REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	7.Sep.00	THESIS

4. TITLE AND SUBTITLE
CONSTRUCTION OF A LIGHTNING INDEX USING INTEGRATED PRECIPITABLE WATER DERIVED FROM THE GLOBAL POSITIONING SYSTEM

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT MAZANY ROBERT A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF HAWAII AT MONOA

8. PERFORMING ORGANIZATION REPORT NUMBER
CY00367

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20000921 075

14. SUBJECT TERMS

15. NUMBER OF PAGES
65

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. 239.18
Designed using Perform Pro, WHS/DIOR, Oct 94

DTIC QUALITY INSPECTED 4

Construction of a Lightning Index using Integrated Precipitable Water Derived from the Global Positioning System

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

METEOROLOGY

DECEMBER 2000

By

Robert A. Mazany

Thesis Committee:

Steven Businger, Chairperson
Gary M. Barnes
Pao-Shin Chu

We certify that we have read this thesis and that, in our opinion, it is satisfactory in scope and quality as a thesis for the degree of Master of Science in Meteorology.

THESIS COMMITTEE

____//Original Signed//___

Chairperson

____//Original Signed//___

____//Original Signed//___

ACKNOWLEDGEMENTS

I would like to begin by thanking the United States Air Force for giving me the opportunity and for supporting me while earning my master's degree in paradise.

The data resources and collection that went into this research were enormous and could not have been accomplished without the special assistance of Seth Gutman from the Field System Laboratory, William Roder from the 45th Weather Squadron, Kennedy Space Center, and also James Foster, Pacific GPS Facility, University of Hawaii.

I would like to thank Gary Barnes and Pao-Shin Chu for their careful reading of this thesis and their suggestion for improvements. I wish to express my sincere appreciation to Professor Steve Businger for his mentorship throughout my studies and his insightful guidance with this research.

Finally I would like to thank my family. My children Rebekah, Britney, Derek, Meagan and my lovely wife Valerie, for which they fill my heart with love and make me want to be a better person. Best of life to you all.

ABSTRACT

A new approach is presented to forecasting the Kennedy Space Center's primary weather challenge, lightning. After examining the first years worth of integrated precipitable water data derived from Global Positioning System (GPS) and surface stations, two periods were chosen to develop a GPS lightning prediction model. Statistical regression methods were used to identify predictors that added skill in forecasting a lightning event. Four predictors proved important in forecasting lightning events; maximum electric field mill values, GPS Integrated Precipitable Water (IPW), nine hour change (∇9-hr) in IPW, and K index. Using the coefficients for these predictors along with a logistic regression equation, a running time series was plotted for the predictand. A common pattern emerged several hours prior to a lightning event. Whenever the predictand log value was 0.7 or below, lightning occurred within the next 12.5 hours. Lightning events were predicted using a logistic threshold value of 0.7 and forecasting time constraints based on the Kennedy Space Center (KSC) criteria. Forecast verification results obtained by using a contingency table

revealed a 26.2% decrease from the Cape's previous season false alarm rates for a non-independent period, and a 13.2% decrease in false alarm rates for an independent test season using the GPS lightning model. Additionally, the model improved the KSC's desired lead-time by nearly 10%. Although a lightning strike window of 12 hours is quit lengthy, forecasters will now have an additional forecasting tool that can be implemented in their lighting forecast process. Once the value of the GPS lightning model has been confirmed using data from the 2000 season, it is anticipated that the model will enhance mission readiness and save valuable time and dollars by helping forecasters anticipate and improve forecast lightning events at the KSC.

TABLE OF CONTENTS

Acknowledgments iii

Abstract iv

List of Tables vi

List of Figures vii

Chapter 1: Introduction 1

1.1 Previous Studies 2

1.2 Role of Water 4

1.3 Global Positioning System and Atmospheric Propagation 6

1.4 Goals 12

Chapter 2: Data Resources 14

Chapter 3: Development of a Lightning Prediction Model. 19

3.1 Logistic Regression 19

3.2 The Predictors. 20

3.3 Relationship between Predictors and Predictand.22

3.4 Assessing the Fit of the GPS Lightning Model. 28

3.5 A GPS Lightning Prediction Model 31

3.6 Categorical Forecasts for the Thunderstorm Season 33

3.7 Categorical Forecasts for the Independent Preseason 34

3.8 Missed Events 37

Chapter 4: Summary and Conclusion 39

4.1 Summary 38

Appendix A: Tables 43

Appendix B: Figures 49

References 61

LIST OF TABLES

Table		Page
1.	Initial Predictors	43
2.	Logistic Regression Table	44
3.	Percent Frequency of Thunderstorms	44
4.	Goodness of Fit Tests	45
5.	Observed and Expected Frequencies	45
6.	Measures of Association	46
7.	Contingency Table	46
8.	Accuracy Measures For Binary Forecasts	47
9.	Thunderstorm Season Test Results	47
10.	Independent Test Case Results	48

LIST OF FIGURES

Figure

Page

1. GPS Sites 49

2. Sliding Window Approach 50

3. Composite Time Series of GPS IPW and Max Electric Field Mills 51

4. Composite of Non-Event Days 51

5. Scatter Plot Diagram Of Average Electric Field Mill Values And GPS IPW 52

6. Scatter Plot Diagram Of Average Electric Field Mill Values And K Index 52

7. Scatter Plot Diagram Of Average Electric Field Mill Values And ∇1 GPS IPW 53

8. Scatter Plot Diagram Of Maximum Electric Field Mill Values And ∇9 GPS IPW 53

9. Composite Time Series Of Average ∇9 GPS IPW 54

10. 5 July 99, GPS Lightning model run 55

11. 15 July 99, GPS Lightning model run 55

12. 18 July 99, GPS Lightning model run 56

13. 3 July 99, GPS Lightning model run 56

14. 10 July 99, GPS Lightning model run 57

15. 1 August 99, GPS Lightning model run 57

16. First-Strike Time Chart 58

17. 19 May 99. GPS Lightning model run 58

18. 13 May 99, GPS Lightning model run 59

19. 20 May 99, GPS Lightning model run. 59

20. 21-22 Jul 99. GPS Lightning model run. 60

21. 9 Aug 99, GPS Lightning model run 60

Chapter 1:
Introduction

Space launches and landings at the Kennedy Space Center (KSC) are subject to strict weather-related constraints. Nearly 75% of all space shuttle countdowns between 1981 and 1994 were delayed or scrubbed, with about one-half of these due to weather (Hazen et al. 1995). Of the various weather constraints, their primary weather challenge is to forecast lightning 90 minutes before a *first-strike* and within a 20nm radius of the complex. The National Lightning Detection Network indicates that this region has the highest lightning flash density in the country, averaging 10 flashes/km^2/yr. Lightning has a huge impact on the Kennedy Space Center. First of all there is the safety of personnel working on the complex. Next, resource protection for over $10 Billion in rocket launching systems and platforms, that includes the Space Shuttle, Athena, Pegasus, Atlas, Trident II, and Titan IV. Finally, delay costs can run anywhere from $90,000 for a 24-hour delay to $1,000,000 if the Shuttle must land at another facility and be transported back to the KSC.

1.1 Previous Studies

Modeling and observational studies conclude that patterns and locations of Florida convection are related to the interaction of the synoptic wind field with the mesoscale sea-breeze (Estoque 1962; Neumann 1971; Pielke 1974; Boybeyi and Raman 1992). The sea-breeze circulation and patterns of convection have different characteristics dependent on the whether or not the low-level flow has onshore, offshore, or an alongshore component with respect to Florida's east coast (Aritt 1993).

Onshore, easterly flow typically generates less vigorous convection than offshore, westerly flow (Foote 1991). However, onshore flow is characterized by a shallow low-level maritime moist layer, capped by a subsidence layer with dry conditions aloft, creating difficulties in predicting convection associated with this type of regime (Pielke 1974; Bauman et al. 1997). Blanchard and Lopez (1985) show that the majority of convection takes place in the sea-breeze and lake breeze convergence zones. They also stated that convection is sparse and requires low-level forcing. Generally, only when the east coast sea-breeze has advanced westward and merges with the west coast sea-breeze is there enough low-

level forcing to generate deep convection. Obviously convection does develop independently of a sea-breeze frontal merge, but it is usually weaker than when the fronts merge.

Reap (1994) found that southwesterly flow tends to be more unstable and produce more lightning strikes along the Florida east coast than easterly flow. The southwesterly flow also contains deeper moisture and accounts for two-thirds of the lightning strikes during the summer at KSC. In contrast, easterly flow only accounts for less than 5% of the total lighting flashes (Watson et al. 1991).

The International Station Meteorological Climate Summary for Cape Canaveral (Mar 68- Feb 78) indicated an annual average of 76 days with thunderstorms per year. Most of the thunderstorms, 81.2%, occur from May through September (inclusive). The 45th Weather Squadron (WS) issues over 1200 lightning watches and warning per year.

The 45th WS uses many observation systems to detect and predict lightning in support of the space center needs.

Along with the use of satellite data, numerical models, weather radars, and rawinsonde data, there are five lightning detection systems. The Lightning Detection And

Ranging (LDAR) is a seven antenna radio-wave time-of-arrival system which provides a three-dimensional picture of in-cloud, cloud-to-cloud, cloud-to-clear air, and cloud-to-ground lightning. The Cloud to Ground Lightning Surveillance System is a 5-antenna magnetic direction finding system. The Launch Pad Lightning Warning System (LPLWS) is a network of 31 surface electric field mills. The National Lightning Detection Network (NLDN) is a national network of magnetic direction finding and time-of-arrival antennas. The A.D. Little Corp sensor is an older system using one antenna to estimate the lightning distance from the magnetic pulse change. Most of these lighting detection systems are more fully described by Harms et al. (1997).

For the 1999 thunderstorm season, the 45th WS's capability of detecting thunderstorms is 97.5%, 79.1% of which meet the desired lead-time. The KSC false alarm rate is 43.2%. There is room for improvement in these statistics, particularly reducing the false alarm rate.

1.2 The Role of Water

The water molecule has a unique structure that results in a permanent dipole moment. This dipole moment

is caused by an asymmetric distribution of charge in the water molecule. Several different mechanisms have been proposed to account for generation of electrical charge separation in clouds. However, only the polarization mechanism has been shown by numerical modeling to be capable of generating the amounts of charge at rates typical of thunderstorms (e.g., Fleagle and Businger 1980, p. 139). When collisions occur between the falling grauple and a cloud droplet or ice pellet, a negative charge is transferred to the grauple, leaving the droplet or ice pellet positively charged. The smaller particle, now with positive charge, is carried upward in the updrafts, while the heavier grauple carries the negative charge downward. This process is reinforcing because as charges are separated, the electric field strength increases, thus increasing both polarization and the transfer of charge occurring at each collision (Fleagle and Businger 1980, p. 139).

Water plays a critical role in a variety of atmospheric processes that act over a wide range of temporal and spatial scales. It is the most variable of the major constituents of the atmosphere. The distribution of water vapor is closely coupled with the

distribution of clouds and rainfall. Due to the large latent heat release of water vapor during a phase change, the distribution of water vapor plays a crucial role in the vertical stability of the atmosphere and evolution of storm systems.

Bevis et al. (1992, 1994) describe the methodology for using GPS to monitor atmospheric water vapor from ground-based GPS sites and explores the error analysis of GPS precipitable water. Duan et al. (1996) provides the first direct estimation of precipitable water by eliminating any need for external comparison with water vapor radiometer observations. Businger et al. (1996) describes meteorological application of atmospheric monitoring by GPS for use in weather and climate studies, and in numerical weather prediction models. In the next subsection, the historical development of GPS meteorology is summarized.

1.3 Global Positioning System and Atmospheric Propagation

The GPS consists of a constellation of 24 satellites that transmit L band (1.228 and 1.575 GHz) radio signals to a number of users equipped with GPS receivers for use

in time transfer, navigation, and relative positioning (Businger et al. 1996). These microwave radio signals transmitted by GPS satellites suffer propagation delays due to the refraction of the signal by the Earth's atmosphere. Meteorologists can exploit these delays to determine the total Integrated Precipitable Water Vapor (IPW) over a particular GPS site (Businger et al. 1996).

The total delay is comprised of two parts, the ionospheric delay and the neutral atmospheric delay. The ionosphere introduces a delay that can be determined and removed by recording both of the frequencies transmitted and exploiting the known dispersion relations for the ionosphere (Spilker 1980). The neutral delay varies according to the angle at which the GPS signal propagates through the atmosphere. The minimum is observed for a vertical path when the satellite is in the zenith position. Therefore a mapping function is used to convert a delay into a single Zenith Tropospheric Delay. The simplest delay models have the form,

$$D = Zm(\theta), \tag{1}$$

where D is the delay along a single path with elevation angle θ, Z is the zenith delay, and $m(\theta)$ is the mapping function where $m(\theta) \approx 1/\sin(\theta)$ (Davis et al. 1985).

