NASA/CR-2012- 216312

Objective Lightning Probability Forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase IV

William H. Bauman III
ENSCO, Inc., Cocoa Beach, Florida
NASA Applied Meteorology Unit, Kennedy Space Center, Florida

Winifred C. Crawford
ENSCO, Inc., Cocoa Beach, Florida
NASA Applied Meteorology Unit, Kennedy Space Center, Florida

August 2012

NASA STI Program ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Fax your question to the NASA STI Information Desk at 443-757-5803

- Phone the NASA STI Information Desk at 443-757-5802

- Write to:
  STI Information Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076-1320

NASA/CR-2012- 216312

Objective Lightning Probability Forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase IV

William H. Bauman III
ENSCO, Inc., Cocoa Beach, Florida
NASA Applied Meteorology Unit, Kennedy Space Center, Florida

Winifred C. Crawford
ENSCO, Inc., Cocoa Beach, Florida
NASA Applied Meteorology Unit, Kennedy Space Center, Florida

National Aeronautics and
Space Administration

Kennedy Space Center
Kennedy Space Center, FL 32899-0001

August 2012

Acknowledgements

The authors wish to thank Mr. William Roeder of the 45th Weather Squadron for his suggestions and many consultations throughout the task as well as his efforts to acquire the National Lighting Detection Network lightning strike data from the 14th Weather Squadron (14 WS) Strategic Climatic Information Service and National Climatic Data Center. Without the individual strike data this task would not have been possible. Additionally, the 14 WS delivered the data in a format that significantly reduced processing time by the Applied Meteorology Unit.

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

This report is also available in electronic form at

http://science.ksc.nasa.gov/amu/

Executive Summary

The 45th Weather Squadron (45 WS) includes the lightning probability forecast in their daily morning briefings. This forecast is most important in the warm season months, May-October, when the area is most affected by lightning. The forecasters use this information for planning daily ground operations on Kennedy Space Center (KSC) and Cape Canaveral Air Force Station (CCAFS) for the current day and the following seven days, and as part of the launch forecast process. The daily lightning probability forecast is based on the output from an objective lightning forecast tool developed in two phases by the AMU that the forecasters supplement with subjective analyses of model and observational data.

The current operational tool was developed in Phase II and consists of a set of equations, one for each month May-September that calculates the probability of lightning occurrence for the day more accurately than previous forecast methods (Lambert and Wheeler 2005, Lambert 2007). The equations are accessed through a graphical user interface (GUI) in the 45 WS primary weather analysis and display system, the Meteorological Interactive Data Display System (MIDDS). The goal of Phase III (Crawford 2010) was to create equations based on the progression of the lightning season as seen in the daily climatology instead of an equation for each month in order to capture the physical attributes that contribute to thunderstorm formation. These Phase III equations did not outperform Phase II. Therefore, the Phase II equations are still in operational use.

For this phase, the 45 WS requested the AMU make another attempt to stratify the data by lightning sub-season, add three more years of May-September data than used in Phase III to expand the period of record (POR) to 23 years (1989-2011) and include data from October. The AMU did this by using lightning observations across central Florida from the National Lightning Detection Network (NLDN) instead of the 45th Space Wing Cloud-to-Ground Lightning Surveillance System (CGLSS) data used in Phase III. The lightning season could start anywhere in central Florida, not just locally at KSC/CCAFS as the CGLSS data would show. In the event that lightning sub-seasons cannot be identified, the 45 WS requested that the AMU create monthly equations with using the expanded 23-year POR.

The AMU established a correlation between the NLDN flash count start/ramp-up dates and the wet-season start dates used by the National Weather Service in Melbourne after removing days with NLDN flashes that were due to large-scale weather patterns. Not all days with NLDN flashes could be eliminated based on these patterns, so determining the start/ramp-up of each lightning season in the POR still had some level of subjectivity. With the start/ramp-up correlation established, the AMU then looked for correlations between the NLDN data and other lightning sub-seasons: lightning (plateau), ramp-down, and post. Determining the start of the sub-seasons using the NLDN data was also subjective. By examining annual charts of daily flash count, the AMU observed there were often multiple sub-seasons of the same type. The AMU presented these findings to the 45 WS personnel who would use the tool. This led to a consensus that it would be difficult for the forecasters to determine the lightning sub-seasons in real-time using NLDN data.

Therefore, the AMU used monthly stratifications to develop and test new equations. The performance of the new equations was compared to that of five other forecast methods including the Phase II equations being used in operations. The new equations outperformed every method except Phase II. Therefore, the Phase IV equations will not replace the Phase II equations in operations. The AMU updated and tested the MIDDS tool by adding the October equations and updating the GUI.

Table of Contents

List of Figures

List of Tables

1. Introduction

The 45th Weather Squadron (45 WS) includes the probability of lightning occurrence in their daily morning briefings. This forecast is important in the warm season months, May-October, when the area is most affected by lightning. The forecasters use this information for planning daily ground operations on Kennedy Space Center (KSC) and Cape Canaveral Air Force Station (CCAFS) for the current day and the following seven days, and as part of the launch forecast process.

The daily lightning probability forecast is based in part on the output from an objective lightning forecast tool developed by the AMU that the forecasters supplement with subjective analyses of model and observational data. The tool developed in Phase II consists of a set of equations, one for each warm season month that calculates the probability of lightning occurrence for the day more accurately than previous forecast methods (Lambert and Wheeler 2005, Lambert 2007). The equations are accessed through a graphical user interface (GUI) in the 45 WS primary weather analysis and display system, the Meteorological Interactive Data Display System (MIDDS). The goal of Phase III (Crawford 2010) was to create equations based on the progression of the lightning season as seen in the daily climatology instead of an equation for each month in order to capture the physical attributes that contribute to thunderstorm formation. Five sub-seasons were discerned from the daily climatology, and the AMU created and tested an equation for each. The Phase III equations did not outperform Phase II. Therefore, the Phase II equations are still in operational use.

For this phase, the 45 WS requested the AMU make another attempt to stratify the data by lightning sub-season and include October data in the analysis. The AMU did this by using lightning observations across central Florida from the National Lightning Detection Network (NLDN) instead of the 45th Space Wing Cloud-to-Ground Lightning Surveillance System (CGLSS) data used in Phase III. The lightning season could start anywhere in central Florida, not just locally at KSC/CCAFS as the CGLSS data would show. By using the lightning season across central Florida and relying on the flow regimes to influence the probabilities of lightning in the KSC/CCAFS area, a more physically representative model may result. In the event that lightning sub-seasons cannot be identified, the 45 WS requested that the AMU create monthly equations with six more years of data than used in Phase II.

2. Data

The AMU collected all the data needed for this task for the period of record (POR) May-October 1989-2011. This includes the NLDN flash data in an area covering central Florida, the CCAFS 1000 UTC soundings (XMR), and the 1200 UTC soundings from Jacksonville (JAX), Tampa (TBW) and Miami (MFL), Fla. The AMU also collected NLDN data for April and November in the POR to observe lightning behavior before and after the defined lightning season of May-October. The NLDN data were provided by the 14th Weather Squadron (14 WS) through Mr. Roeder of the 45 WS and cover the area shown in Figure 1. Mr. Madison of Computer Sciences Raytheon (CSR) provided XMR and CGLSS data, and the AMU downloaded the National Weather Service (NWS) soundings for JAX, TBW and MFL from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory website (http://www.esrl.noaa.gov/raobs/).

2.1 Lightning Data

The AMU wrote scripts in TIBCO Spotfire S+ (TIBCO 2010) statistical analysis software to extract the daily lightning flash counts from the raw NLDN data and export it as Microsoft Excel files. In Excel, they used Visual Basic for Applications (VBA) to write scripts to create annual charts of the daily NLDN flash count for April-November 1989-2011. They also developed VBA scripts in Excel to export the daily NLDN flash locations to Keyhole Markup Language (KML) for display on a map in Google Earth.

The AMU processed the CGLSS data to create the predictand needed for equation development. The predictand is binary and indicates whether lightning occurred within any of the 5 NM warning circles on KSC/CCAFS on each day. The AMU also used the CGLSS data to create the daily climatology and one-day persistence, two of the candidate predictors for the equations. The daily climatology is the percent of days lightning occurred on each date in the warm season POR. One-day persistence is binary like the predictand, and indicates whether lightning occurred on the previous day.

