NASA/CR—2013–217921

Objective Lightning Probability Forecasts for East-Central Florida Airports

Winifred C. Crawford
ENSCO, Inc., Cocoa Beach, Florida
Kennedy Space Center, Florida

May 2013

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—2013–217921

Objective Lightning Probability Forecasts for East-Central Florida Airports

Winifred C. Crawford
ENSCO, Inc., Cocoa Beach, Florida
Kennedy Space Center, Florida

National Aeronautics and
Space Administration

Kennedy Space Center
Kennedy Space Center, FL 32899-0001

May 2013

Acknowledgments

The author thanks Mr. Matthew Volkmer and Mr. David Sharp of the National Weather Service in Melbourne, FL for their assistance in determining the thresholds for the speed-stratified flow regimes, and their input on the design of the graphical user interface used to display the lightning occurrence probabilities. The AMU team also played a significant role in the success of this task by providing algorithms for and testing the GUI to ensure it contained no errors.

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 forecasters at the National Weather Service in Melbourne, FL, (NWS MLB) identified a need to make more accurate lightning forecasts to help alleviate delays due to thunderstorms in the vicinity of several commercial airports in central Florida at which they are responsible for issuing terminal aerodrome forecasts. Such forecasts would provide safer ground operations around terminals, and would be of value to Center Weather Service Units serving air traffic controllers in Florida. To improve the forecast, the AMU was tasked to develop an objective lightning probability forecast tool for the airports. The resulting forecast tool is similar to that developed by the AMU to support space launch operations in previous tasks.

The tool includes a set of equations that forecast the probability of lightning occurrence in the warm season months, defined as May-September, within 10 NM an airfield's center and within four three-hour increments: 1500-1800, 1800-2100, 2100-0000, and 0000-0300 UTC. These are the times when lightning is most likely to occur in east-central Florida. The airfields are Orlando International Airport, Melbourne International Airport, and Space Coast Regional Airport. The 10 NM radius and three-hour time increments reflect the forecast requirements at the airfields for which NWS MLB has forecast responsibility.

The period of record (POR) for this task was the warm season months in the 17 years 1995-2011. The data sources included the National Lightning Detection Network (NLDN) flash data in an area covering central Florida, 1000 UTC CCAFS soundings (XMR), and 1200 UTC Jacksonville (JAX), Tampa (TBW) and Miami (MFL), FL, soundings. Data from NLDN were used to determine lightning occurrence. The 1200 UTC JAX, TBW, and MFL soundings were used to calculate the daily flow regimes, and the 1000 UTC XMR soundings were used to calculate the standard stability parameters and to help determine the flow regime on each day. The AMU processed the datasets to create the predictand and candidate predictors needed for the statistical forecast equation development. The predictand is the element to be predicted from a predictor or group of predictors.

Before developing the equations, the AMU stratified the data into development and verification datasets. The AMU then developed a set of four equations, one for each time period, for each warm season month and each airfield. The performance of the equations was assessed using a technique appropriate for probability forecasts. Although the AMU tested, refined, re-created, and further tested the equations, their performance was mixed with a slight overall improvement over other forecast methods. The AMU concluded that stratifying the data into the shorter time periods reduced the number of lightning occurrences in each time period such that robust statistical relationships could not be realized. Nonetheless, NWS MLB requested that the equation output be provided through a graphical user interface (GUI) along with the other climatological values. They are still interested in seeing the equation output since it will be calculated using parameters from the current sounding and may still provide added value to the forecast.

The forecasters interact with the equations through a Microsoft® Excel© GUI. The GUI accesses data in specific worksheets based on user input, then outputs the equation, the daily climatology, and flow regime lightning probabilities for the forecasters to compare. NWS MLB was involved in the GUI development by providing comments and suggestions on the design to ensure that the final product addressed their operational needs.

Future work should involve collecting more data over a longer POR, developing equations for more airfields, and conducting other verification tests of the equations to determine all aspects of their performance. This will create equations that provide a good estimate of lightning occurrence at the airfields in the time periods of interest and give the forecasters a good first guess from which to build a forecast from other data sources.

Table of Contents

List of Figures

List of Tables

1 Introduction

The forecasters at the National Weather Service in Melbourne, FL, (NWS MLB) identified a need to make more accurate lightning forecasts to help alleviate delays due to thunderstorms in the vicinity of several commercial airports in central Florida at which they are responsible for issuing terminal aerodrome forecasts. Such forecasts would also provide safer ground operations around terminals, and would be of value to Center Weather Service Units serving air traffic controllers in Florida. To improve the forecast, the AMU was tasked to develop an objective lightning probability forecast tool for the airports using data from the National Lightning Detection Network (NLDN). The resulting forecast tool is similar to that developed by the AMU to support space launch operations at Kennedy Space Center (KSC) and Cape Canaveral Air Force Station (CCAFS) for use by the 45th Weather Squadron (45 WS) in previous tasks (Lambert and Wheeler 2005, Lambert 2007). The lightning probability forecasts are valid for the time periods and areas needed by the NWS MLB forecasters in the warm season months, defined in this task as May-September.

1.1 Previous Tasks

The AMU developed lightning forecast tools in previous tasks for use by NWS MLB and the 45 WS that are in current operational use. The data and results from those tasks formed the basis of the work done in this task.

1.1.1 NWS MLB

In the first lightning forecasting task for NWS MLB, the AMU created 6-, 12- and 24-hour gridded lightning density and frequency climatologies (Lambert et al. 2006) based on synoptic-scale flow regimes over the Florida peninsula (Lericos et al., 2002) during the warm season months. To supplement these gridded climatologies, the AMU created climatological soundings of wind speed and direction, temperature and dew point temperature at Jacksonville (JAX), Tampa (TBW), Miami (MFL), and CCAFS for each of the flow regimes (Short 2006). NWS MLB then requested to have these soundings available for display in the Advanced Weather Interactive Processing System (AWIPS) for overlay onto current soundings (Barrett 2008), allowing the forecasters to compare the current state of the atmosphere with climatology. This helped them adjust the lightning probability forecasts in their county warning area (CWA).

As an extension of the first task, the AMU created lightning climatologies based on the flow regimes for 5-, 10-, 20-, and 30-NM circles and 1-, 3-, and 6-hour increments for several airfields in the CWA (Bauman 2007, 2009). In the most recent task, the AMU added more airfields, and included a moisture stratification based on precipitable water and a stability stratification based on Thompson Index (TI) (Bauman 2011). This created lightning occurrence climatologies for environments from moist and unstable to dry and stable, providing valuable insight to how the lightning climatologies can change based on these atmospheric parameters.

1.1.2 45 WS

The AMU developed an objective lightning probability forecast tool for KSC/CCAFS consisting of a set of five equations, one for each warm season month, that calculate 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). These equations showed an improvement in performance over other standard forecast methods in use and were transitioned to operations.

The goal of two more tasks (Crawford 2010, Bauman 2012) was to create new 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. Neither task resulted in equations that outperformed those created in the earlier tasks. Therefore, the MIDDS tool mentioned previously that has an equation for each month is still in operational use, with the addition of an equation for October.

As an extension of the task done for NWS MLB mentioned in Section 1.1.1, the AMU also created lightning climatologies based on synoptic-scale flow regimes for 5-, 10-, 20-, and 30-NM circles and 1-, 3-, and 6-hour increments for the Shuttle Landing Facility, CCAFS Skid Strip and Patrick Air Force Base (Bauman 2009, 2011). The procedures followed for this work were the same as for the NWS MLB task.

1.2 Current Work

This task combines the warm season equation development done for the 45 WS with the area around airfields and time period stratifications done for NWS MLB. Specifically, the equations forecast the probability of lightning occurrence within 10 NM the airfield center and within four three-hour increments in the time period 1500-0300 UTC (1100-2300 local time): 1500-1800, 1800-2100, 2100-0000, and 0000-0300 UTC. These are the times when lightning is most likely to occur in east-central Florida. The airfields include the Orlando International Airport (MCO), Melbourne International Airport (MLB), and Space Coast Regional Airport (TIX). The 10 NM radius and three-hour time increments reflect the forecast requirements at the airfields for which NWS MLB has forecast responsibility. The airfield locations and 10 NM radius circles are shown in Figure 1. The forecasters interact with the equations through a Microsoft® Excel© GUI. They input values for the predictors and the GUI outputs the probability of lightning for the airfield and time period of interest.

This report will describe the data used in Section 2, the data analysis in Section 3, equation development and testing in Section 4, a description of the GUI in Section 5, and conclusions in Section 6.

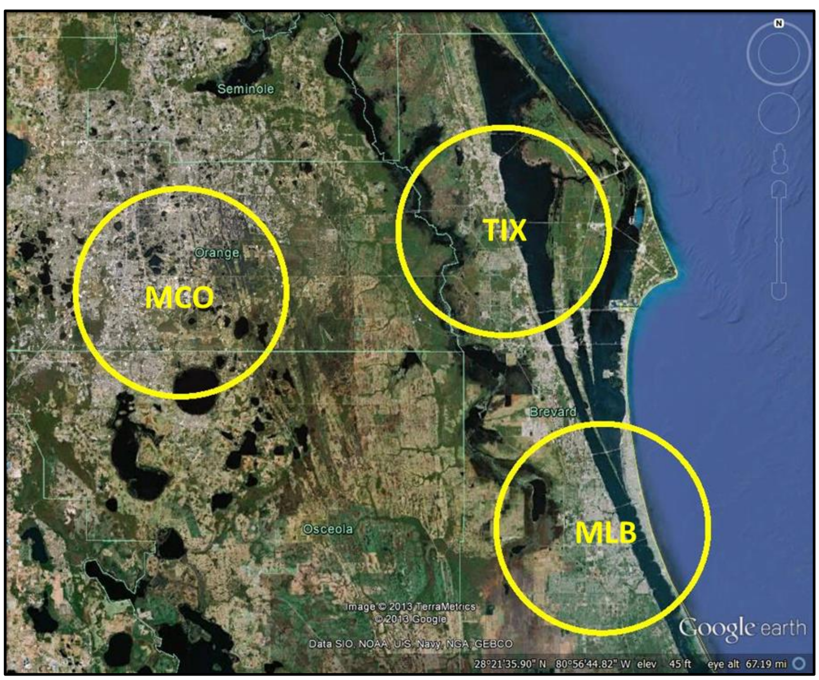


Figure 1. Google Earth image showing the 10 NM radius range rings in yellow for MCO, MLB, and TIX. The county boundaries and names are in green.

2 Data

The period of record (POR) for this task is the warm season months of May-September in the 17 years 1995-2011. The data sources include the

- NLDN flash data in an area covering central Florida,
- 1000 UTC CCAFS soundings (XMR), and
- 1200 UTC JAX, TBW and MFL soundings.

Data from NLDN were used to determine lightning occurrence. The 1200 UTC JAX, TBW, and MFL soundings were used to calculate the daily flow regimes, and the 1000 UTC XMR soundings were used to calculate the standard stability parameters and to help determine the flow regime on each day. The following sections describe these data types and how they were processed prior to the creation of the predictors and predictand for the forecast equations. All data were processed using the TIBCO Spotfire S+® software package (TIBCO 2010).

2.1 NLDN

NLDN is a national network of cloud-to-ground lightning sensors owned by Vaisala (Cummins and Murphy 2009). The NLDN data were provided to the AMU through Mr. Roeder of the 45 WS by the 14th Weather Squadron. The files contain the date, time in UTC, latitude, longitude, polarity, and strength of every strike in an area that encompassed central Florida. These data were used to determine whether or not lightning occurred on each day in the POR. The primary purpose of the NLDN data was to create the binary predictand for the equations. The data were also used to create the daily climatological lightning frequency and persistence forecasts that would be used as candidate predictors and forecast benchmarks against which to test the new equations.

The NLDN data also caused a change in the POR. The original POR was 1989–2011. During a presentation at Vaisala's International Lightning Meteorology Conference in April 2012 (http://www.vaisala.com/en/events/ildcilmc/Pages/ILDC-2012-archive.aspx), the AMU learned the NLDN system underwent a major upgrade in 1994 (Cummins et al. 1998), causing researchers to not use data from before that year (Hodanish 2012). Dr. Ken Cummins provided the AMU with an image of 2°x2° grids containing NLDN detection efficiency (DE) correction values over the U.S. during 1994-1998 relative to 1999 DE, shown in Figure 2. The DE correction values are proportional to the DE in each grid cell. To normalize the flash counts and make them consistent with 1999 performance, the number of strikes detected is divided by the DE correction for that year. For a DE correction < 1, this results in a larger number of strikes.