The delay in the neutral atmosphere is due to the presence of gases composing the Earth's atmosphere. A significant and unique delay is introduced by water vapor because it is the only component that has a permanent dipole moment. This dipole moment is caused by an asymmetric distribution of charge in the water molecule. Per mole, the refractivity of water vapor is 20 times that of dry air (Bevis et al. 1992). For this reason, the neutral atmosphere can be grouped into two categories; Zenith Wet Delay (Z_w), the dipole component of water, and Zenith Hydrostatic Delay (Z_h). Thus, Eq.(1) can now be generalized as

$$D = Z_w m_w(\theta) + Z_h m_h(\theta) \tag{2}$$

where, $m_w(\theta)$ is the wet mapping function and $m_h(\theta)$ is the hydrostatic mapping function. Tralli and Lichten (1990) show that the wet and hydrostatic mapping functions only differ slightly and can be treated together using a single mapping function. Thus, parameterizing the problem solely in terms of the zenith neutral delay Z_n

$$D = Z_n m_n(\theta) \tag{3}$$

where $m_n(\theta)$ is the neutral mapping function and

$$Z_n = Z_w + Z_h. \tag{4}$$

Using Eq. (4) the zenith wet delay can easily be estimated by subtracting from the neutral delay, the hydrostatic delay.

The Z_h is proportional to the total mass of gas encountered along a single path. This in turn is proportional to the surface pressure. So, given surface pressure readings at the GPS receiver, the retrieved Z_h delay can be resolved. Elgered et al. (1991) adopted a model in which

$$Z_h = (2.2779 \pm 0024) P_s / f(\lambda, H) \qquad (5)$$

where P_s is the total pressure in millibars at the Earth's surface, and

$$f(\lambda, H) = (1 - .00266 \cos 2\lambda - .00028H) \qquad (6)$$

accounts for the variation in gravitational acceleration with latitude λ and the height H of the surface above the earth (in km). The troposphere accounts for approximately 75% of the total hydrostatic delay.

The next component of the neutral delay is the wet delay. The zenith wet delay is given by Davis et al. (1992) as

$$Z_w = 10^{-6}[k'_2 \int (P_v/T)\, dz + k_3 \int (P_v/T^2)\, dz] \qquad (7)$$

where k'_2 = (17 $\pm$ 10) k mb^{-1}, k_3 = (3.6776 $\pm$ .004)X 10^5 K^2 mb^{-1}. Both are constants related to the refractivity of moist air. P_v is the partial pressure of water vapor (in millibars), T is the atmospheric temperature (in degrees Kelvin), the integral is along the zenith path, and the delay is given in units of z (millimeters).

To derive the relationship between the vertically integrated water vapor (IWV) and an observed wet delay, Davis et al. (1985) introduced a weighted mean temperature, T_m as:

$$T_m = \frac{\int (P_v/T)\, dz}{\int (P_v/T^2)\, dz} \tag{8}$$

Combining (7) and (8), and the equation of state for water vapor we obtain:

$$\text{Total Mass of Water} = \text{IWV} = \int \rho_v\, dz \approx \kappa\, Z_w \tag{9}$$

where we now combine the various constants and introduce a dimensionless, constant of proportionality:

$$\Pi = \kappa/\rho_w = 10^6 [\rho_w R_v (k'_2 + k_3/T_m)]^{-1}. \tag{10}$$

The water vapor content of the atmosphere can be stated as the height of an equivalent column of liquid water, which

is also know as precipitable water (PW). The IWV is just the product of ρ_w and PW, thus $PW = Z_w\Pi$. Normally, Π is approximately .15, but actual values can vary by as much as 15% depending on the local climate and seasonal variability (Businger et al. 1996). Bevis et al. (1992) concluded that PW typically can be recovered with a root mean square error of less than 2mm +2% of the PW and biases of less than 2mm. The typical values of the GPS signal delay range from 5.16 m to 32.85m. The wet delay portion of the signal is actually very small. Only about 1% of the total delay is used to determine the precipitable water.

The National Oceanic and Atmospheric Administration's (NOAA) Forecast Systems Laboratory (FSL) established the first GPS network dedicated to atmospheric remote sensing in the United States (Fig. 1, Wolf et al. 1998). This network was established to demonstrate the feasibility and utility of surfaced-GPS observations for climate monitoring, satellite sensor calibration/validation, improve weather forecasting using GPS IPW to initialize numerical weather prediction models, and to transfer this new observation system technology to operational use.

Not until recently, did GPS processing algorithms have the capability to process data in real time. James Foster (personal communication) has applied a sliding window technique that provides estimates of IPW in near real-time (Fig. 2). The sliding window solutions are generated using an 8-hour window with zenith neutral delay estimates every half-hour along with 3 gradient estimates. Solutions are run in hourly steps and any singular estimate in a sliding window solution can be given a 95% confidence limit of $\pm$ 2 mm precipitable water (James Foster, personal communication). These real-time sliding window solutions are available, courtesy of the NOAA FSL on the world wide web at http ://ghub1.fsl.noaa.gov/rt/rtlinks/.

1.4 Goals

The purpose of this study is to develop a predictive GPS lightning model that takes advantage of Integrated Precipitable Water from the Global Positioning System (GPS) at the Kennedy Space Center. Additionally, other meteorological variables that may be factors in lightning prediction are investigated. A statistical approach that combines these data is used to develop a new predictor of

thunderstorm activity and, thus increase the skill of forecasting a *first-strike* at the Kennedy Space Center.

Chapter 2:
Data Resources

The thunderstorm season runs from May through September. The data were divided into two periods, a pre-season (14 Apr - 9 Jun 99) and a thunderstorm season (10 Jun - 26 Sep 99). These particular dates were chosen due to the distribution of thunderstorms and data availability associated with instrumentation down time.

The thunderstorm season data were used to create the logistic regression model for lightning prediction. Thunderstorm season data contained 46 event days and provided robust predictor data. During this season GPS IPW values are much higher and show more variability than in the winter and preseason. The preseason data were reserved for an independent test using the logistic regression model results.

Cape Canaveral has a very dense array of weather sensors. One of the challenges of this research was determining which meteorological variables would add skill in a lightning prediction model. Twenty-three potential predictors were initially evaluated (Table 1).

The availability of real time GPS IPW was the primary motivation for the undertaking the research. Since October

1998 GPS IPW data have become available with 30 minute temporal resolution for a GPS site located 28.48 N. latitude, 80.38 W longitude, roughly the center of the Cape just north of the primary landing strip. Data for this research covers one year from October 1998 to October 1999. The GPS site had periods of down time, and the data from any day that contained partial down time was eliminated from the data set.

The electric field mills provided the next source of data investigated. There are 31 field mills that measure the electric potential of the atmosphere in volt/meters (V/m) every five minutes. The maximum field mill value was used for the 30-minute window ranging from the top of the hour until the half-hour mark. This maximum value was assigned to the GPS IPW value taken 15 minutes after the hour. From the half-hour mark to the top of the hour, that value was assigned with the GPS IPW value taken 45 minutes after the hour. Typical fair weather electric field mill values ranged from 70 V/m to 800 V/m. During inclement weather when the potential for lightning existed, values would increase substantially, sometimes reaching values of 12,000 V/m during a lightning event. The only suspect values were around 1000 UTC (0500L).

From a normal field of 100-200V/m just before 1000 UTC, field mill values would jump, sometimes up to 3000 V/m for what appeared to be no meteorological event. Marshall et. al. (1999) explains this *sunrise effect* as the local, upward mixing of the denser, low-lying, electrode-layer charge.

Other variables investigated for the GPS lightning model include 700-mb vertical velocity from the ETA model, Total Totals (TT) index, K-Index (KI), freezing level from rawinsondes, and surface temperature, dewpoint, pressure, and wind direction taken from station observations. Cape Canaveral data resources offered benefits that make this study possible. Upper air soundings are often taken more than just twice a day, pending various launches and weather conditions. So there is a slight increase in the temporal resolution.

The KI considers the static stability of the 850-500-mb layer. The KI is given by the equation

$$KI = T_{850} - T_{500} + T_{d850} - (T_{700} - T_{d700}) \qquad (11)$$

Where T_{850} and T_{d850} are the dry bulb temperature and dewpoint at 850 mb, and T_{500} is the dry bulb temperature 500 mb. The quantity $T_{700} - T_{d700}$ is the 700-mb dewpoint depression. In order for the KI to correspond with the 30

minute GPS temporal resolution, KI values were interpolated linearly. Granted the atmosphere does not operate in a linear fashion and since the KI only varies slightly during a six to twelve hour period, this method was most easily employed with minimal affects. Also, there are observers at the Cape 24-hours a day. Adding a human element to the observation codes, especially in the remark's section, provided an increase in understanding of the meteorological conditions.

Finally, LDAR data were used as *ground truth* to verify when and where a lightning event occurred. LDAR data are voluminous; the sensors detect stepped-leaders. With a time resolution on the order of milliseconds, one lightning flash can have up to 20,000 LDAR points, one thunderstorm can have thousands of flashes, so one thunderstorm can have up to tens of millions of LDAR points. These LDAR points are ranged in meters from a central site in an x, y, and z direction. A point is classified as a new flash if the new point is 300 milliseconds (ms) later or 5000 meters from the previous point. Also, two or more points make a flash. This is same criteria that the National Weather Service, Melbourne FL., uses to actually verify step-leader points as a

lightning flash. For the purpose of the research, the first-strike was verified using LDAR data and matched to the nearest corresponding GPS IPW data. Figures in section 3 used this method to depict the time of the first-strike.

Other predictor variables could have been included in this study. However, potential predictors such as low level divergence, thunderstorm motion, radar data, and other instability indices would be nearly impossible to assimilate into this research in a reasonable time.

Chapter 3

Development of a Lightning Prediction Model

3.1 Logistic Regression

Regression methods provide the best opportunity for data analysis concerned with describing the relationship between a response variable and one or more predictor variables. Since the event to be forecast was the first-strike of a lightning event, a Binary Logistic Regression model was chosen as opposed to a linear regression model.

What distinguishes a logistic regression model from the linear regression model is that the outcome variable in logistic regression is binary or dichotomous (Hosmer and Lemeshow 1989). The two outcomes are *yes* the lightning event occurred or *no* it did not.

The quantity $\pi_j = E(Y|x_j)$ represents the conditional mean of a lightning strike (Y) given a predictor (x) when the logistic distribution is used. The specific form of the logistic regression model is

$$\pi_j = \frac{e^{\beta_0 + \beta x_j}}{1 + e^{\beta_0 + \beta x_j}}, \qquad (12)$$

where π_j is the probability of a response for the j^{th} covariate, β_0 is the intercept, β is a vector of unknown coefficients associated with the predictor, x_j is a predictor variable associated with the j^{th} covariate. Next, Hosmer and Lemeshow (1989) use a logit transformation of π_j defined as:

$$g(\pi_j) = \ln [(\pi_j)/(1-\pi_j)] = \beta_0 + \beta_j X_j \qquad (13)$$

The importance of this link function is that $g(\pi_j)$ has many of the desirable properties of a linear model. The logit, $g(x)$ is linear in its parameters, may be continuous, and may range from $-\infty$ to $+\infty$, depending on the range of x.

3.2 The Predictors

In order to determine what variables contribute significantly in the regression, the initial set of predictors included 23 in all (Table 1). In the table changes (∇) in IPW with time for periods ranging from 1 to 12 hours comprise variables 2 through 13. The purpose of the ∇-hourly changes is to capture the moisture difference in the air mass properties over various periods of time. For example, if the current time was 3:00 PM, a ∇-9hr IPW subtracts the IPW of 6:00 AM from the 3:00 PM IPW value (a

nine hour time difference). Thus, positive values indicate that IPW has increased over that 9-hr time period.

Logistic regression model output shows the estimates of the coefficients, standard error of the coefficients, z-values, p-values, and a 95% confidence interval for the odds ratio. Predictors that did not meet the 99% significance level, sometimes called the *p-value,* in the model results were eliminated.

Model output of the initial 23 variables left only four predictors that met the 99% significance level. These four are electric field mill maximum (V/m), GPS IPW, ∇9 hr IPW, and KI. The coefficient of each predictor is the estimated change in the link function with a one-unit change in the predictor, assuming all other factors and covariates are the same.

Statistical hypothesis testing is carried out by setting up a null hypothesis. If we state the null hypothesis as the coefficients are zero, the estimated coefficients in Table 2, show that the remaining predictors all have a p-value $\leq .01$. This indicates that there is sufficient evidence that the parameters are not zero with a 99% significance level. Thus we can reject the null hypothesis and use our estimated coefficients.

