Global Positioning System (GPS) Precipitable Water in Forecasting Lightning at Spaceport Canaveral

Kristen Kehrer
Mail Stop EA-C,
National Aeronautics and Space Administration, Kennedy Space Center, Florida
Kristen.C.Kehrer@nasa.gov

Brian Graf
National Aeronautics and Space Administration, Kennedy Space Center, Florida

William P. Roeder
45th Weather Squadron, Patrick Air Force Base, Florida

ABSTRACT

This paper evaluates the use of precipitable water (PW) from Global Positioning System (GPS) in lightning prediction. Additional independent verification of an earlier model is performed. This earlier model used binary logistic regression with the following four predictor variables optimally selected from a candidate list of 23 candidate predictors: the current precipitable water value for a given time of the day, the change in GPS-PW over the past 9 hours, the K-Index, and the electric field mill value. This earlier model was not optimized for any specific forecast interval, but showed promise for 6 hour and 1.5 hour forecasts. Two new models were developed and verified. These new models were optimized for two operationally significant forecast intervals. The first model was optimized for the 0.5 hour lightning advisories issued by the 45th Weather Squadron. An additional 1.5 hours was allowed for sensor dwell, communication, calculation, analysis, and advisory decision by the forecaster. Therefore the 0.5 hour advisory model became a 2 hour forecast model for lightning within the 45th Weather Squadron advisory areas. The second model was optimized for major ground processing operations supported by the 45th Weather Squadron, which can require lightning forecasts with a lead-time of up to 7.5 hours. Using the same 1.5 lag as in the other new model, this became a 9 hour forecast model for lightning within 37 km (20 NM)) of the 45th Weather Squadron advisory areas. The two new models were built using binary logistic regression from a list of 26 candidate predictor variables: the current GPS-PW value, the change of GPS-PW over 0.5 hour increments from 0.5 to 12 hours, and the K-index. The new 2 hour model found the following for predictors to be statistically significant, listed in decreasing order of contribution to the forecast: the 0.5 hour change in GPS-PW, the 7.5 hour change in GPS-PW, the current GPS-PW value, and the K-Index. The new 9 hour forecast model found the following five independent variables to be statistically significant, listed in decreasing order of contribution to the forecast: the current GPS-PW value, the 8.5 hour change in GPS-PW, the 3.5 hour change in GPS-PW, the 12 hour change in GPS-PW, and the K-Index. In both models, the GPS-PW parameters had better correlation to the lightning forecast than the K-Index, a widely used thunderstorm index. Possible future improvements to this study are discussed.

1. Background

The 45th Weather Squadron (45 WS) provides comprehensive weather services to America's space program at Cape Canaveral Air Force Station (CCAFS) and Kennedy Space Center (KSC). These facilities are located in east-central Florida near the highest lightning flash densities in North America (Figure 1). The most frequent products of the 45 WS are lightning advisories for 13 different points (Figure 2). These lighting advisories are issued for personnel safety of over 25,000 people and resource protection of facilities worth over 17 billion dollars. A two-tiered advisory process is used. A Phase-1 lightning advisory is issued for a point if lightning of any type is expected within five nautical miles of any of the points with a desired lead-time of 30 minutes. A Phase-2 lightning advisory is issued for a point if lightning of any type is imminent or occurring within 9.3 kilometers (5 nautical miles) of the point. Lightning forecasting is important to other operations supported by the 45 WS, especially major ground processing. For example, transporting the Space Shuttle from the Vehicle Assembly Building to the launch pad requires a less than 10% probability of lightning within 37.0 kilometers (20 nautical miles) during the approximate six hours the Shuttle is being moved with the briefing for the final decision occurring about two hours before transport begins. Lightning is also vitally important to space launches, but a special set of launch commit criteria are used for space launches and lightning support for these operations are not covered in this paper (Roeder and McNamara, 2006). The 45th WS has several techniques for forecasting lightning to support their lightning advisory and ground processing requirements. But the 45th WS is always trying to refine their current techniques and develop new methods to improve their lightning forecasts.

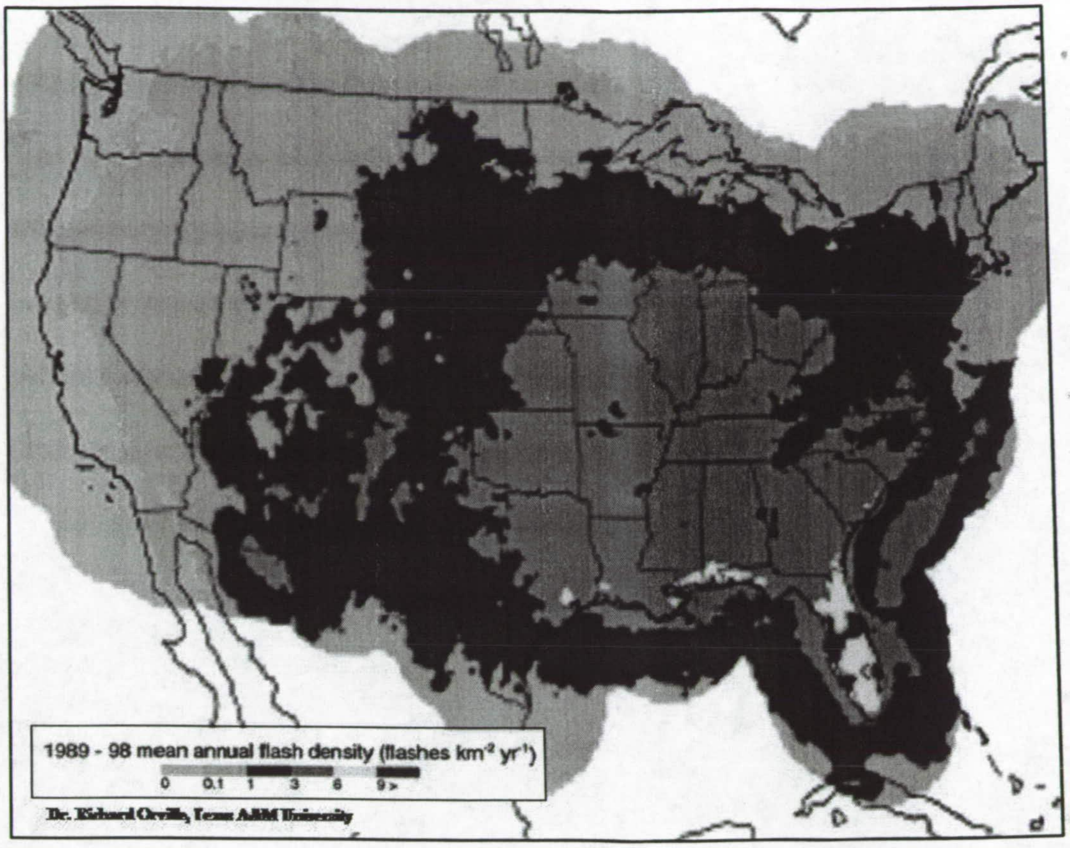

Figure 1: Average cloud-ground lightning flash density for the contiguous U.S. (1989-1998). The largest flash densities occur in central Florida. (Courtesy of Dr. Richard Orville, Texas A&M University)