Since the lightning forecast in this task depends on whether lightning occurred and not the number of strikes, the DE corrections values in Figure 2 were not used. However, the low DE of 0.5-0.6 over Florida in 1994 could indicate strikes were missed by NLDN in the locations and times of interest. After consulting with Mr. Matt Volkmer and Mr. Dave Sharp of NWS MLB, the AMU deleted warm season data from the six years 1989-1994, years in which the DE correction over Florida was < 0.7. This eliminated years with low DE while keeping as many years as possible in the POR for equation development and testing, which is the 17 years 1995-2011.

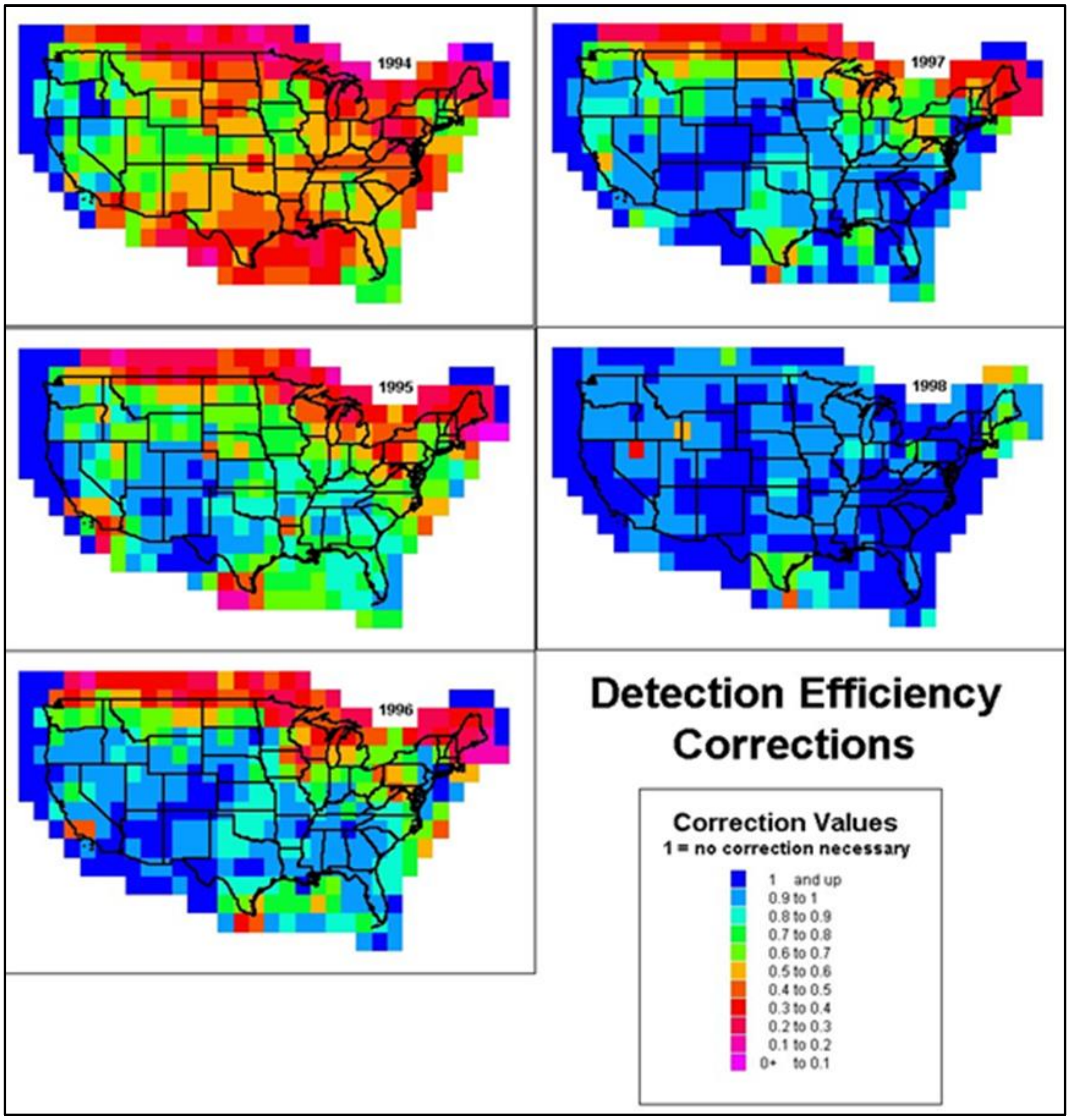


Figure 2. The NLDN DE corrections for 1994-1998 relative to 1999 values. The grid cell size is 2°x2° latitude/longitude (image created by Vaisala).

2.2 Florida 1200 UTC Rawinsondes

The AMU downloaded the 2010 and 2011 soundings for JAX, TBW and MFL from the NOAA Earth System Research Laboratory (ESRL) website (http://www.esrl.noaa.gov/raobs/) and added them to the existing AMU database of soundings from 1995-2009. These soundings were used to calculate the daily flow regimes as described in Lericos et al. (2002). More information about this data set can be found in Lambert (2007).

2.3 XMR 1000 UTC Rawinsonde

Mr. Madison of Computer Sciences Raytheon provided the 2010 and 2011 XMR soundings. The AMU added them to the existing database of soundings from 1995-2009. These data were used to supplement the Florida 1200 UTC soundings in determining the flow regime of the day and to calculate the sounding parameters normally available to the forecasters through MIDDS. The probability of lightning occurrence based on flow regime and the XMR sounding parameters were used as candidate predictors in the equation development.

3 Creation of Equation Predictand and Predictors

The AMU processed the three datasets described in Section 2 to create the predictand and candidate predictors needed for the statistical forecast equation development. The predictand is the element to be predicted from a predictor or group of predictors. There was one predictand value and one set of candidate predictor values per day in the POR. More details of how the data were processed to create these elements are given in Lambert (2007).

3.1 Binary Predictand

The predictand is binary, i.e. 1 or 0, and indicates whether lightning occurred within 10 NM of the centers of MCO, MLB, and TIX, and during the four three-hour time periods consistent with the NWS MLB required forecast products. This determination was straightforward: the predictand value was set to 1 if lightning was detected within each defined time period and spatial area on each day, otherwise a 0 was assigned. The AMU used a binary predictand because the prediction would be for lightning occurrence, not the number of strikes.

3.2 Candidate Predictors

The AMU tested the candidate predictors during equation development to determine which predictors in what combination would provide the best probability forecast of lightning occurrence. The candidate predictor set included one-day persistence and daily climatological lightning frequency calculated from the NLDN binary predictand, the flow regimes determined from the morning rawinsondes, and 12 stability parameters calculated from the XMR morning rawinsonde.

3.2.1 NLDN Predictors

The AMU processed the NLDN data to create the one-day persistence and daily climatology needed for equation development. One-day persistence was binary like the predictand, and just as straightforward to determine. It indicates whether lightning occurred on the previous day for each three-hour period. If lightning occurred on one day in a certain time period, the persistence value for the next day in the same time period was 1. If lightning did not occur, the persistence value was 0. The daily climatology is the percent of days lightning occurred on each date in the POR. The binary predictand values were used to create the daily climatology. Details on how the values were calculated are in the Phase II final report (Lambert 2007).

Figures 3-5 show the raw and smoothed daily climatologies for the four three-hour time periods at each of the stations. These values were calculated using the same 14-day Gaussian smoothing algorithm described in Lambert (2007). For MCO (Figure 3), farther from the coast than the other two stations, the values for 1800-2100 UTC and 2100-0000 UTC are similar as are the values for 1500-1800 UTC and 0000-0300 UTC. For MLB (Figure 4) and TIX (Figure 5), the values for 1800-2100 UTC are highest, and the time periods before (1500-1800 UTC) and after (2100-0000 UTC) have values that are more similar. This may show a ramp up and down of lightning occurrence at these two stations as the sea breeze forms close to the coast in the late morning/early afternoon and moves inland in the afternoon, creating higher values for MCO during the mid- and late-afternoon periods.

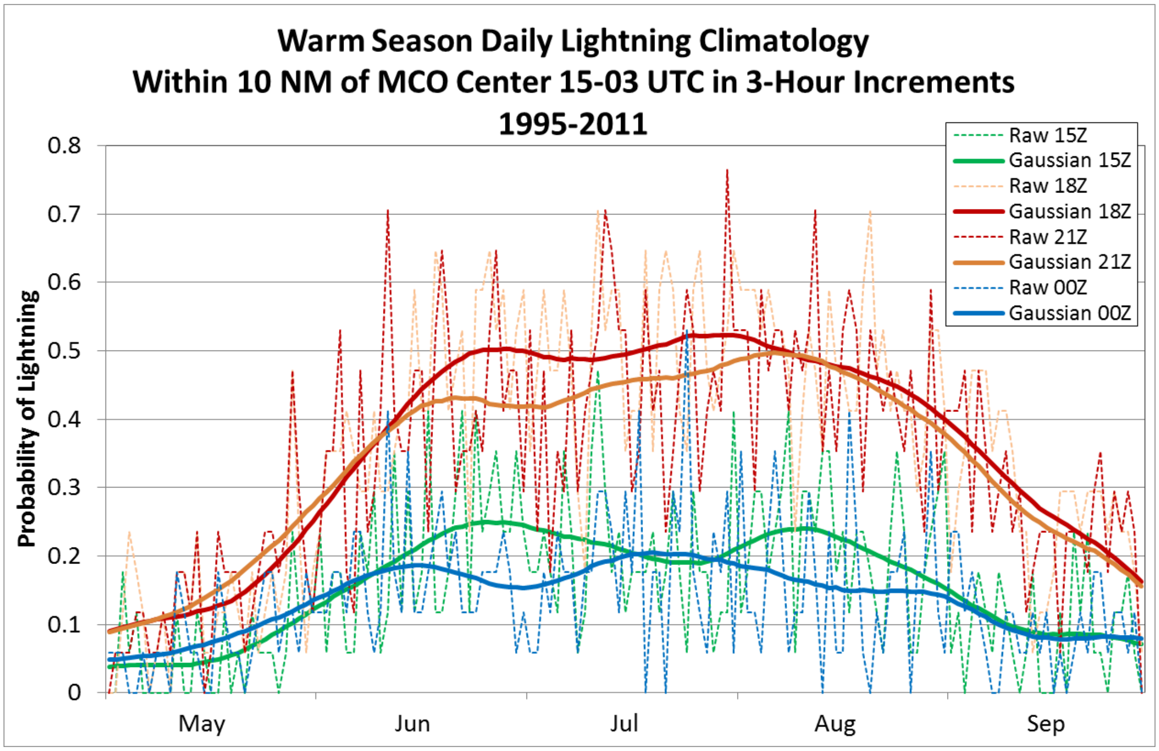


Figure 3. The daily raw and Gaussian-smoothed lightning climatology at MCO for the four three-hour periods 1500-1800, 1800-2100, 2100-0000, and 0000-0300 UTC. The time values in the legend indicate the beginning of the time period.

