AN EXAMINATION OF THE RELATIONSHIP BETWEEN LIGHTNING AND RAINFALL OVER OCEANIC REGIONS USING THE GLOBAL LIGHTNING DATASET (GLD360) AND THE TROPICAL RAINFALL MEASURING MISSION PRECIPITATION RADAR (TRMM PR)

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

MAY 2012

By

Douglas C. Stolz

Thesis Committee:

Steven Businger, Chairperson
John N. Porter
Gary Barnes

ACKNOWLEDGEMENTS

I am indebted to Dr. Steven Businger for his patience and tutelage during my time as a student of meteorology at the University of Hawai‘i. His undying passion for many facets of meteorology incites a strong desire to learn in those who study under his supervision. I'd like to thank Dr. John Porter for helping me during my attempt to demystify the microphysical processes that contribute to the generation of lightning – the time was well spent probing the boundaries of our current understanding. I would like to thank Dr. Gary Barnes for encouraging me to reach my full potential in producing a high quality product. Thank you to all my committee members; I am grateful for the huge investments of time and energy made to help me achieve this end.

I had the privilege of working beside Annick Terpstra for the early portion of this project. Without her technological cunning and ever-curious, brilliant mind, I would have spent many more hours pouring over the finer points of computer programming. I'd like to recognize Tom Robinson and Aaron Levine for helping me develop my own abilities in scientific computing through the years. Thank you to Marco Formenton for sharing his expertise on cloud electrification processes. To my fellow students, working and learning beside each you has been an invaluable experience. I advise you to keep questioning in the face of uncertainty.

My sincerest thank you goes to my parents, Thomas and Joan, and two brothers, Nick and Shane, for their endless support. Each of their unique perspectives helped me through the hardest of times.

ABSTRACT

Data from a precipitation radar (PR) on board NASA's Tropical Rainfall Measuring Mission (TRMM) satellite and data from Vaisala's network of ground-based, long-range lightning detection sensors (Global Lightning Dataset, GLD360) are used to examine the relationship between convective precipitation and lightning over the ocean. Data from May 2011 to February 2012 are analyzed across the North Atlantic Ocean, the Northeast Pacific Ocean, and the Northwest Pacific Ocean. The stroke density and maximum current of lightning strokes detected by GLD360 are compared to reflectivity observed by the TRMM PR.

When the reflectivity is binned according to the stroke density and maximum current, the relationship between lightning and rainfall is found to be log-normal. Lightning occurring over each of the three ocean basins is associated with maximum reflectivity near 2 km altitude within a narrow range of values (45 – 50 dBZ). The reflectivity associated with stroke density and maximum current was shown to be 1 – 4 dBZ higher in winter than in summer. The maximum reflectivity above the freezing level is found to be larger for larger stroke density.

TABLE OF CONTENTS

Acknowledgments …… ii

Abstract ……..... iii

List of Tables ……... v

List of Figures …….. vi

List of Abbreviations …… ix

Chapter 1: Introduction ……. 1

1.1 History and Motivation …… 1

1.2 Background …….. 2
1.2.1 Cloud electrification and the potential influence of aerosols … 2
1.2.2 Atmospheric propagation …….. 8
1.2.3 Lightning detection and location ……... 9

Chapter 2: Data and Methodology ……. 13

2.1. Data……..... 13
2.1.1. The TRMM precipitation radar ……... 13
2.1.2. The Global Lightning Dataset……..... 16

2.2 Methodology ……... 20

2.2.1 Domains of interest and procedure ……..... 20
2.2.2. Sensitivity tests ……... 24

Chapter 3: Results ……..... 27

3.1 Sensitivity to time and space parameters …… 28

3.2 Lightning and rainfall over the ocean ……. 30
3.2.1 Stroke density and maximum reflectivity …… 30
3.2.2 Maximum current and maximum reflectivity …….. 32
3.2.3 Stroke density and average reflectivity …… 33
3.2.4 Maximum current and average reflectivity …….. 34

3.3 Vertical distribution of maximum reflectivity ……..... 35

Chapter 4: Conclusions, Discussion, and Future Work ……..... 38

Tables ……..... 47

Figures ……... 50

References ……. 75

LIST OF TABLES

Table Page

i. Available performance statistics for long-range lightning detection systems that currently operate 47

ii. Goodness-of-fit statistics for least-squares logarithmic fits for stroke density versus maximum reflectivity 47

iii. Goodness-of-fit statistics for least-squares logarithmic fits for maximum current versus maximum reflectivity 47

iv. Logarithmic fits by basin for stroke density versus maximum reflectivity 48

v. Logarithmic fits by basin for maximum current versus maximum reflectivity ... 48

vi. Goodness-of-fit statistics for least-square logarithmic fits for stroke density versus average reflectivity 48

vii. Goodness-of-fit statistics for least-square logarithmic fits for maximum current versus average reflectivity 48

viii. Logarithmic fits by basin for stroke density versus average near-surface reflectivity 49

ix. Logarithmic fits by basin for maximum current versus average reflectivity 49

LIST OF FIGURES

Figure Page

1. Schematic of resulting charge centers in a typical thunderstorm after gravitation separation has occurred. 50
2. "Schematic diagram of the earth-ionosphere waveguide and VLF signal propagation (*Pessi et al. 2009*). 50
3. Schematic diagrams illustrating methods used to locate lightning strokes (*Cummins et al. 2000*) 51
4. A comparison of the Kwajalein Doppler radar and TRMM PR radar data (*Schumcher and Houze 2000*) 52
5. GLD360 estimated stroke DE across the globe for various times (*Vaisala*) 53
6. GLD360's DE as a function of the time of day 54
7. The numbers of observed (matched/unmatched) flashes as a function of maximum current for both NLDN and GLD360 54
8. Climatology of lightning observed a) by NASA's optical transient detector (*Christian et al. 2003*) and b) by GLD360 55
9. GOES infrared imagery with GLD360 flashes overlain (green) a) before reprocessing and b) after reprocessing (*Vaisala*) 56
10. GLD360's LA a) before reprocessing and b) after reprocessing (*Vaisala*) 57
11. The three oceanic domains studied 58
12. A schematic diagram of the procedure for this study 58
13. Mean GLD360 stroke density (+/- σ) as a function of observation time interval and grid box size 58
14. a) GLD360 (May 2011 – August 2011) CG stroke density versus near-surface reflectivity and b) cumulative distribution functions for four different times 59
15. The maximum current amplitude observed by GLD360 versus maximum reflectivity a) before and b) after a CAT has been applied 60
16. The stroke density observed by GLD360 versus maximum reflectivity a) before and b) after a CAT has been applied 61
17. GLD360 stroke density versus TRMM PR maximum reflectivity (NATL) ….. 62
18. GLD360 stroke density versus TRMM PR maximum reflectivity (NWPAC) ... 62