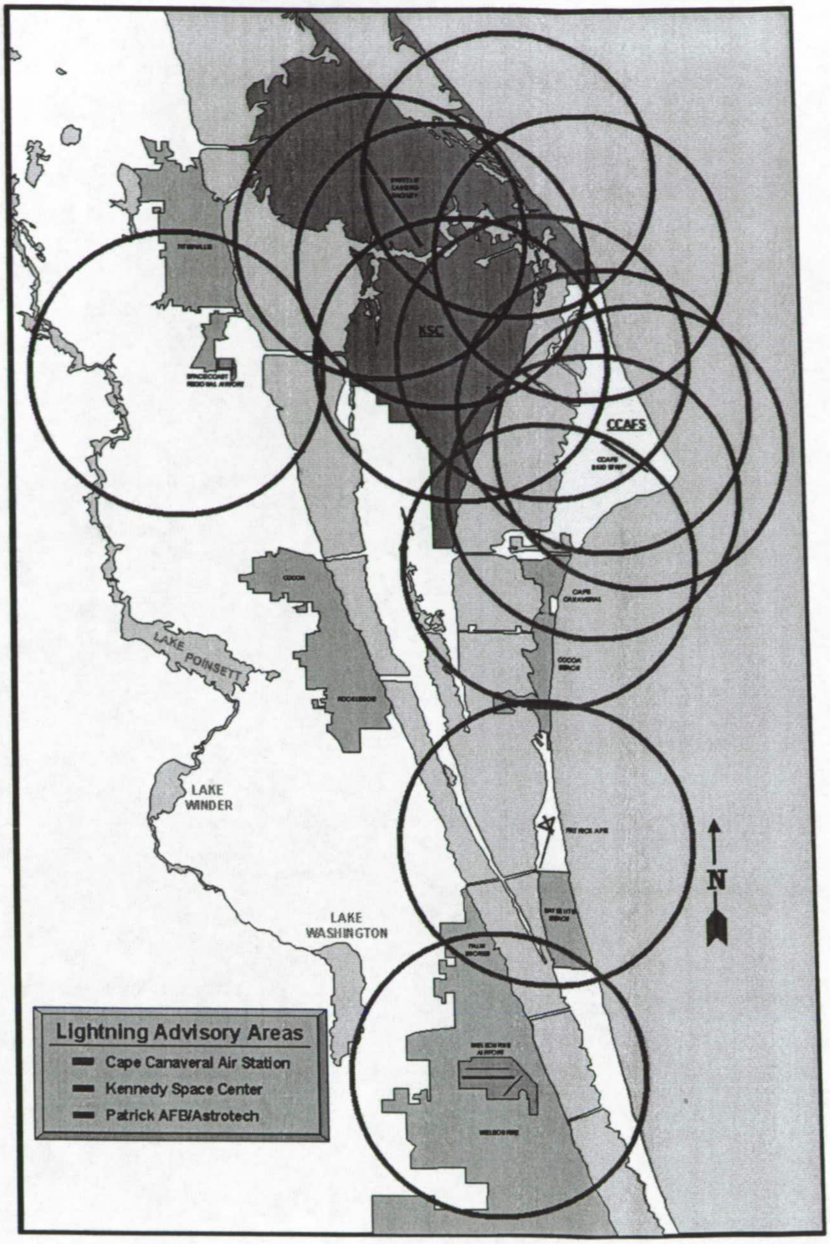

Figure 2. The 45 WS's lightning advisory areas.

This paper explores the use of Global Positioning Satellite (GPS)-based precipitable water (PW) in lightning prediction at CCAFS/KSC. A previous GPS-PW model for lightning prediction at CCAFS/KSC (Mazany et al., 2000) received performance verification on a larger set of independent data. Two new models were developed. The first of the new models was optimized for the relatively short timelines of the 45 WS lightning advisories. The second new model was optimized for the longer timelines of 45 WS major ground processing operations. The performance of the previous GPS-PW lightning prediction model was not duplicated. However, the two new GPS-PW models did show promise.

2. Introduction

a. Description of Global Positioning System Precipitable Water

Precipitable water is traditionally calculated from data obtained by weather balloons. However, it was discovered over a decade ago that PW can be calculated from GPS satellites (Beavis et al., 1992) (Beavis et al., 1994). Previous researchers have called this GPS Integrated Water Vapor (GPS-IWV), but the authors use the term precipitable water rather than IWV since the two terms are equivalent, with precipitable water being the older and better established term. Applications of GPS-PW have been explored by Businger et al. (1996), Bauman et al. (1997), and Wolfe and Gutman (2000). The phase delay of GPS signals passing through the atmosphere depends on the electron density in the ionosphere, the mass of the atmosphere, the amount of hydrometeors in the atmosphere, and the total amount of water vapor in the atmosphere. The delay due to the ionospheric electron density along each GPS line of sight can be calculated from the total electron count, which can be calculated by comparing the L1 and L2 GPS signals. The mass of the atmosphere can be calculated from the surface pressure measured by a barometer at the surface. The GPS phase delay due to hydrometeors is usually insignificant and not considered. Therefore, any GPS propagation delay remaining after accounting for these three delays is attributed to water vapor. GPS-PW is normally measured by averaging the GPS propagation delays over all the GPS satellites more than 15 degrees above the horizon over a 30 minute period. GPS-PW has several important advantages over weather balloons. GPS-PW is as accurate, if not more so, than weather balloons, is available every 30 minutes as compared to twice a day typical of weather balloons, provides a remote all weather capability, and can be automated thereby avoiding the costs of human operated weather balloons.

b. Mazany Model

A model to forecast lightning from GPS-PW was first developed in 2000 (Mazany et al., 2002). This model used binary logistic regression to predict the probability of lightning at CCAFS/KSC using GPS-PW, the K-Index to incorporate atmospheric instability, and the largest value from the network of 31 surface electric field mills at CCAFS/KSC to include the electric signal from developing thunderstorms.

The output of the model is a Lightning Index between 0 (lightning) and 1 (no lightning) that indicates conditions for lightning. The Lightning Index is compared to thresholds to determine if and when lightning will occur. The model is depicted in (1):

$$\hat{y} = \frac{1}{1+e^{(-6.7866+0.001359x_1+0.06063x_2+0.32341x_3+0.06728x_4)}} \tag{1}$$

where

x_1 = electric field mill reading (V m^{-1})

x_2 = PW (mm)

x_3 = 9-h Δ PW (mm)

x_4 = K-Index

$\hat{y}$ = Lightning Index

The Lightning Index was then compared with the onset of lightning and the following thresholds were determined:

0.7–1.0: No lightning in the next ~6 hr;

0.6–0.7: Lightning expected in the next ~6 hr;

0.0–0.6: Lightning expected in the next ~1.5 hr.

The accuracy measures applied to the initial test results of the Mazany Model forecasts were: False Alarm Rate (FAR), Probability of Detection (POD), and Hit Rate (HR). The follow-on independent verification in this new study also included the Kuipers Skill Score (KSS) (Wilks, 1995) and the Operational Utility Index (OUI). The OUI is a non-standard metric developed by 45 WS for comparing lightning forecast tools that gives POD a weight of three, KSS a weight of two, and FAR a weight of negative one, and then normalizes the sum of the weighted metrics by the absolute value of the sum of the weights for easier interpretation; an OUI of one is perfect forecasting, and an OUI of zero is worthless forecasting. The weights were set by the operational importance of the metrics to lightning forecasting by the 45 WS. Since personnel safety is involved, POD is most important. A good level of skill is desired, to provide good service to the customers, but it is less important than POD. A low FAR is also desired, to also provide good service to the customers, but is least important of the three metrics.

$$OUI = \frac{3(POD) + 2(KSS) - 1(FAR)}{6}$$

c. Data sources and validation

Data from four thunderstorm seasons, 1 May to 30 September, 2000 – 2003, were used in the re-verification of the Mazany model. Quality control was performed on the PW and K-Index data. Scatter plots were used to visually identify potential outliers. Statistical Process Control (SPC) charts were also used to more objectively identify potential outliers. The SPC charts identified points outside three standard deviations of the mean of the data set. Specifically, $\bar{x}$ charts were used to identify $\bar{x}$, Upper Control Limits (UCL), and Lower Control Limits (LCL) for each year. SPC $\bar{x}$ charts were created for both PW and K-Index data sets for each year. Points that fell below the LCL and above the UCL were flagged as potential outliers and further examined against meteorological conditions surrounding the data points. One of the authors (Roeder), the meteorologist on the team, reviewed the data from the meteorological perspective and determined that the potential outliers were actually valid data points.