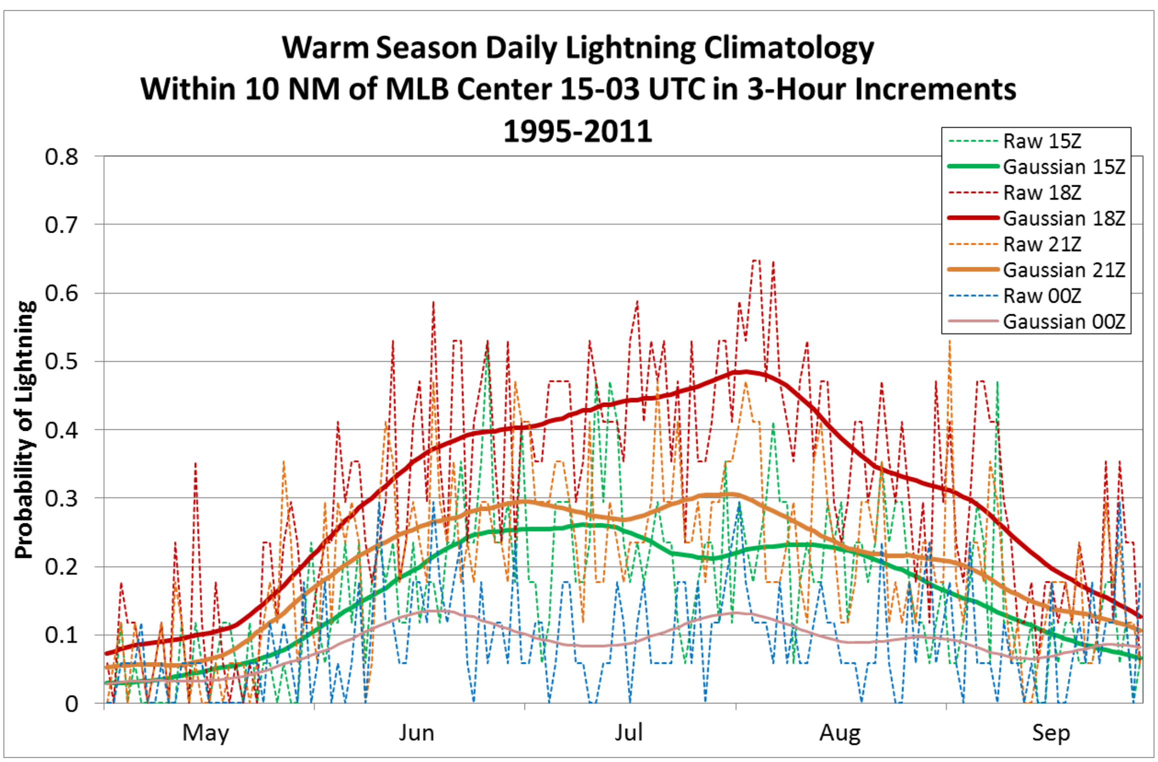


Figure 4. Same as Figure 3 but for MLB.

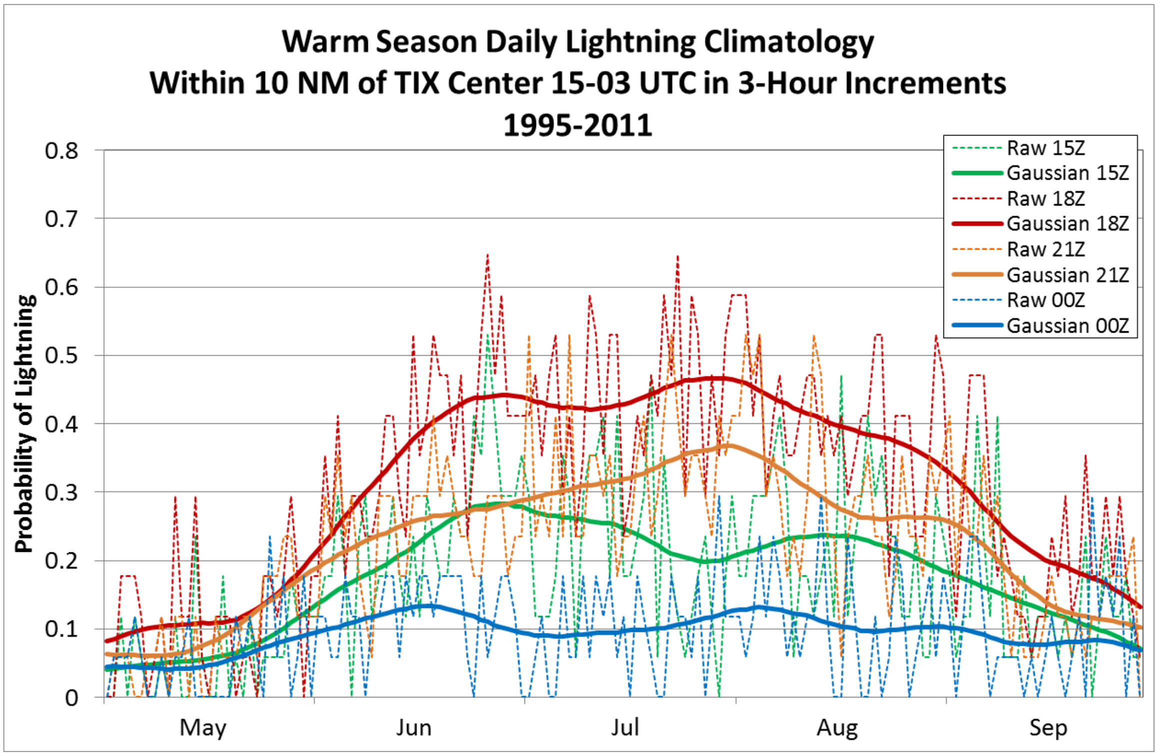


Figure 5. Same as Figure 3 but for TIX.

3.2.2 Flow Regime Probabilities

The AMU used the sounding data from MFL, TBW, JAX, and XMR to determine the flow regime on each day in the 2010 and 2011 warm seasons using the process described in Lambert (2007). The AMU already calculated the flow regimes for the years 1995-2009 for previous AMU tasks. Table 1 contains the flow regime definitions.

Table 1. List of the flow regime names and the corresponding sectors showing the average 1000-700 mb wind directions at each of the stations (Lambert 2007).

Flow Regime Name and Description		***Rawinsonde Station*** MFL	TBW	JAX
SW-1	Subtropical ridge south of MFL Southwest flow over KSC/CCAFS	180°-270°	180°-270°	180°-270°
SW-2	Subtropical ridge north of MFL, south of TBW Southwest flow over KSC/CCAFS	90°-180°	180°-270°	180°-270°
SE-1	Subtropical ridge north of TBW, south of JAX Southeast flow over KSC/CCAFS	90°-180°	90°-180°	180°-270°
SE-2	Subtropical ridge north of JAX Southeast flow over KSC/CCAFS	90°-180°	90°-180°	90°-180°
NW	Northwest flow over Florida	270°-360°	270°-360°	270°-360°
NE	Northeast flow over Florida	0°-90°	0°-90°	0°-90°
Other	When the layer-averaged wind directions at the three stations did not fit in defined flow regime			
Missing	One or more soundings missing			

The AMU calculated the frequencies of lightning occurrence for each warm season month under each daily flow regime and in each three-hour period within 10 NM of MCO, MLB, and TIX. Figure 6 shows the flow regime lightning frequencies for July at MLB. In general, the values are low at the beginning of the day, increase at mid- and late-day, and decrease into the evening. This was the general trend for all months. The values for the flow regimes with an easterly component (NE, SE-1 and SE-2) tend to be highest in the morning and generally decrease through the day. This is consistent with morning showers experienced near the east coast in these regimes. Conversely, the values with a westerly component (NW, SW-1 and SW-2) increase from the morning to mid-day, slightly decrease but remain elevated in the late-day, and then decrease significantly after 0000 UTC (2000 EDT). This is consistent with westerly flow slowing the progression of the sea breeze inland or pinning it at the coast, providing a lifting mechanism for thunderstorm development, and the drop-off in convective activity after sunset.

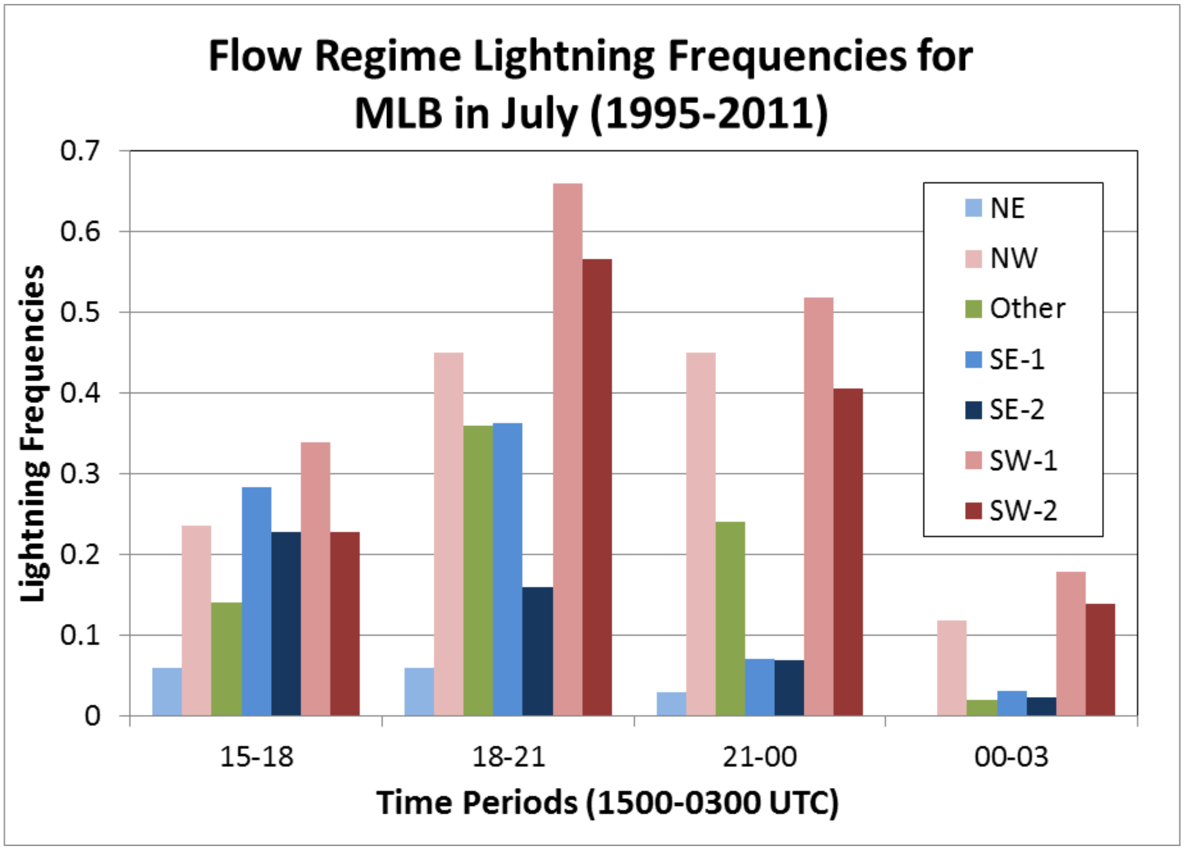


Figure 6. The flow regime lightning frequencies for MLB in July. Flow regimes with an easterly component (NE, SE-1 and SE-2) are in shades of blue and those with a westerly component (NW, SW-1 and SW-2) are in shades of pink. The Other regime is in green.

The AMU used the values in Figure 6 as predictors in developing the equations, which is described in Section 4. The equations did not perform well, so the AMU met with Mr. Volkmer and Mr. Sharp of NWS MLB to determine how to proceed with the task. They decided to combine the low-level mean wind speed with the flow regime on each day to create flow regime/speed stratifications for these predictors. Speed strength influences how far the sea breeze will penetrate inland, which has a direct effect on the location of any storms that form. The AMU used the mean speed in the 1000-700 mb layer from the CCAFS 1000 UTC sounding to create two- and three-speed ranges for each flow regime. For the two-speed ranges, speeds < 10 kt were considered low and speeds ≥ 10 kt were considered high. For the three-speed ranges, speeds < 6 kt were considered light, speeds ≥ 6 and ≤ 14 kt were considered moderate, and speeds > 14 kt were considered strong.

Stratifying the data by flow regime and speed ranges reduced the number of observations in each category such that robust and meaningful lightning occurrence frequencies could not be calculated. In order to increase the number of observations in each stratification, the AMU combined the two southeast (SE-1 and SE-2) and two southwest (SW-1 and SW-2) regimes into one regime each (SE and SW) before stratifying by speed ranges. The lightning frequencies for these two flow regimes and the three-speed ranges at MLB in July are shown in Figure 7. As an example, lightning occurred 62% of the time in the 1800-2100 UTC period when the flow regime was SW and the mean 1000-700 mb wind speed was in the strong range, or > 14 kt. The NW, NE, and Other regimes were also stratified by speed ranges but are not shown in Figure 7 to make the chart less cluttered and easier to interpret. These values were used as candidate predictors in addition to the flow regime probabilities not stratified by speed (Figure 6).

