Performance of a Probabilistic Cloud-to-Ground Lightning Prediction Algorithm

John Cintineo 1,2,3* Valliappa Lakshmanan 1,2, Travis Smith 1,2

Abstract

A probabilistic cloud-to-ground lightning algorithm was created by training a neural network on storm characteristics. The input dataset consisted of all storm cells over the entire coterminous United States on 12 days in 2008-2009 (one day per month). The input characteristics include radar and near-storm environmental parameters and the neural network was set up so that its output is the probability of cloud-to-ground lightning at a grid location 30 minutes in the future. The probabilistic output was evaluated on twelve independent test dates in 2008-2009 and results of that evaluation are presented.

* Corresponding Author: J. Cintineo, 120 David L. Boren Blvd, Norman OK 73072; john.cintineo@ou.edu [1]Cooperative Institute of Mesoscale Meteorological Studies, University of Oklahoma; [2]National Oceanic and Atmospheric Administration / National Severe Storms Laboratory; [3]School of Meteorology, University of Oklahoma

1. INTRODUCTION

Accurate temporal and spatial prediction of lightning is important since cloud-to-ground lightning is hazardous to life and property (Curran and Holle, 1997). Predicting when and where lightning will occur has proven to be a difficult problem, since it is tied to convective initiation, as well as a good spatial and temporal understanding of electrification inside the thunderstorm. A pixel-by-pixel input-output mapping of probability of lightning at 3-hour intervals was successfully developed at a 22 km resolution to train a neural network (Burrows, et al. 2005). At the finer resolution of this study (i.e. about 1 km), however, pixel-by-pixel input-output mapping will not work (explained in Lakshmanan & Smith, 2009), and so it is necessary to treat storms as objects and train the neural network with storm properties, and not simply pixel values. Carey & Rutledge (2003) examined regional variability of cloud-to-ground lightning in severe and non-severe thunderstorms in the north central U.S., and acknowledged the subjectivity of their definition of severe vs. non-severe storms. Another regional analysis investigated lightning initiation around Cape Canaveral, FL, and used the presence of a 10 dBZ echo at the 0^{o} C isotherm as a criterion to judge the onset of cloud-to-ground lightning in a storm (Hondl & Eilts, 1994). These two studies had conflicting thresholds for thunderstorms, due in part to their regional extent. The goal of this study is to further the development of a single, national algorithm to forecast the probability of cloud-to-ground lightning.

A reliable short-term forecast of intense lightning could be a very useful tool for the U.S. National Weather Service. Lakshmanan and Smith (2009) describe a method of predicting the probability of lightning at a given grid location produced by a storm in the next 30 minutes. Their method involved the training of a neural network on input spatial clusters of radar reflectivity, vertically integrated liquid (VIL), and near-storm environment parameters. They explain that it is necessary to treat storms as entities and train the model with storm properties, and not just on pixel values from a grid. This is namely because even though, for instance, strong reflectivity at -10^{o}C is a good indicator of lightning, the cloud-to-ground strikes may not necessarily occur at the location of the maximum reflectivity at -10^{o}C, but may very well occur elsewhere in the storm, such as in the anvil region.

The neural network lightning prediction algorithm was evaluated on all storms over the CONUS on 12 days (one per month) from 2008 and 2009 (three from 2008; nine from 2009), which were independent from the days the algorithm was trained on. Evaluation methods include traditional verification scores, as well as some probabilistic verification techniques such as the Brier Score and Fractions Skill Score.

2. METHOD

A list of potential predictors of cloud-to-ground lightning used to train the neural network are given in Lakshmanan and Smith (2009) and are based on Hondl and Eilts (1994). The predictors include different reflectivity and reflectivity-derived products, as well the current (the time when the forecast is made) lightning density (units of flashes minute^{-1} km^{-2}), and Lagrangian aspects of the storm such as size and speed. The goal is to ultimately provide geometric, spatial, and temporal properties of a storm, in real-time, as

inputs to the neural network to generate a predicted lightning probability field, or cluster, in an operational environment, then advect the lightning probability cluster forward in time by 30 minutes to create a nowcast for lightning. The neural network outputted a cluster of uniform probability of cloud-to-ground lightning every 15 minutes for every "storm" that was tracked. The tracking was done on merged reflectivity clusters, which can include regions of weak and intense reflectivity. Therefore, every output probability region does not necessarily correspond to a thunderstorm, but a region of reflectivity. However, regions of weaker reflectivity would tend to have much smaller probabilities of lightning associated with them, since the algorithm tends to yield higher probabilities for stronger radar reflectivities. Probabilities range from 0.01 to 1.0 incremented by 0.01. Storms are identified using an enhanced Watershed algorithm (Lakshmanan, 2009) and tracked using cross-correlation and Kalman filters (Lakshmanan et al. 2003). Spatial and temporal attributes of clusters are extracted as described in Lakshmanan and Smith (2008). A storm is defined as a group of pixels that meet some size criterion, or "saliency," whose intensity values (in terms of radar reflectivity, here) are greater than a minimum value criterion. The clustering saliency that we chose was 200 km^2, which performed the best in terms of hit rate. The minimum reflectivity value we chose to threshold storms at was 20 dBZ. It is reasonable to expect, however, that different cluster saliencies, and perhaps different storm definitions (in terms of reflectivity) may be more appropriate for different forecast time periods.