19. GLD360 stroke density versus TRMM PR maximum reflectivity (NEPAC) 63
20. Basin comparison of GLD360 stroke density versus TRMM PR maximum reflectivity 63
21. Basin comparison of GLD360 stroke density versus TRMM PR maximum reflectivity for winter and summer 64
22. GLD360 maximum current versus TRMM PR maximum reflectivity (NATL) 64
23. GLD360 maximum current versus TRMM PR maximum reflectivity (NWPAC) 65
24. GLD360 maximum current versus TRMM PR maximum reflectivity (NEPAC) 65
25. Basin comparison of GLD360 maximum current versus TRMM PR maximum reflectivity 66
26. Basin comparison of GLD360 maximum current versus TRMM PR maximum reflectivity for winter and summer 66
27. GLD360 stroke density versus TRMM PR average reflectivity (NATL) 67
28. GLD360 stroke density versus TRMM PR average reflectivity (NWPAC) 67
29. GLD360 stroke density versus TRMM PR average reflectivity (NEPAC) 68
30. Basin comparison of GLD360 stroke density versus TRMM PR average reflectivity 68
31. Basin comparison of GLD360 stroke density versus TRMM PR average reflectivity for winter and summer 69
32. GLD360 maximum current versus TRMM PR average reflectivity (NATL) 69
33. GLD360 maximum current versus TRMM PR average reflectivity (NWPAC) 70
34. GLD360 maximum current versus TRMM PR average reflectivity (NEPAC) 70
35. Basin comparison of GLD360 maximum current versus TRMM PR average reflectivity 71

36. Basin comparison of GLD360 maximum current versus TRMM PR average reflectivity for winter and summer 71
37. Vertical distributions of maximum reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NATL 72
38. Vertical distributions of maximum reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NWPAC 73
39. Vertical distributions of maximum reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NEPAC 74

LIST OF ABBREVIATIONS

Global Lightning Dataset (GLD360)

Tropical Rainfall Measuring Mission (TRMM PR)

Long-range Lightning Detection Systems (LLS)

Convective Area Threshold (CAT)

Very High Frequency (VHF)

Very Low Frequency (VLF)

Magnetic Direction Finding (MDF)

Time-of-Arrival (TOA)

Cloud-to-Ground (CG)

Intra-cloud (IC)

Northwest Pacific Ocean domain (NWPAC)

Northeast Pacific Ocean domain (NEPAC)

North Atlantic Ocean domain (NATL)

Cloud Condensation Nucleus (CCN)

Detection Efficiency (DE)

Location Accuracy (LA)

WorldWide Lightning Location Network (WWLLN)

National Lightning Detection Network (NLDN)

Weather Systems Incorporated Global Lightning Network (WSI GLN)

CHAPTER I: INTRODUCTION

1.1 History and Motivation

In late 2009, Vaisala unveiled an array of ground-based electromagnetic sensors capable of detecting lightning at long ranges (~4000 km). This advanced long-range lightning detection system (LLS) was named the Global Lightning Dataset (GLD360). The GLD360's performance in continuous, long-range lightning detection is unparalleled by any other LLS currently operating (Table 1). Therefore, GLD360 permits a more extensive investigation of lightning and rainfall over historically data-sparse oceanic regions. Here GLD360 is used in conjunction with the Tropical Rainfall Measuring Mission's precipitation radar (TRMM PR) to study lightning and rainfall over three ocean basins in the northern hemisphere.

Nearly thirty years ago, people attempted to correlate lightning with convective activity using prototype networks of ground-based electromagnetic sensors along the east coast of the United States (Orville et al. 1983, Orville 2008). Since that time, people have examined the relationships between lightning and various indicators of convective intensity in many different regions across the world (Zipser and Lutz 1994; Petersen et al. 1996; Cheze and Sauvageot 1997; Toracinta et al. 2002; Cecil et al. 2005). Lightning and rainfall have been studied extensively over land but numerous studies illustrate a strong land-ocean contrast in lightning occurrence (Orville and Henderson 1986; Williams et al. 1992, Zipser 1994; Boccippio et al. 2000; Nesbitt et al. 2000; Toracinta et al. 2002; Christian et al. 2003; Futyan and Del Genio 2007). In the current context this is important because the

results derived from continental lightning studies may not apply over oceanic regions. Results from studies focusing on lightning and rainfall over oceanic regions have shown lightning to be positively correlated with convective rainfall rate, maximum reflectivity, the height of the radar echo top, and ice water path (Carey and Rutledge 2000; Cecil et al. 2005; Pessi and Businger 2009a, here after referred to as PB2009). However, these results should be taken as preliminary due to the shortcomings of the LLS used and the limited domains studied. For example, PB2009 used the Pacific Lightning Detection Network (PacNet), which consisted of four sensors covering a vast expanse of the central North Pacific Ocean. In contrast, GLD360 continuously monitors lightning strokes across a large portion of the Earth's surface. The great precision and high accuracy of GLD360 means that lightning and rainfall may be studied like never before over the open ocean.

The results from several modeling studies showed improvement in initialization and forecasts of extratropical cyclones with the assimilation of lightning data (Alexander et al. 1999; Chang et al. 2001; Papadopoulos et al. 2005; Pessi and Businger 2009b). Thus, better quantification or verification of the relationship between lightning and rainfall over oceanic regions using GLD360's superior lightning detection capability will aid in the continued improvement of weather analysis and forecasting.

1.2 Background

1.2.1 Cloud electrification and the potential influence of aerosols

According to Pruppacher and Klett (1998), a number of charge separation

processes may lead to charge separation within a cloud, but the noninductive mechanism for charge separation is dominant and is able to explain the tripolar charge distribution commonly observed in convective clouds (Fig. 1). Saunders (1993) states that field, laboratory, and modeling studies consistently show that charge separations within convective clouds are largely due to interactions between frozen hydrometeors (e.g. Reynolds et al. 1957; Takahashi 1984; Nelson and Baker 2003). The distribution of charges in an ice crystal that grows by vapor diffusion is dictated by the temperature gradient (Takahashi 1978, Nelson and Baker 2003). Positive ions are able to migrate to the relatively colder interior of the ice crystal while negative charge remains near the surface of the ice crystal during diffusional growth processes. When a small ice crystal (with its generally negative surface charge) comes into contact with a relatively large graupel particle, they may temporarily become fused. When the ice crystal separates from the graupel, ice mass from the surface of the small ice crystal can be left upon the graupel, thereby transferring negative charge to the graupel surface. The loss of the negatively charged ice mass from the surface of the small ice crystal means that the remaining small ice crystal becomes positively charged. A number of studies cite this ice-ice collision process as the main cloud electrification mechanism (Takahashi 1978; Takahashi 1984).

A second noninductive charge separation mechanism involves ice-ice collisions in the presence of supercooled liquid (Saunders 1993). As a graupel particle falls through the mixed-phase region (between roughly -5°C and -20°C), it will collect supercooled droplets that freeze on contact. The latent heat of fusion

released during the freezing process warms the surface of the graupel particle so that it becomes several degrees warmer than the surroundings. When a small ice crystal then comes into contact with the graupel particle, positive ions are transferred to the colder ice crystal and negative ions move toward the warmer graupel particle before separation (Reynolds et al. 1957).

For the noninductive process, the magnitude and polarity of the charge transferred to graupel and small ice crystals respectively during a collision were shown to depend upon the temperature, cloud liquid water content, vertical velocity, and the relative sizes of the interacting hydrometeors (Jayaratne et al. 1983; Takahashi and Miyawaki 2002). At temperatures lower than -10°C, graupel charged positively for high and low cloud liquid water contents and negatively when cloud liquid water was between 0.2 g m^{-3} and 5 g m^{-3} (Takahashi 1978). Later research suggests that for cloud liquid water contents below roughly 0.5 g m^{-3} and temperatures colder than -15°C, graupel may charge negatively (Jayaratne et al. 1983). The results from Jayaratne et al. (1983) were disproved in a follow up study where graupel was shown to charge positively with cloud liquid water content below 0.5 g m^{-3} and temperature at -14°C (Takahashi and Miyawaki 2002). The differences were attributed to potential errors introduced in the study by Jayaratne et al. (1983) as the rotating rod may have led to modified airflow over the riming surface (in laboratory experiments, the riming surface on rotating rods simulates graupel falling at terminal velocity in the environment).