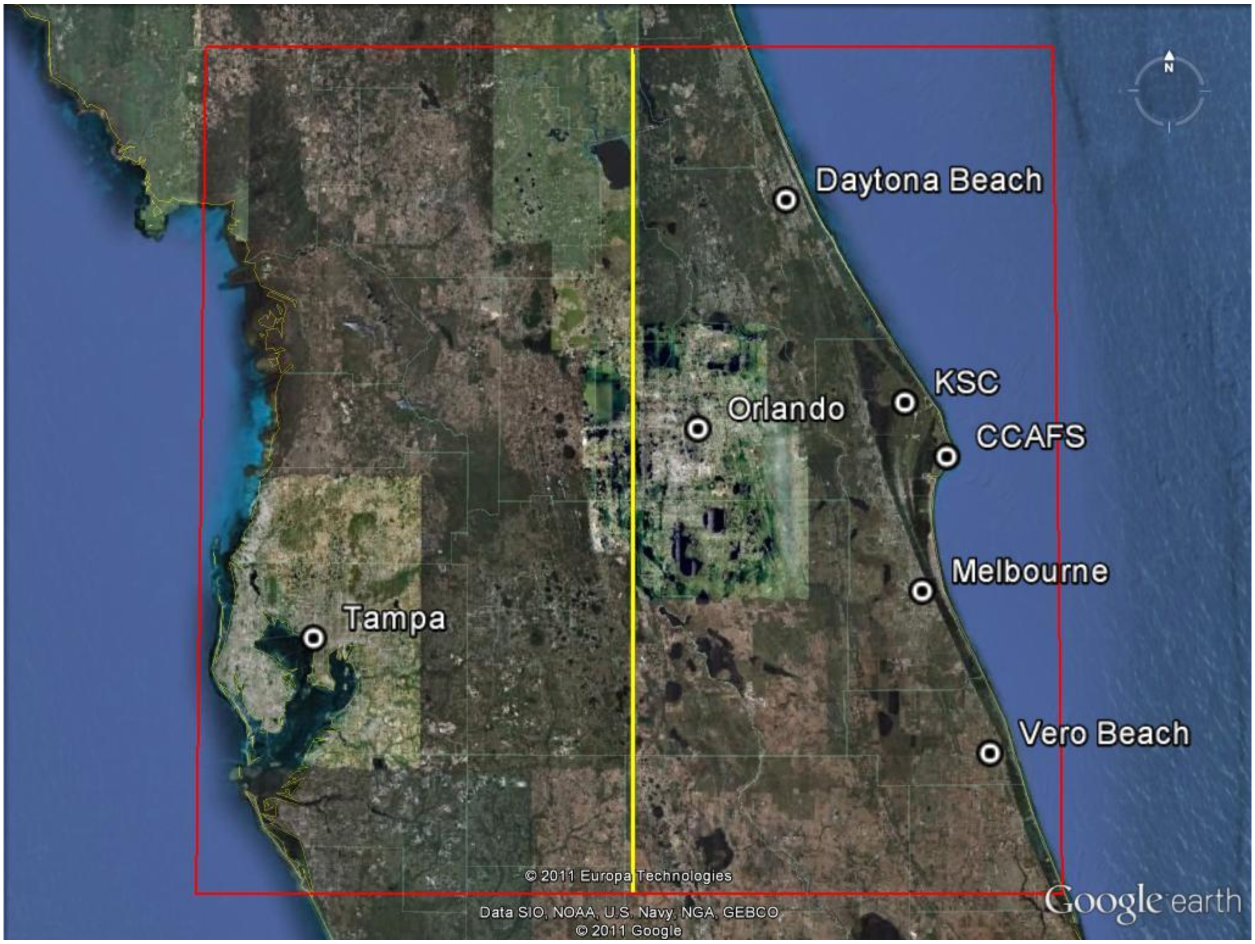

Figure 1. Map of central Florida showing areal coverage (red square) of the NLDN flash count data set. The yellow line is the boundary between east- and west-central Florida.

2.2 Sounding Data

The AMU quality-controlled the XMR sounding data, then determined the flow regimes on each day in the 2010 and 2011 warm seasons and October 1989-2009 using the process described in Lambert (2007). They also calculated the stability indices from the XMR soundings needed as predictors for equation development. The 1989-2009 May-September values were created in earlier AMU tasks.

3. Lightning Season

The approach the AMU used to determine the start of the lightning season consisted of analyzing the NLDN daily flash count charts to determine if these data correlated with known NWS Melbourne (NWS MLB) east-central Florida wet season start dates (Lascody 2002). If a correlation between the NLDN start/ramp-up of the lightning season could be established with the wet-season start dates, the AMU would continue to look for correlations between the NLDN data and other lightning sub-seasons: lightning (plateau), ramp-down, and post. The sub-seasons would be determined by the increase, plateau, and decrease in the number of flashes on each day, not just whether or not lightning occurred. They would then develop logistic regression equations stratified by lightning sub-season to predict the probability of lightning for each day. If a correlation could not be established with the wet season start dates, the 45 WS requested that equations be created for each month as in previous work.

3.1 Lightning Season Start

The chart in Figure 2 shows the daily NLDN lightning flash count across central Florida for April-November 1989. The light blue line shows the number of flashes observed each day while the dark blue line is the 14-day moving average of the daily flashes. The vertical orange and green lines show the 1989 NWS MLB wet season start and end dates for Orlando (MCO) and Melbourne (MLB), respectively. The 1989 wet season start date was 28 May for MCO and 6 June for MLB. The chart shows an increase in the daily flash count beginning in April and peaking in July. It would appear that the start of the lightning season across central Florida preceded the start of the wet season at MCO and MLB. This is possible since the AMU was

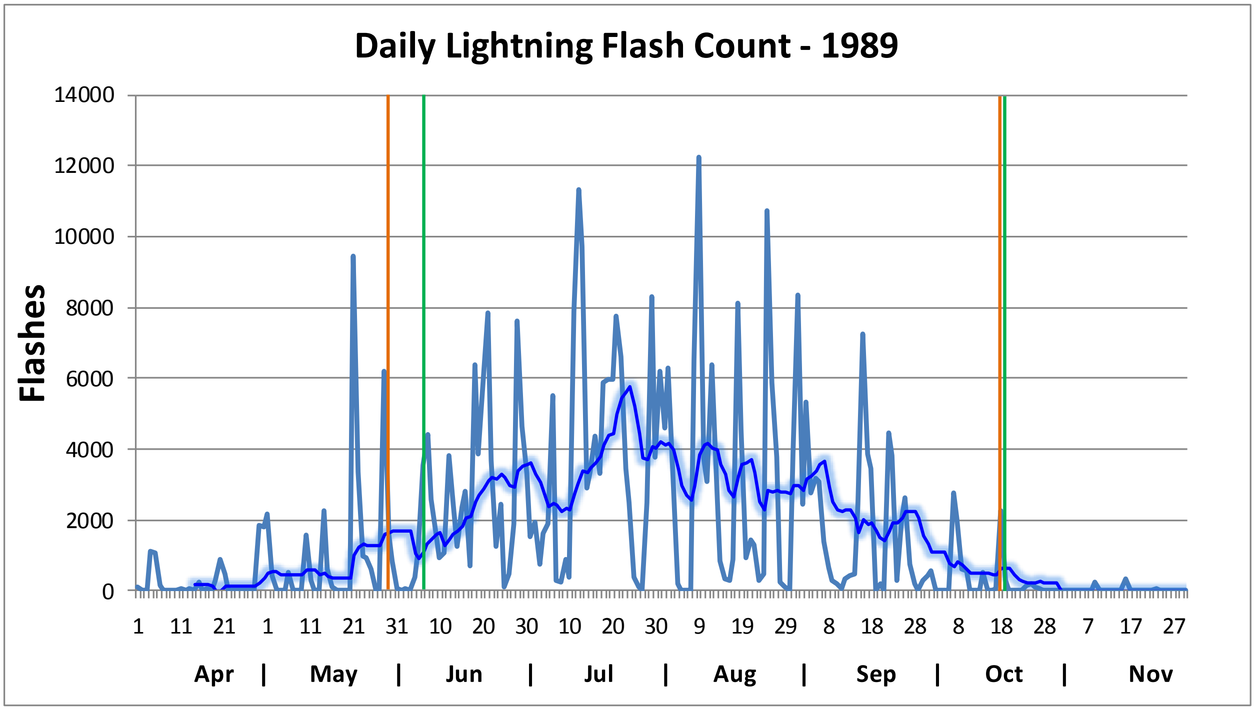

Figure 2. Daily NLDN lightning flash count across central Florida for April-November 1989. The light blue line shows the number of flashes observed each day while the dark blue line is the 14-day moving average of the daily flashes. The vertical orange and green lines show the 1989 NWS MLB wet season start and end dates for Orlando (MCO) and Melbourne (MLB), respectively.

attempting to define the lightning season across central Florida (see Figure 1), not just at MLB and MCO. Most other years in the POR exhibited similar behavior in which the daily NLDN flash count start and ramp-up occurred prior to the MLB and MCO wet season start dates.