Further, review of the odds ratios in Table 2, indicate that some predictors have a greater impact than others. An odds ratio very close to one indicates that a one unit increase minimally affects a lightning event. A more meaningful difference is found if you look at ∇9 hr IPW. An odds ratio of 1.38 indicates that the odds of a lightning event increase by 1.38 times with each unit increase. The z-value is obtained by dividing the coefficient by its standard deviation. Dividing by the standard deviation weighs the accuracy of the coefficient. Smaller standard deviations lead to larger z-values, positive or negative. Table 2 reflects the top four z-values and provides strong evidence that the coefficients are highly accurate and belong in the GPS lightning model.

3.3 Relationship between Predictors and Predictand

The relationship between time series of discrete events can be studied by a technique known as the *superposed epoch* method (Panofsky and Brier 1958). The *first strike* of lightning is defined as a discrete event since it is the critical component of the forecast for the Cape. Since lightning occurs at various times of the day, the superposed epoch method creates composites of the data

surrounding the 27 lightning events during the thunderstorm period. For each lightning event the time of first strike was denoted as T_0. Hours prior to that key time were denoted as T_{0-1}, T_{0-2}, T_{0-3}, etc. Figure 3 depicts the composite GPS IPW values leading up to the first strike. The general increase in IPW suggests a correlation between increasing IPW values and the time of the first strike. Contrary to GPS IPW, electric field mill values show random fluctuations leading right up to the first strike when the field mill values spike up to indicate a lightning event has occurred. In this case, there is very little warning time for forecasting lightning events. This predictor remains in the model due it's relationship with lighting 90-minutes prior to the first strike (Fig. 3). Although this figure shows slight increases up until the first-strike itself, in the 5 minute resolution (raw data), the increases are much more dramatic.

Days containing no lightning at all, non-event days, have IPW values that hover around 35mm. The non-event graph depicts average GPS IPW values for 20 days during the thunderstorm period (Fig 4). As seen in the graph, the 24-hr run is relatively flat exhibiting minor fluctuations

that can easily be attributed to solar heating. This can be seen in the subtle nocturnal decline from 3:15 UTC (2200 EST) until 11:15 UTC (0600 EST) and the rise to 18:15 UTC (13:15 EST).

A series of scatter diagrams were used to document relationships between the predictors and the lightning events. The scatter plots show data for all thunderstorm days. Only the 30-minute data up to the time of the first lightning strike are plotted. No data following the first strike are included in the plots. Lightning events were considered independent if there was a 12-hour period between the end of one lightning event and the first-strike of another. Figure 5, show GPS IPW values > 35mm are more conducive for lightning strikes. Conversely, no first-strike events were noted when the GPS IPW values were < 33mm.

When the KI is plotted with electric field mill data (Fig. 6), a clear bias is present. Lightning events are much more likely to occur when the KI was $\geq$ 26. This stability index proved more relevant than the total totals index. A study on nowcasting convective activity for the KSC conducted by Bauman et al. (1997), concluded that of all the stability indices, only the KI was found to have a

modest utility in discriminating convective activity. This can be attributed to the fact that the KI captures a moisture layer from 850mb to 700mb, as opposed to just one reference point of 850mb dewpoint temperature by total totals. Typical Cape KI values ranging from 26-30 yield an air mass thunderstorm probability of 40% to 60%. A value of 31-35 yields a probability range of 60%-80%. While 36-40 yields a probability of 80%-90%. Values for this study hover around the 50% to 60% range, not giving a forecaster much better odds than flipping a coin.

The relationship between change in GPS IPW and lightning occurrence proves a little more challenging to explain. Initially changes of IPW for one hour were used. Eventually, this was carried out to twelve-hour IPW changes. Again, the superposed epoch method was applied to the data. Use of the hourly ∇-1 to ∇-12 hourly IPW predictors enabled the GPS lightning model to capture the IPW changes of the air mass. As it turned out, (statistically) the ∇9-hr IPW predictor had more impact than the other ∇ IPW predictors. Looking at field mill values and ∇1-hr IPW (Fig. 7), lightning indicators values are scattered on either side of the zero line with the average value around

2.5 mm. When compared to the ∇9-hr IPW plot (Fig. 8), average ∇9-hr IPW values are double the ∇1-hr averages and indicate no lightning strikes occurred below the zero line. Other intervals, such as ∇6 hr IPW (not shown), do show a tendency of increased lightning as IPW increased. However, statistically with the logistic regression model, the ∇9-hr IPW prevails as the best predictor. The ∇9-hr IPW exhibits the most prominent increase of IPW in the 5 hours prior to that first strike (Fig. 9).

The ∇9-hr IPW predictor refers to a timeframe 9 hours prior to the first strike. Therefore, the ∇9-hr IPW examines changes in IPW on meteorological events that span the course of 9 hours. Table 3 shows that most of the lightning events occur in mid to late afternoon. Mechanisms linked to this timeframe may include affects attributed to the diurnal cycle of solar heating and moisture properties associated with the sea breeze and as discussed in the introduction, the impact of synoptic circulations.

Increases in the ∇9-hr IPW indicate an increase in the amount of mid-level moisture that plays an important role in the stability of convective clouds. Moist air (as

opposed to dry air) being entrained into these clouds will result in an increase in their buoyancy.

A possible mechanism for increased midlevel moisture is the interaction of various mesoscale boundaries associated with the geography in central Florida. A strong sea breeze from the western peninsula coast advances and interacts with the eastern coastal sea breeze. The ∇9-hr IPW predictor could be detecting the increased moisture associated with sea breeze fronts. Other mechanism include deeper moisture associated with southwesterly flow regimes, detecting an increase in the maritime moist layer (Reap 1994).

Another important mechanism is dynamics associated with the passage of jet streaks aloft (Bauman et al. 1997). Divergence aloft is associated with jet entrance and exit regions, and draws moisture up to mid-levels in the troposphere.

3.4 Assessing the Fit of the GPS Lightning Model

To determine the effectiveness of the GPS lightning model in describing the outcome variable, the fit of the estimated logistic regression must now be assessed. This is referred to as goodness of fit. Hosmer and Lemeshow (1989) recommend three methods to determine goodness of fit; Pearson residual, Deviance residual, and the Hosmer-Lemeshow test (Table 4). They also introduce a decile of risk method for observed and expected frequencies (Table 5) as well as a measures of association between the response variable and the predicted probabilities
(Table 6).

The p-values range from .605 to 1.000 for the Pearson residuals and Deviance residuals, and the Hosmer-Lemeshow tests (Table 4). This indicates that there is sufficient evidence for the model fitting the data adequately. If the p-values were less than the accepted level (.05), the test would indicate sufficient evidence for a conclusion of an inadequate model fit.

The results of applying the decile of risk grouping strategy (Hosmer and Lemeshow 1989) to the estimated probabilities computed from the model for lightning

strikes are given in Table 5. The data in Table 5 are grouped by their estimated probabilities from lowest to highest in 0.1 increments. Thus, group 1 contains the data with the lowest estimated probabilities ($\leq$0.1) while group 10 contains data with the highest estimated probabilities (>0.9). Since the total number of lightning strikes is 995, each decile group total must be evenly distributed for proper comparison. Therefore, the Hosmer and Lemeshow strategy breaks down each group into a total of 99 or 100 events.

The following will help explain the meaning of this table. The observed frequency in the *yes* (y=0, a lightning strike) group for the seventh decile (<0.7) of risk is 26, meaning that there were 26 lightning events actually observed from the seventh decile group. These are the events that have an estimated probability of occurring of $\leq$0.7. In a similar fashion the corresponding estimated expected frequency for this seventh decile is 25.8, which is the sum of the models estimated probabilities for these lightning events to occur. The observed frequency for the *no* lightning (y=1) group is 99-26 = 73, and the estimated frequency is 99-25.8 = 73.2. Table 4 provides sufficient evidence that the model does

fit the data well because the observed and expected frequencies are very close. Further information on this table can be found in Hosmer and Lemeshow (1989).

Measures of Association (Table 6) display a table of the number and percentage of concordant, discordant, and tied pairs (Hosmer and Lemeshow 1989). These values measure the association between the observed responses and the predicted probabilities. The values in Table 5 are calculated by pairing the observations with different response values. Here you have 221 *yes* lightning strikes and 774 *no* lightning events recorded during the thunderstorm period. This results in 221*774 =171054 pairs with different response values. Based on the GPS lightning model, a pair is concordant if the *yes* lightning event has a higher probability by the sum of their individual estimated probabilities being greater than the observed lightning events and discordant if the opposite is true, and tied if the probabilities are equal. These values are used as a comparative measure of prediction.

3.5 A GPS Lightning Prediction Model

The lightning model was tested on data from the thunderstorm season. In order obtain the proper predictand, Wilks (1995), suggests using

$$\hat{y} = \frac{1}{1+ \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4)}. \quad (14)$$

where $\hat{y}$ is the predictand, β the coefficients for each predictor, x the value of the predictor, and the subscripts indicate which predictor it is for. In this case, using the GPS lightning model coefficients for each predictor, Eq. (14) becomes:

$$\hat{y} = \frac{1}{1+\exp((-6.7866)+(.0011359)x_1+(.06063)x_2+(.32341)x_3+(.06728)x_4)}. \quad (15)$$

The meaning of this equation is most easily understood in the limits, as $(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4) \rightarrow \pm \infty$. As the exponential function in the denominator becomes arbitrarily large, the predicted value approaches zero, a lightning strike. As the exponential function in the denominator approaches zero, the predictand approaches one, a non-event. Thus, it is guaranteed that the logistic regression will produce properly bounded

probability estimates. The predictand value was calculated for the entire data set for both test periods.

After examining every day from the thunderstorm season sample dates, some common recurring patterns are relevant. Non-events day predictand values typically fluctuate very close to 1.0 (Figs. 10, 11, and 12). When predictand values fell below 0.7, lightning events followed. I call this 0.7 level, the Predictand Threshold Value (PTV). This particular level was chosen after reviewing all the predictand value time series for every day in the thunderstorm period. Other predictand levels did not have much forecast ability. The 0.8 level often recovered to the 0.9 level, indicating a non-lightning event, while the 0.6 level often did not offer much lead time before the first strike. The 0.7 level was best suited for a forecast indicator. Creating a running mean time series of predictand values 10-12 hours prior to the first strike, graphically captures the predictive value of the GPS lightning model. Figures 13, 14, and 15, depict typical lightning event days. In these cases, the PTV was reached up to ten hours prior to the first strike.

3.6 Categorical Forecasts for the Thunderstorm Season

Forecast verification is needed to test the predictive accuracy of the GPS lightning model. Anytime the predictand value fell below the PTV and up to 90 minutes prior to first strike (meeting 90 minutes desired lead time), it was counted as a *yes* forecast. A Contingency Table (Table 7) is now set up in order to evaluate the GPS lightning model's prediction capabilities (Wilks 1995). For the thunderstorm period (subscript n), there were a total of 46 days evaluated. Twenty-five thunderstorms days were observed and forecasted by the GPS lightning model (quadrant a). Five thunderstorms days were forecasted to occur but did not (quadrant b). Three storms days were observed to occur but the model failed to respond (quadrant c). The 13 remaining days are when the model did not forecast a lightning event and none was observed (quadrant d). The data from these quadrants are now used to determine the accuracy measures for a binary forecast (Table 8). The results of these calculations comprise Table 9. The GPS lightning model proved its' viability, particularly in the area of False Alarm Rates (FAR).

Although the thunderstorm season data not statistically independent, the models applied to these data reduced the FAR to 16.6%. This is a decrease of 26% of KSC's previous FAR. The Probability of Detection (POD) results, were only 8% less than the KSC POD for last season. In making these comparisons it should be noted that the time window of the GPS lightning model is 12.5 hours. The time window associated with KSC forecasts vary with synoptic situation, but on the average vary 4 to 6 hours.

3.7 Categorical Forecasts for the Independent Preseason

Using the same 90-minute desired lead time and PTV criteria, for the independent preseason (subscript i, Table 7), there were a total of 21 days evaluated. Seven thunderstorms days were observed and forecasted by the GPS lightning model (quadrant a). Three thunderstorms days were forecasted to occur but did not (quadrant b). One storm day was observed to occur but the model failed to respond (quadrant c). The 10 remaining days are when the model did not forecast a lightning event and none was observed (quadrant d). The data from these quadrants are again used to determine the accuracy measures for a binary

forecast (Table 8). The results of these calculations comprise Table 10.

GPS lightning model's results for FARs in the independent preseason, were 13.2% lower than KSC results of 43.2% (Table 10). POD was only down by only 10%. These measures could easily be improved by adding forecaster's input of additional knowledge. Reference to satellite and Doppler radar data would give a forecaster the benefit of knowing the tracks and intensities of thunderstorms moving into the area. The GPS lightning model's capability to improve FARs would enhance mission readiness. Mission functions which cease for lightning, would not be delayed by a forecast of lightning that does occur.

Results for the thunderstorm season are slightly better than the preseason. This is attributed to seasonal availability in the moisture of the atmosphere during the summer season. The GPS trends show an increase in the amount of IPW as well as more fluctuations during the thunderstorm season.