In addition to the noninductive process, several other processes have been hypothesized to contribute to charge separation within a cloud. One such additional

mechanism is the inductive mechanism, which, in contrast to the noninductive mechanism, requires the presence of a background environmental electric field to polarize hydrometeors prior to collisions. The fair weather electric field has negative charge at the earth's surface and more positive charge high above the surface in the troposphere. The background electric field induces negative charge on the underside of hydrometeors and positive charge on their upper surfaces. Thus, a small ice crystal rebounding from the underside of a larger graupel particle would receive net positive charge whereas the graupel particle would charge negatively. Many have questioned to ability of the inductive mechanism to generate sufficient charge separation in clouds based on a number of reasons. First, the collisions between small ice particles and graupel are not long enough to permit charge transfer due in part to the low conductivity of the ice surface (Illingworth and Caranti 1985). Second, the presence of the environmental electric field promotes coalescence or collection, which would not allow for subsequent charge separation (Pruppacher and Klett 1998).

The convective mechanism (neither inductive nor noninductive) for charge separation involves the collection of ions in the immediate vicinity of a cloud and subsequent transport within a cloud (Vonnegut 1982). Positive ions are drawn into the cloud from below and these ions quickly become attached to nascent cloud droplets. Positive charge is then carried to high levels in the cloud as the ionized cloud droplets are transported by the cloud's updraft. At middle and upper levels in the cloud, the accumulation of positive charge attracts negative ions which themselves are entrained into the cloud via turbulent mixing at the cloud boundary. Negative charge can then be transported to low levels in the cloud in downdrafts

associated with the formation of precipitation. Negative charge near the cloud base then attracts additional positive charge from below the cloud and the convective mechanism for charge separation may continue (Saunders 1993). Some doubt the ability of the convective mechanism to generate sufficient charge separation for lightning due to uncertainties surrounding both airflow in the cloud (Latham 1981) and the availability of positive and negative ions in the environment (Wormell 1953).

Gravitational separation of hydrometeors leads to the development of regions of concentrated charge within the cloud. Positive charge accumulates in the upper levels of the cloud due to the vertical advection of small positive ice crystals. Negative charge resides with the larger graupel particles in the middle or lower reaches of the cloud. A small region of positive charge exists near cloud base possibly due to the "charge reversal" of graupel for warmer temperatures (Saunders 1993). In the lower levels of the cloud, graupel may charge positively and small ice crystals negatively (Takahashi 1978). Overall, since ice crystals and graupel may become charged after collisions and accumulate at various levels within the cloud, it appears that lightning requires an appreciable ice concentration aloft. Therefore, any process that affects ice concentrations above the freezing level could influence lighting generation.

What processes could affect ice particle concentrations above the freezing level? Warm-rain precipitation efficiency will affect the amount of liquid water that is carried above the freezing level and aerosol concentrations have been hypothesized to affect warm-rain precipitation efficiency. Assume that a clean air parcel, characteristic of the oceanic environment, ascends until it reaches saturation. At this

point, condensation begins and the available water vapor condenses onto a low number of cloud condensation nuclei (CCN). When few cloud droplets are present, the surface area onto which water vapor condenses is small and therefore the supersaturation becomes large. The increase in supersaturation causes existing droplets to grow faster. Conversely, polluted air results in lower supersaturations and slow droplet growth (Porter 1988). In order to produce rainfall, the collision-coalescense process is important (Pruppacher and Klett 1998). Previous research has shown that once the modal drop size (the most frequently observed droplet size) reaches roughly 30 μm, collision-coalescence processes become very efficient in broadening the drop size distribution (Takahashi 1981; Albrecht 1989). If supersaturations are low (as expected in polluted air conditions), the droplets may never grow up to 30 μm and the warm-rain process may be delayed or inhibited all together. Therefore, it is expected that aerosol concentrations can affect the amount of liquid water carried above the freezing level and consequently ice concentrations there. The microphysical effect of aerosols on cloud processes is recognized to be one of the remaining research frontiers (Rosenfeld 2012).

Are variations in aerosol concentrations sufficiently large to have an impact on warm-rain precipitation efficiency? In-situ observations of aerosol concentrations and size distributions, if they exist continuously in time, are limited over the ocean. However, various attempts to sample marine aerosol populations have been made during field experiments. Observations show that CCN concentrations over the ocean vary between 25 CCN cm^{-3} (extremely clean air) and 350 CCN cm^{-3} (Porter 1988; Clark et al. 1997). Typical marine CCN concentrations are 100 CCN cm^{-3} with

diameters between 0.1 – 10 μm (Takahashi 1981; Porter and Clark 1997). The relatively low CCN concentrations observed over the ocean in these studies means that there is the possibility that warm-rain precipitation efficiency is affected over the ocean (due to clean air conditions), but the current data set does not permit further exploration of this idea.

1.2.2 Atmospheric propagation

If the electric field within a cloud becomes larger than 75 – 100 kV m^{-1} (as a result of the any number of charge separation mechanisms), air is not capable of insulating the separate charge regions and discharge (lightning) occurs. Three types of lightning flashes occur. A cloud-to-ground (CG) stroke may emanate from the negative charge region in the middle to low levels of the ground (- CG) or from the positively charged anvil and travel through clear air to ground (+CG). Intracloud (IC) strokes frequently occur between charge centers in a cloud and may propagate both vertically and horizontally (Fig. 1).

During a given lightning stroke, energy is emitted across many frequencies in the electromagnetic spectrum from the very high frequencies (VHF) to the very low frequencies (VLF). Collectively, the frequencies from 10^4 – 10^{12} Hz are defined as lightning atmospherics, or "sferics" (Pessi et al. 2009). The peak amplitude of the current varies by the type of flash with cloud-to-ground (CG) events typically producing the largest observed currents as opposed to intra-cloud (IC) events, which usually have peak amplitudes less than 15 kA.

The VHF signal depends on direct propagation and is not reflected by the

ionosphere, giving it a maximum range for detection of a few hundred kilometers. In contrast, VLF sferics, with approximate frequencies of 3-30 kHz, easily travel over thousands of kilometers in the Earth-Ionosphere waveguide (Fig. 2). In the ionosphere, the D-region is the layer through which the density of electrons drops by a factor of 1/*e* and its altitude varies on a diurnal timescale. The propagation of VLF sferics depends strongly upon the conductivity of the surface (land vs. ocean) and the time of day since the vertical position of the D-region varies diurnally (Thompson 1993; Thompson et al. 2007). In addition, the VLF signal is subject to attenuation as it propagates great distances (O(10^3) km) across the globe.

1.2.3 Lightning detection and location

In the late 1960's and early 1970's, orthogonal magnetic loop antennas (one oriented in the north/south direction and the other in the east/west direction) were first used to observe changes in the magnetic field that resulted from sferic propagation past the instrument. The changes in the magnetic field through the loop antennas induce voltages in each loop antenna. The magnitude of the voltage is proportional to the change in the magnetic field through the loop; therefore, the components of the magnetic field may be resolved using the voltages measured in each of the orthogonal magnetic loops. The components are then combined to estimate the direction of the stroke's origin (Krider et al. 1980, Orville et al. 1983, Cummins et al. 1998). This technique is called magnetic direction finding (MDF). An estimate of flash location is given by the intersection of vectors from two direction finders. If only two sensors detect a flash that occurs on or near the sensor baseline (the line joining the two

sensors), then the location of the flash is ambiguous. The direction estimates from three or more sensors enclose an area (Fig. 3a) and the estimation of the optimal flash location is that point within the enclosed area for which the direction error of the reporting sensors is minimized (Cummins et al. 2000).

