the test results in this section were obtained from forecasts made on independent data. Figure 5 shows the percent improvement over climate, of the Brier scores for the new AVN-based thunderstorm forecasts and the older NGM-based thunderstorm guidance. The new AVN MOS thunderstorm forecasts are more skillful. Much of this improvement can be attributed to improvements in the observations of the lightning data as well as the greater forecast accuracy of the AVN model. Figure 6 shows the skill of the AVN MOS guidance for each forecast projection. All of the forecasts are skillful. Figure 7 is similar, but shows the skill of the severe thunderstorm forecasts. Forecasting severe weather is a more difficult forecasting problem due to the rarity of the events, but again, compared to forecasts based on climate, the severe thunderstorm forecasts are skillful through the entire period.

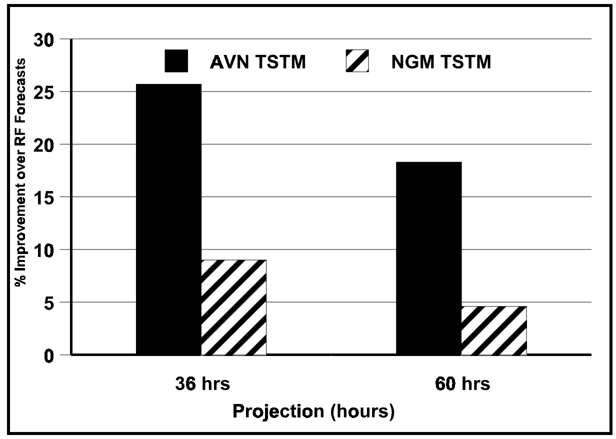

Figure 5. Percent improvement in the Brier score of AVN and NGM MOS 24-h thunderstorm forecasts for the 0000 UTC cycle, summer season 1999.

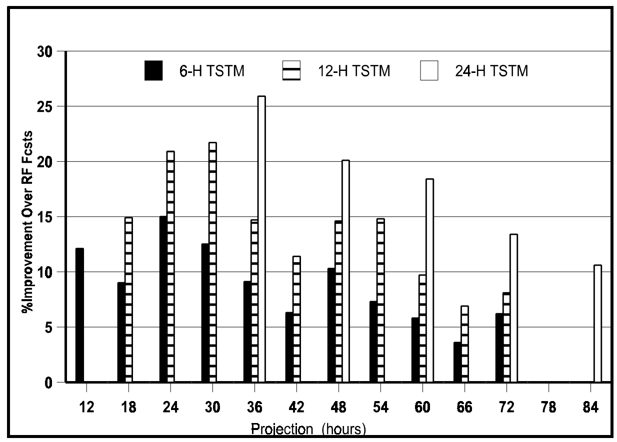

Figure 6. Same as Fig. 5, except for AVN MOS thunderstorm forecasts only.

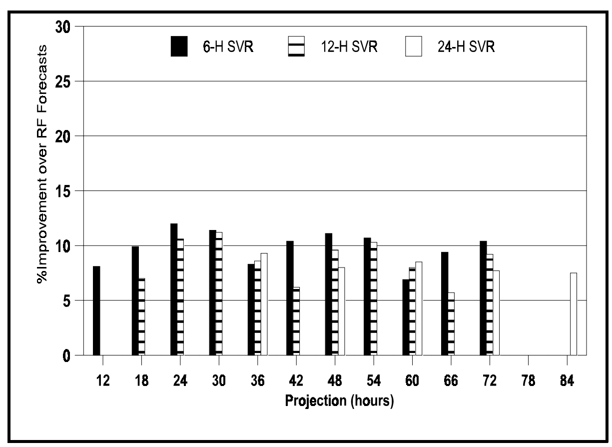

Figure 7. Same as Fig. 6, except for AVN MOS severe thunderstorm forecasts.

Figure 8 shows the skill of the forecasts for the MRF-based MOS thunderstorms for the summer season of 2000. The forecasts were compared to forecasts of climate based on 5 years of relative frequency data for each grid point. The percent improvement over climate of the Brier scores shows skill at every projection, including a slight improvement over climate at 192 hours. The results for spring, and the 12-h spring and summer periods also indicated skill at every projection.