The probability of lightning regions were evaluated using some traditional verification scores, including the probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and Heidke Skill Score (HSS) (e.g. Wilks, 2006). Cloud-to-ground strikes from the National Lightning Detection Network (NLDN) were averaged in space (3 km radius) and time (15 minutes) to create grids of lightning density. A location is said to have experienced lightning if there is a cloud-to-ground strike within a given distance to that point within the past 15 minutes. Fig. 1 shows the correspondence of the lightning density and lightning probability clusters for a given time, as well as how the algorithm was able to predict lightning initiation. The clusters are all created on a grid with resolution of 0.01 degrees latitude and longitude (about 1 km by 1 km in the mid-latitudes).

If any area fraction of a lightning probability cluster, forecasted at time t_0, intersects any area fraction of a lightning density cluster at $t_{0+30min}$, then that particular probabilistic forecast is considered a 'hit.' A 'false alarm' occurs if the probability cluster intersects no area of lightning density, and a 'miss' may also occur, if a lightning density cluster did not have a lightning probability forecast covering any of the area. This 'double penalty' was partially ameliorated in two ways: firstly, a search radius of 10 pixels (about 10 km) for each probability cluster was employed; secondly, probability clusters and lightning density clusters were ignored if they did not meet the size threshold of at least 50 km^2. The search radius helps forecasts that are relatively close, and it also helps capture anvil strikes that are away from the core of the storm. The size threshold helps eliminate single strikes far out from the core, in the outer perimeter of the anvil, which would yield a low lightning density cluster that is spatially unconnected with the rest of the parent storm. The size threshold does not neglect very small storms. Instead it ignores weak storms. A storm that creates a 50 km^2 cluster of lightning density had just one lightning flash since a single

