6.1 PREDICTING THE LOCATION AND INTENSITY OF LIGHTNING USING AN EXPERIMENTAL AUTOMATED STATISTICAL METHOD

Phillip D. Bothwell*
NOAA/NWS/NCEP/SPC, Norman, Oklahoma

1. INTRODUCTION

Previous Cloud-to-Ground (CG) lightning prediction work using the "perfect prog (prognosis)" approach (Bothwell 2002 and Bothwell 2005) discussed the importance of lightning predictions and why it is not only important to predict lightning occurrence but to distinguish between areas with low lightning flash rates and areas with high lightning flash rates. However, all storms producing CG flashes are potentially hazardous. For example, storms producing only one lightning flash are still capable of causing death or injury. Storms that produce only a relatively low number of CG flashes in dry thunderstorms (those producing lightning with little or no rainfall) are often responsible for starting wildfires in the western United States, accounting for over 50 percent of the total acres burned with suppression costs around a billion dollars each year. Finally, the storms (and/or storm complexes) that produce large numbers of CG flashes are a threat to both lives and property and are often associated with severe weather and/or heavy rain.

Herein, the perfect prog method combines with gridded input data from NWP models, lightning climatology, and logistic regression to derive a probability of lightning for each 40 x 40 km grid box. At this resolution, this method does not predict individual storms; rather it predicts areas where lightning is possible, given the available input data. As the name for the method implies, the results will only be as good as the model predictions are "perfect". Unlike previous lightning prediction methods (such as Reap 1986) which have had little success in predicting areas where large flash rates were possible, this perfect prog approach (Bothwell 2005) has been shown to be capable of predicting areas of significant lightning.

*Corresponding author address: Phillip D. Bothwell, NOAA, NWS Storm Prediction Center, 120 David L. Boren Blvd, Suite 2300, Norman, OK 73072; e-mail: phillip.bothwell@noaa.gov.

In addition to the NWP model data as input, the method incorporates a lightning climatology (Bothwell 2005) derived using data from the National Lightning Detection Network (NLDN). This climatology serves two purposes. It provides the forecaster with information about when and where lightning and significant lightning, in particular, is most likely, as well as serving as input to the perfect prog scheme. Since the probability of lightning varies greatly in space and time, it is critical to relate any probability forecasts of CG lightning to the climatology of the predicted event.

Since 2003, these perfect prog lightning forecasts have been available to forecasters at the Storm Prediction Center (SPC) as guidance in SPC thunderstorm outlooks and for fire weather outlooks (which include forecasts for dry thunderstorms). In a collaborative project with the U.S. Forest Service, the method is being expanded to forecast lightning for Alaska, beginning in the summer of 2008, with full implementation by the summer of 2009. The current perfect prog approach used at the SPC has been updated to extend the forecasts from 60 hours out to 84 hours, using the input from the operational NAM-WRF model. The method produces 3 hourly forecasts for 18 regions (shown in Figure 1) across the lower 48 states on a 40 km grid (Bothwell 2005). The NAM forecasts provide three to four day forecasts of one or more CG flashes as well as 100 or more CG flashes, and are updated every 6 hours with each new NAM run. Additionally, while the perfect prog method discussed in this paper includes the results from the NAM forecasts, it also produces forecasts using the RUC model data input and the SPC hourly 3-dimensional gridded analysis (Bothwell et al. 2002). This multi-model system provides forecasts from 84 to 87 hours (NAM) down to 0 to 3 hours (e.g., hourly 3 hour forecasts) using this perfect prog method.

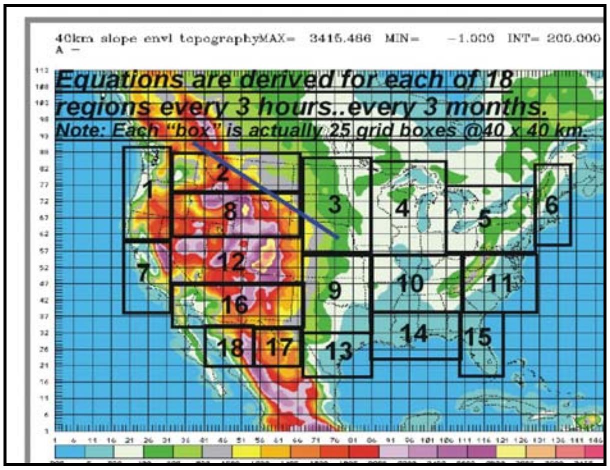

Figure 1. Location of each of the 18 regions that use separate prediction equations. Each of the large grid squares shown above contains 25 of the 40 km grid boxes.

