Prepared for the 17th conference on Weather Analysis and Forecasting, Denver

5.3

AVN-BASED STATISTICAL FORECASTS OF THUNDERSTORMS AND SEVERE THUNDERSTORMS FOR THE CONTIGUOUS U.S.

Kathryn K. Hughes*
Techniques Development Laboratory
Office of Systems Development
National Weather Service, NOAA
Silver Spring, MD

1. INTRODUCTION

In the fall of 1999, the Techniques Development Laboratory (TDL) of the National Weather Service (NWS) plans to implement a statistical weather forecast system based on output from the National Centers for Environmental Prediction's (NCEP) Aviation (AVN) run of the Global Spectral Model (GSM; Kanamitsu 1989). As part of this system, equations to predict the probability of thunderstorms and the conditional probability of severe thunderstorms will be derived for the contiguous U.S. by using the Model Output Statistics (MOS) approach (Glahn and Lowry 1972). The primary focus of the initial AVN-based thunderstorm forecast system will be to provide guidance from 6 to 72 hours in advance for both the 0000 and 1200 UTC cycles. The new MOS forecasts will be valid for 6-h, 12-h, and 24-h time periods.

Future plans include adding forecasts for 3-h time periods, generating forecasts out to 78 hours in advance, and developing AVN-based guidance from the 0600 and 1800 UTC model runs (Iredell and Caplan 1997). We also intend to provide thunderstorm and severe thunderstorm forecasts from the Medium Range Forecast (MRF) run of the GSM for projections out to 144 hours, or as far in advance as skill allows.

Cloud-to-ground lightning data from the national lightning location network is used to define the presence of a thunderstorm in the predictand data. The predictand data for severe weather consists of reports of tornadoes, large hail, and wind gusts or damage, conditional on the occurrence of a thunderstorm. In this paper, we discuss the development and testing of the AVN-based forecast equations for thunderstorms and associated severe thunderstorms.

---

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

---

2. PAST DEVELOPMENTS - EXISTING GUIDANCE

2.1 *MOS Thunderstorm Guidance for 6- and 12-h Periods*

The current thunderstorm and severe thunderstorm MOS forecasts for 6- and 12-h periods (Bower 1993) are generated twice a day from output of the Nested Grid Model (NGM; Hoke et al. 1989) and are available in the FOUS14 (FWC) message for over 500 sites. The 6- and 12-h probability forecasts are valid 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 observation, 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 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.

Forecast equations were developed by relating the observations of thunderstorm or severe thunderstorms (the predictand) to forecast variables (the predictors) from the NGM. A regionalized approach was used for the equation development. Thus, data for a group of stations were combined to increase the sample size and the stability of the equations. Depending on the season and the time of day, as many as six regions, or as few as one region were used. The conditional severe thunderstorm equations were developed by using only those cases where thunderstorms occurred. Finally, the developmental data were stratified into three seasons: spring (March 16 - June 30), summer (July 1 - October 15), and cool (October 16 - March 15).