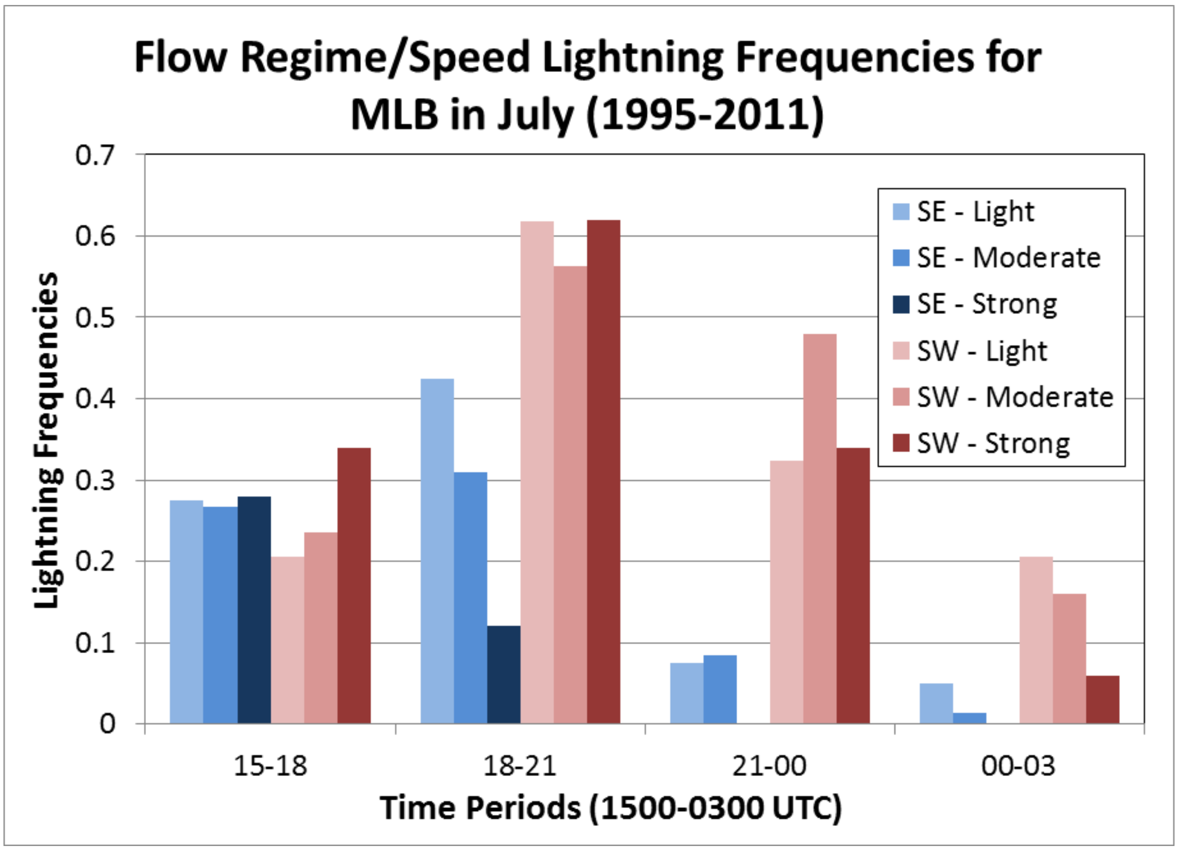


Figure 7. The frequencies of lightning occurrence for the southeast (SE, shades of blue) and southwest (SW, shades of pink) flow regimes for the three-speed ranges of light, moderate and strong winds.

3.2.3 Stability Parameters

The AMU calculated 12 stability indices and moisture parameters from the 2010 and 2011 XMR soundings needed as predictors for equation development. The 1995-2009 warm season values were created for earlier AMU tasks. They are

- Total Totals (TT),
- Cross Totals (CT),
- Vertical Totals (VT),
- K-Index (KI),
- Lifted Index (LI),
- TI (KI – LI),
- Severe Weather ThrEAT Index (SWEAT),
- Showalter Index (SI),
- Temperature at 500 mb (T_{500}),
- Mean Relative Humidity (RH) in the 825–525 mb layer (RH_{85}),
- Mean RH in the 800–600 mb layer (RH_{86}), and
- Precipitable Water up to 500 mb (PW).

The formulas in the MIDDS code used for the indices are standard and can be found in several sources (e.g. Peppler and Lamb 1989; Ohio State University Severe Weather Products web page at http://twister.sbs.ohio-state.edu). The formulations will not be shown here. Only four indices in the above list are not readily available to the forecasters: VT, TI, RH_{85} and RH_{86}. TI is calculated easily with the equation TI = KI – LI, as is VT with VT = $T_{850} - T_{500}$. RH_{85} and RH_{86} were calculated using a log(p)-weighted average described in detail in Lambert (2007: Section 3.5 equations 5 and 6).

3.2.4 Summary of Candidate Predictors

The full list of the candidate predictors, 17 in all, is given here as a reference. They are

- One-day persistence for each time period (Pers),
- Daily climatological lightning frequency for each time period (DClimo),
- Flow regime lightning probability (FR),
- Two-speed stratified flow regime lightning probability (FR2),
- Three-speed stratified flow regime lightning probability (FR3),
- TT,
- CT,
- VT,
- KI,
- LI,
- TI,
- SWEAT,
- SSI,
- T_{500},
- RH_{85},
- RH_{86}, and
- PW.

The values for these candidate predictors were used with the binary predictand in the development of the logistic regression lightning forecast equations.

4 Equation Development and Testing

There were three major steps in this portion of the task:

- Create the development and verification datasets,
- Develop the logistic regression equations, and
- Determine the equation performance.

Before developing the equations, the AMU stratified the data into development and verification datasets. The AMU then developed a set of four equations, one for each time period, for each warm season month. The performance of the equations was assessed using the Brier Skill Score (SS; Wilks 2006), a verification technique appropriate for probability forecasts.

4.1 Development and Verification Datasets

The AMU created the equations using the development dataset, which required enough samples so that the resulting equations performance was stable, i.e. the equations would maintain consistent forecast accuracy on different datasets. A small dataset is less likely to contain a representative set of events. The equations developed from a small set may show wide variations in accuracy on different datasets causing forecasters to not have confidence in the results. The verification dataset was needed for equation testing in order to have a more realistic view of how the equations would perform in operations. It was expected that the equations would not perform as well on the verification data as they would on the data from which they were developed. However, if performance was a great deal worse with the verification data, this would indicate that too many predictors were chosen and the equations were fit too strongly to the development data, the development dataset was too small, or the equations did not have predictive value.

The AMU stratified candidate predictors and predictand for each month into development and verification datasets. Care was taken to ensure there would be at least 250 events in the development dataset (World Meteorological Organization (WMO) 1992), while still having enough events in the verification dataset to make reasonable conclusions about equation performance. Of the 17 warm seasons in the POR, 14 were used for equation development and 3 were set aside for equation verification. This ensured that each month in the warm season was equally represented in both datasets.

The stratification did not involve choosing individual warm season years for each dataset, but rather individual warm season days. Days for the verification dataset were chosen first. Given that there are 153 days in the warm season, the random number generator in Microsoft® Excel© was used to create three sets of 153 numbers representing the years between and including 1995 and 2011. The resulting three sets of years were assigned to each day in the warm season. Thus, each day in the warm season was represented by days from three random "years". The dates were manually checked to ensure there were no duplicate years for each day from the random number generator. For example, the verification dataset contains 1 May 1995/2005/2007, 2 May 1999/2004/2009, etc. All other dates were made part of the development dataset. This random method was chosen to reduce the likelihood that any unusual convective seasons would bias the results. The development datasets had well above the 250 events defined by the WMO needed to develop reliable equations.

4.2 Equation Development

As in previous work, the AMU used logistic regression to create the lightning forecast equations. Due to time constraints, equations were developed for MCO and MLB but not TIX. Four time periods, five warm season months, and two airfields resulted in 40 equations to be developed, 20 for each airfield. The AMU selected predictors using the same method as in previous work for each individual equation. Detailed descriptions of logistic regression and the predictor selection procedure with supporting figures and equations are found in the Phase II final report (Lambert 2007). Those procedures were followed exactly for this task.

4.2.1 Development Issues

The AMU used SS values as the first test in determining equation performance. SS shows the percent improvement or degradation the equations provide compared to other forecast methods, or benchmarks. Positive SS values indicates the equations outperform the forecast benchmarks, negative values indicate that the equations were worse than the forecast benchmarks.

Initially, the AMU began equation development for MCO as requested by NWS MLB. The MCO equations were tested and refined using SS values to determine their performance compared to Pers, DClimo and FR probability values. All equations outperformed Pers by significant amounts, which was also true for the final set of equations. However, results compared to DClimo and FR showed similar performance. More than half of the SS values were less than 10% indicating that the equations performed similarly to the benchmarks, and close to 20% of the values were negative. The equations should provide a good first guess when forecasters are determining the probability of lightning, but it appeared that this set of equations would be ineffective in doing so.

The final equation performance, shown in Section 4.2.2, was also similar to the other forecast benchmarks, with just a slight improvement. The AMU conducted tests to determine the cause for this unexceptional performance and consulted with NWS MLB during the process. These consultations resulted in creating the speed-stratified flow regimes described in Section 3.2.2. Several attempts at developing and re-developing the equations did not improve their performance. After much examination, the AMU concluded that two major issues caused the equations to not outperform DClimo and FR probability values: the percentage of lightning occurrence in the datasets and the operationally defined time and space restrictions for lightning occurrence.

4.2.1.1 Lightning Occurrence Ratios

The unremarkable equation performance prompted the AMU to compare lightning occurrence in the development data set with that in the verification data set. A large percentage of lightning in one data set and a small percentage in the other could cause the bad performance shown by the SS values. Table 2 shows the percentage of lightning days in the MCO development and verification data sets for each month and time period. The equations were developed with the development data and tested on the verification data. The values in Table 2 were calculated to determine if both data sets contained similar ratios of lightning to non-lightning days. That most values are within 10% indicates the data sets were similar with respect to lightning occurrence. The largest outlier of 11% occurred for May 1800-2100 UTC. Note the low frequencies in May and September for all four time periods, and during 1500-1800 and 0000-0300 UTC for the other months. Such low percentages likely made it difficult to determine strong relationships between the predictors and lightning occurrence during these periods. The larger percentages for 1800-2100 and 2100-0000 UTC in June, July, and August, however, appeared to show enough occurrences to develop strong predictand/predictor relationships.

Table 2. The percentage of lightning days in the MCO development and verification data sets used to create and test the equations, respectively. The first column is the month, the second identifies the data set and the number of days it contains not including days with missing data. The last four columns show the percentages of days with lightning in each time period.

Month	***Data Set (# Days)***	***15-18***	***18-21***	***21-00***	***00-03***
May	Development (390)	5	11	12	6
	Verification (85)	12	22	22	13
June	Development (382)	21	45	42	16
	Verification (86)	22	42	40	24
July	Development (394)	21	50	44	17
	Verification (85)	16	45	44	20
August	Development (387)	24	49	45	15
	Verification (84)	17	45	49	18
September	Development (378)	8	28	29	10
	Verification (79)	11	20	22	10

4.2.1.2 Time and Space Filters

The AMU looked at the differences in the area and time filters between this task and the 45 WS's Objective Lightning Probability tool (Lambert 2007). That tool outperformed all forecast benchmarks by significant percentages. The area used for the 45 WS tool was somewhat larger than the 10 NM radius in this task but not significantly so. However, the 45 WS time period was 17 hours (0700 to midnight local time) as opposed to the four three-hour time periods during 1500 to 0300 UTC, which totals 12 hours (1100 to 2300 local time). If lightning occurred a few seconds before or after a time period, it would not be counted as a lightning event for that time period. Predictands that might otherwise indicate lightning occurrence would be paired with a non-lightning event even if lightning occurred just before or after the time period. The result is that a robust relationship between predictors and the predictand would not be fully realized by the equations, resulting in poor equation performance.

The AMU spoke with Mr. Sharp and Mr. Volkmer and suggested increasing the area and/or time period to help improve equation performance. There were similar issues with the time filter in another AMU task to use Global Positioning System (GPS) PW estimates to forecast lightning (Huddleston 2012). Mr. Sharp and Mr. Volkmer said the radius cannot be changed and the time periods should remain the same as they reflect their forecast requirements. They also requested that the equations be developed for MCO and MLB using the original requested area and time period filters, even if they did not perform significantly better than the DClimo and FR probabilities.