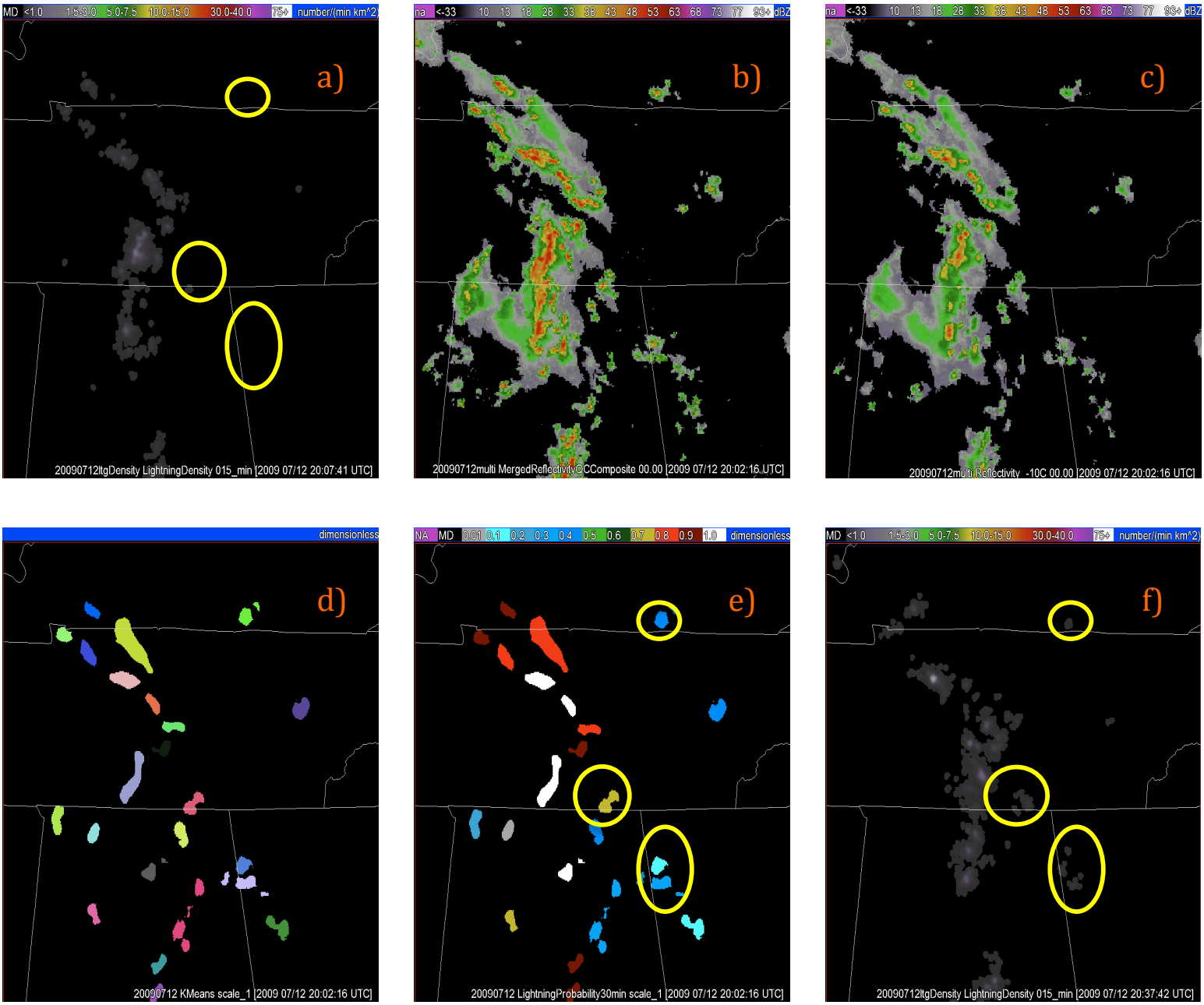

Fig. 1: Representation of how the algorithm tracks storms and captures lightning initiation. The date is 12 July 2009, one of the evaluation dates, and geographically, this region depicts the southeast U.S., including Tennessee, southern Kentucky, northern Alabama and northern Georgia. **a)** Initial observed lightning density clusters at time t_0, which is 2002-2007 UTC (depending on the product). **b)** The composite reflectivity field (lowest tilt) at t_0. **c)** The reflectivity at the -10° C isotherm at t_0. **d)** Clusters identified in the reflectivity composite (at time t_0, as described in Lakshmanan et al. 2003.) **e)** Predicted clusters of probability of cloud-to-ground lightning at t_0+30 minutes (2037 UTC). Yellows, reds, and white indicate higher probabilities (~70-100%), and blues indicate lower probabilities (~20-40%). **f)** The actual observed lightning density clusters, at t_0+30 minutes. Yellow circled regions are example areas of where the algorithm captured new initiation of cloud-to-ground lightning strikes. These are areas where there was no lightning observed at t_0 (in **a**) but it was forecasted at t_0 (in **e**) and it was indeed observed at t_0+30 minutes (in **f**).

lightning flash is averaged in space of a 3 km radius (1 km^2 +/- 3 km in +/- x and y directions yields a 7 km by 7 km box equal to 49 km^2). In essence, we are ignoring storms that produce less than 2 flashes in 15 minutes.

Correct nulls are computed, but not in the traditional manner. Since cloud-to-ground lightning is such a low probability phenomena, climatologically speaking, any areas that do not have lightning density clusters or probability clusters could be considered correct nulls. However, this would saturate the standard 2-by-2 contingency table (e.g. Wilks, 2006) with an enormous amount of correct nulls. In this study, correct nulls are computed based on a probability threshold. Any probability cluster with a probability below that threshold that would normally be scored a false alarm will now be considered a correct null, and any cluster with a probability at or above that threshold will be treated as usual, and scored either a hit or false alarm. Since only probability clusters (and therefore only tracked clusters of merged reflectivity) can be considered, we are essentially rewarding the algorithm for correctly not forecasting lightning in areas of reflectivity where there was no lightning, and not simply counting everywhere in the CONUS where there was not lightning as a correct null. For example, at the 20% threshold, all probability clusters less than 20% that contained no lightning density will be considered a correct null instead of a false alarm. It should be noted that the Heidke Skill Score couldn't be computed at the 0% threshold (all probability clusters) since (by our definition) we have no information about correct nulls.

3. RESULTS

The days evaluated from 2008 were: 9 October, 11 November, and 10 December; from 2009: 3 January, 11 February, 28 March, 10 April, 8 May, 9 June, 12 July, 9 August, and 8 September. Days were chosen such that there was at least some lightning in the CONUS. Probability of lightning forecasts were produced every 15 minutes of each day, and the aggregate results included: 31,285 hits; 28,462 misses; 111,663 false alarms; and 70,934 observations. The number of hits and misses do not add up to the total amount of observations since a probability cluster may contain one or more lightning density clusters, or observations.

Figure 2 displays the aggregate POD, FAR, CSI, and HSS for the twelve evaluation days as a function of probability threshold, in a binary sense. Any forecast probability at or above the threshold indicates a "yes" forecast and below the threshold indicates no forecast. The maximum CSI (0.325) is produced at the 50% threshold, which is of course, the threshold that would produce non-probabilistic forecasts ("yes" or "no") of the more likely of the two events (lightning or no lightning).

The Heidke skill score uses the proportion correct (sum of hits and correct nulls divided by sample size, [e.g. Wilks, 2006]) as the accuracy measure. Thus, skill is relative to that of random chance. HSS = 1 for a perfect forecast; a forecast equivalent to the reference forecast receives a score of zero; and HSS < 0 are given to forecasts worse than random chance. With the exception of the 10% probability threshold, which had negative skill, all other probability thresholds had positive skill with respect to random chance, with a maximum HSS at the 70% threshold, of 0.307.