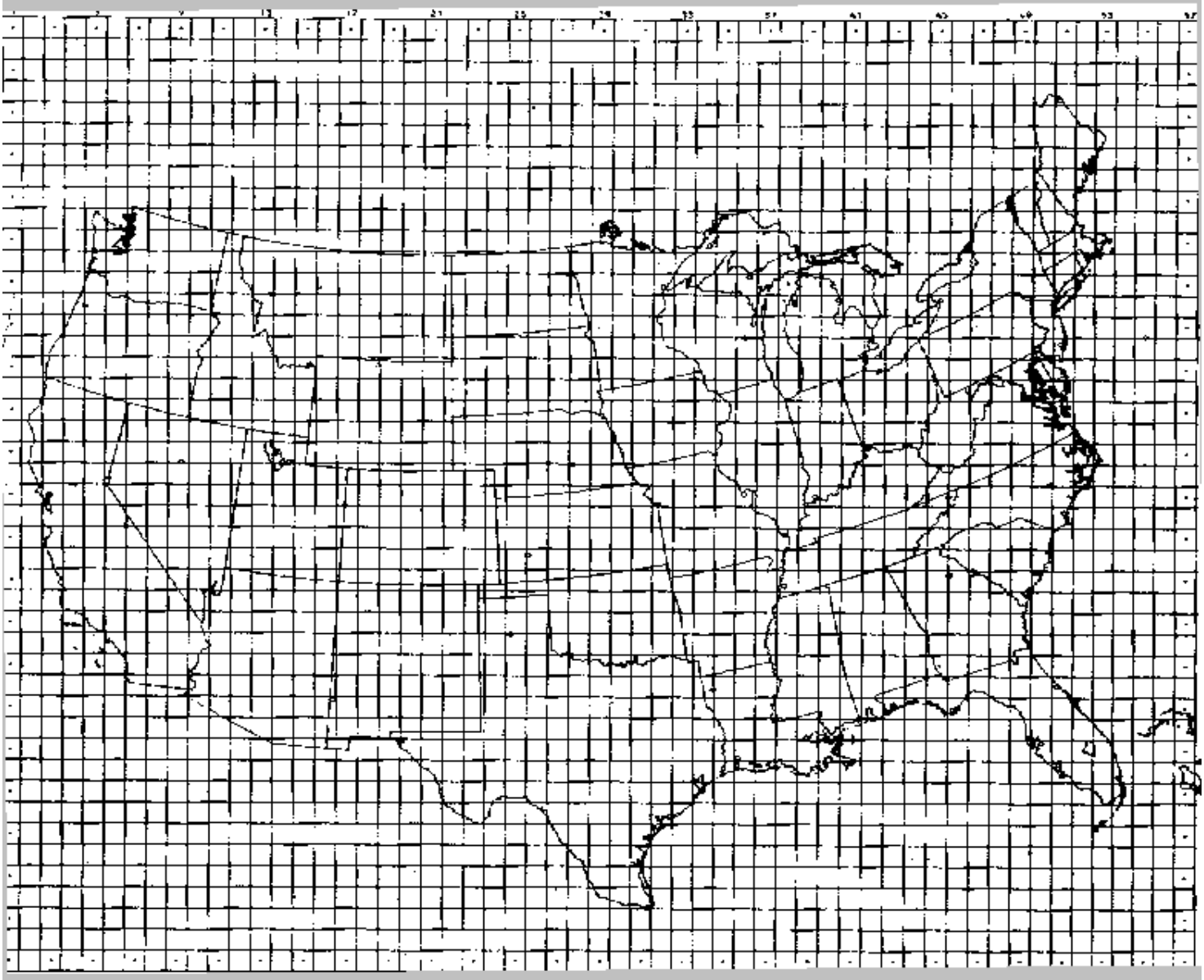

Figure 1. MDR grid with every other grid line shown.

2.2 *NGM-Based Thunderstorm Guidance for 24-h Periods*

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. Forecast products are available only in graphical form and provide guidance for specific areas. The development of the forecast equations and the characteristics of the subsequent forecasts for the 24-h periods differed significantly from the station-oriented guidance for the 6- and 12-h periods described in Section 2.1. First, the data used to define the occurrence of a thunderstorm were extracted from the national lightning location network. In other words, neither hourly surface observations nor radar data were used. Secondly, a thunderstorm was defined by the occurrence of two or more cloud-to-ground lightning flashes in grid blocks approximately 48 km on a side. Figure 1 shows the 89x113 grid (known as the MDR grid) covering the contiguous U.S. and adjacent areas. Similarly, the occurrence of a severe thunderstorm was defined as a report in the NSSFC logs of severe weather in one of the grid blocks. Third, thunderstorm equations for the 24-h products were developed for warm (March 15 - September 30) and cool seasons (October 1 - March 14) while the severe thunderstorm equations were valid for seasons similar to those used in the 6- and 12-h development. Finally, two developmental regions (the eastern and western U.S.) were used for all seasons and projections for the 24-h guidance.