The K-Index was calculated from the CCAFS weather balloons which are usually available at 1000 GMT, 1500 GMT, and 2300 GMT during the summer thunderstorm season (May-September). The time series of K-Index could be used in two different methods. To account for changes in K-Index between weather balloon observations, the K-Index could be interpolated linearly between existing observations. This linear interpolation of the K-Index was used in the re-validation of the Mazany model. However, to mimic how the information would be used operationally, the future value would not be known at forecast time therefore the last K-Index was used unchanged until the next weather balloon observation in developing the two new forecasting tools.

As discussed above, the electric field mill data were not important to the relatively large forecast intervals being verified, when the electric fields could be fair field values. At these fair field values, scale analysis shows that the Mazany model is insensitive to typical variations in electric fields. Therefore, since it would not be important to the forecast intervals being verified, a constant typical fair field value of 300 V/m was used to validate the Mazany Model (Marshall et al., 1999).

3. Categorical verification of Mazany Model

Actual observations of the independent and dependent variables for the four seasons (2000-2003) were used to validate the Mazany Model. Categorical forecasts were created from the continuous predictand of the Mazany model. If the model predictand was 0.7 or less, it was considered a 'yes' forecast for the 6 hour forecast interval. If the model predictand was greater than 0.7, it was considered a 'no' forecast for the 6 hour interval. If the model predictand was 0.6 or less, it was considered a 'yes' forecast for the 1.5 hour forecast interval. If the model predictand was greater than 0.6, it as considered a 'no' forecast for the 1.5 hour interval. Two forecast intervals were verified. A six hour forecast was used to match the verification in the original Mazany study (Mazany et al., 2002). A 1.5 hour forecast was also verified to match part of the operationally focused verification of the new models. Standard 2×2 contingency tables and the metrics discussed above were used to analyze performance. Table 1 shows the results of the predictive capability of the current Mazany Model for each year and a combined performance for years 2000 through 2003.

Table 1: Accuracy Measure and Skill Scores of Forecasts of the original Mazany Model. The 2×2 contingencies tables for the 1.5-hour and 6-hour forecast intervals for each year (2000-2003) and all years are presented, and then a summary table is presented.

Lightning Forecasts in next 6 and 1.5 Hours

Year: 2000

		Observed Yes	Observed No	HR	POD	FAR	KSS	OUI
6 hr Forecast	Yes	186	232	68%	29%	56%	14%	10%
	No	447	1281					
1.5 hr Forecast	Yes	370	946	63%	51%	72%	17%	19%
	No	353	1796					

Lightning Forecasts in next 6 and 1.5 Hours

Year: 2001

6 hr Forecast

Forecast	Observed Yes	Observed No
Yes	344	559
No	646	2291

HR	POD	FAR	KSS	OUI
69%	35%	62%	15%	12%

1.5 hr Forecast

Forecast	Observed Yes	Observed No
Yes	682	2004
No	508	3332

HR	POD	FAR	KSS	OUI
62%	57%	75%	20%	23%

Lightning Forecasts in next 6 and 1.5 Hours

Year: 2002

6 hr Forecast

Forecast	Observed Yes	Observed No
Yes	241	446
No	178	1584

HR	POD	FAR	KSS	OUI
75%	58%	65%	36%	30%

1.5 hr Forecast

Forecast	Observed Yes	Observed No
Yes	345	1240
No	202	2248

HR	POD	FAR	KSS	OUI
64%	63%	78%	28%	28%

Lightning Forecasts in next 6 and 1.5 Hours

Year: 2003

6 hr Forecast

Forecast	Observed Yes	Observed No
Yes	381	582
No	406	1773

HR	POD	FAR	KSS	OUI
69%	48%	60%	24%	22%

1.5 hr Forecast

Forecast	Observed Yes	Observed No
Yes	587	2026
No	354	2788

HR	POD	FAR	KSS	OUI
59%	62%	78%	20%	25%

Lightning Forecasts in next 6 and 1.5 Hours

Combined Performance (Year 2000 – 2003)

6 hr		Observed Yes	Observed No
Forecast	Yes	1152	1819
	No	1677	6929

HR	POD	FAR	KSS	OUI
70%	41%	61%	20%	17%

1.5 hr		Observed Yes	Observed No
Forecast	Yes	1984	6216
	No	1417	10164

HR	POD	FAR	KSS	OUI
61%	58%	76%	20%	23%

Summary Table

	2000		2001		2002		2003		2000-2003	
	6 Hr	1.5 Hr	6 Hr	1.5 Hr	6 Hr	1.5 Hr	6 Hr	1.5 Hr	6 Hr	1.5 Hr
HR	68%	63%	69%	62%	75%	64%	69%	59%	70%	61%
POD	29%	51%	35%	57%	58%	63%	48%	62%	41%	58%
FAR	56%	72%	62%	75%	65%	78%	60%	78%	61%	76%
KSS	14%	17%	15%	20%	36%	28%	24%	20%	20%	20%
OUI	10%	19%	12%	23%	30%	28%	22%	25%	17%	23%

The results of the accuracy measures for the 1.5-hour forecast period were compared to the accuracy measures of the test results of the Mazany Model forecasts performed for the time identified as Period B from the original study (10 June 1999 to 26 September 1999). The Period B forecasts were based upon the index value falling below 0.7 and 1.5 hr prior to the first strike. Table 2 compares 1999 Period B with the four thunderstorm seasons from 1 May to 30 September, 2000–2003 and the combined performance.

Table 2: Comparison of Accuracy Measurements of GPS Lighting Index

1.5-hr Forecast Period	1999 Period B	2000	2001	2002	2003	2000-2003 Combined
FAR	16.7%	71.9%	74.6%	78.2%	77.5%	75.8%
POD	89.3%	51.2%	57.3%	63.1%	62.4%	58.3%
HR	82.6%	62.5%	61.5%	64.3%	58.6%	61.4%

The GPS Lightning Index accuracy measure results for all four thunderstorm seasons were below expectations when compared to the results of the 1999 Period B. On average, the FAR increased by 58%, the POD decreased by 31%, and the HR decreased by 20%. Even though the accuracy measures indicate that the Mazany Model is not reliable for forecasting lightning events 1.5 hr prior to the first strike, the encouraging aspect of the model results was the consistency of the measures. This suggests there is some useful signal in the GPS-PW timelines for forecasting lightning in the study area and that perhaps better performance could be obtained with alternate regression models with other predictors and other lead times.

As was determined in the Mazany paper (2002), the performance of the GPS Lightning Index lead time with regard to the timing of the first strike follows approximately a normal distribution. With the range of lead time before the first strike varying between 0 and 12 hr, the typical lead time is approximately 6 hr. Based upon this finding, the GPS Lighting Index was also applied to the 6-hr forecast period. When compared to the 1.5-hr forecast period, the Mazany Model performed better overall. Table 3 compares the average 1.5-hr forecast period accuracy measures with the 6-hr forecast period accuracy measures the four thunderstorm seasons from 1 May to 30 September, 2000–2003 and the combined performance.

Table 3: Comparison of Accuracy Measurements for Lead Times of 1.5 hr and 6 hr for the Mazany Model

6-hr Forecast Period	Average 1.5-hr Forecast for 2000–2003	2000	2001	2002	2003	2000-2003 Combined
FAR	75.6%	55.5%	61.9%	64.9%	60.4%	61.2%
POD	58.5%	29.4%	34.7%	57.5%	48.4%	40.7%
HR	61.7%	68.4%	68.6%	74.5%	68.6%	69.8%

On average, the Mazany model performed better in the longer 6 hour time period in both FAR and HR. On average, the FAR for the 6 hour forecast period was lower by 15%, and the HR was higher by 7%. However, the Mazany model performed better in the shorter 1.5 hour time period in POD, which was on average higher by 16% over the 6 hour time period.

a. Timeline

The project sought to develop two new forecasting tools for the Spaceport Canaveral. The objective of the first tool was to provide a desired 0.5 hour lead time prior to a lightning event; the second sought a desired 7.5 hour lead time. These lead times were chosen to meet operational requirements. The 0.5 hour lead time is for lightning advisories. The 7.5 hour lead time is for major ground processing operations, such as roll out of the Space Shuttle to the launch pad, transport of major components, and others.