2. SUMMER 2006 AND 2007 FORECASTS

As part of an experimental program for the Western Region (WR) National Weather Service (NWS) meteorologists and the forecasters from various Federal wildland fire agencies in the western United States, a set of automated (perfect prog derived) CG lightning forecasts have been produced during the summers of 2006 and 2007 (during the peak of the fire season) as guidance for the prediction of dry thunderstorms. In addition to the lightning forecasts, separate guidance has been supplied in the form of the Dry Thunderstorm Potential Index. The DTPI (either forecast from the NAM model or an analysis) is a non-dimensional measure of the height of the cloud base above ground level and the dryness of the sub-cloud humidity, both of which strongly influence the evaporation rate of rain falling from a thunderstorm. The DTPI is scaled to produce a number from 0 to 100 with numbers from 50 to 100 generally indicating the greater likelihood of a dry thunderstorm. However, the DTPI is not strictly a probability. Mean storm motion vectors are also shown on the graphics. These experimental/automated forecasts are available to the forecasters at the SPC and in the western U.S. over an SPC internet web page. The forecasts are updated every 6 hours with each successive run of the NAM model.

The prediction method using NAM fields produces 3 hourly forecasts out to 87 hours, but this 3 hourly temporal resolution is displayed only for Day 1. Day 2 forecasts are combined into 6 hour periods for display, and the Day 3 forecasts are combined into 24 hour forecasts. This is accomplished by simply taking the maximum probability from any of the included time periods. In 2007, in response to forecaster requests, the 3 hour time periods were combined (again, by taking the maximum probability from any of the 3 hour periods) to produce "quick look" 24 hour forecasts for each of Days 1, 2, 3 and for a partial Day 4 forecast.

In 2007, experimental guidance was expanded to the Weather Forecast Office (WFO) in Raleigh, NC, at their request, as part of their effort to forecast high flash density CG lightning events for the early morning Hazardous Weather Outlook. The Web Page was updated to include the area in and around Raleigh. The guidance produced was similar to that for the western United States; however, the emphasis was on lightning events of 100 or more CG flashes. These forecasts, along with the forecasts of 1 or more CG, are evaluated in the next sections.

Additionally, an archive web site of all forecasts was made available to forecasters in 2006 to subjectively aid in forecast verification. Beginning with the summer of 2007 (with the permission of the NLDN parent company, Vaisala), the web page forecasts included the actual number of lightning strikes per 40 km grid box for each 24 hour period corresponding to the forecast time period. To provide information about the character of the storms (wet vs. dry), matching precipitation images were available from the archive directory. The number of hours lightning occurred in the grid box was also available from the archive.

3. VERIFICATION - 1 OR MORE CG FLASHES

In order to evaluate how the forecasts (06 UTC NAM input data) for 1 or more CG flashes verified during the summers of 2006 and 2007, each of the grid point 24 hour forecasts and the corresponding 24 hour lightning were statistically verified, and the forecast versus observations were plotted on a reliability diagram (see Fig. 2). The forecasts were divided into the West versus the East to assess forecast performance in different climate regimes. Moisture is normally substantially higher in the Central and Eastern U.S. than the West. The forecasts were grouped and plotted for regions east of 102 degrees longitude (central and eastern U.S.) and west of

102 degrees longitude (western U.S). The 06 UTC perfect prog NAM forecasts were evaluated, rather than the forecast from the 12 UTC run, because the 06 UTC forecasts were available to the forecasters in the early to mid morning at a time when they were evaluating the potential for dry thunderstorms.

Figure 2 shows that under-forecasting was occurring in both regions at low to moderate forecast probabilities, while over-forecasting was evident at the higher probabilities. Under-forecasting occurs when one or more CG flashes are observed more frequently than forecast. Over-forecasting (above 65 to 75 percent) occurs when lightning is not observed as often as forecast. On several occasions of high probability forecasts, it was found that the lower level moisture forecast from the NAM was significantly over-forecast. This was noted during the summer of 2007, across southwest Montana and parts of the Northwest, and could account for some of the over-forecasting.

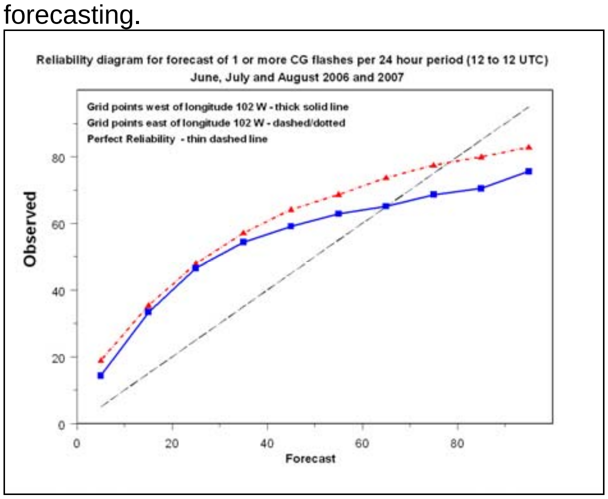

Figure 2. Reliability diagram showing forecasts for western U.S. (thick solid line) and eastern U.S. (dashed-dotted line).