The time-of-group-arrival method for flash location involves a group of synchronized sensors. Assuming the velocity of sferics in the atmosphere is constant, the time it takes for a sferic to reach a sensor is a function of the distance between the stroke and the sensor. If the difference in arrival time at two separate sensors is recorded, then the collection of possible flash locations that guarantees this constant arrival time difference is a hyperbola (Fig. 3b). The intersection of hyperbolas resulting from multiple sensor pairs detecting the stroke provides an estimation of the stroke's location (Lewis et al. 1960, Cummins et al. 1998; Cummins et al. 2000; Pessi et al. 2009).

Detection efficiency (DE) is defined as the ratio of the number of strokes observed, N_o, to the number of strokes that actually occur, N_a, expressed as a percent.

$$DE(x,y,z,t) = \frac{N_o}{N_a} \cdot 100\%$$

Location accuracy (LA) is defined as the error distance between the reported stroke location and the actual stroke location. In order to use an LLS in a quantitative way to improve weather analysis and forecasting, both network DE and LA must be accurately calibrated.

Comparatively few, if any, validation studies have been carried out over the ocean as opposed to over land for LLS, thus empirical models have been developed to

predict network performance over the ocean. Individual sensor DE is dependent upon the signal attenuation over a given distance, propagation characteristics, and the peak current distribution. First, a CG stroke occurs with a peak electric/magnetic field that travels over a certain distance to a sensor in the network. The probability that the sensor detects the CG stroke is given by the sensor's individual DE function. The DE function for the distribution of peak currents is determined empirically and then implemented for all sensors in the network. The minimum detectable peak current amplitude is dictated by the tuning of the sensor, GLD360 has a minimum detection threshold near 5 – 7 kA before a stroke is confirmed. Since each sensor is likely at a different distance from the actual stroke location and the peak electric field is subject to attenuation, the probability of detection for any given sensor is different. Assuming each sensor operates independently, the total probability of detection for a given stroke (the desired network DE) is the product of the individual sensor probabilities of detection:

$$DE_{total} = \prod_{n=1}^{N} DE_n$$

This formulation applies to any number of sensors, n, in an LLS (Pessi et al. 2009).

Location accuracy is modeled according to the semi-major axis of the location error ellipse surrounding the optimal stroke location. It is a function of the DE modeled for individual/groups of sensors, sensor location, and errors in timing and direction (Cummins et al., 1998). Assuming the location errors in latitude and longitude are Gaussian, the model described by Cummins et al. (1998) is set to transect the error surface at a probability level of 0.5. By this method, the error

ellipse then represents the median location errors in both dimensions. The actual 50th percentile error ellipse for any given sensor is determined once the errors are prescribed. Finally, a grid (resolution is arbitrary) is applied over the observation domain and the LA is computed at the center of each grid cell. The LA value reported is taken to be the length of the semi-major axis of the resultant error ellipse (for reporting sensors).

CHAPTER II: DATA AND METHODOLOGY

2.1 Data

2.1.1 The TRMM precipitation radar

The Tropical Rainfall Measuring Mission's Precipitation Radar (TRMM PR) aboard a NASA satellite provides coverage across all longitudes from 38°S to 38°N. The inclination angle of 35° yields roughly 16 earth orbits each day and coverage over any single point within the domain roughly twice each day. Each orbit consists of approximately 9150 scans, where each scan contains 49 rays in the cross track direction. The horizontal resolution of the instrument is approximately 5 km x 5 km and the swath width is about 220 km.

The level 2 TRMM processing algorithm (version 7) Radar Rainfall Rate and Profile (2A25) is used. In the upgraded version 7 algorithm, 80 range bins are sampled approximately every 250 m along the radar beam with the surface of the earth ellipsoid corresponding to bin 79 (B_o). The approximate height, h_a, of the radar return can be found using its reported range bin, B_r, by $h_a = (B_o - B_r) \cdot 250$ m. The zenith angle of the beam is measured according to the local zenith and reaches a maximum value of ~17-18° at the scan edges for TRMM PR's observation mode (Kummerow et al., 1998). Given h_a and the beam zenith angle, ψ, the height of the radar return, h_r, is $h_r = h_a \cdot \cos(\psi)$.

The TRMM PR operates at 13.8 GHz (a wavelength of roughly 2 cm) examining rainfall characteristics over both land and ocean. According to Hitschfeld and Bordan (1954), much of the early work in the radar meteorology was done with

radars using a wavelength, $\lambda = 10$ cm, but the power returned to the radar antenna is proportional to λ^{-4}, so researchers began using radars with $\lambda \sim 3 - 5$ cm. At these shorter wavelengths, the sensitivity and resolution of cloud features is dramatically better, but the signal is strongly attenuated by cloud and precipitation-sized droplets. Models to correct the errors in reflectivity due to attenuation have been developed. By the surface reference method, the total effects of attenuation are determined by the ratio of the power returned to the radar antenna to the power returned to the radar in an adjacent, rain-free area (Hitschfeld and Bordan 1954; Iguchi and Meneghini 1994).

An estimation of the freezing height is made assuming a climatological sea surface temperature and a lapse rate of 6°C/km. The freezing height (freezH) is then the quotient of the $T_{surface}/T_{lapse}$ (km). The 2A25 algorithm requires the bright band (roughly continuous, local maximum in reflectivity in a single layer) be detected within the vertical range freezH +/- 1.5 km.

The algorithm makes the distinction between convective and stratiform precipitation using horizontal and vertical criteria. For the purpose of this discussion, "pixel" refers to one TRMM PR footprint (roughly 5 km x 5 km). In the vertical scheme, if a bright band exists, then the rain is classified as stratiform. If no bright band is detected and the reflectivity surpasses a threshold of 40 dBZ at any range bin, then the precipitation is classified as convective. Ground validation testing indicates that bright band detection (which has implications for rain type classification) is approximately 80% within +/- 7° of nadir and decreases toward the scan edges to values as low as 20% (NASA GES-DISC Mirador).

The horizontal scheme used by TRMM PR for convective/stratiform

precipitation distinction is based on that devised by Steiner et al. (1995). In the horizontal scheme, pixels with reflectivity values of 40 dBZ are immediately labeled convective since it is highly unlikely that rainfall associated with this reflectivity is stratiform (Steiner et al. 1995). Convective rainfall may also be classified by a "peakedness" criterion. For this criterion, the average of the non-zero reflectivity is taken within an 11 km radius of the pixel being sampled. If the reflectivity of the central pixel exceeds the spatial average by a predetermined amount then the pixel is convective. The "peakedness" threshold varies according the intensity of the background average reflectivity. Values of background reflectivity below 20 dBZ require a difference in reflectivity between 5 – 10 dBZ in order to classify central pixel as convective. As the background average reflectivity approaches 40 dBZ, the difference in reflectivity required to invoke the "peakedness" distinction of convective precipitation for the central pixel tends toward zero. Lastly, surrounding pixels are classified as convective by the "surrounding" criterion if they fall within a predetermined radius of a convective pixel, again according to the mean background reflectivity.