After discussing this finding with Mr. Roeder of the 45 WS, the AMU recommended considering two additional stratifications to try and account for the difference: (1) limit the areal coverage of the NLDN data to east-central Florida and (2) eliminate any lightning days due to influences from synoptic weather patterns such as cold fronts that will sometimes penetrate into central Florida as late as May. The first option would determine if the technique chose dates consistent with the NWS MLB start dates for MLB and MCO. If so, it could be used to cover all of central Florida. The second option is consistent with the formation of lightning in the warm season being dominated by the interaction of low level boundaries such as the sea breeze fronts, river breeze fronts, convective outflows, horizontal convective rolls and others. Forecasting lightning early in the season when weak cold fronts play a role is easier for forecasters.

The AMU developed a VBA script to convert Excel NLDN flash data into KML format to display each flash on a map of central Florida in Google Earth. They were then able to visually inspect the NLDN strike locations occurring on each day from 0000-2359 UTC. The map in Figure 3 shows an example of this display for 5 April 1989. By reviewing similar NLDN maps for April and May of all years in the POR, the AMU discovered it was not uncommon for lightning to occur early in the warm season as in Figure 3, and not be confined to the east or west half of the state as is more typical under prevailing westerly or easterly flow warm season regimes.

However, as the warm season progressed, there were clear divisions between east and west coast daily NLDN events dependent on the flow regime. Therefore, the AMU reduced the NLDN data set to flashes only occurring in east-central Florida, which covered the eastern half of the red square in Figure 3. However, further analysis revealed that reducing the NLDN events to east-central Florida did not explain the non-correlation of the beginning of lightning flash count ramp-up with the NWS MLB wet season start dates.

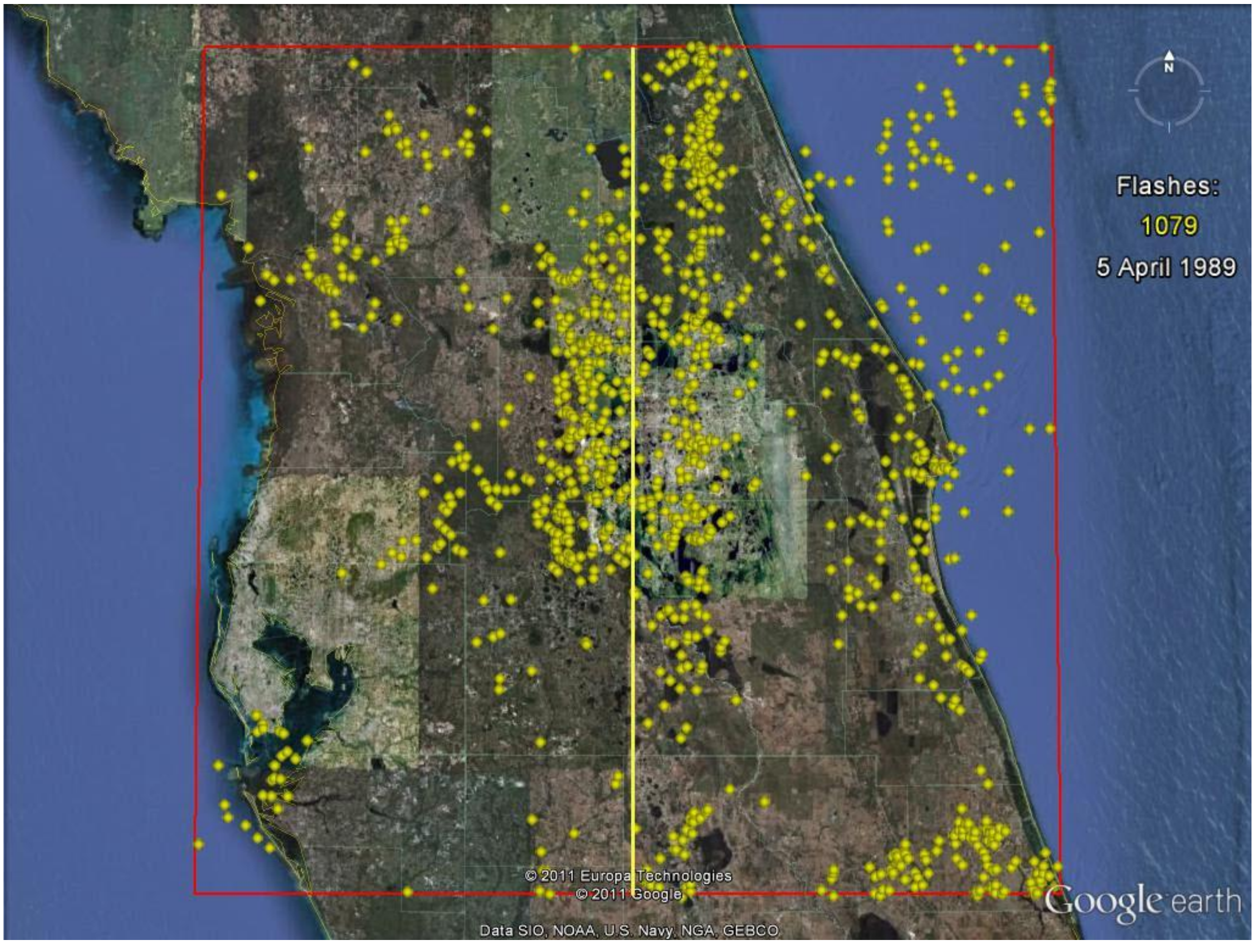

Figure 3. The NLDN flashes across central Florida for 0000-2359 UTC 5 April 1989. Each yellow dot represents one cloud-to-ground lightning flash observed by NLDN. On this day there were 1,079 flashes across central Florida within the area bounded by the red lines.

To account for synoptic weather pattern influences other than the warm season peninsular flow regimes, such as low pressure systems and fronts, the AMU reviewed all April and May daily weather maps (NOAA 2011) for days with NLDN flashes in central Florida to find these events and eliminate them from the data set. As an example of a synoptic weather system influence, the daily weather maps for 7:00 AM EST (1200 UTC) on 5 and 6 April 1989 are shown in Figure 4. On 5 April, a cold front was draped across the southeast United States and a pre-frontal squall line was located in north Florida. By 7:00 AM EST on 6 April, the cold front was located just south of central Florida. It is highly likely that this cold front and/or its associated pre-frontal squall line was responsible for the NLDN flashes shown in Figure 3. Therefore, the NLDN flashes on 5 April were eliminated from the data set. This methodology eliminated most NLDN flash days in April and many in May throughout the POR resulting in a better correlation between the NLDN-based lightning season and NWS MLB wet season start dates.