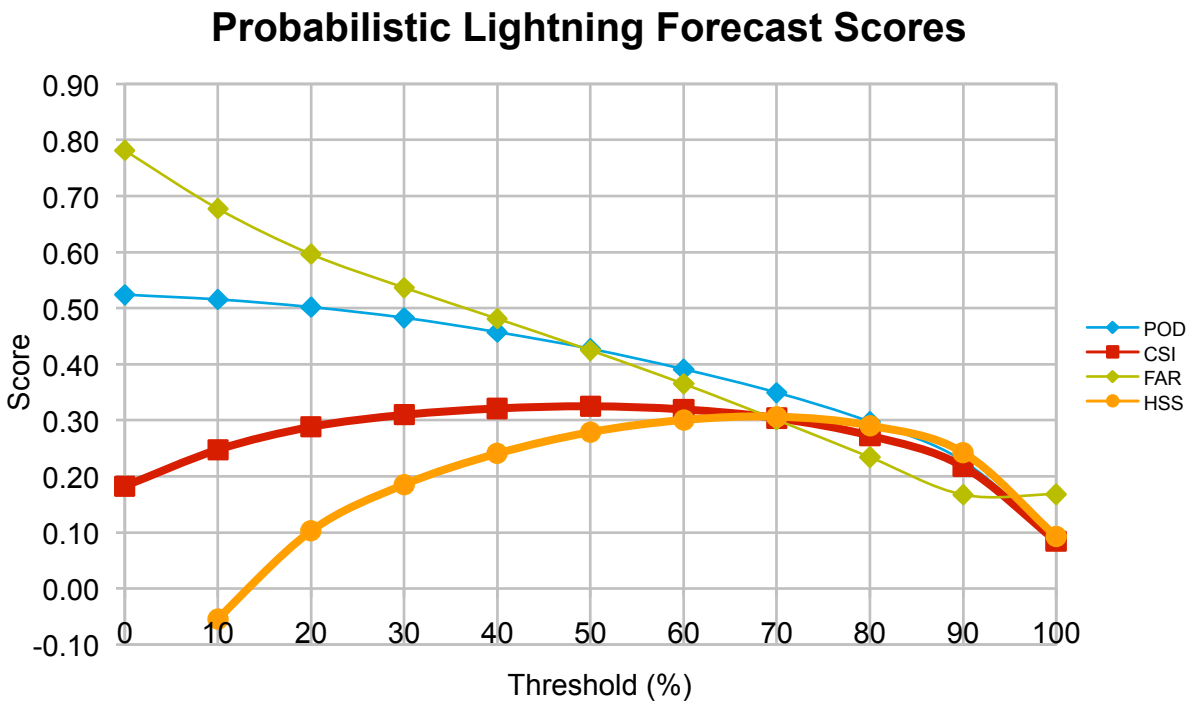

Fig. 2: POD (blue), FAR (green), CSI (red), and HSS (orange) as a function of probability threshold, which yields binary forecasts.

The maximum CSI varied by season and by threshold, as shown in Figure 3. The cool season (December to March, here) had maximum CSI's below 0.2, while the warm season days (April to November) all had a CSI of 0.25 or greater. It should be noted again that only one day per month was evaluated, so these results may not be entirely representative of the monthly population, however, it is interesting to see the seasonal difference in CSI with our small sample size. One can also see how the threshold at where the maximum CSI occurred ranges from the 30% threshold to the 80% threshold, with the cool season months tending to have a lower threshold value than the warm season months.

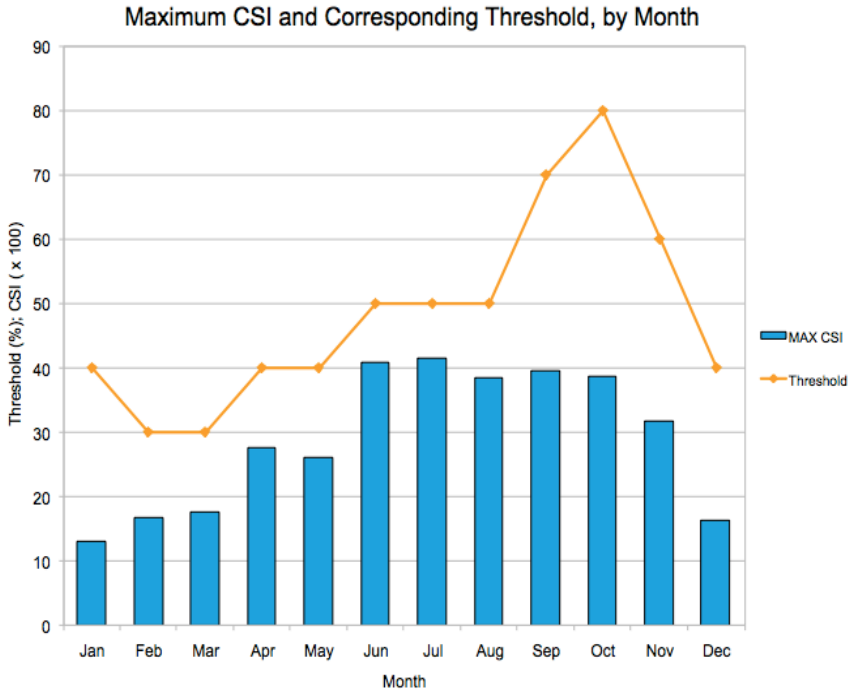

Fig. 3: Maximum CSI for each month (blue bars), and at what threshold that CSI occurred (orange line). CSI is multiplied by 100 to scale it similarly to threshold.