To achieve the desired lead times for both tools, consideration was paid to the process that would be used in the operational implementation of the tools. A 1.5 hour delay or lag time had to be built into the model to account for the process. The initial lag is caused by the PW 30-min dwell time. When the operator receives the PW reading, it is already 30 min behind the center time stamp and 15 min behind the end of the PW dwell time. Another 15-min lag time was added to account for communication of the most recent PW data to the operator. A 45-min delay was added for processing of the model, communication of the results to the 45th Weather Squadron, and comparison of the model output with other weather data to make the forecast decision. This process is depicted in Figure 1.

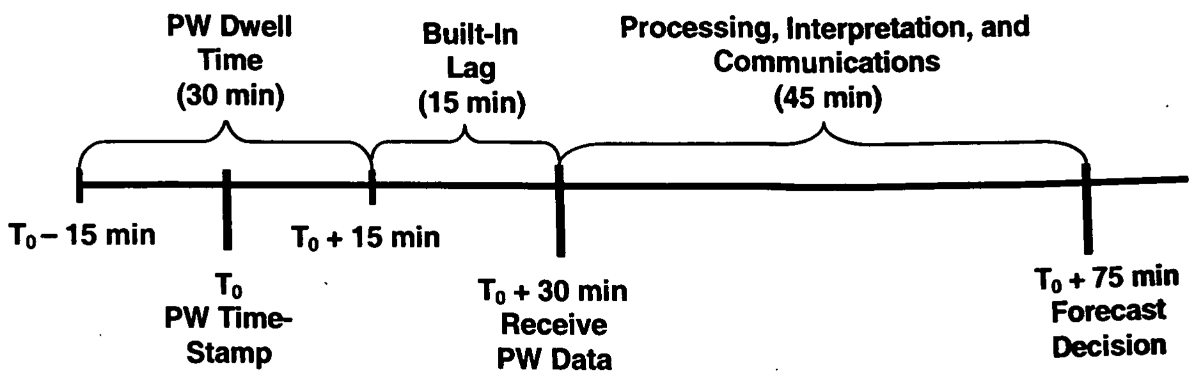

Figure 1: 1.5-hr Lag Time

This turns the desired 0.5 hour lead time into a 2-Hr Forecast Tool and the desired 7.5 hour lead time into a 9-Hr Forecast Tool.

b. Tool development

1) LOGISTIC REGRESSION

Logistic regression was chosen as the tool to develop both models for several reasons. Previous studies conducted for the 45th Weather Squadron indicate the applicability of logistic regression in modeling meteorological data. Logistic regression is constrained to be between zero and one as are probabilities. Linear regression for probability forecasting can predict probabilities greater than one and less than zero. Logistic regression can also model rapid changes in probability as thresholds of predictors are exceeded as often happens in meteorology. This is depicted in Figure 2.

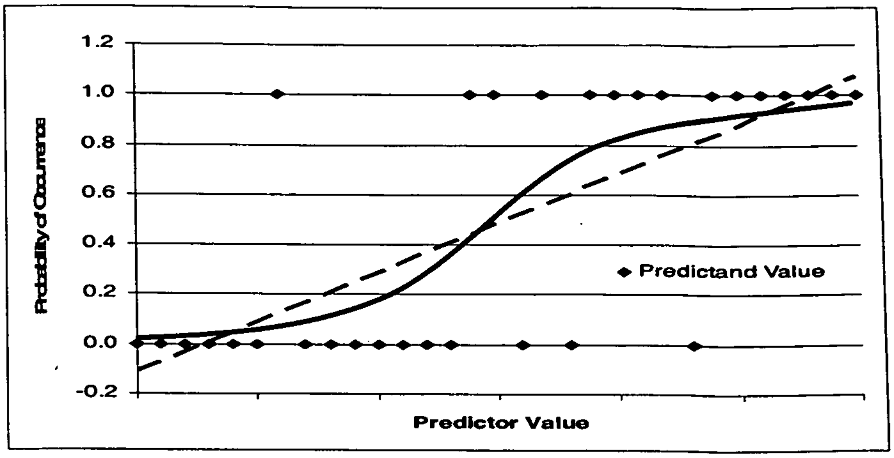

Figure 2: Logistical Regression is usually preferred over linear regression in probability forecasting. Logistic regression can model "threshold" values with rapid transition in predictand better than linear regression and avoids the problem linear regression can have of predicting undefined probabilities less than zero and greater then one.

Finally, the output of the logistic regression model is the probability that the predictand is equal to 1. In the case of both the 2-Hr and 9-Hr Forecast Tools, the output of the model is the probability of lightning. This allowed the creation of a Lightning Index threshold. The Lightning Index threshold is the point at which lightning is predicted when the model output falls above the threshold and not predicted when model output falls below.

2) CANDIDATE PREDICTORS

The data set consisted of 26 candidate independent variables that are shown in Table 4. These were chosen to match the three basic requirements for thunderstorms: 1) moisture, 2) instability, and 3) a trigger of upward motion. The current value of GPS-PW and K-Index directly measure moisture and instability. The K-Index was chosen since it is one of the best of the traditional stability indexes for forecasting thunderstorms at Spaceport Canaveral (Kelly et al., 1998) and to match the Mazany model. The change in GPS-PW indirectly measure thunderstorm triggers by the moisture convergence from vertical motions. The change in GPS-PW over 30 minute interval up to 12 hours as chosen to exceed the lead times of the forecast tool and thus represent the triggers that apply to those lead times.

Table 4: Candidate Regressors

Precipitable Water (PW) (cm)	6.5-hr Δ PW (cm)
0.5-hr Δ PW (cm)	7.0-hr Δ PW (cm)
1.0-hr Δ PW (cm)	7.5-hr Δ PW (cm)
1.5-hr Δ PW (cm)	8.0-hr Δ PW (cm)
2.0-hr Δ PW (cm)	8.5-hr Δ PW (cm)
2.5-hr Δ PW (cm)	9.0-hr Δ PW (cm)
3.0-hr Δ PW (cm)	9.5-hr Δ PW (cm)
3.5-hr Δ PW (cm)	10.0-hr Δ PW (cm)
4.0-hr Δ PW (cm)	10.5-hr Δ PW (cm)
4.5-hr Δ PW (cm)	11.0-hr Δ PW (cm)
5.0-hr Δ PW (cm)	11.5-hr Δ PW (cm)
5.5-hr Δ PW (cm)	12.0-hr Δ PW (cm)
6.0-hr Δ PW (cm)	K-Index (step)

3) METHODOLOGIES FOR REGRESSOR SELECTION

Two basic methodologies were used for regressor selection: a forward method and a backward method. In the forward method, one variable was added at each iteration of the model. In the backward method, all variables were initially entered into the model, with one variable removed at each iteration. In both cases, a regressor was selected only when there was a 95% probability that the regressor was significant in predicting the model outcome.

Goodness-of-fit statistics were initially calculated to evaluate the model fit as different sets of predictors were used. Specifically, Cox & Snell R-squared, Nagelkerke R-Squared, and the Hosmer-Lemeshow goodness-of-fit were produced by Statistical Package for the Social Sciences (SPSS) tool that was used for model development. For both tools, there appeared to be a lack of fit based on the values of Cox & Snell R-squared, Nagelkerke R-Squared, and the Hosmer-Lemeshow goodness-of-fit. While this seemed to indicate a poor fit of the model, further investigation eliminated this concern for several reasons. First, the R-squared values from a logistic regression are not the same as the R-squared values calculated in a linear regression model. Therefore, they are not proven to be good measures of model fit. Second, these values are not relevant in measuring the model's utility because they are evaluated against a threshold of 0.5. This means that outputs above 0.5 were predicted as lightning, whereas values below 0.5 were predicted as no-lightning. The utility of the model will be evaluated based on an optimized threshold.