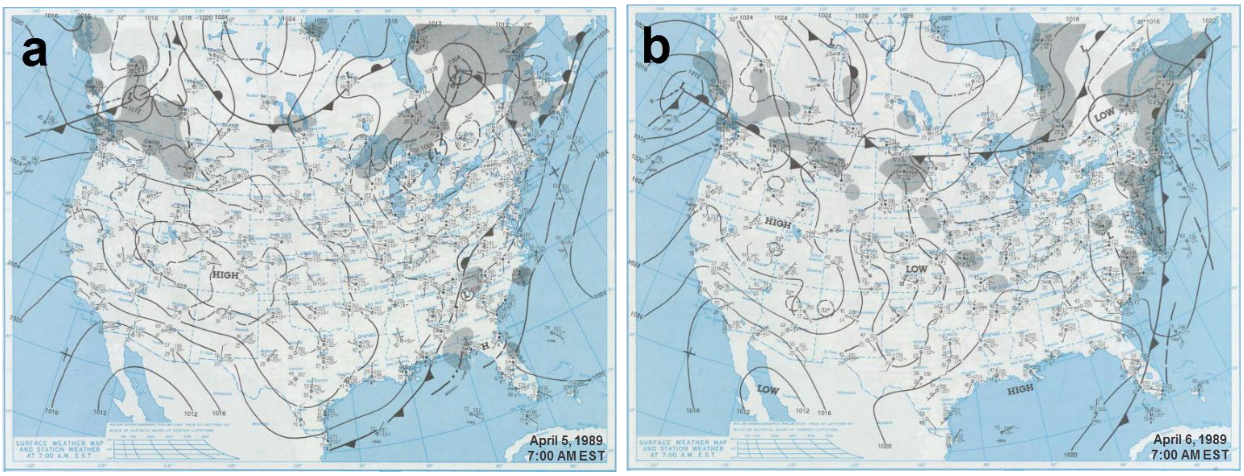

Figure 4. The NOAA daily weather maps (NOAA 2011) from (a) 5 April 1989 and (b) 6 April 1989 showing surface synoptic-scale weather systems.

The daily weather maps did not always show lows or frontal systems in the area on days with lightning in April and May, so the AMU could not conclude that these systems were the cause of the lightning. They did not eliminate these days from the data set. However, some of these fronts may have been weak and not be fully depicted in the NOAA daily weather maps. Since not all days with NLDN flashes could be eliminated based on synoptic patterns, determining the start/ramp-up of each lightning season in the POR still had some level of subjectivity. Figure 5 shows the same time period as Figure 2 except it includes only NLDN flashes from east-central Florida with the synoptic weather system influenced flashes removed. Hence, the 1989 east-central Florida lightning season started within a week of the NWS MLB wet season start at MCO.

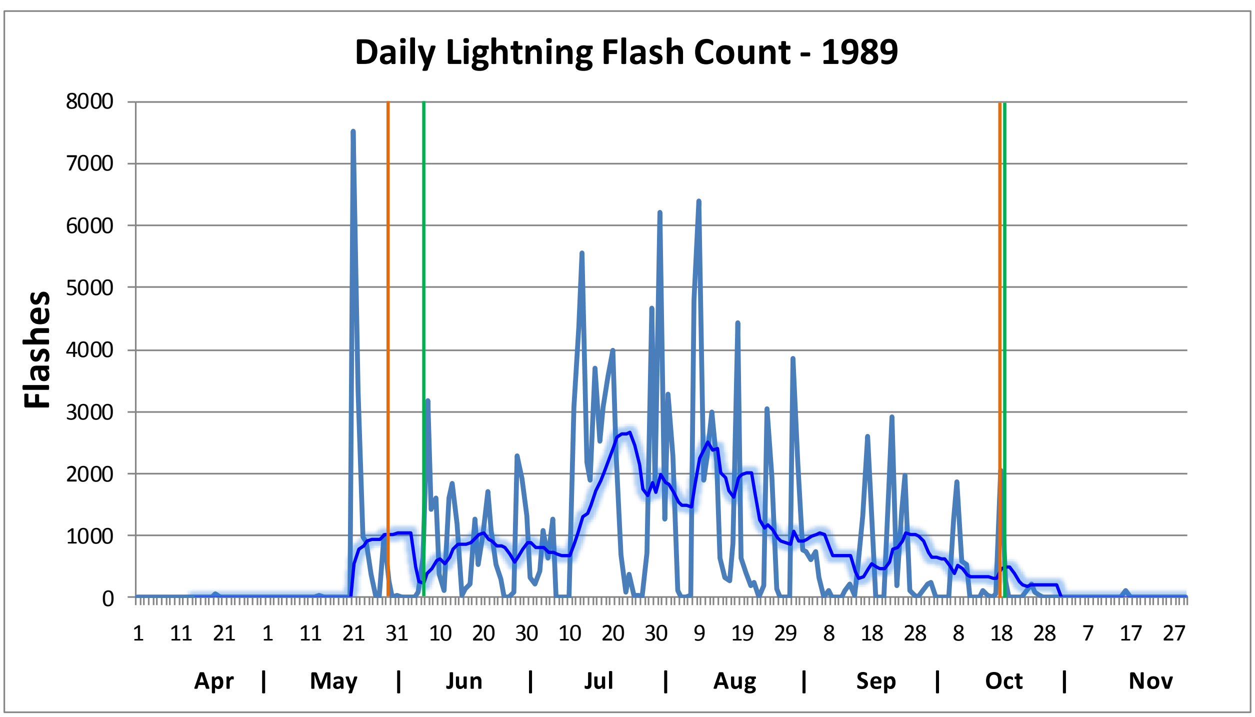

Figure 5. As in Figure 2, but for the east-central Florida region and with NLDN flashes due to synoptic frontal systems removed.

Over the entire POR, the median NLDN flash count ramp-up started six days prior to the mean MCO/MLB wet season start dates. The earliest NLDN start/ramp-up was 21 days prior to wet season start (1999) and the latest was 3 days after (1996 and 2007). In 65 percent of the years (15 of 23) the start dates were within 7 days of each other. Lascody (2002) notes that in determining the start of the wet season, “It must be stated that a purely objective analysis is not possible since the exact onset of the Wet Season is difficult to determine in some years.” Given the subjectivity in determining both the wet season start and NLDN start/ramp-up and the data presented, it appears there is a correlation between the wet season method and the NLDN method. Also, as noted previously, the start date for central Florida could be earlier than for MLB and MCO.

3.2 Lightning Sub-season Stratifications

With a correlation established between the NLDN flash count start/ramp-up dates and the NWS MLB wet-season start dates, the AMU looked for correlations between the NLDN data and other proposed lightning sub-seasons: lightning (plateau), ramp-down, and post. Just as determining the start of the lightning season using the NLDN data was subjective, so was determining sub-seasons. In examination of the annual charts of daily flash count, the AMU observed there were often multiple sub-seasons of the same type. For example, in Figure 6, it appears there were two ramp-up sub-seasons in 1999 followed by a relatively consistent flash count (plateau) from mid-June through early September, a ramp-down sub-season from mid-September to mid-October, and then a post lightning sub-season. In a second example from 2004 (Figure 7), after the ramp-up there were two distinct consistently high lightning flash count periods (plateaus) with a lull in lightning flash count from mid-July through early August followed by a ramp-down sub-season and post lightning sub-season.