Intercomparison between the ground-based, S-band Doppler radar at Kwajalein and the TRMM PR have illustrated that TRMM PR captures the spatial distribution reflectivity in oceanic convection well (Schumacher and Houze 2000). TRMM PR's reflectivity appears to be about 3 – 5 dBZ lower than the reflectivity found by ground-based radar in convective areas (Figs. 4a and 4b). Despite the differences in the intensity of the reflectivity found by each instrument, the spatial distributions of convective and stratiform precipitation identified by the ground-based

Doppler radar and TRMM PR match reasonably well (Figs. 4c and 4d). TRMM PR has an issue deciphering between convective and stratiform precipitation where reflectivity is weak. Weak reflectivity signifies either a low number of precipitation droplets or small droplet sizes and therefore, these areas are unlikely to be convective.

Near-surface reflectivity, the attenuation-corrected reflectivity through the lowest twenty kilometers, and the height of the freezing level are reported for each of the 49 pixels in a scan. Note that over the ocean, TRMM PR has significant issues with ground clutter (i.e., ocean surface roughness and sea spray) and accordingly, the algorithm flags data that is likely to contaminated. The data are generally better closer to the center of the TRMM PR scan (i.e., near nadir), where ground clutter may affect only the lowest 500 m. Toward scan edges, ground clutter is present up to approximately 1.5 km. The near-surface reflectivity parameter samples reflectivity at the lowest possible range bin that is free of ground clutter and therefore is a measure of the reflectivity roughly over the lowest 0.5 – 1.5 km (since the height of ground clutter contamination varies).

2.1.2 The Global Lightning Dataset

GLD360 is composed of a network of ground-based sensors that detect lightning sferics in the VLF portion of the electromagnetic spectrum. These sensors are specially tuned to pick up the relatively weak sferic signal that results from multiple reflections in the waveguide between earth and the ionosphere. The minimum detection threshold is roughly 5 – 7 kA so it is safe to assume that GLD360 predominantly sample CG flashes (Pessi et al. 2009).

PacNet was a predecessor array to GLD360 in that it was among the first systems to provide continuous lightning detection over open ocean on the synoptic scale (Pessi et al. 2009). The development of the DE and LA models specifically for PacNet began by comparing data from a PacNet test sensor positioned in Tuscon, Arizona and data from the National Lightning Detection Network (NLDN). During a week in 2002, a midlatitude cyclone with highly electrified squall lines progressed from the inter-mountain west, across the high plains, and eventually toward the east coast. Lightning was detected by both the PacNet test sensor and NLDN at a number of ranges. As mentioned previously, the VLF signal is strongly dependent upon the attenuation characteristics of the environment. Analysis of the data sample depicted an exponential drop off in signal strength with increasing distance which lead to the attenuation scaling factor, A, modeled as:

$$A = \frac{\alpha_f}{R}\sqrt{\left(\frac{\theta}{\sin(\theta)}\right)}\exp\left(-\frac{R}{\lambda_f}\right)$$

where α_f is a scaling constant, R is the distance between the flash and the sensor location, and $\theta = R/R_e$ where R_e is the Earth's radius. The scaling constant, α_f, is adjusted so that the error between the fit and the actual data is minimized. The space constant, λ_f, depends upon surface characteristics and its proper determination alone yields a complete description of the attenuation as defined above. Note that A is proportional to R^{-1}, which means that the attenuation factor is small for very large values of R. When the attenuation factor is applied to an individual sensor's DE, the end result is that the DE is smaller for strokes occurring at large ranges from the sensor which matches observations. Although the propagation described by this

method is simplified, it provides a reasonable approximation for sferic attenuation given that the distance between GLD360's sensors is generally on the order of 4000 km or less (Pessi et al. 2009).

GLD360 is the only long-range lightning detection network to provide continuous coverage across the majority of Earth's surface with relatively high DE (greater than 60% across the northern hemisphere). When GLD360 observes a flash, it records the date, time (with millisecond precision), latitude, longitude, maximum current, and polarity. GLD360 uses magnetic direction finding and time-of-arrival techniques to locate lightning. Furthermore, each sensor uses recognition algorithms to diagnose sferic waveform variations resulting from propagation over great distances through the Earth-Ionosphere waveguide and this aids in more accurate geolocation (Said et al. 2010). Observations are then transmitted to Vaisala's Network Control Center in Tuscan, AZ for processing and then made available to the data user in near real time (Grogan and Demetriades 2009).

The DE of GLD360 shows fluctuations due to sensor malfunction and generally lower values over the southern hemisphere due to the lack of sensors and their resulting geometry there. DE was generally greater than 50% across the northern hemisphere at the start of 2011 and at present the region surrounding the Hawaiian Islands remains the only weakness across the northern hemisphere. By September 2012, the DE will be approximately 70% across the whole northern hemisphere and portions of the southern hemisphere with a general decrease in performance poleward of roughly 45°S (Fig. 5).

Recently, NLDN was used to validate GLD360's performance across the

conterminous United States and immediately surrounding regions. The results of a study by Demetriades et al. (2010) showed that GLD360 had DE of 86-92% relative to NLDN across the domain (Fig. 6). The agreement in number of observed flashes between NLDN and GLD360 was found to be good for maximum currents larger than about +/- 20 kA (Fig. 7). GLD360 overestimated maximum current below 20 kA, but was generally accurate between 20 kA and 160 kA.

The spatial pattern of stroke density is well captured by GLD360's normalized annual climatology of lightning stroke density (Fig. 8). Specifically, maxima in lightning stroke density over the southeast/midwestern United States, Columbia, Venezuela, sub-Saharan/central Africa, regions of India adjacent to the Himalaya Mountains, and Southeast Asia match the observations of Christian et al. (2003) reasonably well. There are a number of fundamental differences between the climatologies provided in Fig. 8 (e.g. resolution, flashes versus strokes counted, total observation period), and therefore, only a rough comparison is possible between these two data sets. However, an assumption made here is that the GLD360 data used in this study are representative of lightning across the oceanic domains considered.

GLD360 began reporting in late 2009/early 2010 so preliminary data analysis started from January 2010 and continued through May 2011. Upon close inspection, it became clear that a non-negligible amount of noise existed in this dataset on the large scale (Fig. 9). In early May 2011 and again in November 2011, Vaisala implemented new processing algorithms to improve data quality. The DE improvements resulted from adjustments to reduce noise (such as that caused by electrical power lines) and new time filtration techniques in the algorithm guaranteed

a reduction in the number double-counted events. In addition, it is reported that individual sensors operate without malfunction for 99% of the time on average (Vaisala) – the more sensors involved in lightning detection generally yields higher DE and better LA. Continuous improvement of GLD360's LA results from the use of waveform recognition techniques at each sensor site. It is well known that variations in sferic waveforms result from propagation through the Earth-Ionosphere waveguide (Said et al. 2010). The variations in sferic waveforms received by a GLD360 sensor are compared to a stored "catalog" of waveform variations as a function of distance. The end result is then a better estimate of the distance between the stroke location and sensor location (Ryan Said, personal communication).

The reprocessing algorithms coupled with the growing network dramatically increased DE and LA across all domains. Intercomparison studies completed after reprocessing have demonstrated GLD360's overall LA improvement with the lowest values resulting when more sensors are involved in stroke location (Fig. 10). At present, GLD360 median LA is 2 – 5 km across the global domain.

2.2 Methodology

2.2.1 Domains of interest and procedure

The global coverage of GLD360 allows for an expanded analysis over multiple oceanic domains in the present study (Fig. 11). Although analysis of this data began when GLD360 began reporting in January 2010, ten months of data from GLD360 and the TRMM PR (May 2011 to February 2012) constitute the final data set in this study. This time period was chosen following improvements in GLD360's

performance and data processing as well as an upgrade (version 7 algorithm) to TRMM PR data processing.