Since these are forecasts of one or more CG flashes per grid box, part of the reason for the under-forecasting in the lower probability range is that low numbers of isolated flashes (e.g., 1 or 2 CG flashes per grid box) are often in areas that are weakly forced and more subject to topographical influences. Further, since the moisture for storms in the West is often reduced due to mountains blocking any appreciable moisture, small changes in the moisture field are difficult for any operational model to correctly resolve. Both weak forcing and unresolved topographical influences are difficult to resolve for many forecast models. In addition, the forecast equations were developed with only two years of training data. It is likely that results will improve with additional years of training data.

Figure 3 is one example of how the forecasts and verification lightning (at 40 km resolution) appear. Ideally, the observed lightning will be overlayed where there are forecast probabilities and no lightning will be observed where the probabilities are low (e.g. less 1 percent in this case). The lowest forecast probabilities from 1 to 10 percent were plotted in order to show where some isolated flashes might still occur, although given a probability from 1 to 10 percent, very few flashes would be expected. Secondly, these low-end probabilities were displayed to demonstrate the skill in this method by focusing in on small areas with low probabilities. The low probabilities do not vary widely. That is to say, they exhibit good continuity in both space and time. Finally, it was found that the NAM forecasts changed only slightly from one model run to the next. As might well be expected, forecasts normally converged to the "correct" forecast/solution as the event became closer in time. Figures 4 and 5 are shown to illustrate these results from the 06 UTC (Fig. 4) to the 12 UTC (Fig. 5) run.

4. VERIFICATION - 100 OR MORE CG FLASHES

In much the same manner as discussed in Section 3, the forecasts for the areas in and around Raleigh, North Carolina were evaluated, for 100 or more CG flashes. The number currently chosen here (i.e., 100 or more) can be adjusted for other CG flash densities. The selected flash density should be determined by the lightning climatology for the particular area. That is to say, for example, a high flash density for Montana will be significantly different from almost any area in Florida on any given day. Most often, the synoptic and mesoscale (e.g., sea breeze) forcing is stronger, and the available moisture is greater, for the higher flash events.

Figure 6 shows that the forecasts for the Carolinas and Virginia areas have good reliability up to 25 percent, which is well above the climatology for the area. Above that, there are significant overforecasts with limited resolution.

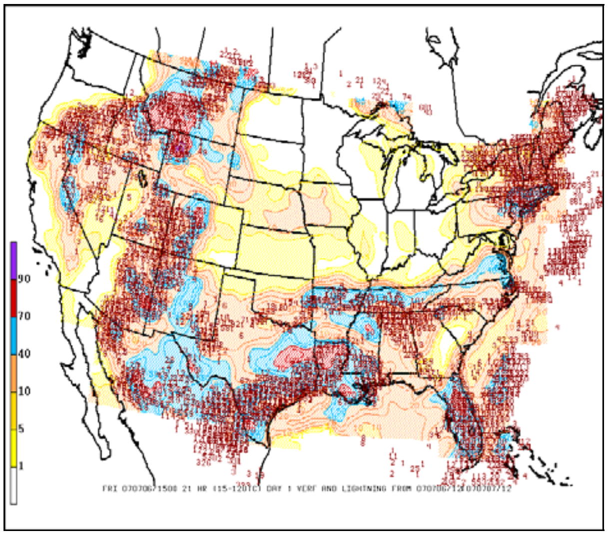

Figure 3. Nationwide probability forecast (contours/color filled image) for 1 or more CG flashes (40 km resolution) and number of lightning flashes plotted for each grid box for 24 hour period from 12 UTC 6 July to 12 UTC 7 July 2007. (First 3 hour forecast period from 12 to 15 UTC is not computed/included in the time period for the forecast because of timing constraints).