In selecting the model, the most weight was given to the values of the accuracy measurements and skill scores discussed above. These performance metrics included HR, FAR, POD, KSS, and OUI.

The forward and backward methods of model selection chose different regressors for both tools. For the 2-Hr Forecast Tool, the most recent PW measurement, the latest reading of the K-Index, and the 0.5-hr Δ PW were all selected to be significant in predicting lightning in a 0.5 hour time period by both methods. However, the forward model selection process selected an additional regressor, the 7.5-hr Δ PW, while the backward model selection process picked two additional regressors, the 4.5-hr Δ PW and the 5.5-hr Δ PW. For the 9-Hr Forecast Tool, both the forward and backward model selection methodologies chose the most recent PW measurement, the latest reading of the K-Index, the 3.5-hr Δ PW, the 8.5-hr Δ PW, and the 12.0-hr Δ PW. The backward model selection method picked an additional regressor the 6.5-hr Δ PW that was not chosen by the forward model selection method.

The predictors selected as having the most independent signal in predicting lightning may provide insight as to the physical mechanisms causing the lightning for the respective time period being forecast. The 2-Hr Forecast Tool selected the 0.5-hour change in GPS-PW as the most important predictor. The authors speculate that this predictor represents the local moisture convergence of the developing thunderstorms. Indeed, detecting this mechanism was the original inspiration by one of the authors (Roeder) for using timelines of GPS-PW in local lightning forecasting and as hypothesized as the top predictor for such short term forecasts. The second most important predictor for the 2-Hr Forecast Tool is the 7.5-hour change in GPS-PW. The mechanism associated with this predictor is not very obvious. Possibilities may include general convergence over the Florida Peninsula due to solar heating (sunrise to typical thunderstorms forming at 2000 UTC is about 8 hours), or perhaps a dynamic trigger in the asynoptic upward motion in the right entrance and left exit regions of weak jet streaks over the forecast area (Uccellini and Kocin, 1987), or moisture convergence under flow with a southerly component, or other mechanisms. The 2-Hr Forecast Tool selected the current PW and the K-Index as the third and fourth most important predictors, respectively. These predictors are likely due to the fact that thunderstorms require moisture and instability to form, respectively.

The 9-Hr Forecast Tool selected four predictors in the following order of statistical importance: current GPS-PW, 3.5-hour change in GPS-PW, 8.5-hour change in GPS-PW, and K-index. The current GPS-PW and K-Index likely have the same meteorological explanation as for the 2-Hr Forecast Tool. The 8.5-hour change in GPS-PW may be due to the same reasons speculated for the 2-Hr Forecast Tool. The 3.5-hour change in GPS-PW is also not obvious. The authors speculate it may be due to some trigger of upward motion, perhaps from approaching sea breeze fronts that are known to be important in thunderstorm formation in east central Florida.

It is interesting that the K-Index was the least important predictor for both tools. The K-Index has been shown to be one of the best performing of the traditional thunderstorm indexes in east central Florida (Kelly et al., 1998).

4) Model optimization

For the 2-Hr Forecast Tool, the model was optimized based on the value of the OUI. The OUI is a locally developed performance metric to optimize personnel safety. The OUI is considered the most critical factor because this metric applies more weight to the POD, which is critical when personnel safety is at stake. The 2-Hr Forecast Tool is meant to support the Phase 1 Lightning Advisory, which is issued to ensure personnel working outdoors have adequate time to seek shelter. The equation for the OUI is

$$\mathrm{OUI} = \frac{(3\times\mathrm{POD})+(2\times\mathrm{KSS})+(1\times\mathrm{FAR})}{6} \tag{2}$$

For the 9-Hr Forecast Tool, the model was optimized based on the KSS. The 9-Hr Forecast Tool supports major ground processing operations where personnel safety is less of an issue. The KSS is a more traditional measure of skill, which is appropriate for this application.

5) Lightning Index threshold

As mentioned previously, the output of the logistic regression model for both the 2-Hr and the 9-Hr Forecast Tools is the probability of lightning. The statistical software package used in model development defaults to a threshold of 0.5, meaning that when the probability is 0.5 (or 50%) or greater, lightning is predicted. Conversely, when the probability falls below 0.5, lightning is not predicted. A Lightning Index threshold was established by varying the default value and recalculating the various skill scores and accuracy measures. The Lightning Index threshold was established to optimize the OUI for the 2-Hr Forecast Tool and the KSS for the 9-Hr Forecast Tool.

6) TEST AND VERIFICATION

A key last step to developing the two forecast tools was to independently test and verify the model. This ensures that the model results are repeatable. A verification data set was created for each model using a random sampling of 10% of the initial data points. This verification data set was not used in the development of the model, and thus was not included as part of the development data set. Its sole purpose was to serve as a check after the models were developed to compare results of the verification data set to the development data set.

4. 2-Hr Forecast Tool

The 2-Hr Forecast Tool was designed to support the current Phase 1 Lightning Advisory System at the Spaceport Canaveral by providing a 0.5-h lead time prior to a lightning event. The accuracy measurements and skill scores were calculated based on varying the Lightning Index threshold from 0.0 to 1.0. Table 5 depicts the accuracy measures and skill scores of both forward and backward model selection methodologies at various levels of the Lightning Index threshold. Again, the objective is to maximize the OUI. The model and Lightning Index that produced the highest OUI is shaded in the Table 5.

Table 5: Comparison of Accuracy Measurements and Skill Scores for the 2-Hr Forecast Tool

Model Selection Method	Index	Hit	POD	FAR	KSS	OUI
Forward Method	0.0	49.7%	100.0%	50.3%	0.0%	**41.6%**
Backward Method	0.0	49.7%	100.0%	50.3%	0.0%	**41.6%**

Model Selection Method	Index	Hit	POD	FAR	KSS	OUI
Forward Method	0.1	50.2%	100.0%	50.1%	0.9%	41.9%
Backward Method	0.1	50.0%	100.0%	50.2%	0.5%	41.8%
Forward Method	0.2	52.6%	99.5%	48.8%	5.8%	43.6%
Backward Method	0.2	53.1%	99.4%	48.5%	6.8%	43.9%
Forward Method	0.3	57.4%	96.9%	46.0%	15.2%	45.8%
Backward Method	0.3	57.3%	96.5%	46.1%	15.0%	45.6%
Forward Method	0.4	59.8%	87.8%	43.9%	19.9%	43.3%
Backward Method	0.4	60.1%	87.2%	43.6%	20.6%	43.2%
Forward Method	0.5	58.3%	64.2%	42.9%	16.6%	30.5%
Backward Method	0.5	60.8%	66.5%	40.5%	21.7%	33.7%
Forward Method	0.6	55.5%	28.9%	39.1%	10.6%	11.5%
Backward Method	0.6	57.4%	32.6%	35.8%	14.6%	15.2%
Forwards Method	0.7	51.5%	4.8%	33.3%	2.4%	–2.4%
Backwards Method	0.7	52.6%	7.6%	28.2%	4.7%	0.7%
Forward Method	0.8	50.4%	0.4%	25.0%	0.3%	–3.9%
Backward Method	0.8	50.5%	0.6%	28.6%	0.4%	–4.3%
Forward Method	0.9	50.3%	0.0%	0.0%	0.0%	0.0%
Backward Method	0.9	50.3%	0.0%	0.0%	0.0%	0.0%
Forward Method	1.0	50.3%	0.0%	0.0%	0.0%	0.0%
Backward Method	1.0	50.3%	0.0%	0.0%	0.0%	0.0%

Table 5 shows how the various accuracy measurements change as the Lightning Index threshold changes. Setting the Lightning Index threshold at 0.0 means that when the model outputs the probability of lightning at greater than 0.0, the model predicts that lightning will occur. Because the output of a logistic regression equation is between 0.0 and 1.0, lightning will always be predicted at a Lightning Index threshold of 0.0. Conversely, setting the Lightning Index threshold at 1.0 means that when the model outputs the probability of lightning at greater than 1.0, the model predicts that lightning will occur. The latter case is impossible, because the output of a logistic regression is always between 0.0 and 1.0. Therefore, lightning will never be predicted when the Lightning Index threshold equals 1.0. POD is at its highest when the Lightning Index threshold equals 0.0 because lightning is predicted every time. This results in a very high OUI because the OUI places the most weight on POD. A lower Lightning Index threshold drives a higher FAR because lightning is falsely predicted more often. At a Lightning Index threshold of 1.0, the POD becomes 0.0 because lightning is never detected. This also produces a 0.0% FAR because if lightning is never predicted, naturally it is never falsely predicted.