In the development of the 6-h and 12-h NGM-based thunderstorm guidance, the lightning data were not available as predictand data for the time period used in the development. Unfortunately, the use of the MDR data to define the occurrence of a thunderstorm likely caused overforecasting in the cool season when heavy rains often had VIP levels resembling those of a thunderstorm. The MDR data covering the western U.S. also suffered from limitations caused by the blocking of the radar beam by the surrounding terrain.

3. DEVELOPMENT OF AVN-BASED THUNDERSTORM GUIDANCE

While a package of MOS guidance has been available from the AVN model since 1994 (Jensenius et al. 1994), the guidance is only available for limited weather elements and for approximately 250 stations in the contiguous U.S. and Alaska. The new AVN-based MOS package will contain forecasts for many more elements for over 1000 stations in the contiguous U.S., Alaska, Hawaii, and Puerto Rico. The probability of thunderstorms and the conditional probability of severe thunderstorms will be included for stations in the contiguous U.S..

3.1 *Predictand Definition*

In this AVN-based development, the cloud-to-ground lightning data from the national lightning detection network, maintained by Global Atmospherics, Inc., will be the only source used to define the presence of a thunderstorm for the 6-, 12-, and 24-h periods. This approach will eliminate inconsistencies in the guidance and provide better spatial resolution. The lightning data are available for the April 1994 - March 1999 period and were obtained from the Global Hydrology Resource Center (GHRC) at NASA's Global Hydrology and Climate Center, Huntsville, Alabama. Because the data are random in time and place, for development of the forecast equation, the lightning data are placed on the grid shown in Fig. 1. All strikes for a given hour within a given box are added up and assigned to the grid box labeled by the lower left-hand corner. Hours with no reports of lightning are simply considered non-thunderstorm events. The hourly thunderstorm reports are then summed for the appropriate 6-h, 12-h, or 24-h period. Similarly to the 24-h NGM thunderstorms, any boxes with two or more cloud-to-ground lightning strikes for a given period are considered thunderstorm events. Two or more flashes were required because previous studies (Reap and Orville 1990) have shown that isolated, single flashes can be unreliable indicators of thunderstorms.

The predictand data for the severe weather consist of individual reports of tornadoes, large hail, and thunderstorm wind gusts or damage, compiled by the Office of Meteorology from Storm Data reports. These Storm Data reports are sent by NWS Forecast Offices and are stored in the NWS severe weather database. Data are also available from April 1994 through March 1999. As in the thunderstorm development, the severe thunderstorm reports are assigned to specific grid boxes (Fig. 1) and are summed over the appropriate period. The forecasts for the severe thunderstorms will be conditional based on the occurrence of a thunderstorm.

Initially, we are developing equations to predict the probability of a thunderstorm and conditional probability of severe thunderstorms for 6-h, 12-h, and 24-h time periods from both the 0000 and 1200 UTC AVN runs of the GSM. Currently, we are developing equations to generate forecasts from 6-h periods valid from 6-12, 12-18, 18-24,..., 60-66, and 66-72 hours after the initial model time. The equations for 12-h periods will be developed for 6-18, 18-30, 30-42, 42-54, and 54-66 hours after the initial model time. The equations for 24-h periods will be developed for periods of 12-36 and 36-60 hours after 0000 UTC and for the 24-48 and 48-72 hour periods after 1200 UTC.

3.2 *Predictors*

The predictor data are comprised of forecast variables generated by the operational AVN model, by specialized predictors derived from the model output, and by climatic predictors. One type of specialized predictor is sometimes referred to as an interactive predictor. Interactive predictors are especially important when forecasting rare events such as thunderstorms and severe thunderstorms. One example of an interactive predictor used to forecast thunderstorms is the KF predictor (Reap and Foster 1979) which is the K stability index multiplied by the thunderstorm relative frequency (F). This predictor forces the climatology (the relative frequencies) to be more responsive to the synoptic situation. We used 5 years (April 1994 - March 1999) of lightning data to create the monthly relative frequencies of thunderstorms. These monthly relative frequencies are interpolated to daily values before they are offered to the regression analysis. As explained in Section 3.1, the relative frequencies are valid for the grid boxes shown in Fig. 1. A list of potential predictors is shown in Table 1.