As with any attempt to forecast a meteorological event, timing is critical. The lightning model output indicates a potential first-strike when the PTV is met. Timing of this first-strike is now important. Due to the

small sample size of first-strike events the bar chart shows the combination of the two test periods (Fig. 16). A bimodal shape curve can be identified with a wide range of timing profiles (0-12 hours). The average timing profile is between 3-7.5 hours. Since the GPS lightning model results are limited to the model only without benefit of additional data, the human element must now be added. A forecaster using the GPS lightning model now can add additional tools to tailor the timing of this first-strike, knowing that they now only have a 30% chance of a false alarm as opposed to a 43.2% chance.

Using the data from figure 16, the timing range of 12 hrs could be cut down to 7.5 hrs. The GPS lightning model (for the thunderstorm season) would still accurately forecast most of the lightning events while maintaining a 40% FAR. The key to the model's success is not timing of the event, but to alert a forecaster to the possibility of lightning. Armed with the GPS lightning model, a forecaster now knows he has a better chance of properly forecasting the event (meaning a reduction in the possibility of a false alarm). Now other tools such as radar and satellite can be used to time the event.

3.8 Missed Events

During the independent test the first missed event occurred on 19 May 99 (Fig. 17). The GPS lightning model did forecast the lightning event, but only 1 hour prior to first strike, thus not meeting the 90-minute desired lead time. Another event on 13 May 99, is a prime example of a false alarm. On this day, the PTV was met, but a lightning strike never occurred (Fig. 18). In this case, the observations show early morning fog and light winds.

The GPS lightning model was checked to see how it handled rainshowers *with no lightning*, nocturnal events, and back to back events. Figure 19, depicts a day marked by distant nocturnal lightning (>30nm from KSC), morning fog, and afternoon towering cumulus in all quadrants. Although the log value showed fluctuations, the PTV was never met; the lightning model correctly handled this event. For a nocturnal example (Fig. 20), the PTV was met seven hours prior to the first strike around midnight. Finally, the lightning model captured back to back events (Fig. 21). In this case, a nocturnal thunderstorm ends just around midnight. Nine hours later the PTV is met and the first strike follows 4.5 hours later. Another

interesting phenomena that occurred frequently in the GPS lightning model runs, was after the PTV was met, there was sometimes a flatness or increase, in the predictand value, prior to the first strike. This may be attributed to compensating mesoscale subsidence associated with the developing thunderstorm may be drying out the atmosphere above the GPS site.

Chapter 4:

Summary and Conclusions

4.1 Summary

A GPS lightning prediction model is presented that provides a tool for forecasting the Kennedy Space Center's primary weather challenge. After examining a year's worth of operational GPS Integrated Precipitable Water (IPW) data and based on the climatology of lightning occurrence in southern Florida, a thunderstorm season (6/10/99 - 9/26/99) and a preseason (4/14/99 - 6/9/99) were chosen to evaluate the GPS lightning prediction model. A binary logistic regression model was used to identify which of a set of 23 predictors had the most influence in forecasting a lightning event. Four predictors proved important for forecasting lightning events; maximum electric field mill values, GPS IPW, the nine hour change (∇9-hr) of IPW, and KI.

Maximum electric field mill values increased substantially during inclement weather when the potential for lightning existed. Values would increase, sometimes reaching values of 12,000 V/m during a lightning event. But, this variable lacked any long-term (90-minute plus) predictability.

Composites of GPS IPW and ∇9-hr IPW several hours prior to an initial lightning strike show an increase of precipitable water for the site. By using current GPS/IPW and ∇9-hr IPW values, the model not only captures the current IPW of the atmosphere but also the IPW changes in mid-level moisture associated with diurnal and synoptic scale circulations. The KI diagnoses convective activity by examining the moisture in the layer from 850mb to 700mb and the stability of the lapse rate. Given the fact the three of the factors are sensitive to moisture, it important to note that they are not well correlated. The highest correlation coefficient was .47 between IPW and KI.

Using the coefficients for the four best predictors along with a logistic regression equation, a running time series was plotted for the predictand. A common pattern emerged several hours prior to a lightning event. Whenever the Predictand Threshold Value (PTV) was 0.7 or below, lightning was forthcoming. Lightning events were forecasted when the predictand value fell below the PTV. Forecast verification results obtained by using a contingency table revealed a 26.6% decrease during the thunderstorm season, while obtaining, a 13.2% decrease in

false alarm rates for the preseason, compared to the KSC results of 43.2% from the 1998-1999 season. For the KSC, these decreases in false alarm rates mean that missions will not be stopped for a forecast lightning event that does not occur. Additionally, last year the Cape met their desired lead time (90-minute notification) 79.1% of the time. Because the way the forecast verification was set up with reference to the 90-minute desired lead time criteria, Probability of Detection results from the thunderstorm test period (89.2%) and the preseason (87.5%) also reflect the desired lead time statistic. In this research, if a storm failed to meet the desired lead it was counted as a missed event. Not only does the GPS lightning model improve the false alarm rate, but it also improves the previous lead-time at KSC by 10%.

Nevertheless, if the GPS Lightning Model is integrated as a forecasting tool to assist a forecaster along with other tools based on satellite and Doppler radar data, forecasting lightning events at the Kennedy Space Center will likely improve.

Future work will consist of testing this model on data from the 2000 thunderstorm season. Local observations suggest there is a relationship between increases of GPS

IPW and precipitation. This relationship needs to be investigated. The success of this model suggest opportunities to investigate the development of statistical models that target related weather phenomena, for e3xample heavy rain events.

APPENDIX A: TABLES

Table 1: Initial Predictors

Predictors	Source of data
1. GPS IPW	GPS Sensor
2-13. ∇* IPW from 1 to 12 hrs	GPS Sensor
14. 30 min Field Mill Averages	Electric Field Mill
15. 30 min Field Mill Maximums	Electric Field Mill
16. Temperature	GPS Sensor
17. Pressure	GPS Sensor
18. Dewpoint Temperature	Surface Observation
19. Total Totals	RAOB
20. Freezing Level	RAOB
21. Wind Direction	Surface Observation
22. K Index	RAOB
23. 700mb Vertical Velocity	ETA Model

*∇ means change

Table 2: Logistic Regression Table

Predictor	Coefficient	Standard Deviation	Z-Value	p-Value	Odds Ratio
Constant	-6.7866	.7208	-9.42	.000	
Max V/m	.0011359	.0002923	3.89	.000	1.00
IPW	.06063	.01467	4.13	.000	1.06
∇9 IPW	.32341	.02961	10.92	.000	1.38
K-Index	.06728	.02081	3.23	.001	1.07

Table 3: Time of occurrence of thunderstorms from Cape Canaveral. The table combines the first-strike times from both test periods.

HOURS (LST)	Number Lightning Events	Percent
00-02	0	0
03-05	1	2.7
06-08	1	2.7
09-11	2	5.6
12-14	6	16.7
15-17	12	33.3
18-20	10	27.8
21-23	4	11.1

Table 4: Goodness of Fit Tests

Method	p-Value
Pearson Residual	1.000
Deviance Residual	1.000
Hosmer-Lemeshow	.605

Table 5: Observed and Expected Frequencies

	Group										
Value	1	2	3	4	5	6	7	8	9	10	Total
Yes											
Obs	0	0	2	6	16	15	26	38	42	76	221
Exp	.3	1.1	2.7	6.2	11.2	17.5	25.8	35.2	47.6	73.3	
No											
Obs	99	100	97	94	83	85	73	62	57	24	774
Exp	98.7	98.9	96.3	93.8	87.8	82.5	73.2	64.8	51.4	26.7	
Total	99	100	99	100	99	100	99	100	99	100	995

Table 6: Measures of Association

Pairs	Number	Percent
Concordant	146424	85.6%
Discordant	24299	14.2%
Ties	331	.2%
Total	171054	100%

Table 7: categorical Forecast of Discrete Predictands

Contingency Table:

		Observed	
		Yes	No
Model Forecast	Yes	$a_n = 25$ $a_i = 7$	$b_n = 5$ $b_i = 3$
	No	$c_n = 3$ $c_i = 1$	$d_n = 13$ $d_i = 10$

Relationship between counts (letters a-d) of forecasts event pairs for the dichotomous categorical verification. Quadrant *a* denotes the occasions when the lightning was forecast to occur and did. Quadrant *b* denotes the occasions when the lightning was forecast to occur but did not. Quadrant *c* denotes the occasions when the lightning was not forecast to occur but did. Quadrant *d* denotes the occasions when the lightning was not forecasted to occur and did not. Subscripts n and i indicate non-independent test (thunderstorm period) and independent test (preseason).

Table 8: Accuracy Measures for Binary Forecast

False Alarm Rate: Portion of forecast events which fail to materialize.

$$FAR = \frac{b}{a + b}$$

Probability of Detection: Forecast event occurred When it was also forecasted.

$$POD = \frac{a}{a + c}$$

Table 9 Thunderstorm Season Test Results

(10 Jun to 26 Sep 99)

	GPS Lightning Results	KSC Results From 1999 Season
False Alarm Rate:	16.6%	43.2%
Hit Rate:	82.6%	N/A
Threat Score:	75.6%	N/A
Probability of Detection:	89.2%	97.5%

Table 10 Independent Test Case Results (Preseason, 14 Apr 99 to 9 Jun 99).

	GPS Lightning Results	KSC Results From 1999 Season
False Alarm Rate:	30%	43.2%
Hit Rate:	80.9%	N/A
Threat Score:	63.6%	N/A
Probability of Detection:	87.5%	97.5%

APPENDIX B: FIGURES

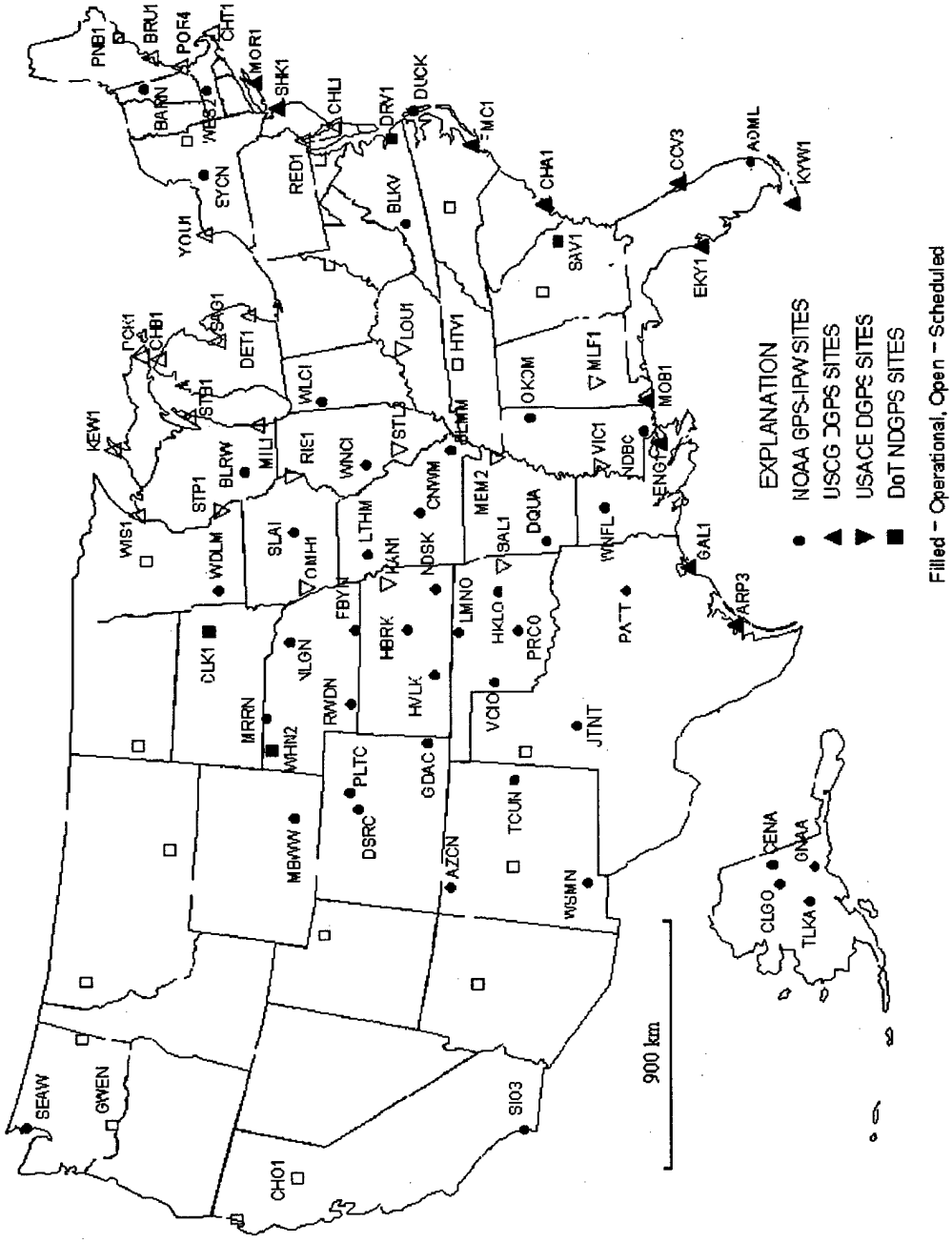

Figure 1: Current GPS Site Configuration 1999. Stations ion Hawaii, Puerto Rico, and the U.S. Virgin Islands are omitted from this illustration.Figure courtesy of Seth Gutman.