The forward and backward model selection methods performed similarly at all levels of the Lightning Index threshold; however, the model developed using the forward selection process produced the highest OUI. The OUI is maximized at 45.8% at a Lightning Index threshold of 0.3. However, thresholds of 0.2 or 0.4 both provide a favorable OUI ranging from 43.2% to 43.9%. The OUI was calculated for indexes in 0.01 increments around the optimal range of 0.25 to 0.45 to fine tune the optimal index.

Figure 3 shows that the OUI increases slightly up to a peak of 0.32 and then begins to fall steadily at a threshold of 0.38. The highest OUI results from the forward selection model process at a Lightning Index threshold of 0.32. This OUI is 46.3%. Figure 3 also shows that there is not much sensitivity in the OUI in this range of Lightning Index thresholds. The OUI ranges between 38.5% and 46.3% when the Lightning Index threshold is varied between 0.25 and 0.45.

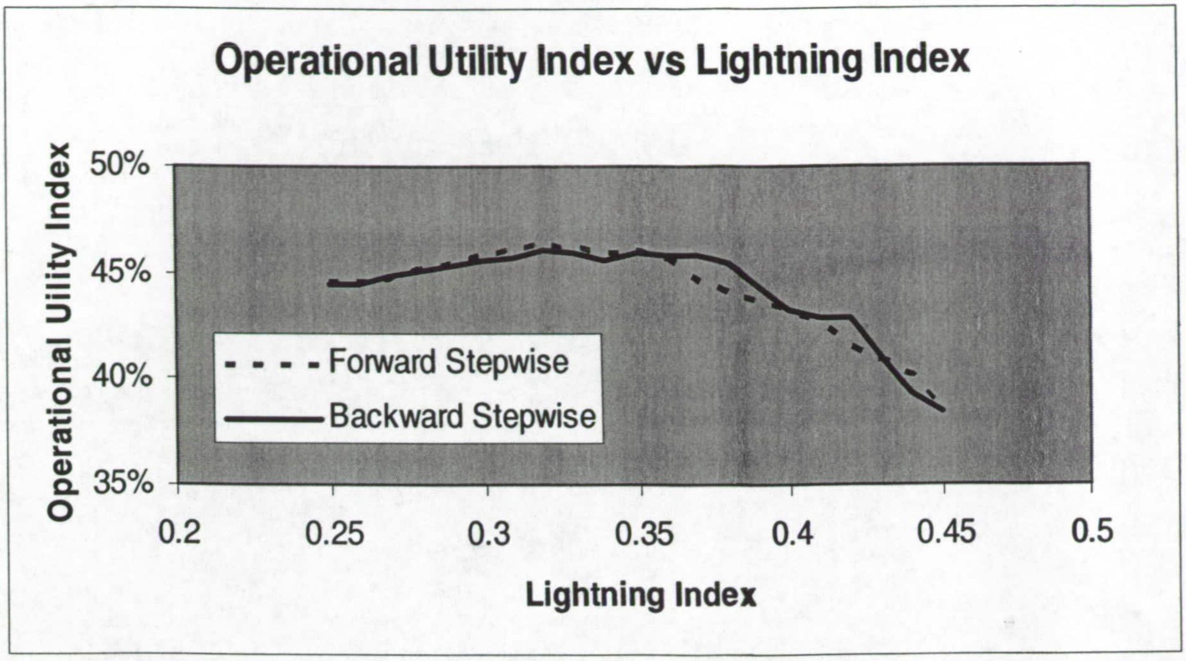

Figure 3: Operational Utility Index vs. Lightning Index for the 2-Hr Forecast Tool Lightning Index threshold

Lowering the Lightning Index threshold from the default value of 0.5 to 0.32 will adjust the mix of lightning forecast/not forecast and lightning observed/not observed, as shown in Table 6.

Table 6: Contingency Tables for the 2-Hr Forecast Tool

Lightning Index Threshold = 0.5		Observed Yes	Observed No	Lightning Index Threshold = 0.32		Observed Yes	Observed No
Forecast	Yes	512	384	Forecast	Yes	768	637
	No	286	423		No	30	170

Decreasing the Lightning Index threshold increases the number of forecast lightning strikes from 896 to 1405 and decreases the number of times that lightning is not forecasted from 709 to 200. The number of missed lightning events decreases from 286 to 30, yet the number of falsely predicted lightning events increases from 384 to 637. Changing the Lightning Index threshold from the default 0.5 to 0.32 changes accuracy measurements as shown in Table 7.

Table 7: Comparison of Accuracy Measurements and Skill Scores at Lightning Index Thresholds 0.5 and 0.32 for the 2-Hr Forecast Tool

Lightning Index Threshold	**Hit**	**POD**	**FAR**	**KSS**	**OUI**
0.5	58.3%	64.2%	42.9%	16.6%	**30.5%**
0.32	58.4%	96.2%	45.3%	17.3%	**46.3%**
Percent Increase	0.3%	50.0%	5.8%	4.4%	**52.1%**

Lowering the Lightning Index threshold means that more lightning is detected. This results in a higher HR, POD, KSS, and OUI. However, this also increases the FAR. This result is acceptable because detecting lightning is much more important than falsely warning of a lightning strike when lives are at stake. Whereas a Lightning Index threshold of 0.32 maximizes the OUI, other thresholds provide a higher HR and KSS and a lower FAR while still maintaining a relatively good OUI.

a. New logistic regression equation

The models perform differently at different Lightning Index threshold levels; therefore, the model selected will vary based on the Lightning Index. At a Lightning Index threshold of 0.32, the model generated using the forward model selection process was selected as the new logistic regression equation because this model maximized the OUI. The logistic regression takes the form of

$$f(z) = \left(1 + e^{-z}\right)^{-1} \tag{3}$$

where

$$z = \alpha' + \Sigma\beta_i x_i \tag{4}$$

and

$\acute{\alpha}$	=	−2.366	
β_1	=	2.053	$x_1 = \Delta$ 0.5 h PW
β_2	=	−0.538	$x_2 = \Delta$ 7.5 h PW
β_3	=	0.031	x_3 = K-Index
β_4	=	0.322	x_4 = PW (cm)

This translates to

$$\frac{1}{1 + e^{-(-2.366 + 2.053x_1 + 0.538x_2 + 0.031x_3 + 0.322x_4)}} \tag{5}$$

The most significant independent variable in the model is the 0.5-h change in PW. A 0.5-hr change in PW will have the greatest effect on the outcome of the model. The least significant variable in the model is the K-Index, which is surprising since it is a traditional tool for forecasting thunderstorms and their associated lightning.

b. Test and verification

An independent data set was used to validate the model. The results are shown in Table 8. For all skill scores and accuracy measures, the verification data set performed closely to the development data set validating the model results.