Following PB2009, a 0.5° x 0.5° grid is applied across the domains of interest taking into account the latitudinal dependence of grid box area and the time interval is set to 1800 s (+/- 900 s from TRMM PR observation time). The choice of grid box size and observation time following PB2009 allows for comparison between the current results and those found previously, however the choice of grid box size and observation time should be made according to a number of considerations.

First, the time that a convective feature may spend within a grid box should be considered. Note that the latitudinal domain (0° - 40°N) studied here spans the tropical, subtropical, and extratropical regimes. Phenomena of differing characteristic time and velocity scales exist within each of the individual regimes. Convection in a passing extratropical system may be in a 0.5° grid box for over an hour while a tropical convective cloud may spend an even longer time within the grid box due to slower horizontal motion (~5 m/s).

Second, the typical life cycle of an airmass thunderstorm is on the order of 30 minutes. It is unlikely that lightning will be present at all times during the life cycle; the proper time interval for an analysis of this type should encompass the majority of the life cycle of convective clouds to allow for the observation of lightning (MacGorman and Rust 1998). However, given that the cloud and precipitation fields evolve during the life cycle of a thunderstorm, the representative quality of the TRMM PR observation may be brought into question. The NASA satellite which carries the TRMM PR instrument orbits at approximately 400 km atlitude with a

speed of 6.9 km s^{-1} and therefore observations in a given grid box are taken in a matter of seconds.

Third, the application of the results must be taken into consideration in the definition of time and space parameters. In the case of PB2009, the result was used to improve the performance of a global numerical weather prediction model. Global models (which have coarse resolution generally greater than 0.5°) do not explicitly resolve individual convective features; therefore, a large area and observation time may be used to get the average or large-scale relationship between lightning and rainfall.

Finally, the choice of time and space parameters may be dictated by the instrumentation used in the study. PB2009 used PacNet, which had DE of 40-60% and LA of 14 – 60 km. The low DE of PacNet (relative to GLD360's DE) means that a longer time period is necessary in order to obtain lightning observations. In addition, if the horizontal dimension of the grid box chosen to be smaller than the LA of the LLS, then it is difficult to discern whether a stroke reported by the LLS actually occurred within a given grid box. There is the possibility of poorly located strokes from grid boxes that are adjacent to the grid box in question being counted which introduces error in the final data set.

The data for this study were obtained by the following steps (Fig. 12). If TRMM PR determines that rain is present in within a grid box, then the 'rainType' parameter of 2A25 identifies these TRMM PR pixels as stratiform or convective precipitation based on horizontal and vertical sampling methods outlined above. If the precipitation is convective, a window of fifteen minutes before and fifteen

minutes after (+/- 900 s) the TRMM PR overpass time is allotted during which TRMM PR's observed reflectivity parameters and the number of lightning strokes and maximum current detected by GLD360 are recorded within the grid box. Strokes occurring over land are not counted in this study. The number of strokes is then normalized to an hour period and an area (km^2) dependent upon the size of the grid box used - here the normalization is 2500 km^2 following PB2009. A winter subset (December 2011, January 2012, and February 2012) and a summer subset (June 2011, July 2011, and August 2011) were extracted from the full data set to investigate the potential annual variability of the relationship between lightning and rainfall.

The frequency distribution for an initial GLD360 data sample (January 2010 - mid 2011) showed a strong concentration at lower stroke density, $O(10^0 - 10^2)$, but with the maximum stroke density near 10^4 strokes 2500 km^{-2} hr^{-1}. The majority of observations of maximum current from GLD360 were concentrated between 10^2 kA and 10^3 kA. This suggests a natural logarithm binning scheme to examine both lightning parameters. The stroke density domain is 10^{-1} - 10^3 strokes 2500 km^{-2} hr^{-1} which corresponds approximately to bins 1 through 7 (e^1 fl - e^7 fl; where e = 2.71828). A least-squares fit is then applied to the bin medians of (maximum and average) reflectivity for each stroke density and maximum current. The least-squares fit has the general form $f(x) = a \cdot \ln(x) + c$. In this case, the stroke density or maximum current would be substituted in for x to determine the fit for the reflectivity value.

Following the above discussion, large concentrations of ice above the freezing level are necessary for lightning to occur. Although TRMM PR doesn't explicitly provide details about the phase of hydrometeors, a large concentration of ice

theoretically will produce larger reflectivity values. Accordingly, the vertical distribution of the maximum attenuation-corrected reflectivity is determined for various categories of lightning stroke density and maximum current: for the total data set, for the 50th percentile, for the 90th percentile, and for the 99th percentile. The data is analyzed for the full ten-month period in each 0.5° grid box across each basin. The maximum value of the attenuation-corrected reflectivity (values between 0.00 dBZ and 100 dBZ) is recorded for each of the eighty range bins between the surface and 20 km altitude for each observation in a grid box. Then the vertical distributions of maximum attenuation-corrected reflectivity within each grid box in the domain are averaged to obtain a final basin wide average vertical distribution.

2.2.2 Sensitivity tests

Data subsets are utilized to conduct various qualitative sensitivity tests for time and space parameter choices. A preliminary assessment of the sensitivity of stroke density was done using data from July 2010 over the NWPAC domain for combinations of 0.1° x 0.1°, 0.5° x 0.5°, or 1.0° x 1.0° areas and various intervals. Data from May 2011 – August 2011 (after GLD360 data improvements) over the NATL domain was used to test for sensitivity in the lightning versus rainfall relationship using four different observation times (+/- 60 s, +/- 120 s, +/- 450 s, and +/- 900 s) on a 0.5° x 0.5° grid.

In Fig. 5 of LeMone and Zipser (1980) the upper 10% of updraft core diameters observed in convective clouds during GATE were 3 km in diameter. But the updraft core may account for only a fraction of the total cloud diameter.

Individual cumuliform clouds may have diameters of 10 km or more. By the method described above, there exists the potential for a single TRMM PR pixel to be identified as convective and therefore the grid box as a whole is admitted into the observation set. But is a ~100 km^2 area representative of the conditions within the larger grid box (2500 km^2 in this case) as a whole?

Archived loops GLD360 data overlain on geostationary operational environment satellite (GOES) infrared imagery (channel 4 at 10.7 μm) are available at weather.hawaii.edu. Review of these loops illustrates that single, isolated lightning strokes occasionally are correlated with cloud tops that are likely to be below the freezing level. Since lightning requires high concentrations of frozen hydrometeors above the freezing level, lightning associated with warm cloud tops is unlikely. Although charge transfer is possible between liquid droplets in warm clouds (i.e., clouds whose tops do not extend above the freezing level), the total charge separation does not get large enough to generate lightning (Takahashi 1981). Data points that represent lightning associated with warm cloud tops may then contribute to the inaccuracy of the final result.

By May 2011, both the maximum detection efficiency and maximum number of lightning observations by GLD360 were found over the NATL domain. A temporal subset of the data from May 2011 – August 2011 is analyzed with and without a convective area threshold (CAT). In general, TRMM PR may pass over an individual grid box within the domain many times during the full analysis period. The CAT mandates that the total convective area within a grid box identified by TRMM PR (number of convective pixels multiplied by the area of TRMM PR's 5 km

x 5 km footprint) be at least 25% of the total grid box area. The hypothesis here is that events which are poorly located by GLD360 contribute to noise for the lowest values of both reflectivity and stroke density or maximum current. The inclusion of a CAT will limit the amount of these outliers (general noise in the dataset) in the final data and ensure that the lighting is being correlated more closely with convective precipitation. After the implementation of the GLD360 reprocessing algorithm, four months of data (May 2011 - Aug 2011) are used in a test of the influence of a convective area threshold (CAT) on the relationship between reflectivity and lightning.