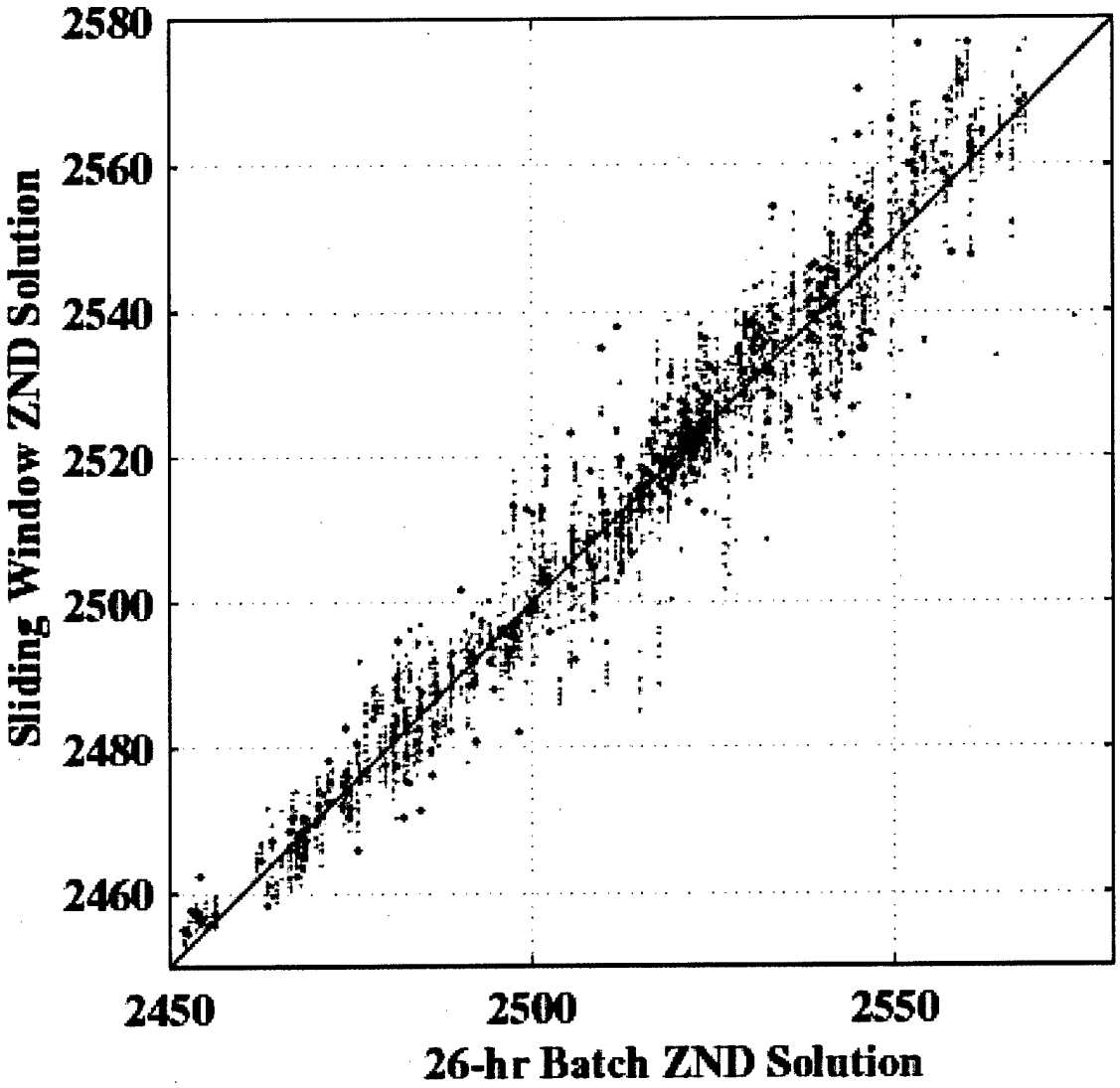

Figure 2: Shows a comparison of GPS zenith neutral delay (ZND)data processed using a sliding window approach and a standard batch approach. A13 mm ZND ~ 2 mm error in precipitable water. The red dots show the real time solution, blue and gray dots show remaining solutions for the window. Figure courtesy of James Foster (personal communication).

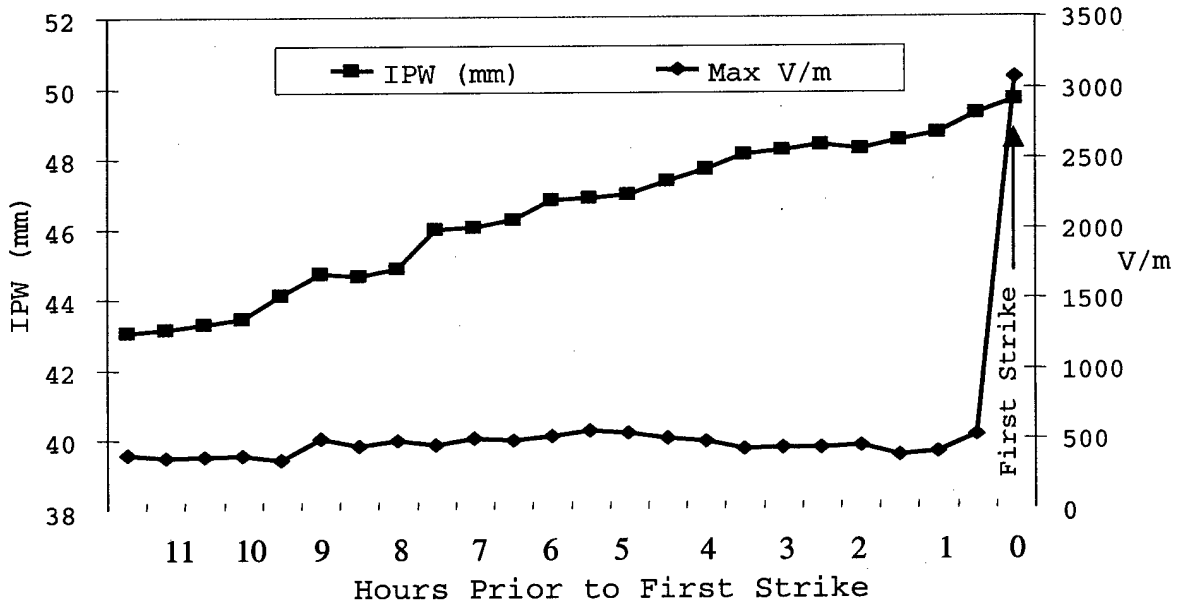

Figure 3: Average time series of GPS IPW and maximum electric field mills values for the hours leading up to a lightning strike.

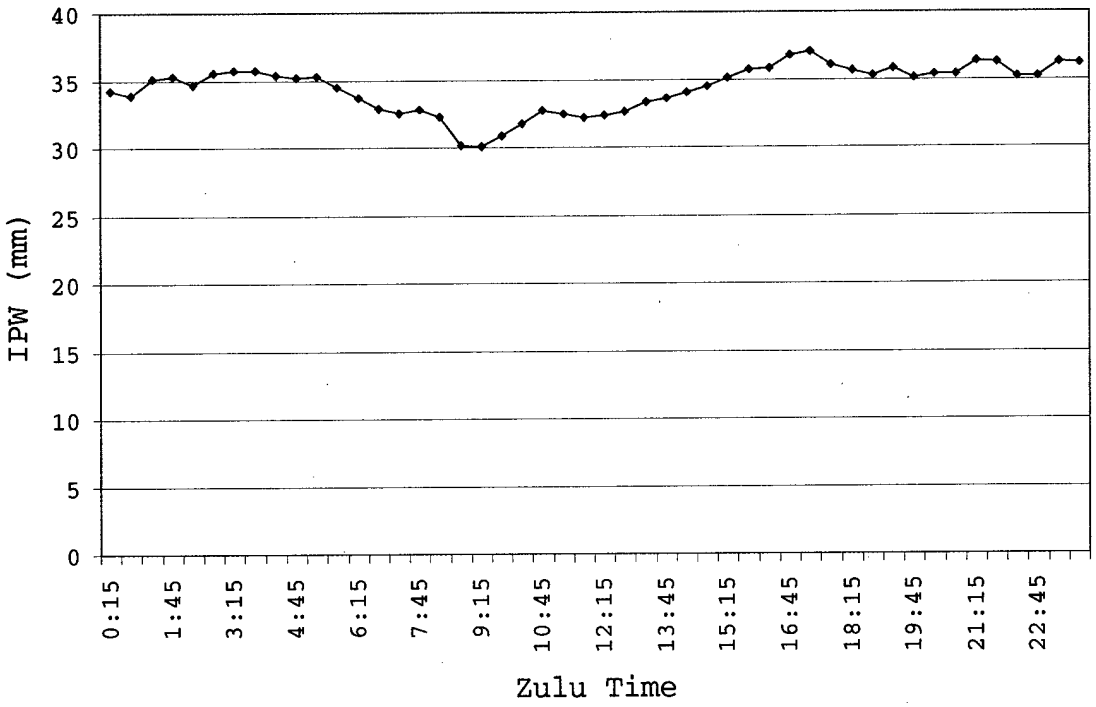

Figure 4: GPS IPW 24-hour average time series for non-event weather days.

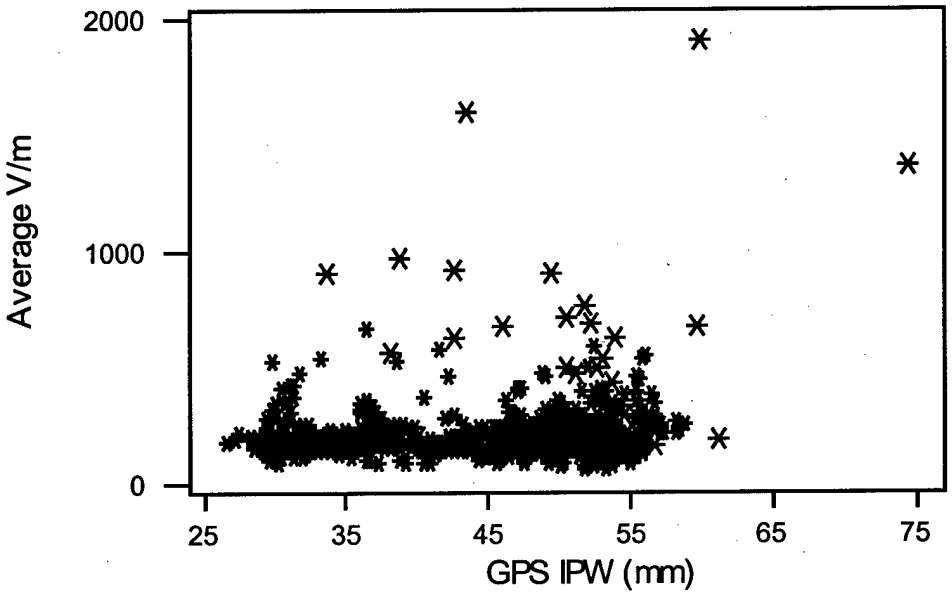

Figure 5: Scatter plot of average electric field mill values and GPS IPW. Red asterisks denotes a *yes* for a lightning flash. Black asterisks denotes *no* lightning flash detected.

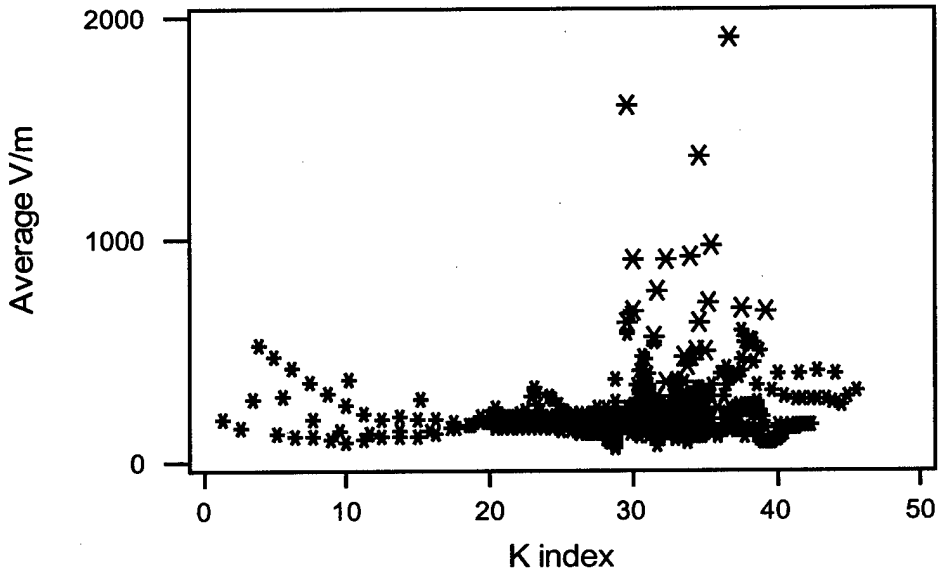

Figure 6: Scatter plot of average electric field mill values and K index. Red asterisks denotes a *yes* for a lightning flash. Black asterisks denotes *no* lightning flash detected.