4.2.2 Final Equations

As described in Lambert (2007), the AMU chose the predictors for each equation in rank order of importance. Two stations, five months in the warm season (May-September), and four 3-hour time periods resulted in 40 equations. The number of predictors in the equations ranges from two to five, with an average of three. In lieu of showing the predictors for each equation, it is more informative to show which candidate predictors were chosen for what number of equations. Table 3 shows all the predictors used in the equations and the number of equations in which they appear, both for the 20 MCO and 20 MLB equations separately and for all 40 equations together. They are listed in order from the greatest number of equations to the least. The first and second ranked predictors for both the MCO and MLB equations are FR3 and TI. Thirty-nine equations have one of the flow regime predictors (FR3, FR2, and FR). This shows the importance of the flow regime in predicting when and where convection and lightning will occur. Representing stability and moisture, 19 equations have TI, 6 have TT, and 6 have PW. The rest of the predictors were not chosen often, but were still important to forecasting lightning in their specific airfield 10 NM radius and time period. Note that the only two candidate predictors never chosen for the equations were T_{500} and SWEAT (Section 3.2.4).

Table 3. The number of times each predictor was used in the 20 MCO and 20 MLB equations, and the total number of times they were used in both sets of equations.

Predictor	MCO	MLB	TOTAL
FR3	15	11	26
TI	11	9	20
FR2	3	7	10
Pers	5	3	8
DClimo	4	3	7
PW	3	3	6
TT	3	3	6
KI	3	2	5
FR	1	2	3
LI	1	2	3
VT	2	1	3
RH_{85}	1	2	3
CT	0	2	2
RH_{86}	1	0	1
SSI	0	1	1

Developing and testing the equations involved creating a base equation, and then adding and eliminating predictors from it to find the combination that performed best using the verification data. The AMU calculated the SS values showing the percent improvement or degradation in skill of the final equations over five forecast benchmarks: Pers, DClimo, FR, FR2, and FR3. Table 4 contains these SS values showing the skill of the MCO and MLB equations relative to the other forecast methods using the verification data. The SS values for Pers are not shown since all equations outperformed that forecast benchmark by a significant amount.

The positive SS values in Table 4 indicate the equations had more skill than the corresponding forecast method, and negative values indicate less skill. Values with magnitudes within 10% of 0, positive or negative, are highlighted in yellow. The results for DClimo and the three FR values were mixed. Not considering the values within 10% of 0, the MCO equations outperformed all of the benchmarks 81% of the time, and the MLB equations 64% of the time. However, when assuming that values within 10% of 0 indicate similar performance, the MCO and MLB equations outperformed the benchmarks only 26% and 30% of the time, respectively.

While there were some time periods in which the equations showed a significant improvement over the forecast benchmarks, most of the values in Table 4 are > -10% and < 10%: 73% for MCO and 70% for MLB. Thus, through all the modifications and equation development iterations, the final equations still do not outperform the forecast benchmarks by a significant amount.

Table 4. The SS values (positive: black; negative: red) for the MCO and MLB equations for the forecast benchmarks of DClimo, FR, FR2, and FR3. The time periods are in UTC. Yellow shading indicates values within 10% of 0.

Station	***Forecast Benchmark***	***15-18***	***18-21***	***21-00***	***00-03***
MCO					
May	DClimo	1	25	16	2
	FR (Flow Regime)	6	25	18	-1
	FR2 (FR 2-Speed)	3	23	17	-1
	FR3 (FR 3-Speed	3	22	19	1
June	DClimo	16	-1	20	8
	FR (Flow Regime)	6	4	7	-2
	FR2 (FR 2-Speed)	4	1	1	-1
	FR3 (FR 3-Speed	7	6	6	3
July	DClimo	-4	7	28	4
	FR (Flow Regime)	-6	5	23	-2
	FR2 (FR 2-Speed)	2	5	20	-5
	FR3 (FR 3-Speed	-1	4	12	-6
August	DClimo	-1	11	6	11
	FR (Flow Regime)	0	8	1	10
	FR2 (FR 2-Speed)	-7	9	1	8
	FR3 (FR 3-Speed	-10	7	2	-2
September	DClimo	24	8	6	8
	FR (Flow Regime)	20	4	0	3
	FR2 (FR 2-Speed)	15	4	0	4
	FR3 (FR 3-Speed	13	1	5	3
MLB					
May	DClimo	2	26	16	-5
	FR (Flow Regime)	3	25	10	-6
	FR2 (FR 2-Speed)	0	24	9	-6
	FR3 (FR 3-Speed	-1	22	9	-9
June	DClimo	17	21	11	10
	FR (Flow Regime)	17	8	-7	-1
	FR2 (FR 2-Speed)	15	4	-4	-2
	FR3 (FR 3-Speed	14	7	-6	-5
July	DClimo	-1	-4	12	3
	FR (Flow Regime)	-2	-4	5	-1
	FR2 (FR 2-Speed)	-1	-3	4	0
	FR3 (FR 3-Speed	-2	-6	1	-2
August	DClimo	11	13	9	1
	FR (Flow Regime)	9	8	3	-1
	FR2 (FR 2-Speed)	10	6	4	-2
	FR3 (FR 3-Speed	7	8	2	-3
September	DClimo	28	29	9	4
	FR (Flow Regime)	19	14	4	-2
	FR2 (FR 2-Speed)	19	8	1	-5
	FR3 (FR 3-Speed	14	11	3	-4

5 Graphical User Interface

Even though the equations did not perform well, NWS MLB requested that their output be provided through a GUI along with the other forecast benchmarks DClimo, FR, FR2, and FR3. The forecasters are still interested in seeing the equation output since it is calculated using parameters from the current sounding and may still provide added value to the forecast. NWS MLB was involved in the GUI development by providing comments and suggestions on the design to ensure the final product addressed their operational needs.

The AMU built the GUI within an Excel workbook using Microsoft Excel Visual Basic for Applications (VBA). It accesses data in specific worksheets based on user input. The GUI itself has three basic dialog boxes. The first asks for the general information such as the airfield and time period of interest, the second asks for equation predictor values, and the third displays the equation and climatological probability values.

5.1 Excel Workbook

The Excel workbook in which the GUI resides contains 14 worksheets. The first worksheet contains brief instructions on how to start and use the GUI. The AMU recommends first-time users read these instructions in their entirety before using the GUI. The next three worksheets contain Excel PivotTables of the DClimo, FR, FR2, and FR3 values. The other 10 worksheets contain information for each station/month combination:

- Predictor names for each time period and their coefficients in the equations, and
- Minimum, maximum, median, mean, standard deviation, and first and third quartile values of the observed sounding stability indices for the month.

The first worksheet, named Introduction, is displayed automatically upon opening the Excel file. The user opens the GUI by clicking on the yellow button on the Introduction sheet. When the GUI is initiated, the first form requesting station, month, time period, and other input is displayed. This will define which equation or equations the user needs to access for the lightning probabilities. After choosing all of the parameters in this form and continuing, another form is displayed that allows the user to input the predictor values for the equation. Once all of these values are chosen, the final form displays the requested output. When the user is finished and exits out of all the dialog boxes, the Introduction sheet will be displayed again before closing the file.

5.2 Initial Input Form

Figure 8 shows the first form in the GUI to input initial values needed to output the lightning probability values. The month, day, station and time are needed for the daily climatology. These values plus the flow regime are needed for the flow regime probabilities, and all values including the mean speed in the flow regime layer are needed for the speed-stratified flow regime probabilities. The user must input the speed value manually, which is the average speed in the surface to 700 mb layer as observed in the XMR 1000 UTC sounding. The initial speed value is -999. If this value is not changed and the user clicks the "Continue…" button, the error message form on the right side of Figure 8 is displayed.

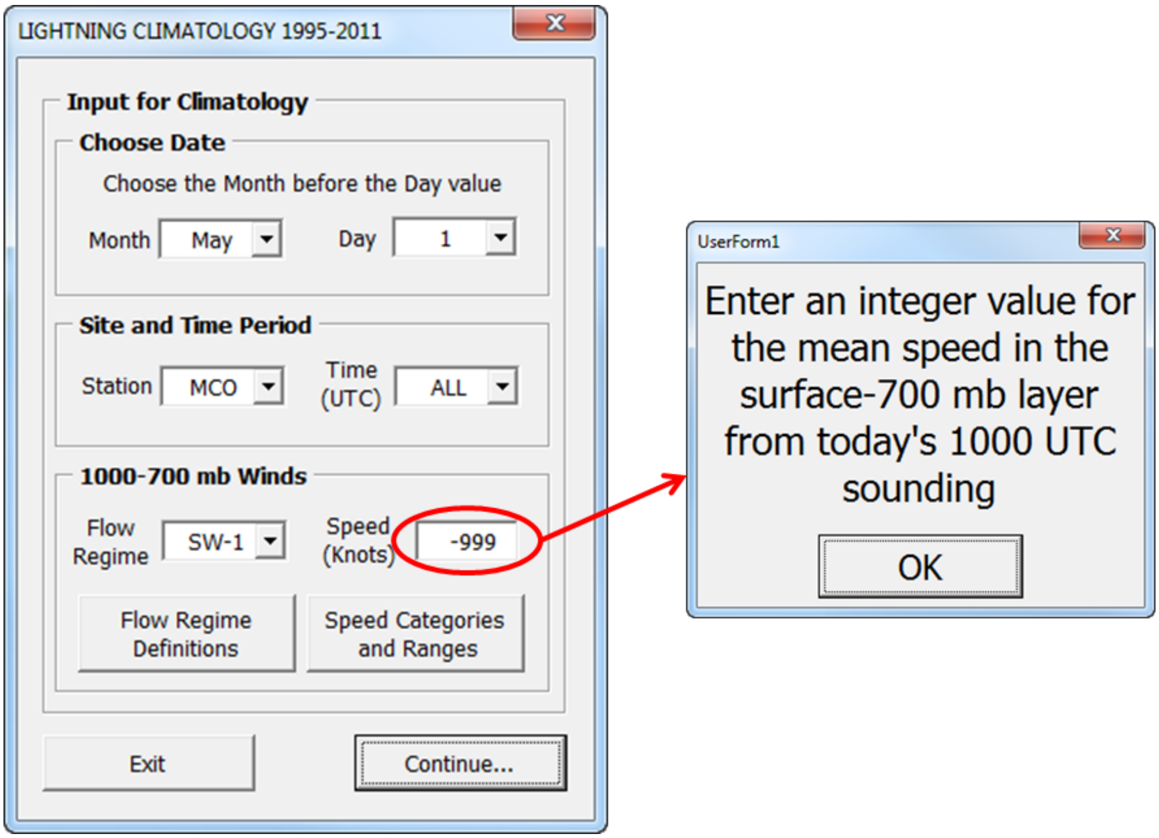


Figure 8. The initial input form showing the default values after starting the GUI (left). The user makes choices through the drop-down lists in the boxes with down arrows, manually inputs a value for speed and then clicks the "Continue…" button to get the predictor input form. If a speed value is not input and the user clicks the "Continue…" button, the form on the right is displayed reminding the user to input a speed value.

Figure 9 shows the initial input form after the user has made choices and with the time period drop-down list displayed. There are two other buttons on the initial input form that provide the user with information. The "Flow Regime Definitions" and the "Speed Categories and Ranges" buttons are surrounded by red rectangles in Figure 9. Arrows point to their respective message forms that provide definitions for all the flow regimes and speed categories. They can only be displayed one at a time and are closed by clicking their "Dismiss" buttons.