CHAPTER III: RESULTS

As mentioned, both stroke density and maximum current are concentrated roughly between magnitudes of 10^0 and $10^2/10^3$. Furthermore, an exponential decrease in the number of observations with increasing stroke density and maximum current is apparent (for all time periods studied). A climatology (January 2010 – December 2010) of strokes in each basin shows the maximum number of flashes occurs during the summer in each case, but observations of lightning in the NEPAC domain are very limited. A subtropical high is present in climatology over much of the NEPAC basin and subsidence associated with the subtropical high inhibits the development of deep convection in that basin. A temperature inversion is typically present between 2 and 3 km altitude in the vicinity of Hawaii in vertical temperature profiles from Hilo, HI and Lihue, HI (radiosondes are launched twice each day at 0 UTC and 12 UTC). The average height of the freezing level found by TRMM PR was near 4 – 5 km in each basin, which matches well with available atmospheric soundings (University of Wyoming). Although the inversion layer itself is not an impenetrable layer, the presence of the temperature inversion in the NEPAC basin means that cloud tops will be below about 3 – 4 km in most cases. If cloud tops are generally located below the freezing level in the NEPAC basin, it is unlikely that appreciable cloud liquid is transported above the freezing level, which explains the lack of lightning observed there.

3.1 Sensitivity to time and space parameters

The stroke density decreases as the grid box size and observation interval increase in general. For a 1° x 1° grid box resolution, the range of observed stroke density is about 8-16 strokes 2500 km^{-2} hr^{-1}. This range increases approximately exponentially as the grid box size decreases; for the 0.1° x 0.1° grid box, the range is roughly 350 - 725 strokes 2500 km^{-2} hr^{-1} (Fig. 13). Note that the vertical axis in the figure is a logarithmic scale. Although the error bars (one standard deviation) span a smaller vertical distance for the smallest grid box size, the actual standard deviation there is about an order of magnitude larger than that for the largest grid box for the same [half] time window.

At first glance, the lightning versus near-surface reflectivity relationship appears to be invariant with respect to the choice of observation time (Fig. 14a). The range of reflectivity values across all lightning values is approximately 6 – 8 dBZ (between ~45 dBZ and ~53 dBZ). However, for a two-minute total observation time (60 s), the sensitivity in the relationship is slightly larger. The range of reflectivity over the domain of stroke density is about 10 dBZ (yellow line showing a range of about 44 – 54 dBZ). Longer observation times generally allow more observations of lower stroke density (Fig. 14b). Additionally, it would appear that the lowest detectable stroke density is a function of the observation time, as a 30 minute total observation time resolves stroke density near 0 whereas a total observation time of 2 minutes truncates at about 25 strokes 2500 km^{-2} hr^{-1}. This result – that stroke density tends to increase overall for shorter observation times – is in line with the preliminary analysis done for the first year of GLD360 data, from January 2010 to December

2010.

The inclusion of a CAT during data processing is shown to reduce the number of data points (specifically the outliers, which are attributed to noise in the data set) for the relationship between maximum current or stroke density and reflectivity. First, the inclusion of a CAT has a profound effect on the slope of the logarithmic fit in the case of the maximum current amplitude. In Fig. 15a, the binned data for maximum current exhibit large spread for the lowest three bins (a given maximum current spans many maximum reflectivity values). Also, the lower bins are skewed, which is manifest by the large distance between the median (red) and the lower quartile (lower bound of the blue box) relative to the distance between the median and the upper quartile (upper bound of the blue box). The bins for larger maximum current are more symmetric about their respective median values. Many outliers (for values of maximum reflectivity generally lower than 35 dBZ) are present for larger maximum currents. The natural logarithmic fit applied to the bin medians has a range of roughly 7 dBZ (~38 - 45 dBZ).

Like the maximum current, the stroke density versus maximum reflectivity initially shows considerable spread in the lowest bins and a large number of outliers (Fig. 16a). The logarithmic fit has a range of about 11 dBZ as well (~38 – 49 dBZ). With the inclusion of a CAT in the data analysis for maximum current (Fig. 15b) and stroke density (Fig. 16b), the number of outliers is drastically reduced across each domain, but nearly all sensitivity in the relationship is lost for maximum current. The logarithmic fit to the bin medians of stroke density versus maximum reflectivity retains a range of about 5 dBZ across the domain of stroke density with a CAT

applied. The spread in each bin is noticeably less in Fig. 16b and the sensitivity of the logarithmic fit is decreased. The higher median values for the lowest three bins of stroke density result from the reduction in noise from the data set with the inclusion of the CAT.

3.2 Lightning and rainfall over the ocean

Following the sensitivity studies, the full time period and winter/summer subsets were analyzed across the three oceanic domains with a CAT applied. Logarithmic models are found to approximate the relationship between lightning detected by GLD360 and rainfall observed by TRMM PR in some cases. The variability in the results between the different ocean basins is generally minimal, but some element of annual variability is found for all basins. The diagnosis of annual variability in these relationships is significant because it is indicative of bias in the occurrence of lightning as a function of location and season. These details may be taken into account when applying the results found here to future studies of oceanic convection.

From these models, proxy reflectivity or rainfall rate data sets may be developed and the proxy data may then be assimilated into numerical weather prediction models or used in real-time weather analysis over oceanic regions where observations are lacking.

3.2.1 Stroke density and maximum reflectivity

Figs. 17, 18, and 19 illustrate that over the NATL, NWPAC, and NEPAC (to a

lesser degree) the maximum reflectivity increases logarithmically with increasing stroke density. The result for the NEPAC shown in Fig. 19 is found using significantly fewer observations in comparison to Figs. 17 and 18 for reasons previously described. Fig. 20 shows the range of maximum reflectivity is roughly 6 dBZ (spanning the range from ~ 46 – 52 dBZ) across the domain of stroke density for all basins, which aligns with PB2009. The NEPAC exhibits the largest values across the domain of stroke density and the NATL shows the lowest values – approximately 1-2 dBZ less than the NEPAC. The NWPAC shows the greatest sensitivity as it tends towards the NATL for the lowest stroke density and the NEPAC for the highest flash density.

A difference in the maximum reflectivity is shown to exist between winter and summer for all three basins (Fig. 21); furthermore, this difference varies in magnitude as a function of stroke density. Below 10^1 strokes 2500 km^{-2} hr^{-1}, the difference is 1 – 2 dBZ, yet above 10^2 strokes 2500 km^{-2} hr^{-1} the difference is more like 4 – 5 dBZ which differs from the range found by PB2009. For winter, the NWPAC consistently has the highest reflectivity values, but the NEPAC and NATL show the largest sensitivity. For the summer subset, the NEPAC exhibits the largest sensitivity. The maximum near-surface reflectivity is found over the NWPAC and NATL for lower stroke density in the summer and over the NEPAC for the highest stroke density values (greater than 10^2 strokes 2500 km^{-2} hr^{-1}).

There is high confidence in the logarithmic least-squares regressions applied to the data since R^2 values range from 0.68 (NEPAC) – 0.89 (NWPAC). There is a great range of R^2 values for seasonal results, but the minimum R^2 is still larger than

0.5. Root mean-square errors are generally less than 1 dBZ. Finally, the errors in TRMM PR's reported reflectivity are about 0.5 dBZ for each data set (Table 2). Overall, the NWPAC stroke density versus maximum reflectivity showed the largest slope and the reflectivity intercepts were larger than 45 dBZ (Table 4).