Table 1. Sample of variable offered as predictors to the screening regression anlysis.

K index
SWEAT index
Total Totals index
Lifted index
K index times lightning relative frequency
Equivalent Potential temperature
"u" and "v" wind components at various levels
Vertical velocities times relative humidity
Moisture convergence
Temperature, Dew point
Vorticity advection
Sine and Cosine of the day-of-year
Temperature lapse rates (850 mb to 500 mb, 850 mb to 700 mb, 700 mb to 500 mb
Surface CAPE

3.3 *Seasonal Stratification*

The developmental data are stratified into three seasons: spring (March 16 - June 30), summer (July 1 - October 15), and cool (October 16 - March 15). Because of changes in the TDL archive of AVN model data, the predictor sample begins in April 1997. Thus, at this point, 2 seasons of data are available to develop the spring, summer, and cool season thunderstorm and severe thunderstorm equations.

3.4 *Regions and Stations*

The equations are to be developed by using the regional approach in which data for a group of stations, or a group of grid boxes, are combined in order to increase the sample size. At this time, we have just begun equation development. For our first effort, all sites in the entire U.S. will be combined into one region. The site-specific monthly relative frequencies should allow the developmental procedure and the subsequent forecast equation to distinguish among sites. A separate forecast equation will be developed for each projection and forecast period.

4. VERIFICATION

The performance of the thunderstorm and severe thunderstorm prediction equations will be measured by testing on independent data. Because of the relatively small sample size available from the AVN model for development, a jackknifing technique will be employed to test our predictors, whereby a portion of the developmental data is left out, and the equations are redeveloped. The new equations are then used to forecast for the data that had previously been omitted. At the conference, we will show the results of our tests and will compare the AVN MOS probabilities to the NGM MOS probabilities. The probabilistic forecasts will be evaluated in terms of Brier score and percent improvement over climate.

5. FUTURE PLANS

Recently, TDL began archiving the AVN model output to 78 hours for the 0000 and 1200 UTC runs, as well as the 0600 and 1800 UTC model runs. The next planned enhancement to the thunderstorm and severe weather guidance will be the addition of forecasts for 3-h periods, and the use of AVN predictors valid at 75 and 78 hours. Eventually, we plan to develop guidance from both the 0600 and 1800 UTC cycle model runs.

The MRF-based MOS package is also being revised at this time (Erickson and Carroll 1999). We plan to provide thunderstorm and severe weather forecasts for the 12- and 24-h periods from the MRF run of the GSM out to 144 hours, or as far in advance as skill allows. This work will begin after development of the AVN MOS package.

Acknowledgments

Special thanks are extended by the author to Jennifer Offutt of the Office of Meteorology for providing the severe weather data, to the Global Hydrology Resource Center (GHRC) at the Global Hydrology and Climate Center, Huntsville, Alabama, for providing the lightning data, and finally to Ron Reap for his guidance and effort in setting up this development.

7. REFERENCES

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.

Erickson, M. E., and K. L. Carroll, 1999: Updated MRF-based MOS guidance: Another step in the evolution of objective medium-range forecasts. Preprints 17th Conference on Weather Analysis and Forecasting, Denver, Amer. Meteor. Soc; (this volume, paper 5.1).

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.

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.

Jensenius, J. S., Jr., J. P. Dallavalle, and S. A. Gilbert, 1994: The AVN-based statistical guidance message. NWS Technical Procedures Bulletin No. 415, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 13 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.

, and R. E. Orville, 1990: The relationships between network lightning locations and surface hourly observations of thunderstorms. Mon. Wea. Rev., **118**, 94-108.

Return to Short-Range Guidance page.