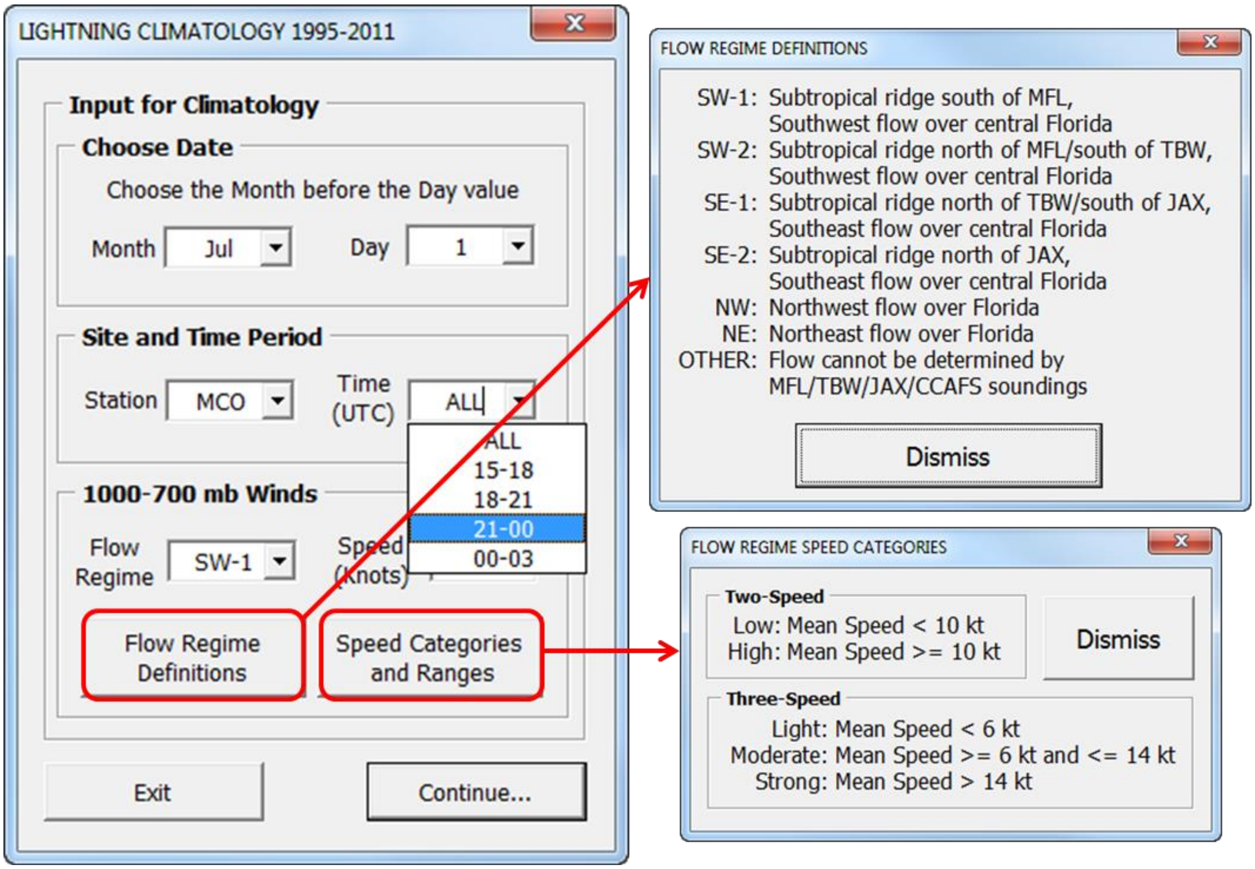


Figure 9. The GUI initial input form after the user has made choices from the drop-down lists (left). The drop-down list for the time period choices is displayed as an example. The wind speed value, covered by the drop-down list, is 10 kt. Two buttons in the initial input form (red rectangles) provide the user definitions of the flow regimes and speed categories. Clicking on each button will display the definitions in separate message forms (right).

5.3 Predictor Input Forms

If the user makes all the choices shown in the initial form in Figure 9, including “21-00” for the time period, and then clicks the “Continue…” button, the predictor input form for the 2100-0000 UTC equation in July at MCO is displayed (Figure 10). The list of predictors in each equation is at the top of the predictor input form. In this example, there are four predictors: FR3, PW, Pers, and DClimo. The FR3 and DClimo probabilities are retrieved from a table by the GUI using the information provided in the initial input form (Figure 9, left). The user inputs the PW from the 1000 UTC CCAFS sounding and chooses Yes or No for Pers indicating whether lightning occurred the previous day during the same time period. Clicking the “Display Probabilities” button will produce the output form (Section 5.4). As with Speed in the initial input form, the initial values for the sounding stability predictors are -999. If any of these values are not changed, an error message form for that predictor will be displayed (Figure 10, right). The “New Input” button closes the form and returns control to the initial input form.

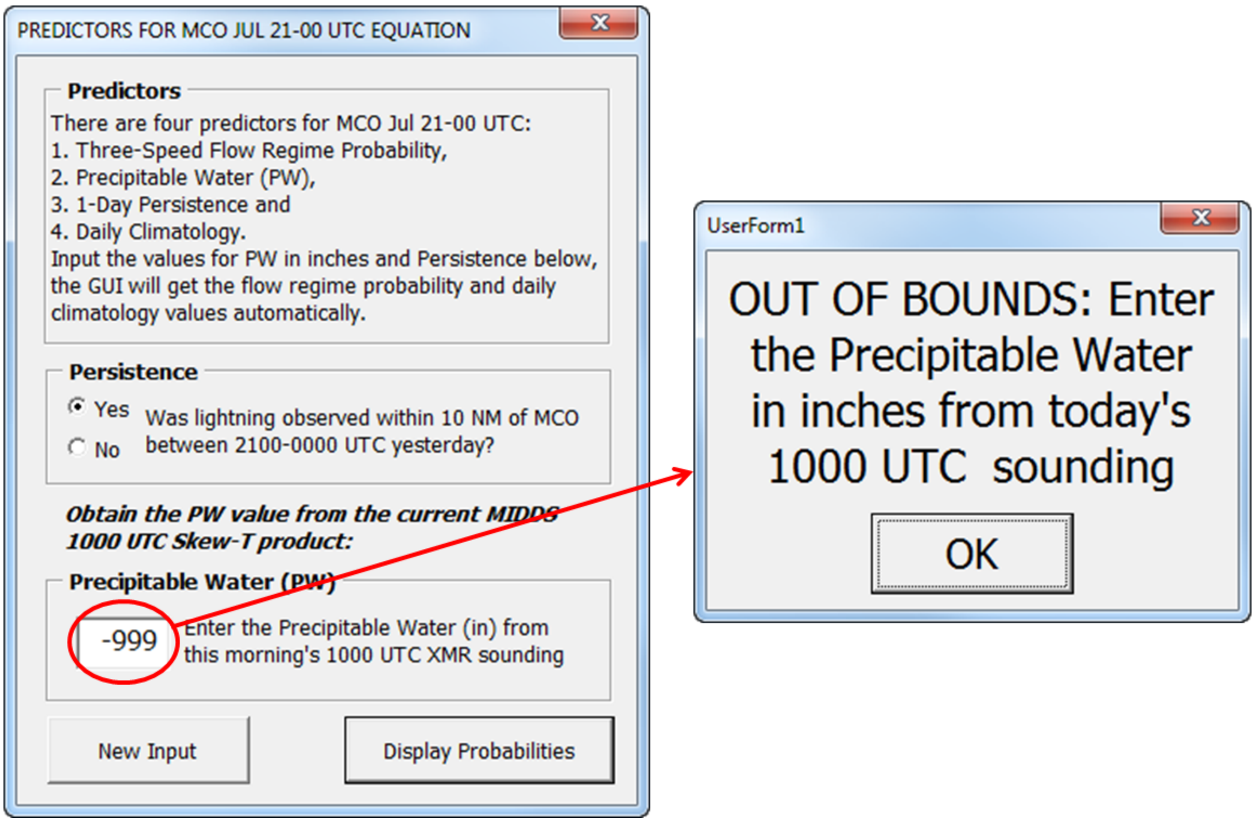


Figure 10. The predictor input form for the July MCO 2100-0000 UTC equation (left). The user inputs values for the predictors, in this case Yes or No for Pers and the PW value in inches, and then clicks the “Display Probabilities” button to get the output form. If predictor values are not input and the user clicks the “Display Probabilities” button, the form on the right is displayed reminding the user to input a predictor value. The “New Input” button closes the form and returns control to the initial input form.

Mr. Volkmer asked for an option to display the probabilities for all time periods in the same output form. The AMU added “ALL” as a choice in the time period drop-down menu on the initial input form and made it the default choice (Figure 8). The predictor input form associated with choosing all time periods for the MCO July example is shown in Figure 11. The user chooses the values for the predictors in all four equations at once. As with the initial input and single equation predictor input forms, the initial values of the sounding predictors are -999. An error message form, similar to that in Figure 10, will be displayed for each value not entered. Figure 11 shows values that were manually input. The user can use the up-down arrows or enter a value directly in the box. In order to not crowd the form, the specific predictors for each time period are not listed as in the individual time period equation input form (Figure 10). The specific predictors for each time period are

- 1500-1800 UTC: TI and FR2;
- 1800-2100 UTC: TI and Pers;
- 2100-0000 UTC: FR3, PW, Pers, and DClimo; and
- 0000-0003 UTC: FR (not speed related), DClimo, and LI.

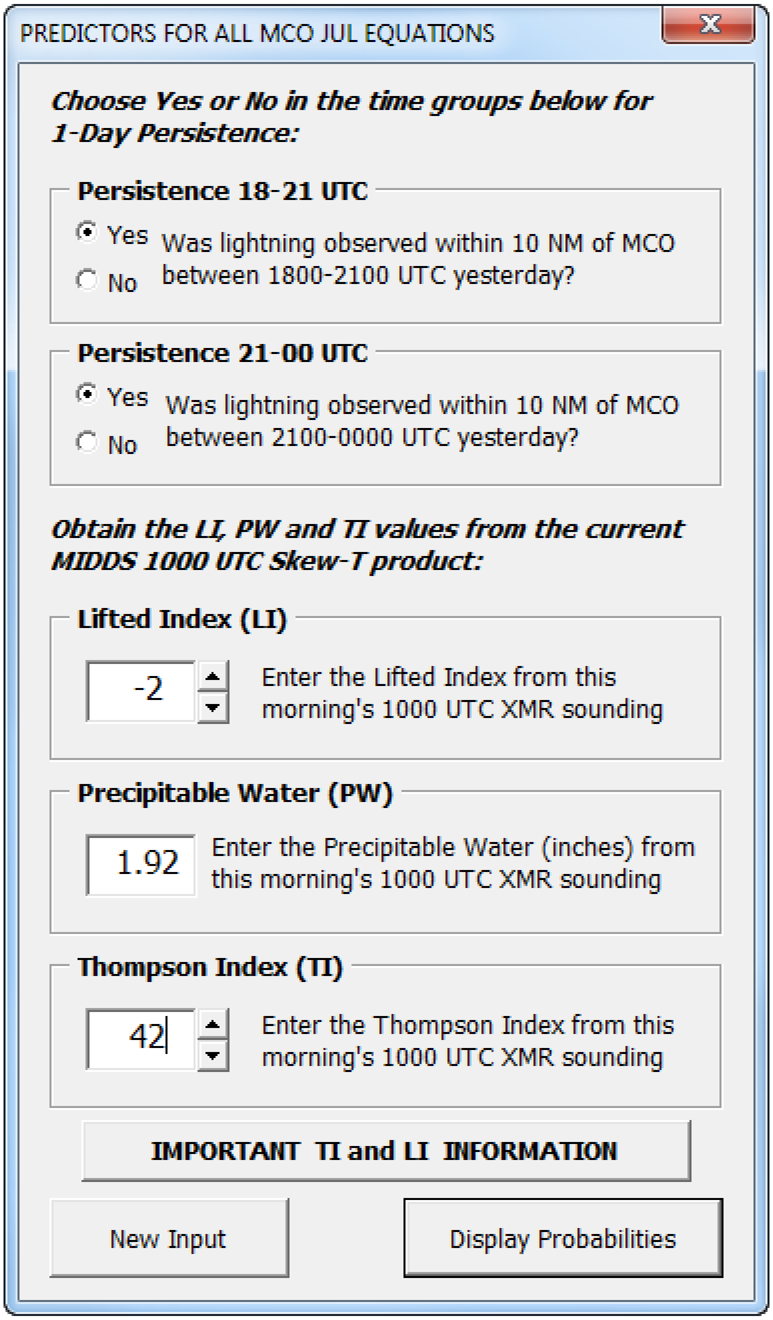


Figure 11. The GUI input form for the predictor values in the July MCO equations for all four time periods. The user makes choices in each category and clicks the "Display Probabilities" button to open the output form. The "New Input" button closes the form and returns control to the initial input form (Figure 8). The "IMPORTANT TI and LI INFORMATION" button provides information about LI calculation differences in AWIPS and MIDDS.