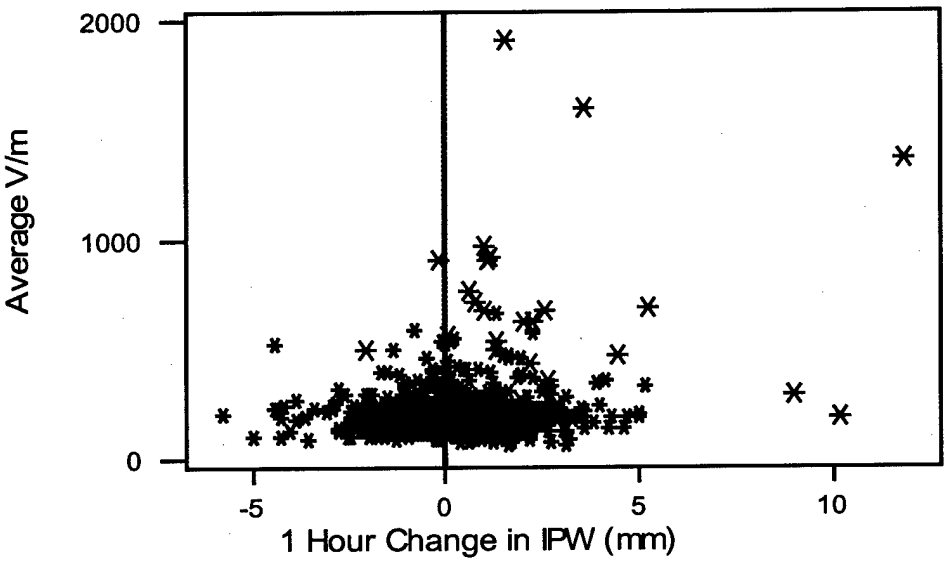

Figure 7: Scatter plot of average electric field mill values and ∇1 GPS IPW. Red asterisks denotes a *yes* for a lightning flash. Black asterisks denotes *no* lightning flash detected.

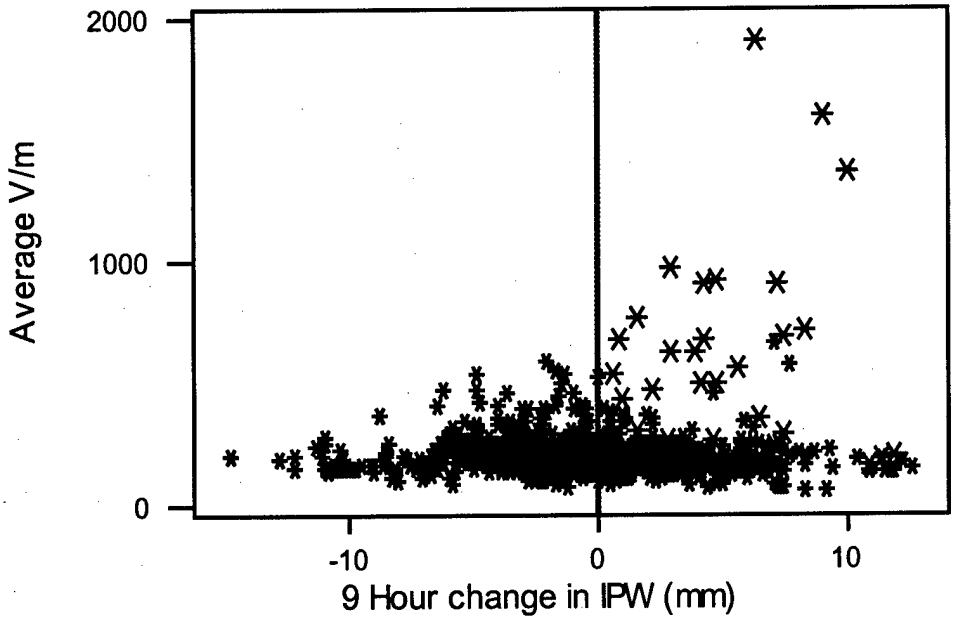

Figure 8: Scatter plot of average electric field mill values and ∇9 GPS IPW. Red asterisks denotes a *yes* for a lightning flash. Black asterisks denotes *no* lightning flash detected

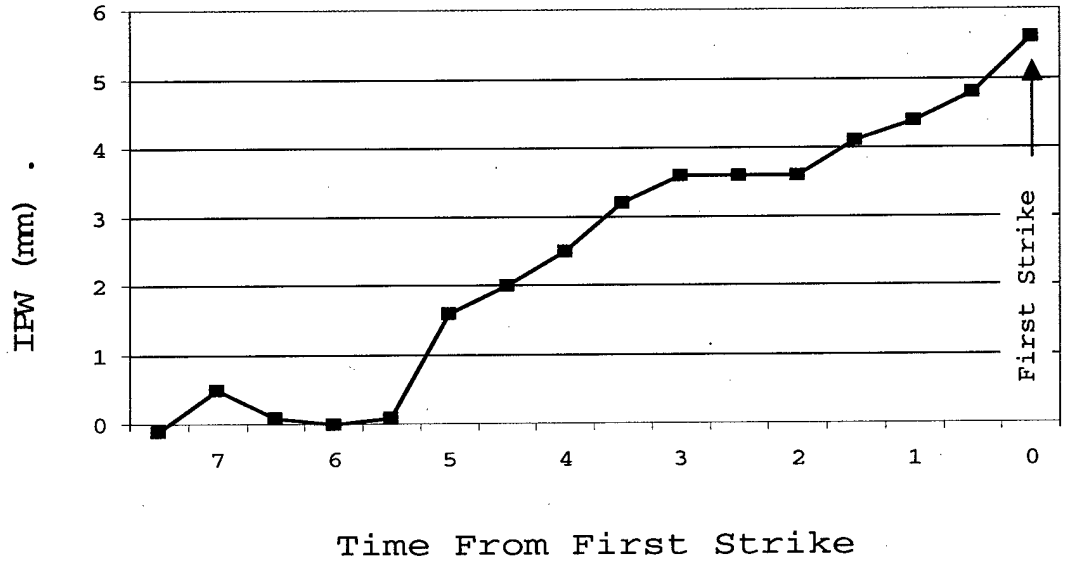

Figure 9: Composite time series of the average change in ∇9-hr GPS IPW from one hour to the next, using the superposed epoch method with the time of the first strike defined as the zero hour.

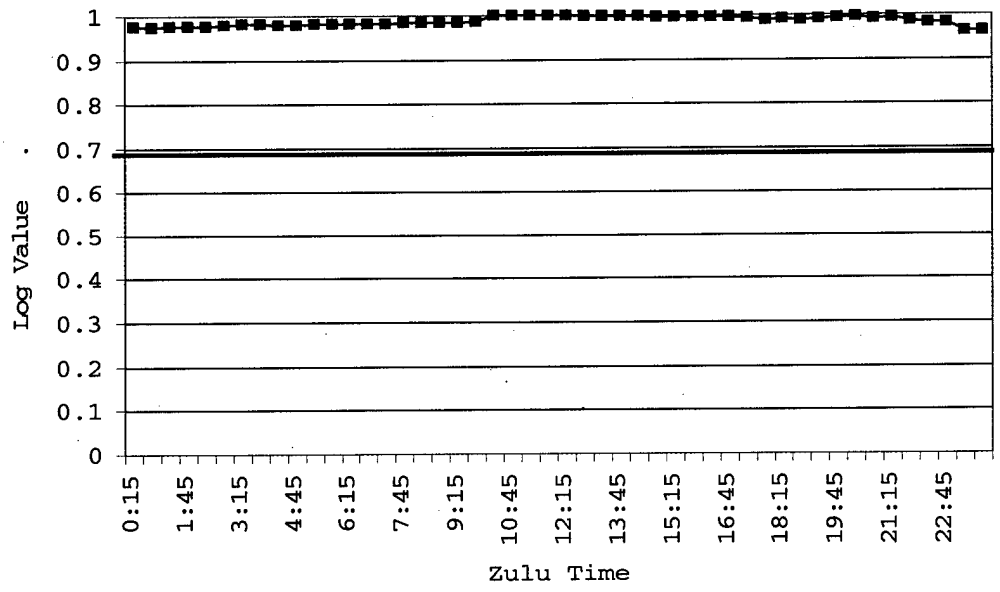

Figure 10: GPS Lightning Model run for a non-event weather day, 5 July 1999.

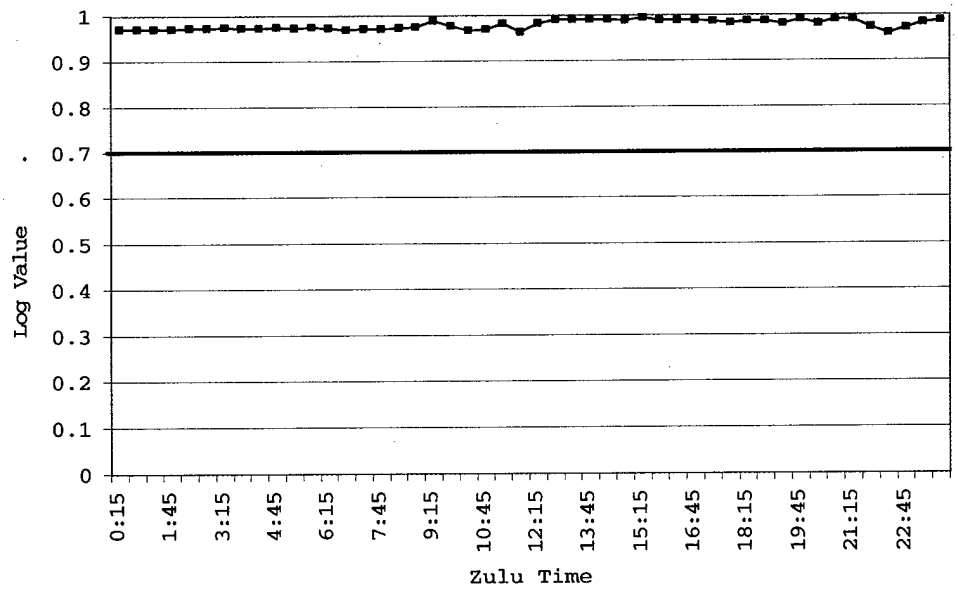

Figure 11: GPS Lightning Model run for a non-event weather day, 15 July 1999.

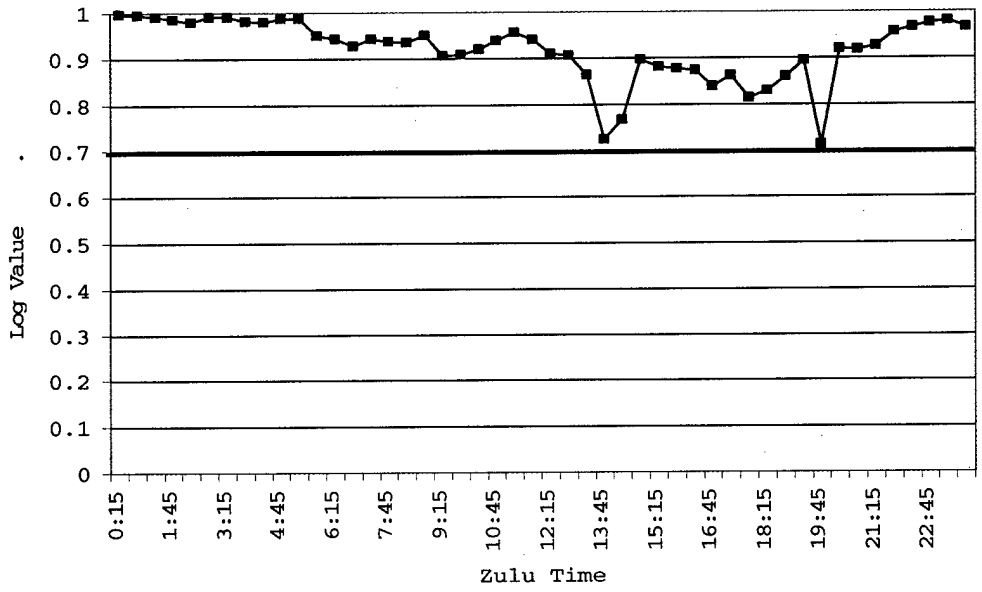

Figure 12: GPS Lightning Model run for a non-event weather day, 18 July 99.