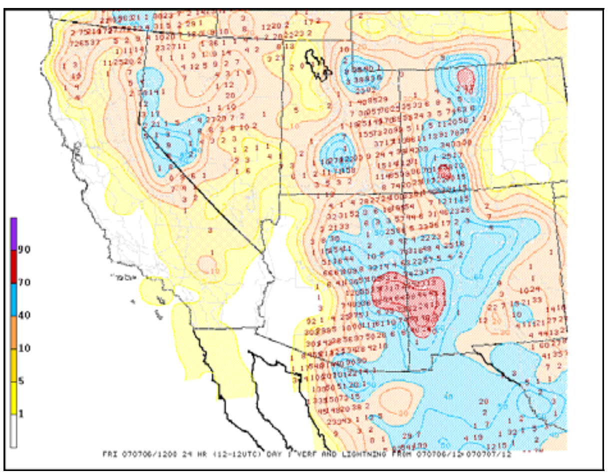

Figure 4. 06 UTC NAM 24 hour forecast for 1 or more CG flashes and plotted number of CG flashes. Same time period as Fig. 3. The largest wildfire in Utah history was started by single strike in southwest Utah.

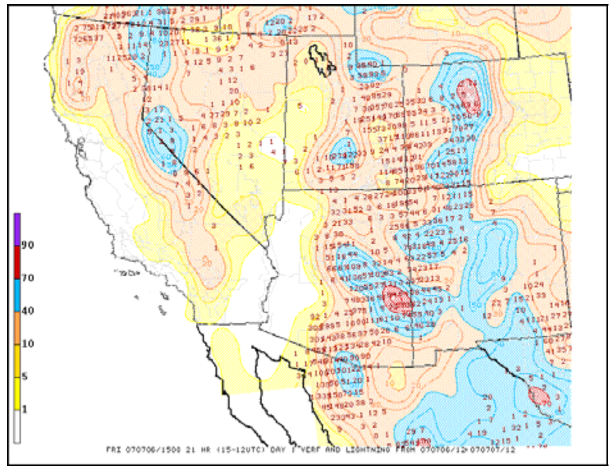

Figure 5. Same as in Figure 4, but produced from the 12 UTC NAM run. (First 3 hour forecast period from 12 to 15 UTC is not computed/included in the time period for the forecast because of timing constraints).

A sample 24 hour forecast of 100 or more CG flashes is shown in Fig. 7. This was a major lightning event for eastern parts of Virginia, North Carolina and South Carolina. The contoured plot of lightning flashes for the same time period, along with the severe weather reports for the 24 hours, are shown in Figs. 8 and 9. The pseudo-log lightning contouring scale of 1, 3, 10, 30, 100, 300, 1000, 3000 and 10000 has been found to be a valuable way to graphically highlight significant lightning events.

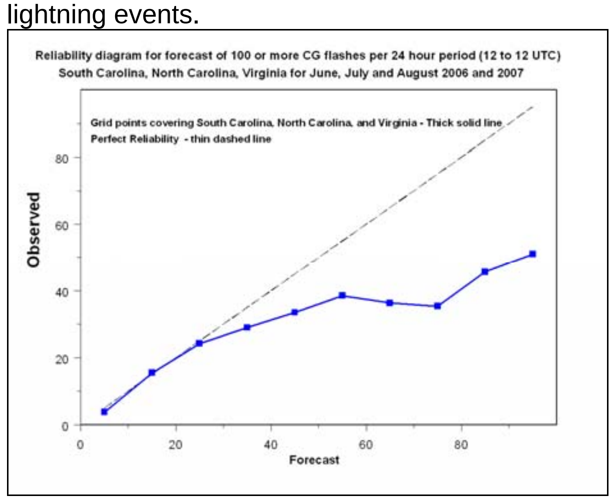

Figure 6. Reliability diagram showing forecasts for East Central U.S. (thick solid line).

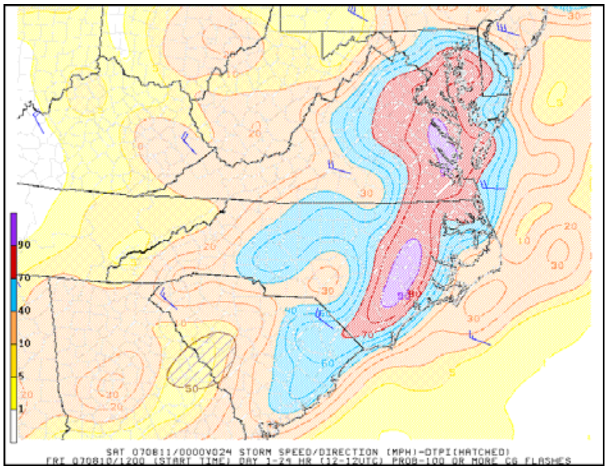

Figure 7. 24 hour forecast for 100 or more CG flashes from 12 UTC 10 August to 12 UTC 11 August 2007. This forecast was produced from the 06 UTC NAM run on 10 August 2007.