The “IMPORTANT TI and LI INFORMATION” button in Figure 11 outputs a message form (Figure 12) that provides information about differences in how LI is calculated between AWIPS used by NWS MLB and MIDDS used by the 45 WS. AWIPS uses the forecast maximum temperature for the surface parcel and MIDDS uses the 1000 UTC observed values. The equations were developed using the MIDDS value. The difference in the calculations would result in a lower, or more unstable, LI value in AWIPS. Use of the AWIPS LI would result in higher lightning probability values than would be calculated with the MIDDS value. Since TI is the difference between K-Index and LI, it would behave similarly. This is important information for forecasters when choosing LI and TI values. The form is closed by clicking the “OK” button and control is returned to the predictor input form.

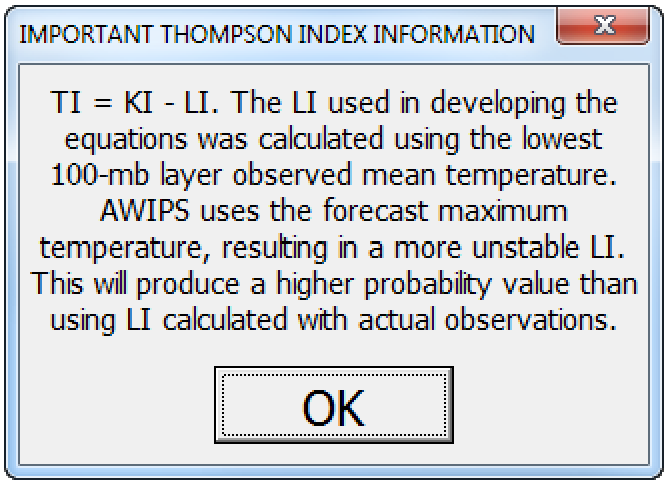


Figure 12. The message form displayed when the “IMPORTANT TI and LI INFORMATION” button is clicked. It is closed by clicking the “OK” button.

5.4 Output Forms

After making choices in either predictor input form, the user clicks the “Display Probabilities” button to display the probability output form. There are two output forms, one for a single time period choice and the other for the “ALL” time period choice.

The single time period output form for 1 July at MCO under 10 kt of SW-1 flow during 2100-0000 UTC is shown in Figure 13. The left panel repeats the information from the initial input form, and the DClimo, FR, FR2, FR3, and equation lightning occurrence probabilities are shown in the right panel. The first four values in the right panel are extracted from Excel PivotTables and are always the same for the same values chosen in the initial input form. DClimo will change with the month, day, station, and time period. The three flow regime probabilities will change with the month, station, flow regime, and speed for FR2 and FR3. The equation output will change as the sounding predictor values from the predictor input form are likely to be different on each individual day. The output form also contains the same two buttons found in the initial input form that display descriptions of the flow regimes, and the wind speed categories and their ranges (Figure 9).

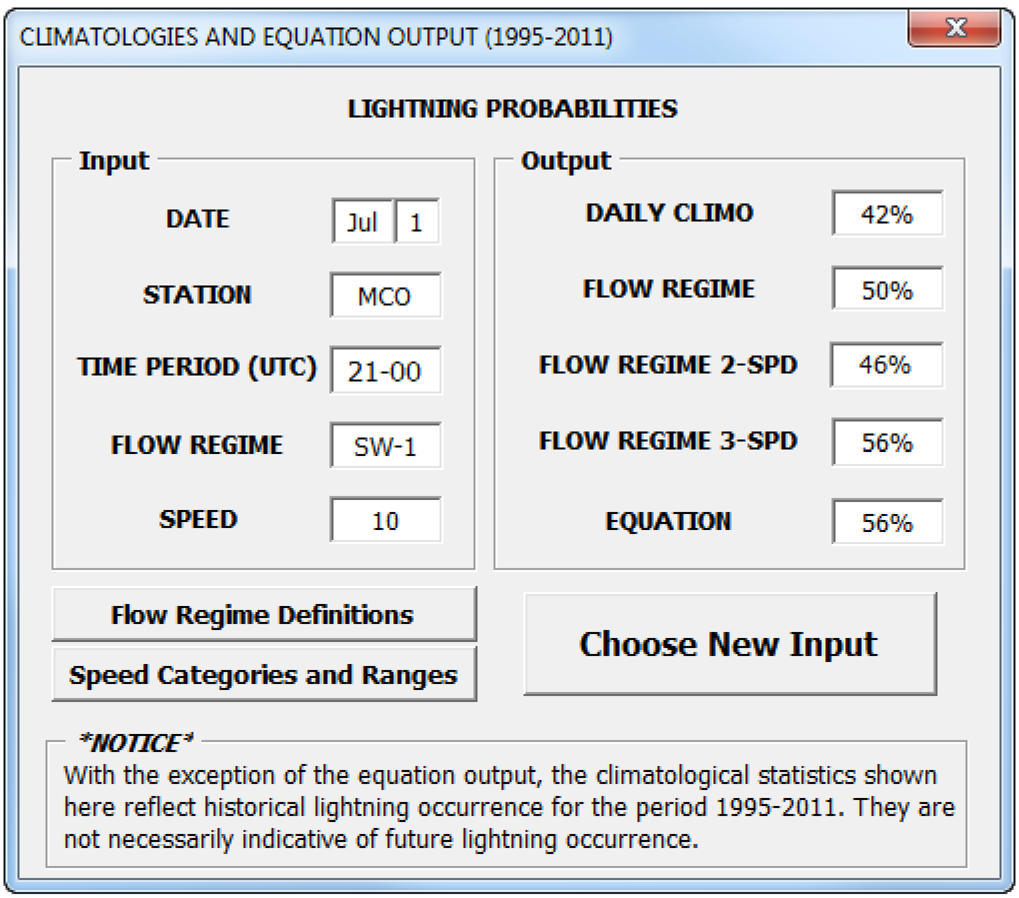
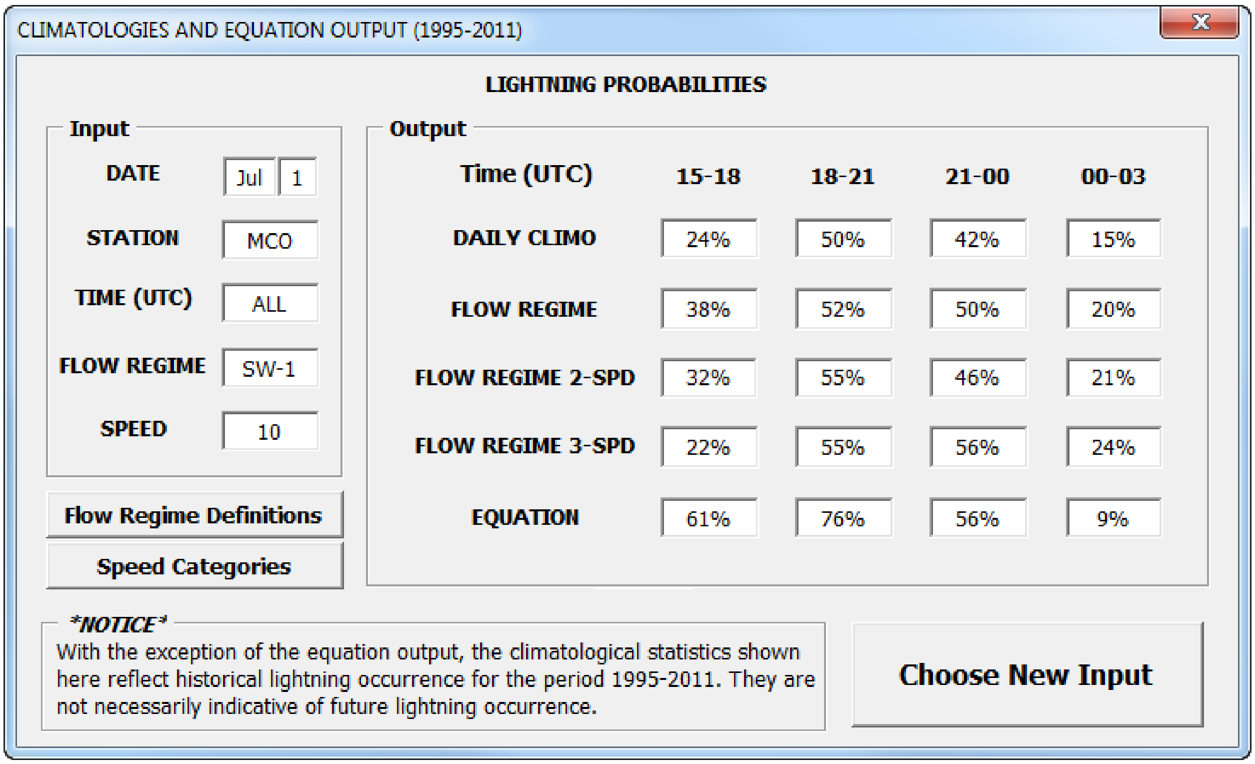


Figure 13. The single time period output form for July 2100-0000 UTC at MCO. The input choices from Figure 9 (left) are in the left panel and the climatological and computed probabilities are in the right panel. The equation probability was calculated from the values chosen in Figure 10 (left) with PW=1.92. The “Choose New Input” button closes the form and returns control to the predictor input form. The “Flow Regime Definitions” and “Speed Categories and Ranges” buttons open message forms that provide definitions of these parameters (see Figure 9).

Figure 14 shows the output form after choosing “ALL” for the time period in Figure 9 and the predictor values in Figure 11. It is similar to the single time period output form, the left panel repeats the information from the initial input form. The right panel, however, displays the probabilities for all four time periods. This allows the forecasters to see the progression of probabilities throughout the day. In this example, the probabilities increase from the first time period to the second (1100-1400 to 1400-1700 EDT), remain relatively steady or decrease a small amount from the second to the third time period (1400-1700 to 1700-2000 EDT), and drop dramatically into the last period, which occurs past sunset (2000-2300 EDT). Also, while the equation probability is similar to the climatological probabilities in the third time period, the equation probabilities for the other three time periods are significantly different than the climatological values. This could provide added value for the forecasters. The equation probabilities are calculated from current day values and would be more representative of the current environment than the climatological values.


CLIMATOLOGIES AND EQUATION OUTPUT (1995-2011)

LIGHTNING PROBABILITIES

Input

DATE	Jul 1
STATION	MCO
TIME (UTC)	ALL
FLOW REGIME	SW-1
SPEED	10

Flow Regime Definitions

Speed Categories

Output

Time (UTC)	15-18	18-21	21-00	00-03
DAILY CLIMO	24%	50%	42%	15%
FLOW REGIME	38%	52%	50%	20%
FLOW REGIME 2-SPD	32%	55%	46%	21%
FLOW REGIME 3-SPD	22%	55%	56%	24%
EQUATION	61%	76%	56%	9%

NOTICE
With the exception of the equation output, the climatological statistics shown here reflect historical lightning occurrence for the period 1995-2011. They are not necessarily indicative of future lightning occurrence.

Choose New Input

Figure 14. The output form as in Figure 13 but for all four time periods. The input choices from Figure 9 are in the left panel and the climatological and computed probabilities for all four time periods are in the right panel. The equation probabilities were calculated from the values chosen in Figure 11. The “Choose New Input” button closes the form and returns control to the predictor input form for the four time periods (Figure 11). The “Flow Regime Definitions” and “Speed Categories” buttons open message forms that contain definitions of these parameters as in Figure 9.

6 Summary and Conclusions

The AMU developed a tool to assist NWS MLB forecasters in making the lightning probability forecasts for airports in east-central Florida. The tool is a GUI developed using VBA in Excel that outputs climatological and forecast equation probabilities. Forecasters will input values for the predictors and the GUI outputs the probabilities of lightning occurrence for the airfield and time period of interest. The GUI also outputs the DClimo, FR, FR2, and FR3 climatological probabilities for comparison.

Work in this task combined the warm season equation development done for the 45 WS with the area around airfields and time period stratifications for NWS MLB. Specifically, the equations forecast the probability of lightning occurrence within 10 NM the airfield center and within four three-hour periods: 1500-1800, 1800-2100, 2100-0000, and 0000-0300 UTC. These are the times when lightning is most likely to occur in east-central Florida during the warm season. The airfields are MCO, MLB, and TIX. The 10 NM radius and three-hour time increments reflect the requirements at the airfields for which NWS MLB has forecast responsibility. An equation was developed for each warm season month, station, and three-hour time period. With five months, two stations (MCO and MLB), and four time periods, that resulted in 40 equations. Equations were not developed for TIX due to time constraints.