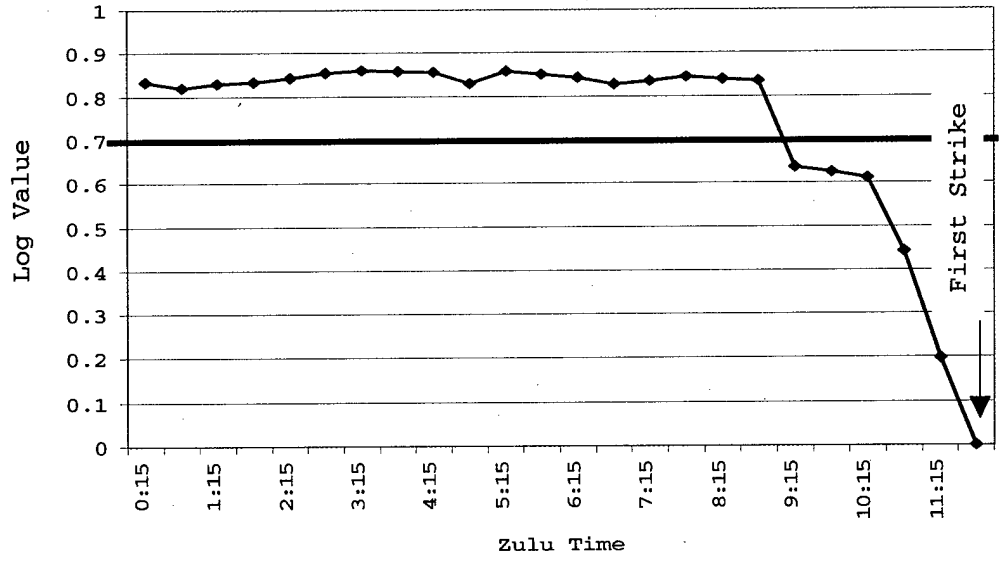

Figure 13: GPS Lightning Model for run a lightning event, 3 July 99.

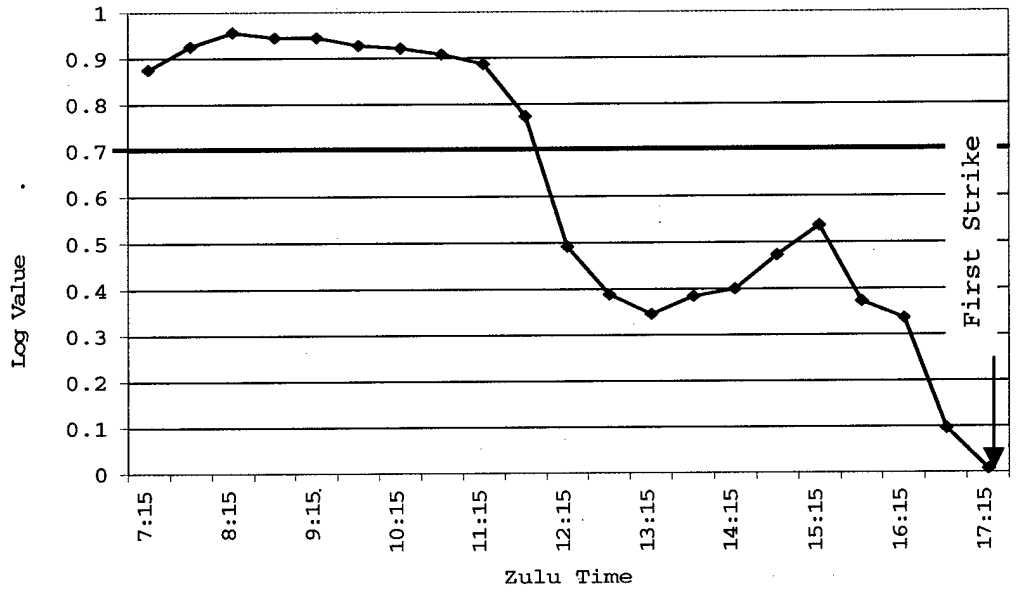

Figure 14: GPS Lightning Model for run a lightning event, 10 Jul 1999.

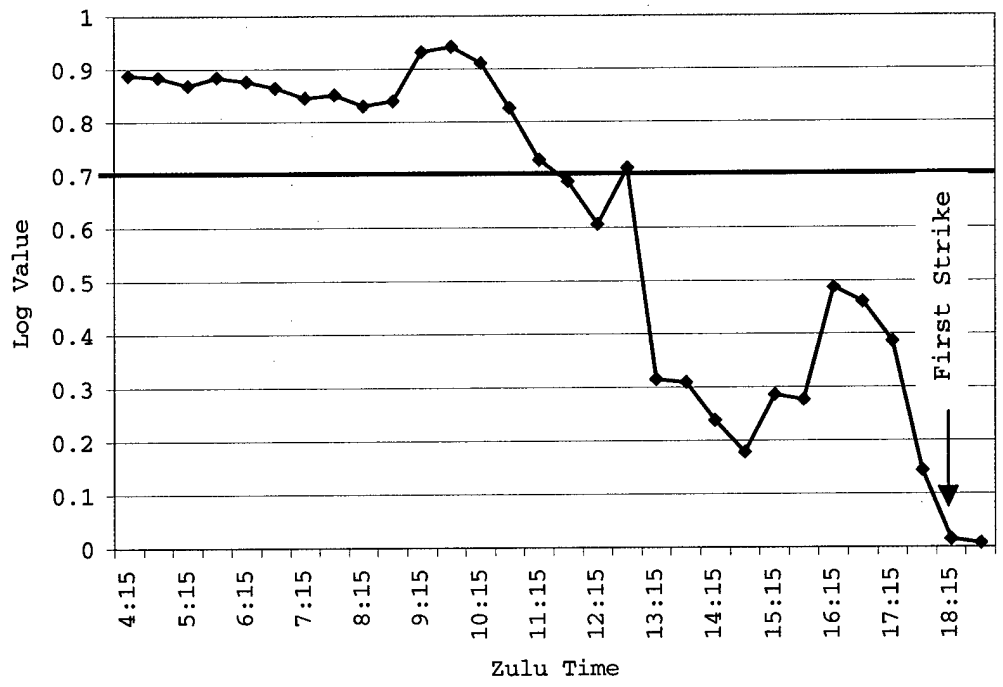

Figure 15: GPS Lightning Model for run a lightning event, 1 Aug 1999.

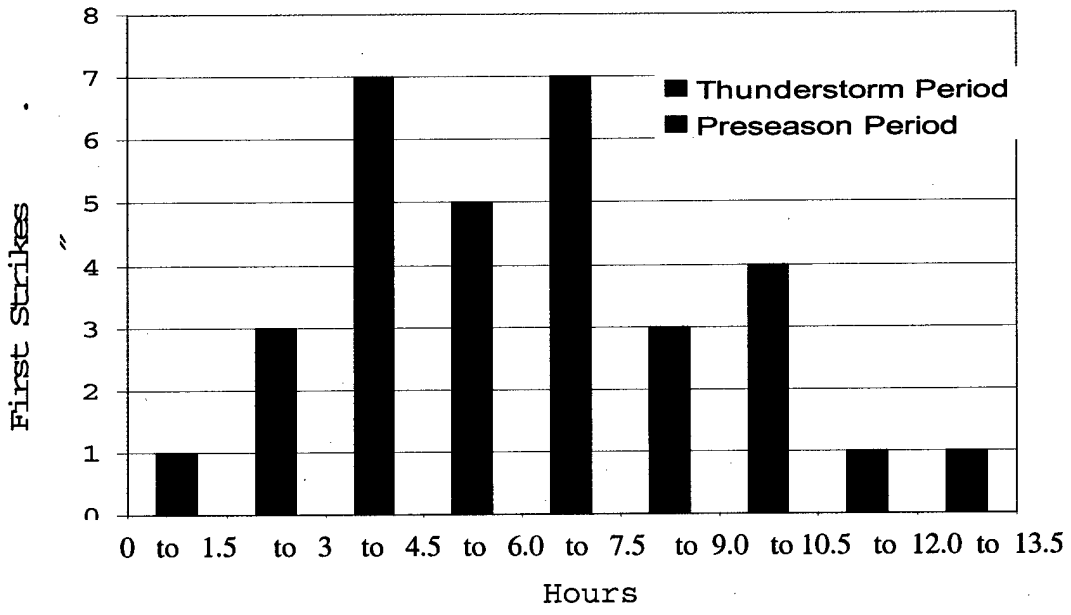

Figure 16: Time in hours from first-strike when LTV was met for all thunderstorms.

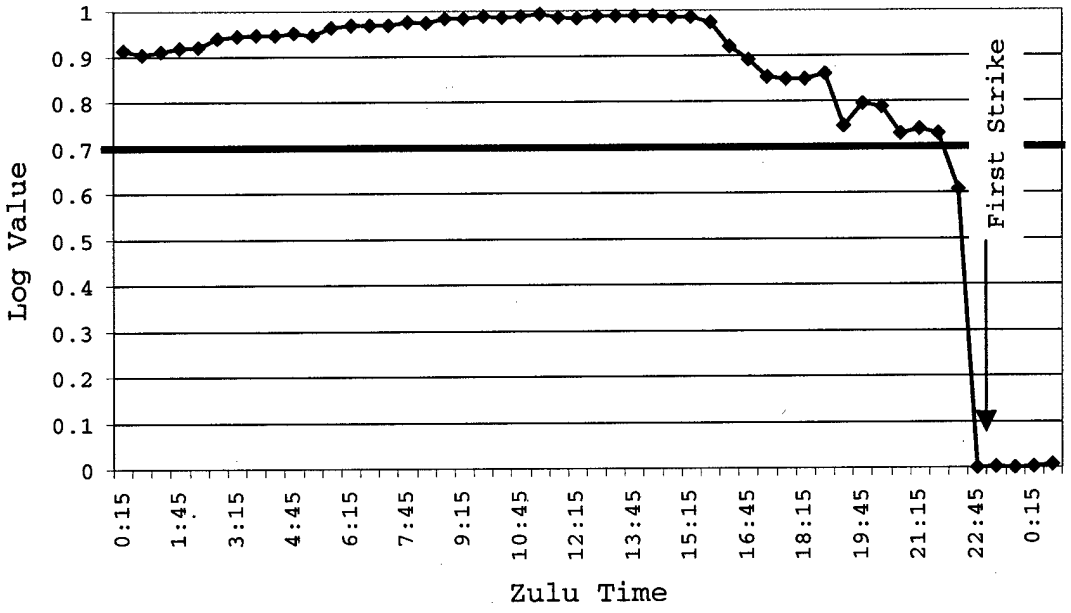

Figure 17: GPS lightning model run for a lightning event in where the desired led time was not met, 19 May 99.

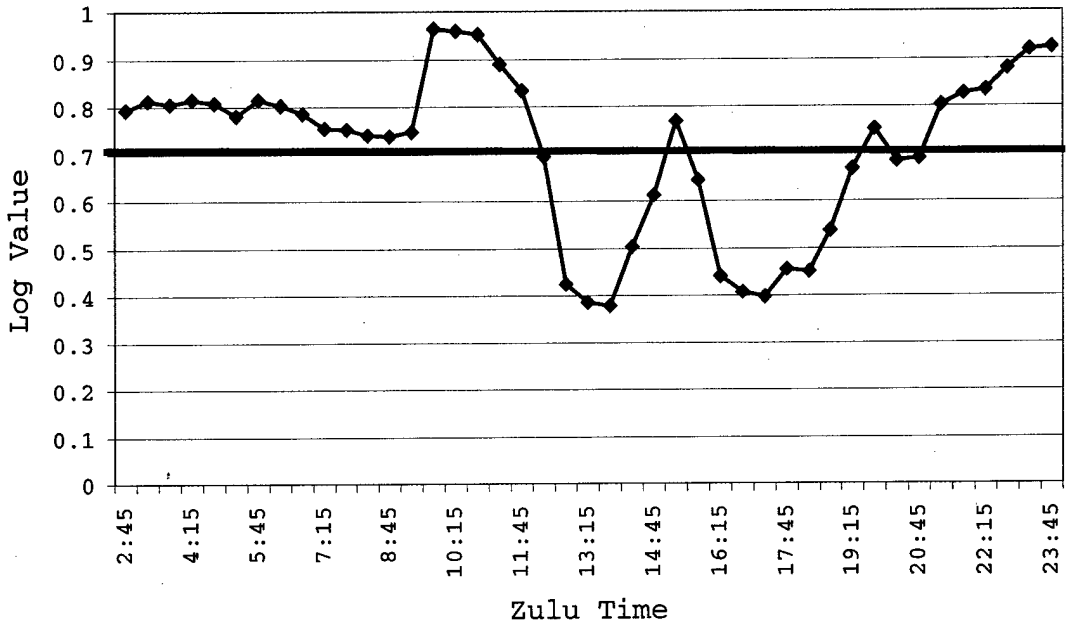

Figure 18: GPS lightning model run for a false alarm event, 13 May 99.