Table 8: Develop and Verification Results for 2-Hr Forecast Tool

(Lightning Index Threshold = 0.32)

	HR	POD	FAR	KSS	OUI
Development	58%	96%	45%	17%	46%
Verification	57%	95%	48%	18%	45%

5. 9-Hr Forecast Tool

The 9-Hr Forecast Tool seeks to provide a 7.5 hour lead time prior to a lightning event to support major outdoor operations, such as Space Shuttle roll-out from the Vehicle Assembly Building to the launch pad. Prior to beginning an extended outdoor activity, it is essential to know the probability of lightning. If lightning has a high probability of occurrence, the outdoor operations will be postponed or rescheduled until weather conditions are more favorable.

The accuracy measurements and skill scores were calculated based on varying the Lightning Index threshold from 0.0 to 1.0. Table 9 depicts the accuracy measures and skill scores of both forward and backward model selection methodologies at various levels of the Lightning Index threshold. For this Tool, the objective is to maximize the KSS. The models and Lightning Index thresholds that produce the highest KSS are shaded in Table 9.

Table 9: Comparison of Accuracy Measurements and Skill Scores for the 9-Hr Forecast Tool

Model Selection Method	Index	Hit	POD	FAR	KSS	OUI
Forward Method	0.0	38.4%	100.0%	61.6%	0.0%	39.7%
Backward Method	0.0	38.4%	100.0%	61.6%	0.0%	39.7%
Forward Method	0.1	51.8%	98.8%	55.7%	21.3%	47.2%
Backward Method	0.1	51.9%	98.8%	55.7%	21.4%	47.2%
Forward Method	0.2	59.2%	95.8%	51.5%	32.2%	50.1%
Backward Method	0.2	59.3%	95.7%	51.5%	32.4%	50.1%
Forward Method	0.3	63.1%	89.1%	48.9%	35.9%	48.4%
Backward Method	0.3	63.1%	88.9%	48.9%	35.8%	48.3%
Forward Method	0.4	66.2%	75.6%	45.7%	35.9%	42.2%
Backward Method	0.4	66.2%	75.4%	45.7%	35.9%	42.0%
Forward Method	0.5	67.1%	51.8%	41.9%	28.5%	28.4%
Backward Method	0.5	67.2%	51.5%	41.8%	28.5%	28.3%
Forward Method	0.6	64.9%	24.8%	39.5%	14.7%	10.8%
Backward Method	0.6	64.8%	24.5%	39.5%	14.5%	10.5%
Forward Method	0.7	62.7%	5.9%	33.7%	4.0%	–1.3%
Backward Method	0.7	62.7%	6.0%	33.7%	4.1%	–1.2%
Forward Method	0.8	61.6%	0.1%	44.4%	0.1%	–7.3%
Backward Method	0.8	61.6%	0.1%	57.1%	0.0%	–9.5%
Forward Method	0.9	61.6%	0.0%	0.0%	0.0%	0.0%

Model Selection Method	Index	Hit	POD	FAR	KSS	OUI
Backward Method	0.9	61.6%	0.0%	0.0%	0.0%	0.0%
Forward Method	1.0	61.6%	0.0%	0.0%	0.0%	0.0%
Backward Method	1.0	61.6%	0.0%	0.0%	0.0%	0.0%

The KSS varies significantly with changes in the Lightning Index threshold, although the values at thresholds of 0.3 and 0.4 are similar. The KSS is maximized at 35.9% at a Lightning Index threshold of 0.3 for the forward method and 0.4 for both the forward and backward methods. Both forward and backward selection models performed similarly at all levels of the threshold.

The KSS is at its highest at Lightning Index thresholds of 0.3 and 0.4; therefore, the Lightning Index threshold was further refined around these two points by calculating KSS for 0.01 increments of the Index from 0.30 to 0.45. Figure 4 shows how the KSS varies with the Lightning Index threshold for both forward and backward models between 0.25 and 0.45.

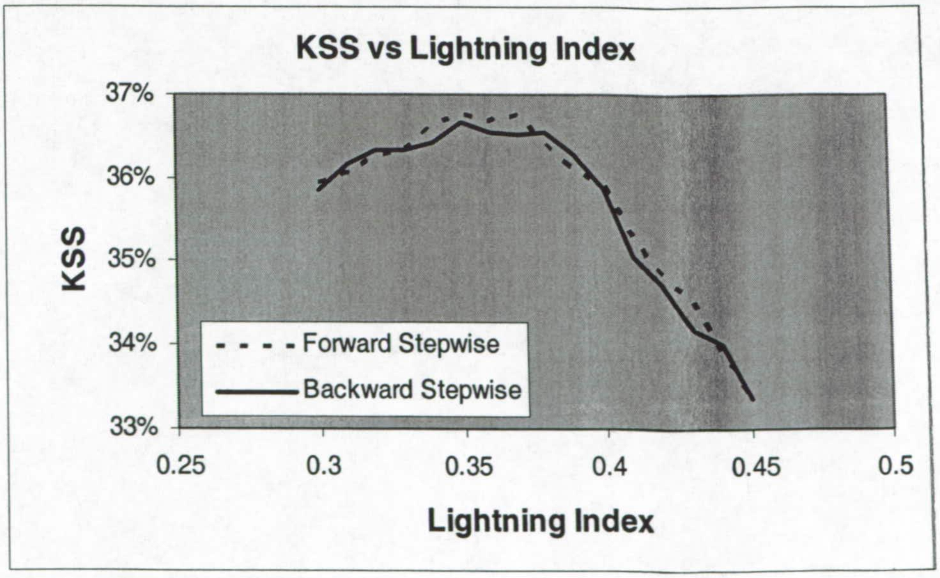

Figure 4: KSS vs. Lightning Index for the 9-Hr Forecast Tool

Figure 4 shows that KSS increases slightly up to a maximum of 0.35 and begins to fall steadily at a Lightning Index threshold of 0.38. The maximum KSS of 36.8% results from the forward selection model process at Lightning Index thresholds of 0.35 and 0.37. In this range of the Lightning Index threshold, the KSS changes minimally, with the largest value being 36.8% and the smallest being 33.4%.

a. Lightning Index threshold

The KSS was maximized at Lightning Index thresholds of both 0.35 and 0.37, indicating an optimal range for the Lightning Index threshold between 0.35 and 0.37. Because the thresholds are being refined to an accuracy of 0.01, a Lightning Index threshold of 0.36 was selected as optimal.

The output of the logistic regression model is the probability that the outcome is equal to 1. In this case, the output is interpreted as the probability of lightning in a continuous 7.5 hour period. The model is designed to predict lightning when the probability of lightning is greater than 50%. Lowering the Lightning Index threshold from 0.5 to 0.36 will increase the amount of lightning that is detected, which will adjust the mix of lightning forecast/not forecast and lightning observed/not observed.

Table 10: Contingency Tables for the 9-Hr Forecast Tool

Lightning Index Threshold = 0.5		Observed		Lightning Index Threshold = 0.36		Observed	
		Yes	No			Yes	No
Forecast	Yes	2393	1727	Forecast	Yes	3790	3357
	No	2227	5677		No	830	4047

Decreasing the Lightning Index threshold increases the number of forecast lightning strikes from 4,120 to 7,147, and decreases the number of times that lightning is not forecasted from 7,904 to 4,877. The number of missed lightning events decreases from 2,227 to 830, yet the number of falsely predicted lightning events increases from 1,727 to 4,047. When the Lightning Index threshold is decreased from 0.5 to 0.36,the skill and accuracy measurements of the model change as shown in Table 11.