The performance of the equations compared to the forecast benchmarks of DClimo, FR, FR2, and FR3 was mixed with just a slight improvment. They did outperform Pers by 20% to over 50%. Even after several attempts to improve performance by modifying predictors and changing the POR, equation performance did not improve. Nonetheless, NWS MLB requested that equations be developed and their output displayed since the observed sounding parameters used in the equations may provide added value to the equation output. The equation probability value is displayed in the GUI along with the associated DClimo, FR, FR2, and FR3. The AMU delivered the GUI to NWS MLB and they are now using it in daily warm season operations.

6.1 XMR 1000 UTC Sounding

In late March 2013, the Air Force stopped releasing rawinsondes at 1000 UTC from CCAFS due to budget issues. This is effective until the end of the 2013 fiscal year on 30 September. The forecasters at NWS MLB understand that the tool may not perform optimally with other data, but they plan to use high temporal resolution model analysis soundings as close to 1000 UTC and CCAFS as possible, and will also use the Cape Canaveral Global Positioning System PW data available from the site http://www.suominet.ucar.edu/.

When the 1000 UTC sounding is reinstated, they will resume using that data. It is important to note the climatological values output with the equation probabilities are still valid as they do not depend on current sounding data.

6.2 Future Work

It is possible more data from a longer POR would improve equation performance. The POR for this task was 17 years, 1995-2011. The first tool the AMU created for the 45 WS was also developed using a warm season POR of 17 years, and was successful in significantly outperforming every forecast benchmark. The main difference between the two tasks was the time periods. The spatial area used for the 45 WS tool was larger, but not significantly. In the 45 WS tool, the forecast was valid anytime between 0700 to midnight local time, a 17-hour period. In this task, the data was stratified into four three-hour time periods in a 12-hour period. The lightning occurrences had to be separated into the four time periods, reducing the number with which to develop relationships between the predictors and lightning occurrence within each time period. The AMU also found this result in their task to relate GPS PW to lightning occurrence (Huddleston 2012) within shorter time periods. The data were stratified to the point of reducing the number of observations in each stratification such that robust statistical relationships could not be realized. A longer POR would provide more lightning occurrences for each stratification, and it may also help to increase the time periods.

Due to the amount of time the AMU spent trying methods to improve equation performance, including changing the POR that resulted in the need to recalculate some of the predictors, there was not enough time to create equations for TIX or other airfields. The AMU also did not have time to conduct other equation performance metrics such as reliability and the ability of the final equations to distinguish between lightning and non-lightning days. The AMU did create charts to determine the ability of the original MCO equations to distinguish between lightning and non-lightning days. In order to have enough samples, the AMU combined data from all months to evaluate each time period and created the charts in Figure 15. For good performance, the blue curves should have a maximum occurrence in the lower probability values decreasing to a minimum at higher probability values, and the red curves should have a minimum in the lower probability values increasing to a maximum at the higher values. The equations were able to distinguish non-lightning days in all periods, but performance was best in the 1500-1800 and 0000-0300 UTC periods representing the beginning and end of the day, respectively. They were able to distinguish lightning days better in the afternoon time periods of 1800-2100 and 2100-0000 UTC than during 1500-1800 and 0000-0300 UTC. Still, the curves were spread over most of the range of forecast probabilities indicating ambiguous performance.

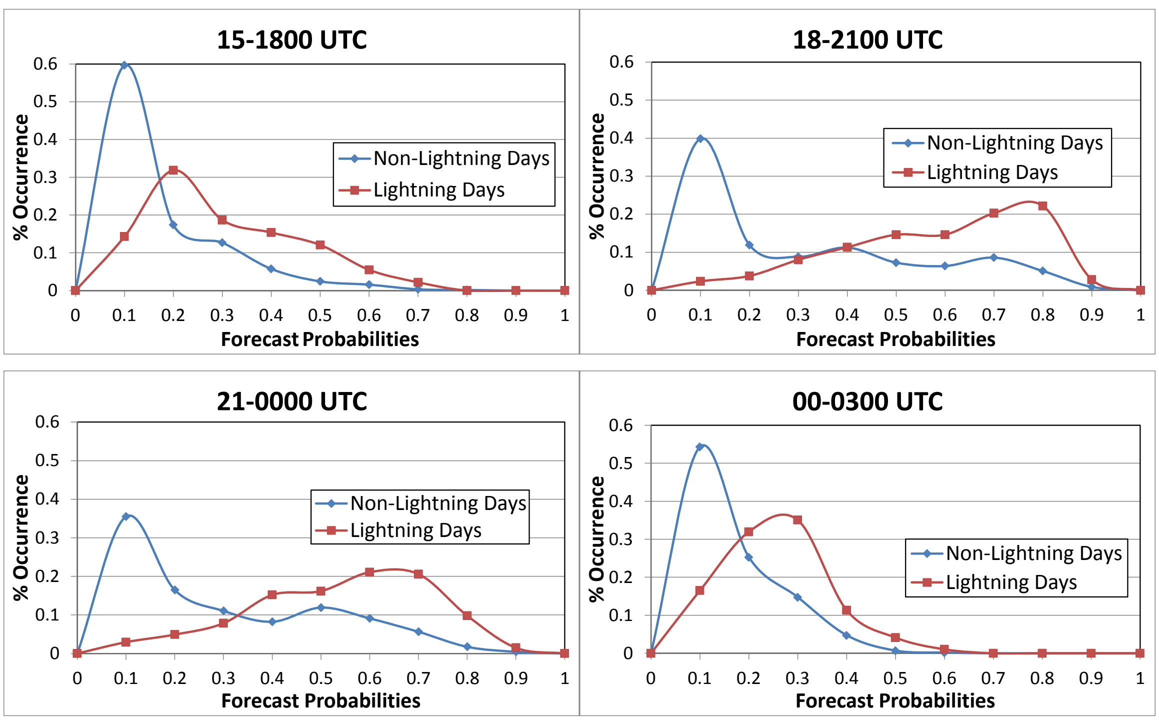


Figure 15. Forecast probability distributions for lightning (red) and non-lightning (blue) days in the verification data in the four time periods for MCO, all months, and for the original POR 1989-1995. The y-axis values are the frequency of occurrence of each probability value, and the x-axis values are the forecast probability values output by the equations.

Based on these results, future work should involve collecting more data over a longer POR, developing equations for more airfields, and conducting other verification tests of the equations to determine all aspects of their performance. This will create equations that provide a good estimate of lightning occurrence at the airfields and in the time periods of interest and give the forecasters a good first guess from which to build a forecast from other data sources.

References

Barrett, J., 2008: Situational Lightning Climatologies for Central Florida: Phase III. NASA Contractor Report CR-2008-214742, Kennedy Space Center, FL, 31 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL 32931 and http://science.ksc.nasa.gov/amu/final-reports/lightning-climo-phase3.pdf.]

Bauman, W., 2007: Situational Lightning Climatologies for Central Florida: Phase II, Part 3. NASA Contractor Report CR-2007-214736, Kennedy Space Center, FL, 25 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL 32931 and http://science.ksc.nasa.gov/amu/final-reports/lightning-climo-phase2.pdf.]

Bauman, W., 2009: Situational Lightning Climatologies for Central Florida: Phase IV. NASA Contractor Report CR-2009-214763, Kennedy Space Center, FL, 39 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL 32931 and http://science.ksc.nasa.gov/amu/final-reports/lightning-climo-phase4.pdf.]

Bauman, W., 2011: Situational Lightning Climatologies for Central Florida: Phase V. NASA Contractor Report CR-2011-216303, Kennedy Space Center, FL, 29 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 830, Cocoa Beach, FL 32931 and http://science.ksc.nasa.gov/amu/final-reports/lightning-climo-phase5.pdf.]

Bauman, W., and W. Crawford, 2012: Objective Lightning Probability Forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase IV. NASA Contractor Report CR-2012-216312, Kennedy Space Center, FL, 29 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-phase4.pdf.]

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.]

Cummins, K., and M. Murphy, 2009: An overview of lightning locating systems: History, techniques, and data uses, with an in-depth look at the U.S. NLDN, *IEEE Trans. Electromagn. Compat.*, **51**, 499-518, doi:10.1109/TEMC.2009.2023450.

Hodanish, S., and P. Wolyn, 2012: Lightning climatology for the State of Colorado. Preprints, *4th International Lightning Meteorology Conference*, Broomfield, CO, Vaisala, 12 pp. [Available online at http://www.vaisala.com/en/events/ildcilmc/Documents/Climatology/Lightning%20Climatology%20for%20the%20State%20of%20Colorado.pdf.]

Huddleston, L., 2012: Using Global Positioning System Integrated Precipitable Water Vapor to Forecast Lightning on KSC/CCAFS. NASA Technical Memorandum TM-2012-216314, Kennedy Space Center, FL, 51 pp. [Available online at http://science.ksc.nasa.gov/amu/final-reports/gps-ipw-ltg.pdf.]

Lambert, W., 2007: Objective Lightning Probability Forecasting for Kennedy Space Center and Cape Canaveral Air Force Station, Phase II. NASA Contractor Report CR-2005-214732, Kennedy Space Center, FL, 57 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-phase2.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-reports/objective-ltg-fcst-phase1.pdf.]

Lambert, W., D. Sharp, S. Spratt, and M. Volkmer, 2006: Using cloud-to-ground lightning climatologies to initialize gridded lightning threat forecasts for east central Florida. Preprints, *Second Conf. on Meteorological Applications of Lightning Data*. Atlanta, GA, Amer. Meteor. Soc., 4 pp. [Available online at https://ams.confex.com/ams/pdfpapers/104395.pdf.]

Lericos, T., H. Fuelberg, A. Watson and R. Holle, 2002: Warm season lightning distributions over the Florida Peninsula as related to synoptic patterns. *Wea. Forecasting*, **17**, 83–98.

Peppler, R. A. and P. J. Lamb, 1989: Tropospheric static stability and Central North American growing season rainfall. *Mon. Wea. Rev.*, **117**, 1156 – 1180.

Short, D., 2006: Situational Lightning Climatologies for Central Florida, Phase II. Applied Meteorology Unit Memorandum, 8 pp. [Available from ENSCO, Inc., 1980 N. Atlantic Ave., Suite 230, Cocoa Beach, FL 32931.]

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

World Meteorological Association (WMO), 1992: *Methods of Interpreting Numerical Weather Prediction Output for Aeronautical Meteorology.* Technical Note No. 195, ISBN 92-63-10770-X, 89 pp.

Acronyms

45 WS	45th Weather Squadron
AMU	Applied Meteorology Unit
AWIPS	Advanced Weather Interactive Processing System
CCAFS	Cape Canaveral Air Force Station
CT	Cross Totals
CWA	County Warning Area
DClimo	Daily Climatology
DE	Detection Efficiency
FR	Flow Regime
FR2	2-speed FR
FR3	3-speed FR
GUI	Graphical User Interface
JAX	Jacksonville, FL 3-letter identifier
KI	K-Index
KSC	Kennedy Space Center
LI	Lifted Index
MCO	Orlando International Airport
MFL	Miami, FL 3-letter identifier
MIDDS	Meteorological Interactive Data Display System
MLB	Melbourne International Airport
NE	Northeast FR
NLDN	National Lightning Detection Network
NW	Northwest FR
NWS MLB	National Weather Service in Melbourne, FL
Pers	One-Day Persistence
POR	Period of Record
PW	Precipitable Water
RH	Relative Humidity
RH_{85}	Mean RH in the 825-525 mb layer
RH_{86}	Mean RH in the 800-600 mb layer
SE	Southeast FR
SE-1	Southeast 1 FR
SE-2	Southeast 2 FR
SI	Showalter Index
SS	Skill Score
SW	Southwest FR
SW-1	Southwest 1 FR
SW-2	Southwest 2 FR
SWEAT	Severe Weather ThrEAT Index
T_{500}	Temperature at 500 mb
TBW	Tampa, FL 3-letter identifier
TI	Thompson Index
TIX	Space Coast Regional Airport
TT	Total Totals
VT	Vertical Totals
WMO	World Meteorological Organization
XMR	CCAFS rawinsonde 3-letter identifier

NOTICE