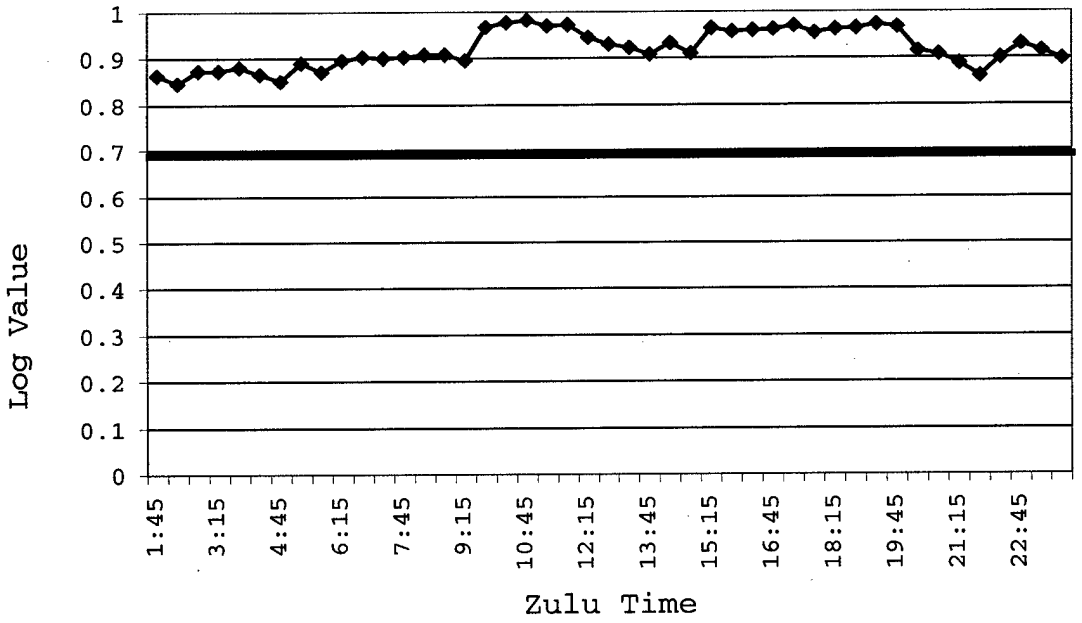

Figure 19: GPS lightning model run for a no lightning event day, 20 May 99.

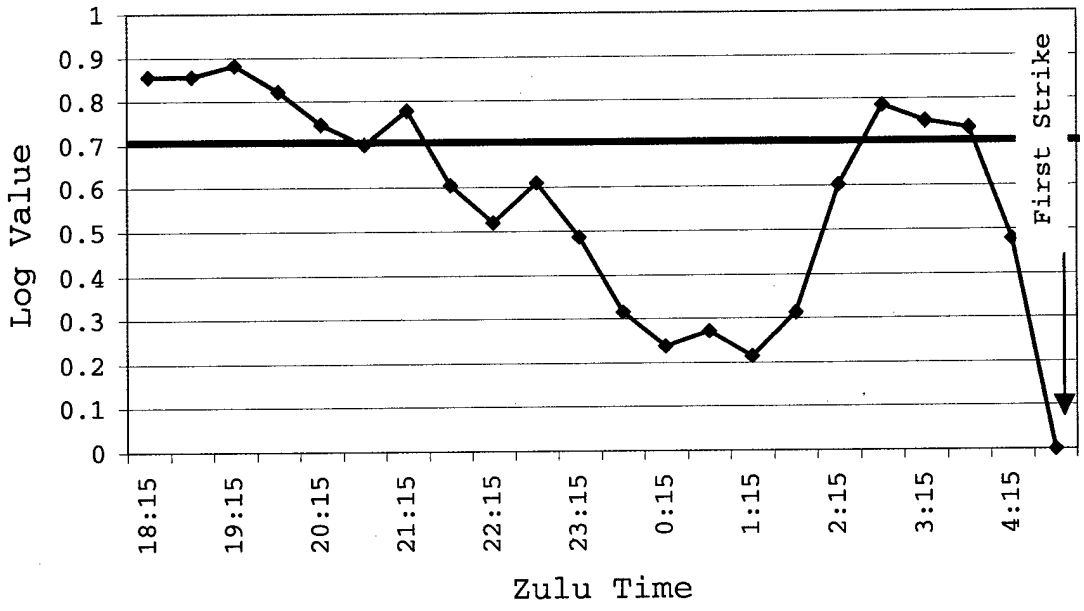

Figure 20: GPS lightning model run for a nocturnal lightning event, 21-22 Jul 99.

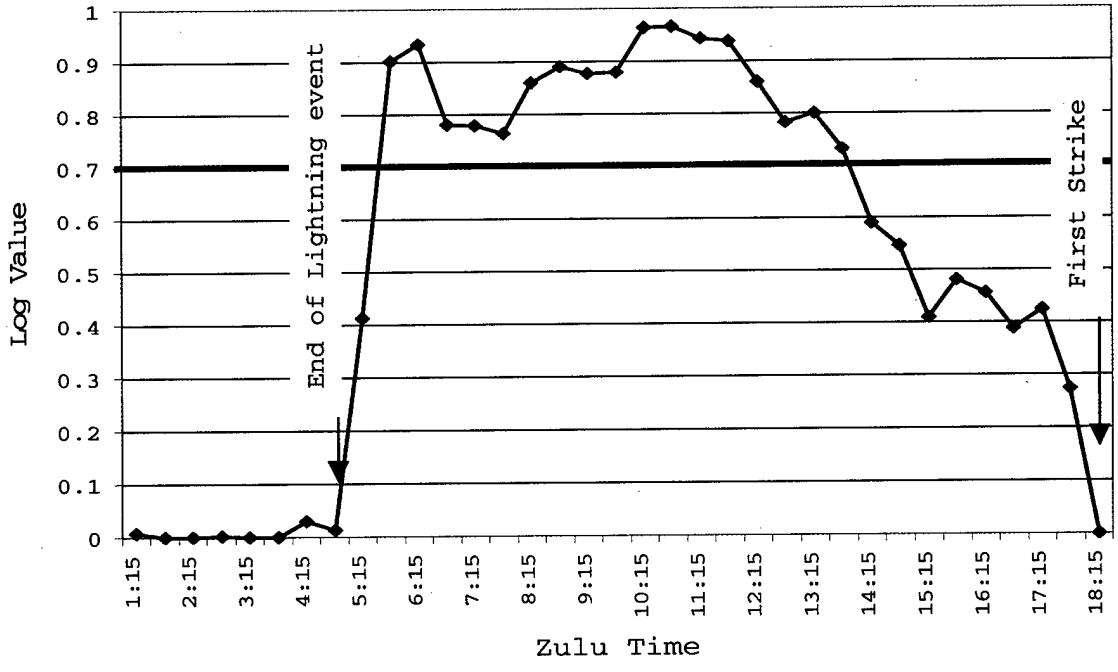

Figure 21: GPS lightning model run for back to back lightning events, 9 Aug 99.

References

Arritt, R. W., 1993: Effects of large-scale flow on characteristic features of the sea-breeze. *J. Appl. Meteor.*, **32**, 116-125.

Bauman, W. H., and S. Businger, 1996: Nowcasting for Space Shuttle landings at Kennedy Space Center, Florida. *Bull. Amer. Meteor. Soc.*, **77**, 2295-2305.

_____, W. H., S. Businger, and M.L. Kaplan, 1997: Nowcasting Convective Activity for Space Shuttle Landings during Easterly Flow Regimes. *Wea. Forecasting*, **12**, 78-107.

Bevis, M., S. Businger, T. A. Herring, C. Rocken,S. R. A. Anthes, and R. H. Ware, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *J. Geophys. Res.*, **97**, 15,787-15,801.

_____,_____, S. Chiswell, T. A. Herring, R. A. Anthes, C. Rocken, and R. H. Ware, 1994: GPS meteorology: Mapping zenith wet delay onto precipitable water. *J. Appl. Meteor.*, **33**, 379-386.

Blanchard, D. O., and R. E. Lopez, 1985: Spatial patterns of convection in south Florida. *Mon. Wea. Rev.*,**113**, 1282-1299.

Boybeyi, Z., and S. Raman, 1992: A three-dimensional numerical sensitivity study of convection over the Florida peninsula. *Bound.-Layer Meteor.*, **60**, 325-359.

Businger, S., M. Bevis, S. Chiswell, M. Bevis, J. Duan, R. Anthes, and R. Ware, C. Rocken, M. Exner, T. VanHove, and F. Solheim, 1996: The Promise of GPS in atmospheric monitoring. *Bull. Amer. Meteor. Soc.*, **77**, 5-17.

Byers, H., and H. R. Rodebush, 1948: Causes of thunderstorms of the Florida Peninsula. *J. Meteor.*, **5**, 275-280.

Davis, J. L.,T. A. Herring, I. I. Shapiro, A. E. Rodgers, and G. Elgered, 1985: Geodesy by radio interferomtry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio. Sci.*, **20**, 1593-1607.

Duan, J., S. Businger, M. Bevis, S. Chiswell, M. Bevis, J. Duan, R. Ware, C. Rocken, P. Fang, Y. Bock, T. VanHove, F. Solheim S. McClusky, T. Herring, and R. King, 1996: GPS Meteorology: Direct estimation of the absolute value of perceptible water. *J. Appl. Meteor.*, **35**, 830-838.

Elgered,G.,J.L. Davis, T.A. Herring, and I.I. Shapiro, 1991: Geodesy by radio interferometry: Water vapor radiometry for estimation of wet delay. *J. Geophys. Res.*, **96**, 6541-6555.

Estoque, M. A.,1962: The sea breeze as a function of the prevailing synoptic situation. *J. Atmos. Sci.*, **19**, 24-25.

Fleagle, R. G., and J. A. Businger, 1980: *An Introduction to Atmospherics Physics*. Academic Press, New York. pp 432.

Foote, G. B., 1991: Scientific overview and operations plan for the convection and precipitation/ electrification program. National Center for Atmospheric Research, Boulder, CO, 145 pp.

Harms, D., B. Boyd, Rlucci, M Hinson, and M Maier, 1997: Systems used to evaluate the natural and triggered lightning threat to the eastern range and Kennedy Space Center. *28th Conference on Radar Meteorology*, 240-241

Hazen, D. S., W. P. Roeder, B. F. Lorens, and T. L. Wilde, 1995: Weather impacts on launch operations at the Eastern Range and Kennedy Space Center. Preprints, *Sixth Conf. On Aviation Weather Syatems*, Dallas, TX, Amer. Meteor. Soc., 270-275.

Hosmer, D. W., and S. Lemeshow, 1989: *Applied Logistic Regression*. Wiley, New York, 307 pp.

International Station Meteorological Climate Summary, 1996: Federal climate Complex, Ashville, North Caroline, Ver 4.0, September.

Marshall, T. C., W. D. Rust, M. Stolzenburg, W.P. Roeder,and P. R. Krehbiel, 1999: A study of enhanced fair-weather electric field occurring soon after sunrise. *J. Geophys. Res.*, **104,** 24455-24469.

Neumann, C. J., 1971: The thunderstorm forecasting system at the Kennedy Space Center. *J. Appl. Meteor.*, **10**, 921-936.

Panofsky, H. A., and G. W. Brier, 1958: *Some Applications of Statistics to Meteorology*. University Park, Pennsylvania, 224 pp.

Pielke, R., 1974: A three-dimensional numerical model of the sea-breeze over south Florida. *Mon. Wea, Rev.*, **102**, 115-139.

Reap, R. M., 1994: Analysis and prediction of lightning strike distributions associated with synoptic map types over Florida. *Mon. Wea. Rev.*, **122**, 1698-1715.

Spilker, J. J., 1980: Signal structure and performance Characteristics. *Global Positioning System,* Vol.1, The institute of navigation, 246 pp.

Tralli, D.M., and S.M.Lichten,1990:Stochastic estimation of tropospheric path delays in Global Positioning System geodetic measurements. *Bull.Geod.*, **64**, 127-159

Watson, A. I., R.L. Holle, R. E. Lopez, R. Ortiz, and J. R. Nicholson, 1991: Surface eind convergence as a short term predictor of cloud-to-ground lightning at Kennedy Space Center. *Wea. Forecasting*, **6**, 49-64.

Wilks, D., 1995: *Statistical Methods in Atmospheric Sciences*. Academic Press, 464 pp.

Wolfe, D. E., S. I. Gutman, R.B. Chadwick, P. Fang, and Y. Bock, 1998: Development of an operational surface-based GPS water vapor observing system for NOAA: Network design and project status. *2nd Symp. on Int. Obs. Sys.*, AMS, Phoenix.