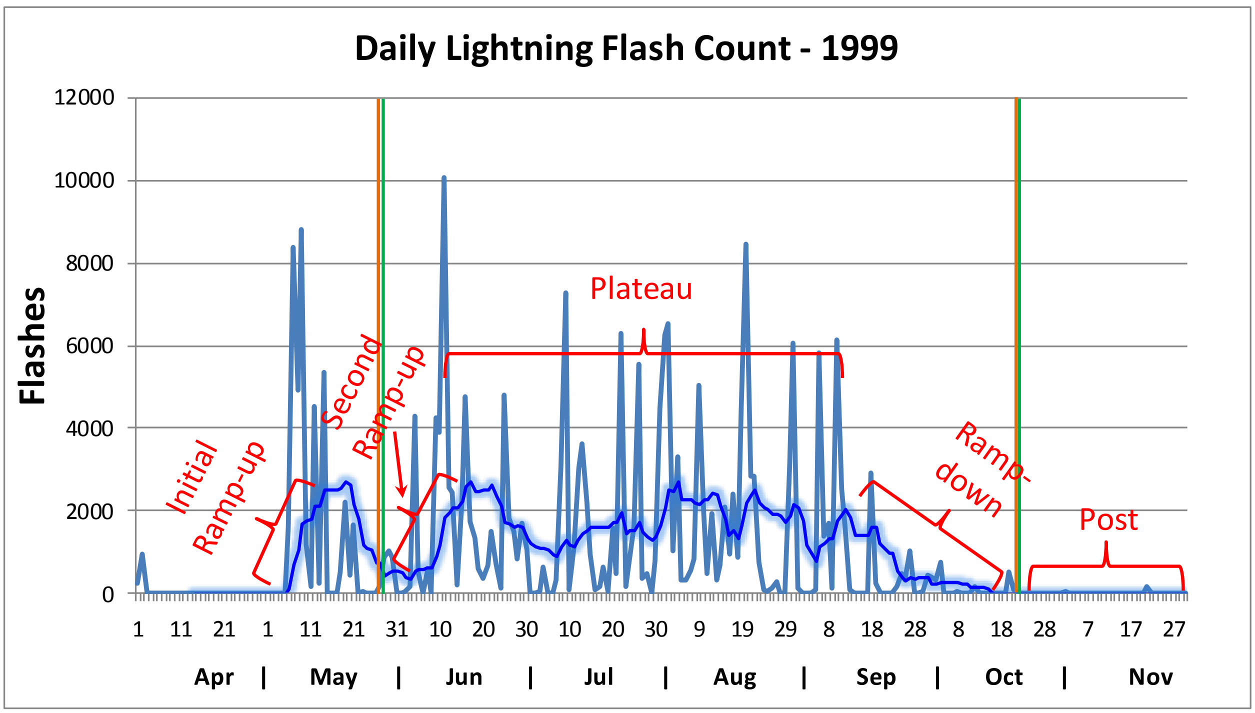

Figure 6. As in Figure 5, but for 1999 and with the lightning sub-seasons highlighted by red brackets and text describing each.

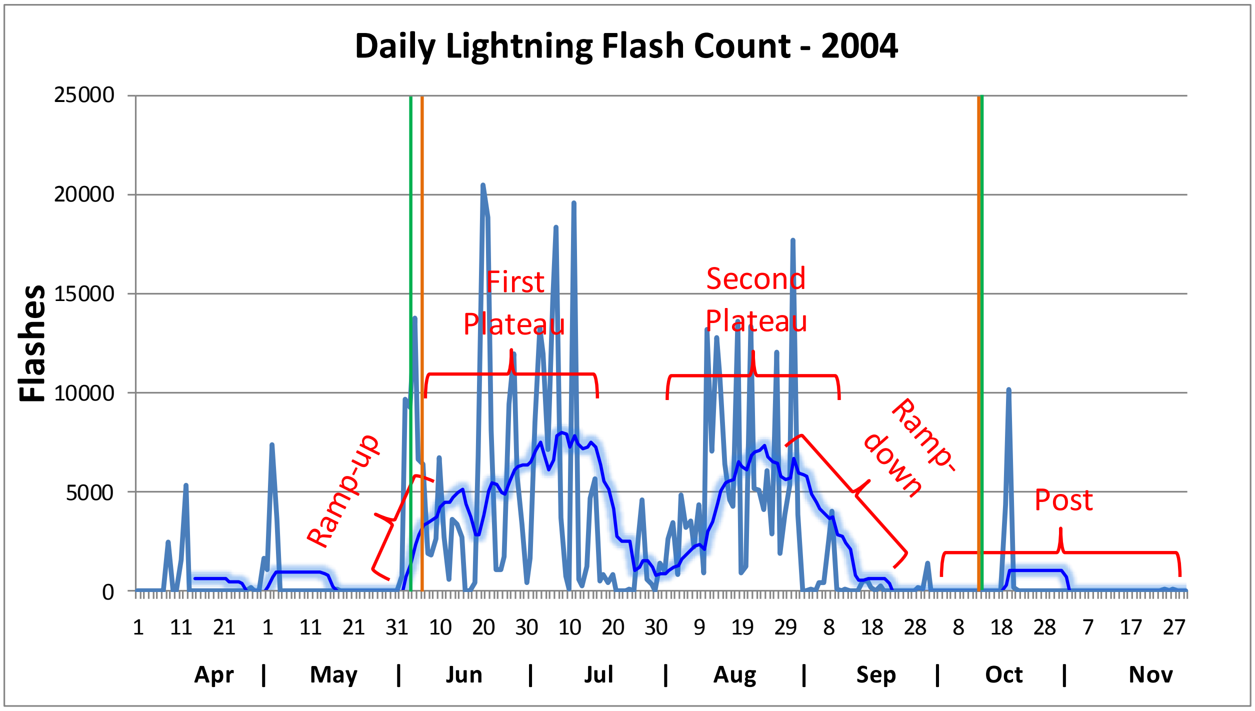

Figure 7. As in Figure 6, but for 2004.

A question that surfaced while assessing the sub-season stratifications was how the operational forecasters would be able to declare the lightning season start and the sub-seasons in real-time in order to know which equation to use in the tool. Since it was subjective and somewhat difficult to make those determinations using climatological data, the AMU consulted with Mr. Roeder on how to address this issue. They decided to present the findings to the 45 WS personnel who would use the tool. The discussion led to a consensus that it would be difficult for the forecasters to determine the lightning sub-seasons in real-time using NLDN data. Determining lightning sub-seasons to stratify the data more on the basis of physical processes than calendar dates is desirable, but the AMU showed how difficult it can be. Therefore, with agreement from Mr. Roeder, the AMU proceeded with this task using monthly stratifications to develop the new equations.

4. Equation Development

The AMU developed and tested six new lightning probability forecast equations, one for each month May-October, using a development dataset consisting of 19 warm seasons (83% of total seasons available) and a verification dataset consisting of 4 warm seasons (17% of total). The verification and development datasets did not consist of individual warm season years, but rather individual warm season days chosen through a random process described in Lambert and Wheeler (2005). They also followed the same iterative procedure for choosing the predictors in the development dataset as outlined in Lambert and Wheeler (2005). The verification data set was not used in developing the equations so it would provide an independent assessment of the expected performance when the equations are used with new data in operational forecasting.

4.1 Predictors

The AMU used 12 stability and moisture parameters from the 1000 UTC XMR soundings as candidate predictors in the equation development:

- Total Totals,
- Cross Totals,
- Vertical Totals (VT),
- K-Index (KI),
- Lifted Index (LI),
- Thompson Index (TI; KI – LI),
- Severe Weather Threat Index,
- Showalter Index,
- Temperature at 500 mb,
- Mean Relative Humidity in the 825–525 mb layer,
- Mean Relative Humidity in the 825–600 mb layer,and
- Precipitable Water up to 500 mb (PW).

These were added to the three predictors shown below, described in the AMU Quarterly Report Q4 FY06 and updated using warm season data in the years 1989-2011, for a total of 15 candidate predictors:

- Daily climatological lightning frequency (Climo),
- 1-day persistence (Pers), and
- Flow regime lightning probability (FRProb).

4.2 New Equations

The procedure to develop the new equations involved adding one predictor at a time and checking the associated reduction in residual deviance. A large reduction in residual deviance meant that a predictor accounted for a large percentage of the variance in the predictand. Therefore, the AMU chose the predictors that effected the largest reduction. They stopped adding predictors as soon as a candidate predictor accounted for < 0.5% of the reduction in residual deviance. This ensured predictor variables were selected in the optimal order and avoided statistical over-fitting of the prediction equations.

Figure 8 shows the percent reduction in residual deviance from the NULL model as each predictor was added for the May equation. TI reduced the residual deviance the most (14.31%)

and was, therefore, the first predictor in the May equation. The second predictor was Pers, which accounted for an additional 6.14% reduction. FRProb was the third predictor, reducing the residual deviance by 4.42%. For May, Climo reduced the residual deviance by 0.26%, therefore, Climo was not chosen. The May equation consists of the predictors TI, Pers and FRProb.