A maximum CSI at a higher threshold may indicate that the lightning density clusters tend to be stronger that day, such that it is likely that more high-probability clusters would be forecast than

normal. In this case, our CSI should improve, since we would have fewer low-probability clusters (and therefore fewer false alarms), as well as more hits (since the higher probability clusters would tend to capture lightning density clusters more often). Both of these results would produce a better CSI at a relatively higher threshold, which is what we see in the warm season months. The cool season months have maximum CSI at lower thresholds (30-40%) since those dates tend to have an abundance of low-probability clusters, and as you increase in threshold, the number of false alarms would increase greatly, reducing the CSI.

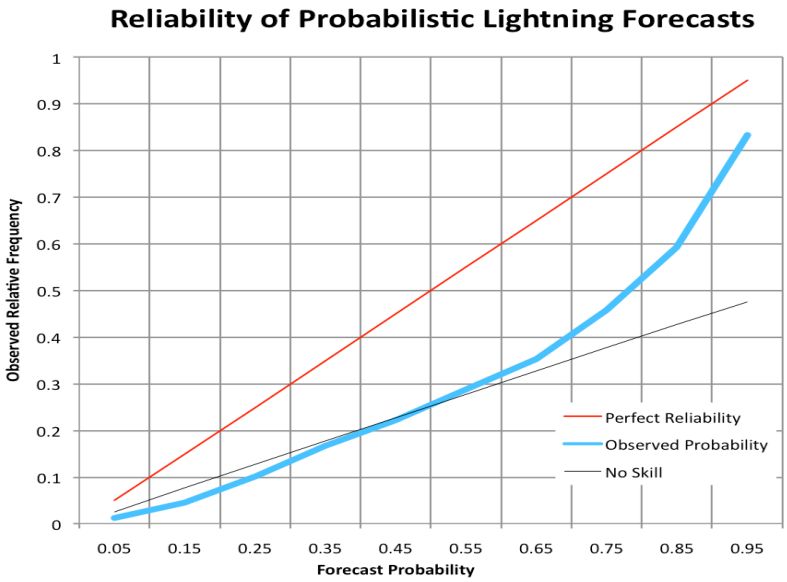

Fig 4: Reliability diagram of the probabilistic lightning forecasts. The blue line plots the observed probability (frequency) of lightning against the forecasted probability. The red line indicates perfect reliability.

The reliability and sharpness of the lightning prediction algorithm are shown in Figures 4 and 5, respectively. The observed probabilities are binned into 10% increments, centered around the values on the abscissa, and inclusive on the upper bound. However, the forecast probability bin of 1 includes only probabilistic lightning forecasts equal to unity. Ideally, we would like to see the observed probability along the line y=x, corresponding to perfect reliability, that is, the probability of the observations given the forecast should equal the actual forecasted probability. Towards both extremes, the algorithm appears to be rather reliable, with the observed probability lining up close to the perfect reliability line. In the middle forecast probability bins, the algorithm is clearly over-forecasting, yielding observed probabilities less than the actual observed frequencies of lightning, i.e. the forecasted probabilities of lightning underestimate the observed frequencies of lightning.

A sharp algorithm is desirable as well, since it tends to forecast extreme values, which are 0.01 and 1 in our case. Keep in mind that the sharpness of a set of forecasts does not say anything about the correspondence between the forecasts and observations, but is only a metric of the forecasts themselves. The algorithm tends to