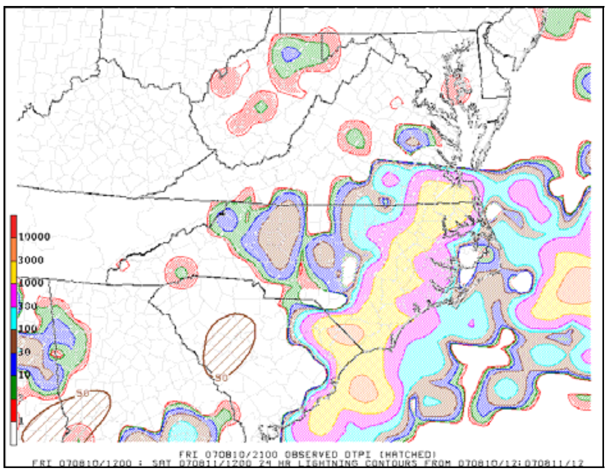

Figure 8. Contours of number of CG lightning flashes per 40 km grid box for same time period as forecast in Fig. 7.

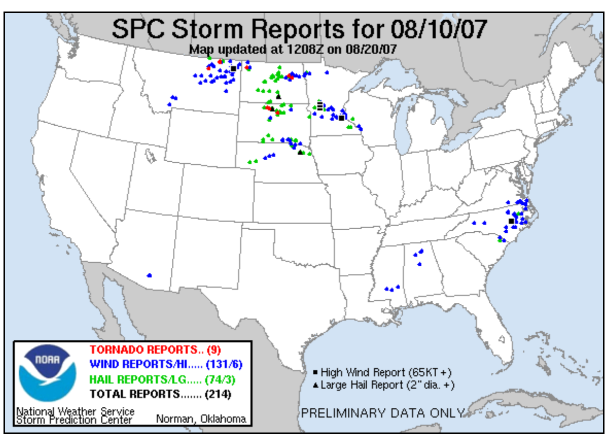

Figure 9. Rough log of severe reports for same time period in Fig. 7.

5. SUMMARY AND FUTURE WORK

Verification efforts are on-going. In particular, the actual 3 hour forecasts that comprise the 24 hour forecasts will be verified in the future. The initial training/development data set that produces the original perfect prog forecast equations were derived from only two years of model data. This was due to a lack of longer time periods of archived model data available at the time, and associated computer processing limitations.

Both of these issues no longer represent a problem to development, and it is anticipated that upwards of 10 years of development data will be available in the near future to update the current equations. Subjectively, it appears that the primary limiting factor affecting the quality of the lightning predictions is associated with model forecasts of the lower level moisture. If the model(s) do not accurately forecast this lower level moisture, the predictions suffer. This is one of the reasons the addition of the GFS model to the perfect prog forecasts is planned. Model Ensemble forecast output also represents a potentially powerful data set that can be incorporated by the perfect prog method. Ensembles can mitigate many of the problems with a single deterministic model prediction of the lower level moisture. Despite the current limitations, in nearly all cases, the perfect prog forecasts converge toward the correct areas as the lightning event approaches.

This work is being expanded to include predictive equations for CG lightning over Alaska in 2008 and 2009. For 2009, the goal is to have predictive equations at resolution around 15 km vs. the 40 km reported in this paper. Also, the on-going work will redesign the perfect prog equations to run on the NAM and GFS models and eventually, the Rapid Refresh model. This will be at a resolution around 15 km over the CONUS, similar to that for Alaska. For both Alaska and the lower 48 states, the development data (sample size) will be greatly increased from the limited two-year training data set used on the current equations. This should result in much more robust equations for both the low-end and high-end lightning events. Finally, a longer term goal is to integrate the perfect prog lightning prediction method with the NWP Ensembles.

6. REFERENCES

Bothwell, P.D., 2002: Prediction of Cloud-to-Ground Lightning in the Western United States, Ph.D., University of Oklahoma, 178pp.

___, J.A. Hart, and R.L. Thompson, 2002: An integrated three-dimensional objective analysis scheme, 21st Conf. on Severe Local Storms, Aug. 12-16, San Antonio, TX, Amer. Meteor. Soc., 4 pp.

___, 2005: Development of an operational statistical scheme to predict the location and intensity of lightning. Conference on Meteorological Applications of Lightning Data, San Diego, CA, Amer. Meteor. Soc., 6 pp.

Reap, R.M. 1986: New 6-h thunderstorm probability forecasts for the West. NWS Technical Procedures Bulletin No. 362, NOAA, U.S. Department of Commerce, 8pp.