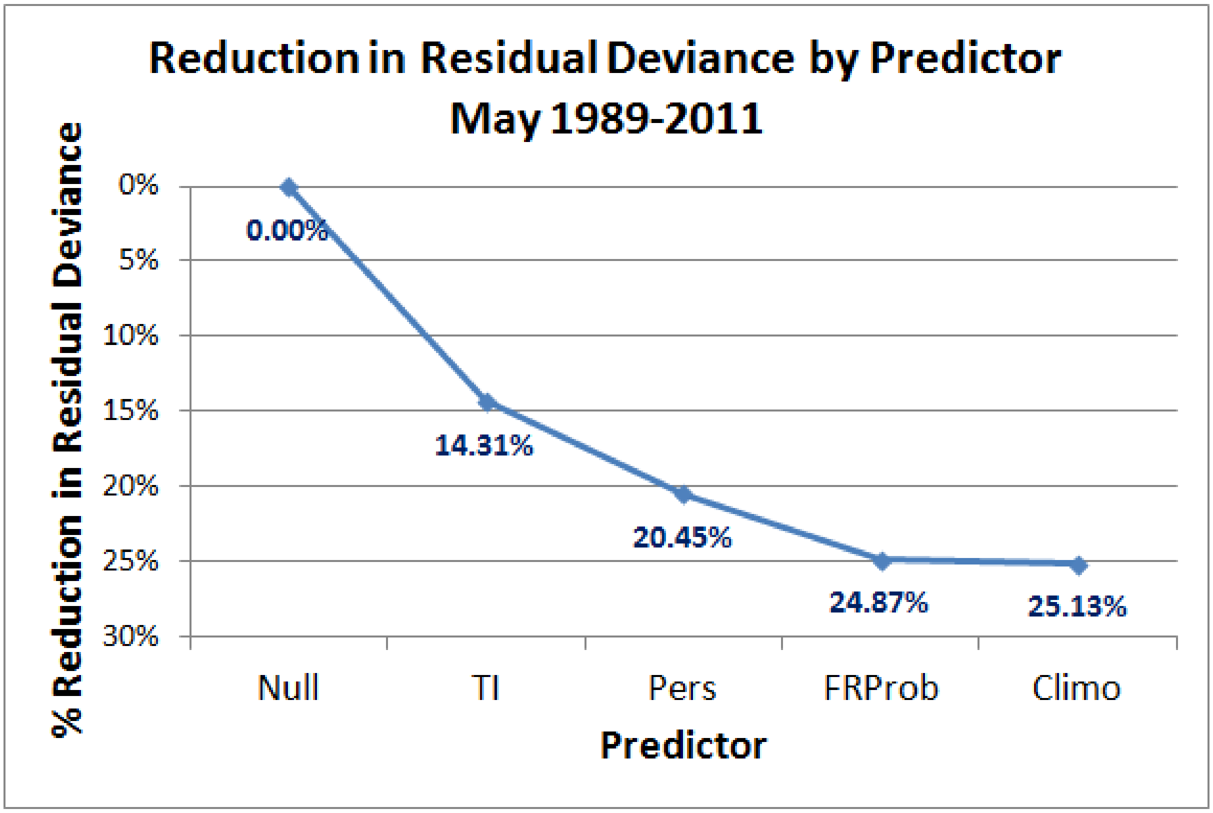

Figure 8. The total percent reduction in residual deviance from that of the NULL model as each predictor was added to the equation using the May development dataset.

Table 1 shows the final predictors for each of the monthly equations in rank order of their reduction in residual deviance. FRProb was the only predictor in all six months. The two strongest predictors in June, July, August and September were TI and FRProb, varying between first and second. FRProb and TI were also strong predictors in May but Pers was more important than FRProb. October is the outlier with Lifted Index and Mean Relative Humidity in the 825–600 mb layer as the top two predictors.

Table 1. The final predictors for each monthly equation, in rank order of their contribution to the reduction in residual deviance. The predictors in red were in every equation, the predictors in blue were in five of the six equations, the predictors in green were in four of the six equations, the predictors in orange were in three of the six equations and the predictors in black were in two or less equations.

May	***June***	***July***	***August***	***September***	***October***
Thompson	Thompson	Flow Regime	Flow Regime	Thompson	Lifted Index
Persistence	Flow Regime	Thompson	Thompson	Flow Regime	825–600 RH
Flow Regime	825–525 RH	Total Totals	825–525 RH	Persistence	Flow Regime
	Persistence	Persistence	Total Totals		Total Totals
			Daily Climo		Daily Climo

For comparison with the results in this task, Table 2 shows the final predictors for each of the Phase II monthly equations currently used in operations in rank order of their reduction in residual deviance. Predictors for October were not assessed in Phase II.

Table 2. The final predictors for each monthly equation from Phase II, in rank order of their reduction in residual deviance. The predictors in red were in every equation, the predictors in blue were in four of the five equations, the predictors in green were in three of the five equations, and the predictors in black were in only one equation.

May	***June***	***July***	***August***	***September***	***October***
K-Index	Thompson	Thompson	Thompson	825–525 RH	N/A
Flow Regime	Flow Regime	Flow Regime	Flow Regime	Flow Regime	
Vertical Totals	Persistence	Total Totals	Daily Climo	Persistence	
Daily Climo	Vertical Totals	Persistence	825–525 RH	Vertical Tot	
Persistence	825–525 RH		Vertical Totals	Daily Climo	

4.3 Equation Testing

The AMU tested the performance of the equations using the verification dataset, which consisted of four warm seasons. None of the days in the verification set were contained in the development set to allow for an independent evaluation of performance. The first step was to determine if the new equations showed improvement in skill over five simple forecast benchmarks. Four of the benchmarks were the same as those in the Phase I task: Pers, Climo, FRProb, and monthly climatology, each evaluated as a separate single predictor. The fifth was the forecasts from the equations developed in Phase II of this work (Lambert 2007) and currently used in operations, hereafter designated as the operational equations.

The skill test began by first calculating the mean squared error (MSE) between the forecasts and observations for all forecast methods. The MSE was calculated using the equation

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2 \text{ (Wilks 2006),}$$

where n is the number of forecast/observation pairs, p_i is the probability associated with the forecast method, and o_i is the corresponding binary lightning observation (Wilks 2006). Then the skill of the new equations was calculated over the five forecast benchmarks using the equation for the Brier Skill Score (SS):

$$\mathrm{SS} = \left(\frac{\mathrm{MSE}_{eqn} - \mathrm{MSE}_{ref}}{\mathrm{MSE}_{perfect} - \mathrm{MSE}_{ref}}\right) * 100 \text{ (Wilks 2006),}$$

where MSE_{eqn} was the MSE of the new equations, MSE_{ref} was the forecast benchmark against which the new equations were tested, and $\mathrm{MSE}_{perfect}$ was the MSE of a perfect forecast, which is always 0. The SS represents a percent improvement or degradation in skill of the equation over the reference forecast when it is positive or negative, respectively.

The SS values for each of the monthly equations are shown in Table 3. The predictors in the equations used to calculate the skill scores in Table 3 produced the best results with the verification dataset and were chosen using the method described in the previous section. The new equations show a double-digit improvement in skill for the first four benchmarks in the table except for October daily climatology. The results for the operational equations also showed double-digit improvement over these four benchmarks (Lambert 2007). The new equations show degradation in skill compared to the operational equations for May, July, August and September. The values of -12% for May and -19% for September are a significant reduction in performance. The values of +3% for June, -2% for July and -1% for August are almost negligible, show similar skill between the new and operational equations for these months, and are likely not statistically significantly different. There is no value for October because there is not a current operational equation for October. The AMU created and tested equations with varying sets of predictors for the four months with a degradation in skill in an attempt to improve the skill of the new equations, but none was realized. The AMU discussed the results with Mr. Roeder of the 45 WS and all agreed not to transition the new May-September equations into operations, but keep the current operational equations in place and add the equation for October.

Table 3. The percent (%) improvement (degradation) in skill of the new equations over the reference forecasts of persistence, daily and monthly climatologies, flow regime probabilities, and the operational equations developed in Lambert (2007). These scores were calculated using the verification data for each month.

Forecast Method	***May***	***June***	***July***	***August***	***September***	***October***
Persistence	52	48	48	59	37	37
Daily Climatology	31	16	10	17	14	7
Monthly Climatology	36	37	13	35	24	16
Flow Regime	34	35	12	34	12	16
Operational Equations	(-12)	3	(-2)	(-1)	(-19)	N/A

5. MIDDS GUI

The AMU updated the MIDDS GUI developed in Phase II (Lambert 2007) with the October equation and delivered it to the 45 WS after testing to ensure proper performance. They modified the existing GUI using the Tool Command Language (Tcl)/Toolkit (Tk) capability in MIDDS.