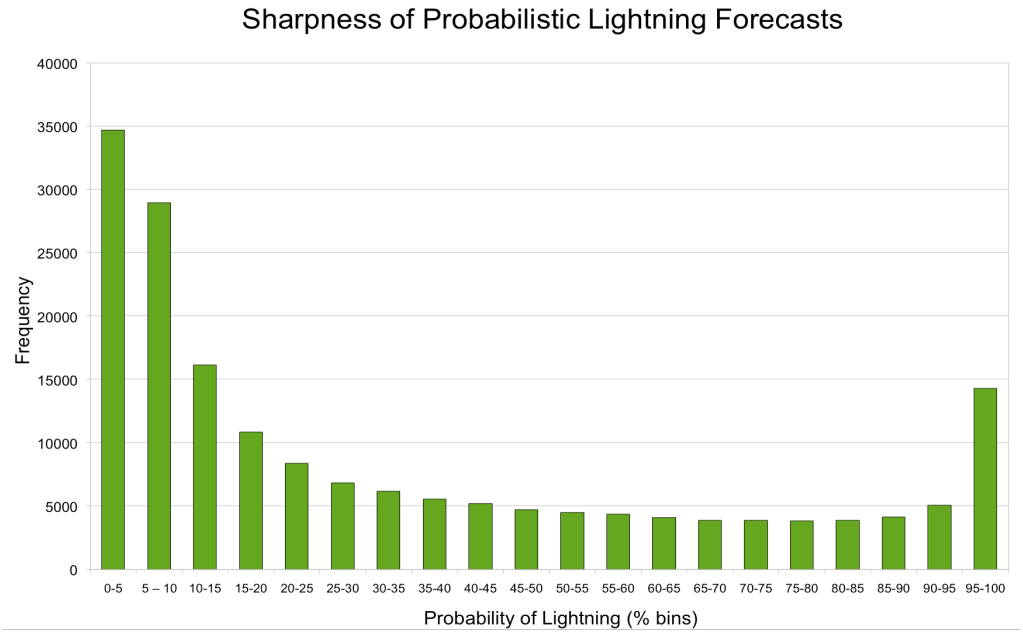

Fig. 5: Sharpness diagram for the probabilistic lightning forecasts. The forecasts are binned into 5% increments, inclusive on the upper bound.

produce more extreme-probability forecasts and less medium-probability forecasts, which is desirable (Figure 5).

The Brier Score (Brier, 1950) is a summary measure for probabilistic forecasts for dichotomous events. A single number will certainly not give a complete picture of forecast performance, but the Brier Score is an excellent choice from a practical standpoint, given its decomposition into reliability, resolution, and uncertainty (Wilks, 2006). It is given by:

$$BS = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2$$

where y_k is the k^{th} forecast and o_k is the corresponding k^{th} observation (either 0 if lightning did occur, or 1 for no lightning, in our case). The Brier Score is basically the mean squared error for probabilistic forecasts, and therefore, the values range from $0 \leq BS \leq 1$, 0 being perfect. Since the algorithm never forecasts 0% probability of lightning, y_k was always between 0.01 and 1.0. Calculating the aggregate *BS* in this fashion for all 142,948 forecasts, we receive a BS of 0.1038, which is quite good. This is in part due to the large number of low-probability forecasts (see Fig. 5). However, this method of computing the *BS* does not take into account misses, or "forecast probability of 0% when lightning did occur." Enter misses into the equation and our $BS = BS_{misses} = 0.2269$, which isn't as skillful as before, but is still below 0.25, which is the *BS* resulting from uniform probability forecasts of 0.50. This still doesn't seem quite fair though, since correct nulls have not been considered. In order to do so, we need to calculate the *BS* at some probability threshold. Figure 6 displays the Brier Score as a function of probability threshold. The relatively large drop in *BS* from the 0 to 10% threshold is do to the fact that we have no correct null concept at a threshold of 0 (i.e. all forecasts). The Brier Score achieves a minimum of 0.2115 at the 100% threshold. Again, false alarms at forecast probabilities below the threshold now become correct nulls, and help lower the Brier Score. Since we had many false alarms below 100%, we might expect this threshold to be our minimum, or best *BS*.

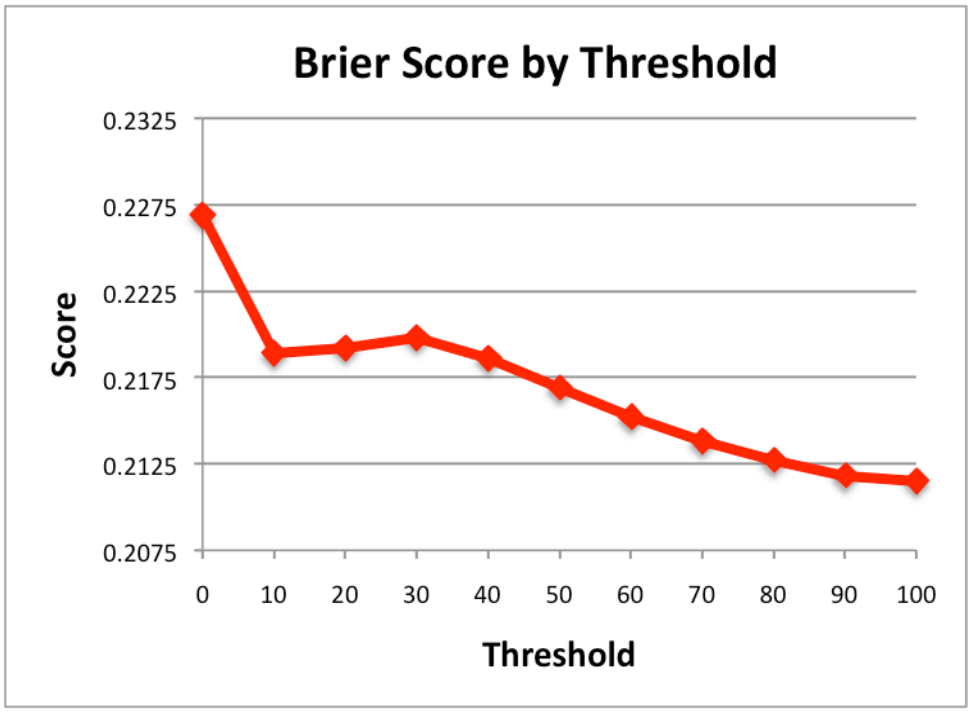

Fig. 6: The Brier Score, threshold by forecasted probability of lightning.