Table 11: Comparison of Accuracy Measurements and Skill Scores at Lightning Index Thresholds of 0.5 and 0.36 for the 9-Hr Forecast Tool

Lightning Index Threshold	**HR**	**POD**	**FAR**	**KSS**	**OUI**
0.5	67.1%	51.8%	41.9%	28.5%	28.4%
0.36	65.2%	82.0%	47.0%	36.7%	27.1%
% Increase/Decrease	-2.8%	58.3%	12.2%	28.8%	90.8%

Lowering the Lightning Index threshold means that more lightning is detected. This results in a higher POD, KSS, and OUI. However, this also increases the FAR. Whereas a Lightning Index threshold of 0.36 maximizes the KSS, other thresholds provide a lower FAR while still providing an acceptable KSS.

b. New logistic regression equation

The model generated using forward model selection process was selected as the new logistic regression equation because this model produced the highest KSS. The logistic regression takes the form of

$$f(z)=\frac{1}{1+e^{-z}} \tag{6}$$

where

$$z = \alpha' + \Sigma\beta_i x_i \tag{7}$$

and

$\acute{\alpha}$ = −4.885

β_1 = 0.541 x_1 = PW (cm)

β_2 = 0.346 $x_2 = \Delta$ 3.5-h PW

β_3 = −0.446 $x_3 = \Delta$ 8.5-h PW

β_4 = 0.235 $x_4 = \Delta$ 12-h PW

β_5 = 0.071 x_5 = K-Index

This translates to

$$\frac{1}{1+e^{-(-4.885+0.541x_1+0.346x_2+-0.446x_3+0.235x_4+0.071x_5)}} \tag{8}$$

The most significant independent variable in predicting lightning is the current PW level. The second most significant variable is the 8.5-hr change in the PW level in the atmosphere. This is similar to the Mazany Model, which determined that the 9-hr change in PW was most significant in predicting lightning. The K-Index was the least statistically significant variable.

c. Test and verification

An independent data set was used to validate the model. The results are shown in Table 12. For all skill scores and accuracy measures, the verification data set performed closely to the development data set validating the model results.

Table 12: Development and Verification Results for 9-Hr Forecast Tool

	HR	POD	FAR	KSS	OUI
Development	65%	82%	47%	37%	45%
Verification	64%	81%	49%	34%	44%

6. Recommendations for future research

This project lays the foundation for demonstrating the utility of GPS-PW timelines in forecasting lightning during the East Central Florida summer season. Several opportunities exist to build upon this research and continually improve forecasting accuracy.

This research focused on only two main variables: PW (current level as well as the change in half-hour increments over a 12-h period) and the current reading of the K-Index. A multitude of meteorological factors (e.g., Lifted Index and Thompson Index) could be added to the list of candidate predictors and tested for significance. While K-Index is among the best indexes for the summer lightning season, more recent studies have indicated that indexes optimized by month provides better skill (Lambert et al., 2006). This same research showed that the monthly indexes, flow regime, 1-day persistence, and daily climatology contributed significantly to the probability of lightning. Perhaps the lightning probability that combines these other predictors should be used as a predictor with GPS-PW based parameters. In addition, the influence of changes in PW over extended periods (beyond 12 h) on lightning could be tested. Evidence of this impact is shown in the 9-Hr Forecast Tool, in which a significant model variable was the 12-h change in PW. This suggests that the real optimal change in GPS-PW likely occurs at a larger time increment.

The research very specifically targeted East Central Florida during the summer thunderstorm season. The potential exists to extend the application of this model to other areas and other seasons. Future research should investigate GPS-PW timelines for lightning prediction at Cape Canaveral Air Force Station/KSC during the winter frontal regime. Also, future research includes investigation of GPS-PW timelines for lightning prediction at other locations across Florida and especially areas outside the subtropics.

Neural network forecast modeling is a sophisticated tool that provides the potential to support forecasting. Selection of nonlinear optimal predictor variables and with eventual integration of all lightning precursors, such as electric field mills, local boundary layer convergence, flow regime, daily persistence, daily climatology, numerical model inputs, into the final answer will improve modeling capabilities.

7. Conclusions

In conclusion, the extended validation of the Mazany Model underperformed expectations in POD, HR, FAR, and KSS. The two new tools developed show promise in supporting forecasting at the Spaceport Canaveral during the summer thunderstorm season. The new 2-Hr Forecast Tool will support the current Phase 1 Lightning Advisory System, and the new 9-Hr Forecast Tool will support major, extended outdoor operations. Both tools will help to improve forecasting accuracy, thus improving personnel safety and reducing costs.

This research concludes that GPS-PW time lines may have utility in forecasting lightning in East Central Florida during the summer thunderstorm season. While this is not the final answer to the task of lightning forecasting at the Spaceport Canaveral, it is another key tool in the forecaster's toolbox.

8. Acknowledgements

The data resources and collection that went into this research were substantial and could not have been accomplished without the assistance of several outside parties. The Air Force Combat Climatology Center provided the K-Indexes from the Cape Canaveral Air Force Station radiosondes used in this research. Dr. Gutman from the National Oceanic & Atmospheric Administration Forecast Systems Laboratory (now Earth System Research Laboratory Global Systems Division) provided the GPS-PW data from the Cape Canaveral Coast Guard Station. This research partially satisfied the M.S. Industrial and Systems Engineering degree from the University of Florida for two of the authors (Graf and Kehrer). Dr. H Edwin Romeijn from the University of Florida supported the project as the advisor and provided invaluable support regarding statistical analysis and model development. This research was supported by NASA-Kennedy Space Center through the funding of the degree and the statistical software used, as well as providing time and place to conduct the research.

8. References

Beavis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware, 1992: GPS Meteorology: Remote Sensing Of Atmospheric Water Vapor Using The Global Positioning System, *Journal Geophysical Research*, **97**, 15787-15801

_____, _____, S. Chriswell, T. A. Herring, R. A. Anthes, C. Roken, and R. H. Ware, 1994: GPS Meteorology: Mapping Zenith Wet Delay Onto Precipitable Water, *Journal Applied Meteorology*, **33**, 1282-1299

Bauman, W. H., S. Businger, and M. L. Kaplan, 1997: Nowcasting Convective Activity For Space Shuttle Landings During Easterly Flow Regimes, Weather and Forecasting, **12**, 78-107

Businger, S., S. R. Chiswell, M. Bevis, and J. Duan, R. A. Anthes, C. Rocken, R. H. Ware, M. Exner, T. VanHove, and Fredrick S. Solheim, 1996: The Promise Of GPS In Atmospheric Monitoring, Bulletin American Meteorological Society, **77**, 5-17

Kelly, J. L., H. E. Fuelberg, and W. P. Roeder, 1998: Thunderstorm Predictive Signatures For The East Coast Sea Breeze At Cape Canaveral Air Station And The Kennedy Space Center, *19th Conference On Severe Local Storms*, 14-18 Sep 98, 677-680

Lambert, W., M. M. Wheeler, and W. P. Roeder, 2006: Lightning Probability Forecasts at Kennedy Space Center/Cape Canaveral Air Force Station, Florida, 1st International Lightning Meteorology Conference, 26-27 April 2006, 6 pp.

Mazany, R. A, S. Businger, S. I. Gutman, and W. Roeder, 2002: A Lightning Prediction Index That Utilizes GPS Integrated Precipitable Water Vapor. *Wea. Forecasting,* **17**, 1034–1047.

Roeder, W. P., and T. McNamara, 2006: A Survey Of The Lightning Launch Commit Criteria, 2nd Conference on Meteorological Applications of Lightning Data, 29 Jan-2 Feb 06, Paper 1.1, 16 pp

Uccellini, L. W., and P. J. Kocin, 1987: The Interaction Of Jet Streak Circulations During Heavy Snow Events Along The East Coast Of The United States. *Wea. Forecasting*, **2**, 289-308

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences, Academic Press, A Ahrcourt Science and Technology Company, 525 B Street, Suite 1900, San Diego, CA 92101-4495, ISBN 0-12-751965-3, 467 pp.

Wolfe, D. E., and S. I. Gutman, 2000: Development Of The NOAA/ERL Ground-Based GPS Water Vapor Demonstration Network: Design And Initial Results, Journal Atmospheric Oceanic Technology, **17**, 426-440