5.1 Testing

The AMU ran the Objective Lightning tool for two weeks prior to the warm season start to make sure the parameters derived from the morning sounding were correctly displayed in the October tab of the GUI. Then at the beginning of the 2012 warm season, the AMU staff began running the Objective Lightning tool and the AMU-developed Severe Weather Forecast tools daily and discovered several errors in the Tcl/ Tk code for the GUI in both tools.

5.1.1 Vertical Totals

The normal range of the VT stability parameter is 20-35 but the Objective Lightning tool was intermittently displaying a value of 0. The AMU and Mr. Madison of CSR began troubleshooting the Objective Lightning tool Tcl/tk GUI code and tracked down the error to a file that was common between this and the Severe Weather Forecast tool. Both tools were creating a file with the same name to read in the date and time. However, the output format of the file being created was not the same. Once the file was created on any given day, it would not be overwritten. Therefore, if the Severe Weather Forecast tool was run first, the Objective Lightning Tool would input the date and time incorrectly causing the VT to be 0. To solve this dilemma, Mr. Madison updated the code in both tools such that each one created a different filename for the date and time. He then moved the updated Tcl/Tk GUI code from the AMU MIDDS to the operational MIDDS.

5.1.2 Lifted Index

The only month that directly outputs LI in the Objective Lightning tool is October, although other months use LI to calculate TI. TI is calculated by subtracting LI from KI. While running the Objective Lightning tool, the AMU noticed the TI value was always 0 while the TI value in the sounding was not 0. The error was in the Objective Lightning tool Tcl/Tk code that called the KI instead of LI to compute TI. Therefore, TI was being calculated by the difference of KI and KI, resulting in a value of 0. The AMU updated the code by changing KI to LI, which solved the problem. Mr. Madison again moved the updated Objective Lightning tool code to the operational MIDDS.

5.1.3 Relative Humidity

The AMU noticed that the relative humidity (RH) value was the same in the Objective Lightning tool and the Severe Weather Forecast tool. This is unlikely because the Objective Lightning tool calculates the layer-averaged RH in the vertical layer from 825-525 mb and the Severe Weather Forecast tool calculates layer-averaged RH in the vertical layer surface-700 mb. Upon review of the Tcl/Tk code, they found that both tools retrieved their respective layer-averaged RH values from the sounding and then wrote out a file containing the RH to be used later in each program. The error in the code was that both tools called the same file containing the layer-averaged RH but the file was only generated once on any given day. Therefore, the tool that was run first created the file that both tools then used for their RH values. They changed the name of the output file in the Objective Lightning tool code, which solved the problem. After further testing, Mr. Madison moved the updated Objective Lightning tool code to the operational MIDDS.

5.2 GUI Update

The user accesses the GUI through the MIDDS Weather menu by clicking on the ‘FCST Tools’ button and choosing ‘Lightning Forecast Tool’ from the drop-down list (Figure 9). This activates the GUI Tcl/Tk code to determine the date and gather the appropriate data for the equation from MIDDS. The code checks the time and date of the most recent CCAFS sounding (XMR). If it does not match the current day and is not within the time period 0900–1159 UTC, an error message dialog box is displayed. This ensures that data from the previous day and data from sounding times other than 1000 UTC are not used in the equations. The 0900–1159 UTC period allows for the fact that not all 1000 UTC soundings are released precisely at 1000 UTC.

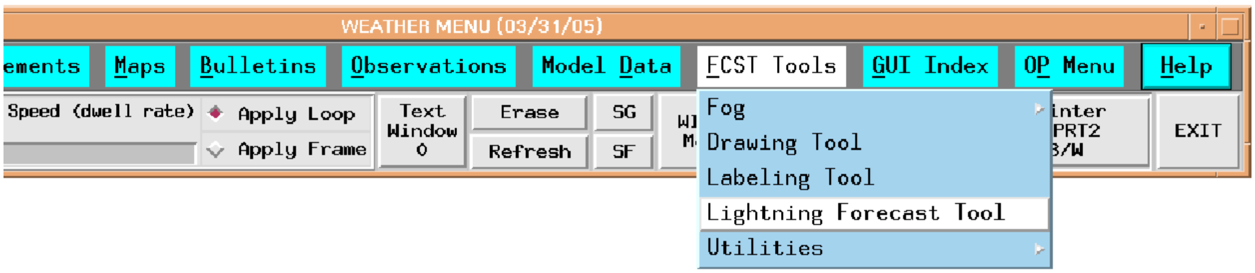

Figure 9. The MIDDS Weather Menu showing the ‘FCST Tools’ button drop-down menu with ‘Lightning Forecast Tool’ highlighted.

Whether or not the 1000 UTC XMR sounding for the current date is available, the equation predictor dialog box is displayed (Figure 10). This will allow the forecasters to use the GUI to create their seven-day forecasts even if data for the current day are not available. The dialog box has six tabs, one for each month. The tab of the current month is displayed initially if the GUI is run between May and October, otherwise 1 May is displayed. The current month, day and sounding time are printed along the top of the dialog box. If the current day’s sounding is not available, ‘No Current Sounding’ will be displayed in place of the date and time in the upper right. The day value can be changed by the up/down arrows or by entering a value manually in the text box. This allows forecasters flexibility when making the seven-day Weekly Planning Forecast. The sounding date and time is formatted by year, day of year, and UTC time.

Forecasters begin by choosing a flow regime. They do not have to enter the sounding parameters as those values are already input by the GUI code and are displayed in their associated text boxes. If there is not a current sounding, the text boxes will be populated with the values from the most recent sounding available. The ‘No Current Sounding’ message in the top right corner will inform the forecaster that this is the case. If the routines cannot find a sounding file of any kind, the text boxes will be populated with the extreme low value in the range of available values for each sounding parameter.

The final step is to click on the ‘Calculate Probability’ button in the lower right corner of the dialog box. The ‘Dismiss’ button in the lower left closes the GUI. If the forecaster does not choose a flow regime, an error message dialog box is displayed telling the forecaster to make a choice.

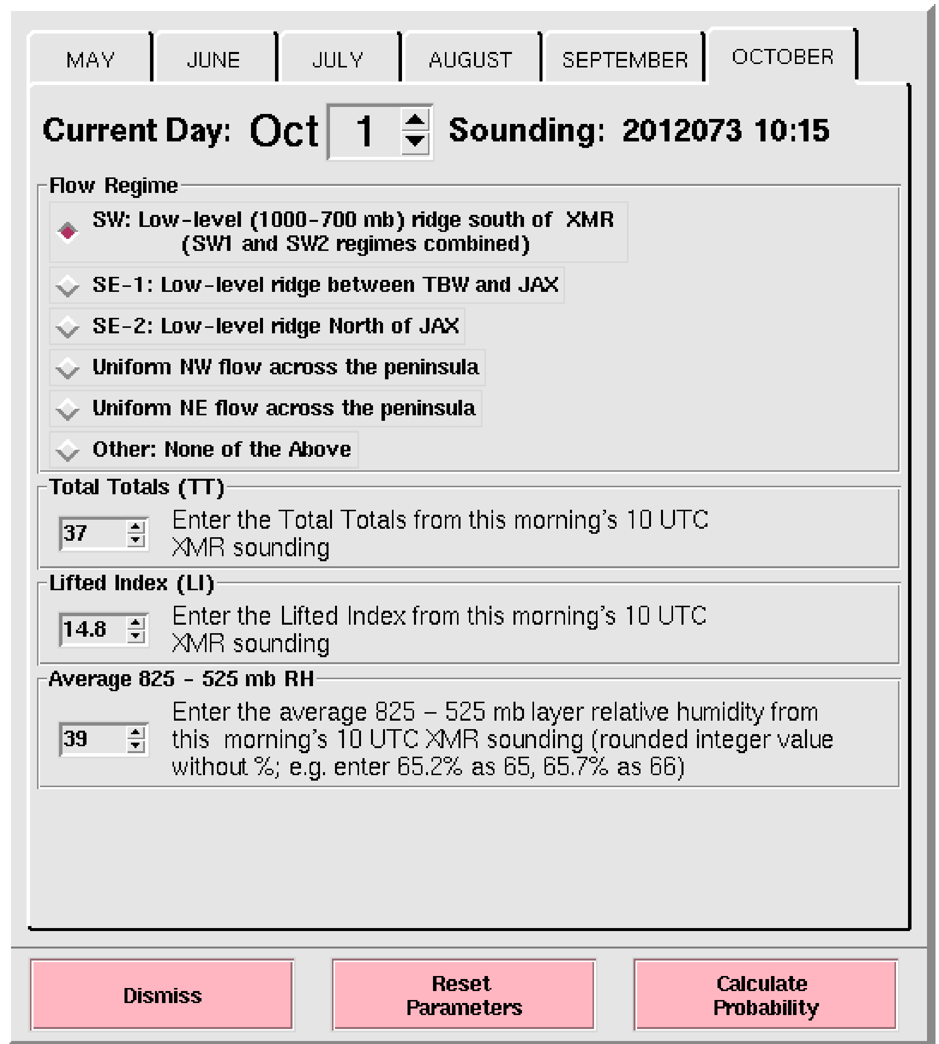

Figure 10. The predictor dialog box for October. A tab for each month is at the top, followed by the date and sounding time, then the predictor values. The 'Dismiss' button closes the GUI, the 'Reset Parameters' button resets the sounding parameters to the original values, and the 'Calculate Probability' button displays the probability output dialog box (Figure 11).