The Fractions Skill Score (*FSS*) (Ebert 2009; Roberts and Lean 2008) compares the forecast and observed lightning density coverage. In our case, the forecast lightning density can be considered as the probability clusters output from the algorithm. The *FSS* gives the evaluator an idea of

how 'realistic' the forecast looks to the observations. It is typically used for rain coverage, but will suffice for lightning coverage. It is given by:

$$FSS = 1 - \frac{FBS}{\frac{1}{N}\left[\sum_N P_{fcst}{}^2 + \sum_N P_{obs}{}^2\right]} \qquad \text{where} \qquad FBS = \frac{1}{N}\sum_N (P_{fcst} - P_{obs})^2$$

FBS is the Fractions Brier Score, and is completely analogous to the traditional *BS*. P_{fcst} and P_{obs} are the fractional forecast and observed lightning areas. Ebert (2009) and Roberts and Lean (2008) show that the value of *FSS* above which the forecasts are considered to have "useful" skill (better than a uniform probability forecast of f_{obs}, which is the observed lightning fraction over the domain) is given by:

$$FSS_{useful} = 0.5 + \frac{f_{obs}}{2} \approx 0.5$$

which we approximate to 0.5 in our study, since the fractional coverage of lightning over the CONUS is typically less than 0.01. A search radius from the centroid of each probability cluster was employed, in order to capture the centroid of the corresponding cluster of lightning density. A search radius of 0.25° of latitude and longitude (about 25km) was used. Since some probability clusters were much larger than the lightning density clusters that they overlapped (and vice versa), we were not able to capture every hit, as the centroids of the two clusters were simply too distant from each other. The 0.25° search radius captured most of the hits (80%), while not encroaching on the overlapping lightning density clusters of other probability forecasts. In this case, our *FSS* = 0.6289, which is greater than FSS_{useful}. We expect that the *FSS* would score quite well however, since the forecasts are only for thirty minutes into the future, and we are only scoring the hits. Scoring misses and false alarms may result in a double penalty, since there is no corresponding probability cluster or lightning density cluster, respectively. If the domain were divided into neighborhoods (as in Ebert, 2009), then we could measure the fractions of each cluster in each neighborhood. On the contrary, since most of the domain does not contain either probability or lightning density clusters at any given time, the *FSS* would be biased with "correct nulls" of fractional coverage (i.e. most neighborhoods would score perfectly since they have neither cluster). So both misses and false alarms were neglected when computing the score. Thus, the *FSS* computed here should be considered for what it is: a score that examines how well the forecast probability clusters match the lightning density clusters for (most of) the hits.

4. DISCUSSION

This neural-network trained probabilistic lightning prediction algorithm is an important step in accurately now-casting cloud-to-ground lightning, in a probabilistic fashion. However, there is much work yet to be done before such an algorithm can be used operationally as a NWS product. Firstly, we recognize that the probability clusters are uniform probability, clustered around the core of a storm. Lightning often strikes outside

of the core, in the anvil region. Lightning that occurs away from the core of the storm is often more dangerous, since victims can be unprepared or unaware of their proximity to danger. Therefore, the probability clusters that the algorithm produces should have contours of probability, and not be uniform. The core of a storm would still most likely have the highest probability of cloud-to-ground lightning, but the anvil regions of storms should have some lesser probability associated with them. There are several ways to consider this procedure. One way is to track the cores of storms on reflectivity of the -10°C isotherm surface (the importance to lightning prediction demonstrated by Hondl and Eilts, 1994 and Vincent et al. 2003), or perhaps -20°C isotherm surface, and then track storms on merged reflectivity at some lower threshold of reflectivity, to capture the anvil regions of the storms. A second approach is to use satellite imagery, similar to the methods proposed by Mecikalski and Bedka (2006), MB06 hereafter.

In MB06, Geostationary Operational Environment Satellite (GOES) imagery is used to determine convective initiation of storms (as opposed to cloud-to-ground lightning), as well as mature cumulonimbus clouds. In particular, they employ the extensive use of brightness temperature, T_B, visible satellite imagery (VIS) and different bands of infrared (IR) imagery. MB06 use gradients of T_B and VIS to first "mask" a section of a GOES-11 or -12 image in order to compute IR-based products only where convective clouds are present. These products include the 10.7μm T_B, multispectral channel differences (e.g. 13.3-10.7μm difference), and time derivatives of multispectral channel differences (e.g. ∂[13.3-10.7μm]/∂t), all in order to predict convective initiation, defined in MB06 as first reflectivity ≥35 dBZ produced by convective clouds. Using eight different IR-based products or "interest fields," MB06 claim their method (which builds on prior work) achieves ~60-70% accuracy, and lead times ~30-45 minutes.