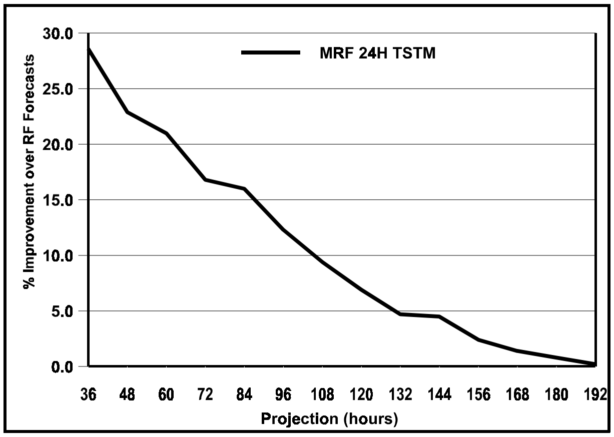

Figure 8. Same as Fig.2, except for the MRF MOS 24-h thunderstorm forecasts for summer 2000.

6. FUTURE PLANS

Work has begun to develop an Eta-based (Black 1994) MOS system for predicting thunderstorms and severe thunderstorm. Thunderstorms and severe thunderstorm will be defined by using the same predictand data described in section 3. The Eta-based forecasts will be available out to 60 hours in advance for 6-, 12-, and 24-h periods.

In the fall of 2001, MDL will initiate a guidance package from the AVN runs at 0600 and 1800 UTC (Iredell and Caplan 1997). We are planning to add thunderstorm and severe thunderstorm guidance in the 0600 and 1800 UTC packages during the spring of 2002.

Work is in progress to generate graphics products of the thunderstorm guidance. We also hope to develop some specialized products for the NCEP Storm Prediction Center, including the probability of individual severe thunderstorm events for the 12-36 h period, and the probability of thunderstorms or severe thunderstorms during 3-h forecast periods. In addition, we have been asked to develop thunderstorm guidance over the oceans, and extend the thunderstorm forecasts beyond the 192-h projection. Available resources will determine some of our future developmental work.

Acknowledgments

Special thanks are extended by the author to OCWWS for providing the severe weather data, and to GHRC for providing the lightning data.

7. REFERENCES

Black, T. L.., 1994: The new NMC mesoscale Eta model: Description and forecast examples. Wea Forecasting , **9**, 265-278.

Bower, J. B., 1993: NGM-based MOS thunderstorm and severe thunderstorm probability forecasts for the contiguous United States. NWS Technical Procedures Bulletin No. 407, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 16 pp.

Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer, 1998: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. J. Geophys. Res., **103**, 9035-9044.

Dallavalle, J. P., and M. C. Erickson, 2000: AVN-based MOS guidance - The alphanumeric messages. NWS Technical Procedures Bulletin No. 463, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Erickson, M. C., and J. P. Dallavalle, 2000: MRF-based MOS guidance - The alphanumeric messages. NWS Technical Procedures Bulletin No. 460, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., **11**, 1203-1211.

Hoke, J. E., N. A. Phillips, G. J. Dimego, J. J. Tuccillo, and J. G. Sela, 1989: The regional analysis and forecast system of the National Meteorological Center. Wea. Forecasting, **4**, 323-334.

Huffines, G. R. and R. Orville, 1999: Lightning ground flash density and thunderstorm duration in the continental United States: 1989-96. J. Appl. Meteor., **38**, 1013-1019.

Iredell, M., and P. Caplan, 1997: Four-times-daily runs of the AVN model. NWS Technical Procedures Bulletin No. 442, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 3 pp.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. Wea. Forecasting, **4**, 335-342.

Reap, R. M., 1994: 24-h NGM based probability and categorical forecasts of thunderstorms and severe local storms for the contiguous U.S. NWS Technical Procedures Bulletin No. 419, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 14 pp.

_____, and D. S. Foster, 1979: Automated 12-36 hour probability forecasts of thunderstorms and severe local storms. J. Appl. Meteor., **18**, 1304-1315.