When the user clicks the 'Calculate Probability' button in the equation predictor dialog box, the probability of lightning occurrence for the day is displayed in a dialog box (Figure 11). The GUI code also outputs a file that contains all of the parameter values input by the user to calculate the probability. This file is currently named LtgProb.txt and resides in the MIDDS data directory.

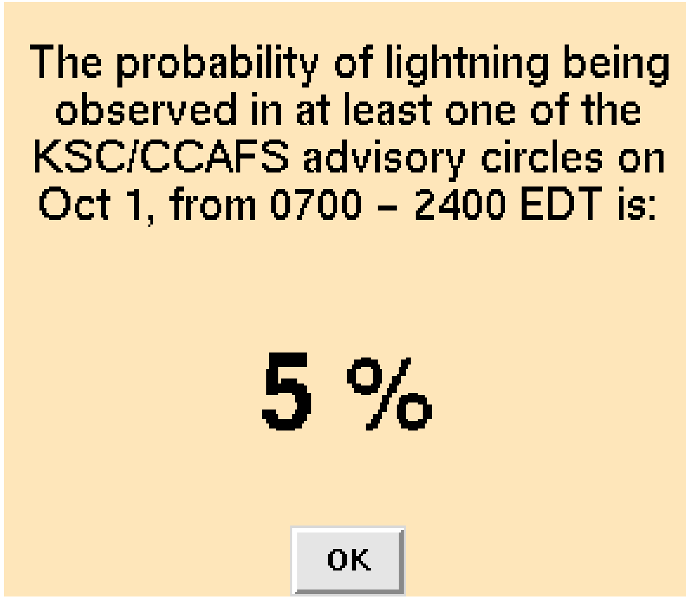

Figure 11. The output dialog box showing the probability of lightning occurrence for the day as calculated by the equation. The ‘OK’ button closes the box.

6. Summary

The AMU created new logistic regression equations in an effort to increase the skill of the Objective Lightning Forecast Tool developed in Phase II (Lambert 2007). The 45 WS requested the AMU make another attempt to stratify the data by lightning sub-season, add three warm seasons to expand the (POR) to 23 years (1989-2011) and include data from October. The AMU did this by using lightning observations across central Florida from the NLDN instead of the CGLSS data used in Phase III. In the event that lightning sub-seasons could not be identified, the 45 WS requested that the AMU create monthly equations with six more years of data than used in Phase II.

Although the AMU was able to establish a correlation between the NLDN flash count start/ramp-up dates and the wet-season start dates used by the NWS MLB by removing days with NLDN flashes that were due to large-scale weather patterns, not all days with NLDN flashes could be eliminated based on these patterns. Determining the start/ramp-up of each lightning season in the POR still had some level of subjectivity. With the start/ramp-up correlation established, the AMU then looked for correlations between the NLDN data and other proposed lightning sub-seasons: lightning (plateau), ramp-down, and post. Just as determining the start of the lightning season using the NLDN data was subjective, so was determining sub-seasons. After the AMU presented these findings to the 45 WS personnel who would use the tool, all agreed it would be difficult for the forecasters to determine the lightning sub-seasons in real-time using NLDN data. Therefore, the AMU completed with this task using monthly stratifications to develop the new equations.

The procedures used to create the predictors and develop the equations were identical to those in Phase II. The equations were made up of one to five predictors. Flow regime probability was the only predictor in all six months. The two strongest predictors in June, July, August and September were TI and flow regime probability, varying between first and second most important. Flow regime probability and TI were also strong predictors in May but persistence was more important than flow regime probability. The performance of the new equations was compared to that of five other forecast methods including the Phase II equations being used in operations. The new equations outperformed four other forecast methods by 7–59% using the verification dataset, but the new equations were outperformed by the Phase II equations in May, July, August and September. The new equations outperformed the Phase II equations in June by only 3%. Since there were no previous equations for October, no comparison could be made. Based on these results, the Phase IV equations did not replace the Phase II equations in operations.

The AMU updated the MIDDS tool by adding the October equations and updating the GUI. Through extensive testing, the AMU discovered three errors caused by interactions between the Objective Lightning Probability tool and Severe Weather Forecast tool. The errors were fixed and the updated tool was implemented on the operations MIDDS.

The AMU did not conduct formal training since the only change to the GUI was the addition of October equations. The discovery of the errors and subsequent fixes were communicated to the 45 WS by the AMU.

References

Crawford, W., 2010: Objective Lightning Probability Forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase III. NASA Contractor Report CR-2010-216292, Kennedy Space Center, FL, 34 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL, 32931, and http://science.ksc.nasa.gov/amu/final-reports/objective-ltg-fcst-phase3.pdf.]

Lambert, W. and M. Wheeler, 2005: Objective lightning probability forecasting for Kennedy Space Center and Cape Canaveral Air Force Station. NASA Contractor Report CR-2005-212564, Kennedy Space Center, FL, 54 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL, 32931, and http://science.ksc.nasa.gov/amu/final-stfi.html]

Lambert, W., 2007: Objective lightning probability forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase II. NASA Contractor Report CR-2007-214732, Kennedy Space Center, FL, 59 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL, 32931, and http://science.ksc.nasa.gov/amu/final-stfi.html]

Lascody, R., 2002: The Onset of the Wet and Dry Seasons in East Central Florida – A Subtropical Wet-Dry Climate?, http://www.srh.noaa.gov/mlb/?n=wetdryseason.

NOAA, 2011: NOAA Central Library U.S. Daily Weather Maps Project, http://docs.lib.noaa.gov/rescue/dwm/data_rescue_daily_weather_maps.html.

TIBCO, 2010: TIBCO Spotfire S+® 8.2 Programmer's Guide, TIBCO Software Inc., Seattle, WA, 532 pp.

Wilks, D. S., 2006: Statistical Methods in the Atmospheric Sciences. 2d ed. Academic Press, Inc., San Diego, CA, 467 pp.

List of Acronyms

14 WS	14th Weather Squadron
45 WS	45th Weather Squadron
AMU	Applied Meteorology Unit
CCAFS	Cape Canaveral Air Force Station
CGLSS	Cloud-to-Ground Lightning Surveillance System
CSR	Computer Sciences Raytheon
GUI	Graphical User Interface
JAX	Jacksonville, FL 3-letter identifier
KI	K-Index
KML	Keyhole Markup Language
KSC	Kennedy Space Center
LCC	Launch Commit Criteria
LI	Lifted Index
MCO	Orlando International Airport, Fla. 3-letter identifier
MFL	Miami, Fla. 3-letter identifier
MIDDS	Meteorological Interactive Data Display System
MLB	Melbourne, Fla. 3-letter identifier
MSE	Mean Squared Error
NLDN	National Lightning Detection Network
NOAA	National Oceanic and Atmospheric Administration
NWS MLB	National Weather Service Melbourne, Fla.
POR	Period of Record
PW	Precipitable Water
RH	Relative Humidity
SS	Skill Score
TBW	Tampa, FL 3-letter identifier
Tcl/Tk	Tool Command Language/Toolkit
TI	Thompson Index
VBA	Visual Basic for Applications
VT	Vertical Totals
XMR	CCAFS rawinsonde 3-letter identifier

NOTICE