Utilizing different IR-satellite products could help to improve the algorithm. Potentially, the algorithm could apply MB06's method to capture *only* convective clouds, and combine this with radar reflectivity. Therefore, the algorithm would no longer be looking at stratiform complexes, which have little chance of producing lightning. The MB06 method could also increase lead times (which are either 30 minutes, or 0 minutes in this paper) for the lightning prediction algorithm. One critical drawback to the method described in MB06 is the reliance on VIS imagery to mask the GOES image. Doing so saves much time in computing, but the obvious problem is that the "mask" can only be employed during the day. The authors acknowledge the shortcoming, and mention that this problem is an area of active research.

Near storm environment parameters from numerical models, such as CAPE, CIN, mixing ratio, etc. coupled with radar and perhaps satellite data may also improve the performance of the algorithm. For instance, a cumulus cloud will not evolve into a cumulonimbus unless it continually ingests a sufficient amount of moisture to initiate ice crystal growth (MB06), which is necessary for electrification inside a thunderstorm. Atmospheric stability metrics, such as CAPE, can serve to identify atmospheric regions capable of supporting further convective growth for identified cumulus clouds. The use of CAPE may also help improve lead times for lightning forecasts, seeing that it can be forecast by numerical models multiple hours in advance.

6. SUMMARY AND CONCLUSION

A probabilistic lightning prediction algorithm using a trained neural network was evaluated on 12 independent (from the training days) days (1 per month). Scores were threshold to yield binary forecasts of "yes" or "no" if lightning would occur. The optimal probability threshold (in terms of CSI) for the combined 12 days was at 50%, with a score of 0.325, or at the 70% threshold (in terms of HSS, which takes random chance into account), with a score of 0.307. The algorithm achieved a Brier Score of 0.1038 (excluding misses) and 0.2269 (including misses, at the 0% threshold). A Fractions Skill Score was also computed, which compares the areas fractions for probability clusters and lightning density clusters (but only for hits). The *FSS* was 0.6289, which is greater than FSS_{useful} = 0.5. Thus, the clusters do have decent correspondence, but that is somewhat expected since the probability of lightning forecasts is only for the next 30 minutes. The algorithm is fairly reliable towards both probability extremes, but it is also over-forecasting in the middle probability bins. The algorithm is also fairly sharp, considering the low climatological probability of lightning.

Several measures have been identified to improve the algorithm, among which are tracking radar reflectivity at -10°C or -20°C iso-surfaces, or using GOES satellite data to capture convective initiation. Near storm environment data from numerical models may also aid in better probabilistic forecasts, as well as vertical gradients of reflectivity inside the storm (Vincent et al. 2003). These avenues will be pursued in further studies of this algorithm.

REFERENCES

Brier, G. W., 1950: Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review,* **78**, 1-3.

Burrows, W., C. Price, and L. Wilson, 2005: Warm Season Lightning Probability Prediction for Canada and the Northern United States. *Weather and Forecasting,* **20**, 891-988.

Carey, L. D., and S. A. Rutledge, Characteristics of cloud-to-ground lightning in severe and nonsevere storms over the central United States from 1989–1998, J. Geophys. Res., 108(D15), 4483, doi:10.1029/2002JD002951, 2003.

Curran, E.B., Holle, R.L., 1997. Lightning fatalities, injuries and damage reports in the United States, 1959–1994, NOAA Technical Memorandum NWS SR-193. NOAA Scientific Services Division, Fort Worth, Texas.

Ebert, E. E., 2009: Neighborhood Verification – a Strategy for Rewarding Close Forecasts. *Weather and Forecasting*

Hondl, K. D. and M. D. Eilts, 1994: Doppler Radar Signatures of Developing Thunderstorms and Their Potential to Indicate the Onset of Cloud-to-Ground Lightning. *Monthly Weather Review,* **122**, 1818-1836.

Lakshmanan, V. "Tuning an algorithm for identifying and tracking cells," in *Southern Thunder 2009 Workshop*, (Cocoa Beach, FL), p. CD, Nat. Aero. Space Admin., July, 2009.

Lakshmanan, V. and T. Smith, 2008: Data Mining Storm Attributes From Spatial Grids. *J. Oceanic and Atmospheric Technology.*

Lakshmanan, V. and T. Smith, 2009: An Algorithm to Nowcast Lightning Initiation and Cessation in Real-time, *4th Conference on the Meteorological Applications of Lightning Data*, Amer. Meteor. Soc.

Lakshmanan, V., R. Rabin, and V. DeBrunner, 2003: Multiscale storm identification and forecast, *J. Atm. Res.*, **67**, pp. 367-380.

Mecikalski, J. R. and K. M. Bedka, 2006: Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Daytime GOES Imagery. *Monthly Weather Review,* **134**, 49-78.

Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review,* **136**, 78-97.

Vincent, B.R., L.D. Carey, D. Schneider, K. Keeter and R. Gonski, 2003: "Using WSR-88D Reflectivity Data for the Prediction of Cloud-to-Ground Lightning: A Central North Carolina Study", December 2003, 27, 35-44.

Wilks, D. S., 2006: Statistical Methods in the Atmospheric Sciences. Academic Press, Volume 91 in International Geophysical Series, 627 pp.