3.2.2 Maximum current and maximum reflectivity

The NATL (Fig. 22), NWPAC (Fig. 23), and NEPAC (Fig. 24) all show concentration of data in the middle of the maximum current domain. The range in reflectivity from (the minimum to maximum reflectivity per maximum current bin) is about 10 – 15 dBZ, but the interquartile range is much more compact (approximately 4 – 5 dBZ). Variability in spread and the skewness of the data within separate maximum current bins is the greatest where the number of observations is the smallest (e.g., for low and high maximum currents).

The maximum current versus maximum reflectivity relationship has approximately zero slope in the NATL and NEPAC, while the NWPAC shows a range of about 5 dBZ (46 – 51 dBZ). Winter values of maximum reflectivity are systematically higher (for various maximum currents) than summer maximum reflectivity by about 2 – 3 dBZ. In both seasons the NATL domain shows the largest sensitivity (Fig. 26). Note that the sensitivity shown in the NATL winter case is about 6 dBZ, which is dramatically larger than all of the other fits and is comparable to the sensitivity for stroke density.

There is low confidence in these logarithmic models for maximum current versus maximum reflectivity since R^2 values range from about 0 - 0.17 for the full ten

months of data separated by basin. Seasonal R^2 values are much higher in some cases; the NATL summer and winter subsets have R^2 values on the order of 0.6 (Table 3). Root mean-square errors range from about 0.56 dBZ (comparable to TRMM PR errors in reflectivity) to 2.3 dBZ. The largest slope in this case was found in the NWPAC and all intercepts are above 46 dBZ (Table 5).

3.2.3 Stroke density and average reflectivity

The average reflectivity is much lower than the maximum reflectivity in a given grid box – for all basins, the median values of average reflectivity are about 20 dBZ lower than the maximum values discussed above. The bins of stroke density in the NATL (Fig. 27) and NWPAC (Fig. 28) show considerable skewness. The upper bound for a given grid box is further from the median than the lower bound. However, in both of these domains, the interquartile ranges are symmetric where data are numerous. The NEPAC domain shows higher average reflectivity for the lower bins of stroke density than both the NATL and NWPAC as medians of individual bins are generally above 20 dBZ (Fig. 30). There are no data in the NEPAC beyond about 10^2 strokes 2500 km^{-2} hr^{-1}.

Similar to the stroke density versus maximum reflectivity case, the relationship here shows an annual separation. The average reflectivity associated with a given stroke density is higher in winter than in summer. Additionally, the difference is larger for larger values of stroke density (Fig. 31). In both the winter and summer seasons, the NEPAC exhibits the highest sensitivity, but again, it is important to note that large stroke densities are not observed in the NEPAC in this

analysis – the model extrapolates average reflectivity from the stroke density beyond about 10^2 strokes 2500 km^{-2} hr^{-1}.

The R^2 values for the full time period across the three basins are between roughly 0.07 and 0.67 (Table 6). Seasonal subsets have R^2 values between 0 and 0.68. Root mean-square error in all cases ranges between 1.5 dBZ and 3.5 dBZ. Combined with TRMM PR's average error of about 0.5 dBZ, the errors are comparable to the separation between winter and summer data.

The greatest sensitivity in the relationship between stroke density and average near-surface reflectivity is found in the NWPAC where R^2 is equal to 0.67 (Table 8). The slope is more than two times the slope found in other basins. The reflectivity intercepts are near 20 dBZ in all cases, which is substantially below PB2009 (~33 dBZ).

3.2.4 Maximum current and average reflectivity

As with the stroke density versus average reflectivity, the average reflectivities are generally low, near 20 dBZ across the domain of maximum currents (Fig. 35). The maximum currents are again concentrated between 10^2 and 10^3 kA. For the full time period, the NEPAC shows the largest sensitivity and the highest average reflectivity for a given maximum current. The range for the NEPAC regression is 18 – 25 dBZ across the domain of maximum current (Table 7). The NATL shows the least sensitivity and the lowest values for average reflectivity out of the three basins studied.

Seasonal fits show considerable variability. The NEPAC summer and

NWPAC winter show decreasing average near-surface reflectivity with increasing maximum current (Fig. 36). The NEPAC winter and NATL winter show a large amount of sensitivity in comparison to other seasonal fits. The ranges for these fits are approximately 15 dBZ across the domain of maximum current. The NWPAC summer and NATL summer also have considerable positive slope, but have smaller ranges of average reflectivity relative to the winter subsets.

There is a large range of reported R^2 values in the case of maximum current versus average reflectivity. The NEPAC summer subset has an R^2 equal to 0.0002, which is indicative of a very poor fit (Table 7). Note that the slope of this fit was negative. On the complete opposite end of the spectrum is the NATL winter basin with R^2 equal to 0.71. However, the larger R^2 is a result of good agreement between the model and the bin medians over a small portion of the maximum current domain where the NATL winter data is concentrated. The root mean-square errors ranged from 1.19 dBZ (NATL winter) to 5.37 dBZ (NEPAC summer). The latter is comparable to the range of many of the logarithmic fits. The reflectivity intercepts for the full data sets are between 17 dBZ and 19 dBZ (Table 9).

3.3 Vertical distributions of maximum reflectivity

The maximum attenuation-corrected reflectivity distributions consistently show higher reflectivity for larger stroke density at all levels (Figs. 37a, 38a, and 39a). The total average reflectivity in the levels for the NWPAC, NEPAC, and NATL is near 20 dBZ (in line with the above results) and the maximum reflectivity in the vertical is between 2 km and 3 km (Solid black lines in Figs. 37a, 38a, and 39a).

There is considerable variability in the vertical distributions of maximum reflectivity for individual months.

For stroke density greater than or equal to the median and 90th percentile, the maximum reflectivity is below the freezing level (maxima located between 2 km – 3 km), but the magnitude is larger. The maxima in the median reflectivity profiles range from 44 dBZ (NEPAC) – 46 dBZ (NATL). The reflectivity drops off rapidly above the freezing level. By 6 km, the reflectivity ranges from 32 dBZ (NEPAC) – 35 dBZ (NATL) for the median profiles and 35 dBZ (NEPAC) – 41 dBZ (NATL) for the highest 10% of the stroke density.

The reflectivity is nearly uniform at about 52 dBZ over the lowest 2 – 3 km in the NWPAC (Fig. 38a) for the highest 1% of stroke density (99th percentile). The maximum reflectivity for the NEPAC and NATL are slightly lower at about 49 dBZ and 50 dBZ respectively. At 6 km, both the NWPAC and NATL retain reflectivity values above 40 dBZ (45 dBZ and 44 dBZ respectively). There is only a small difference between the reflectivity values for the highest 10 % and 1% in the NEPAC basin. Overall, the decrease in reflectivity above the freezing level is more gradual for higher stroke density consistent with previous studies (e.g. Zipser and Lutz 1994).

The vertical distribution of maximum attenuation-corrected reflectivity are similar for varying maximum current as indicated by the close agreement between the median, 90th, and 99th percentile lines being within about 3 dBZ of one another (Figs. 37b, 38b, and 39b). In all three basins, the reflectivity is maximum below the freezing level in the vicinity of 2 – 3 km. The maximum reflectivity values near the surface are 43 dBZ (NEPAC) – 45 dBZ (NATL) for the median profile, 43 dBZ

(NEPAC) – 45 dBZ (NATL) for the top 10% of maximum currents, and 43 dBZ (NEPAC) – 45 dBZ (NWPAC) for the top 1% of observed maximum currents.

The reflectivity drops off above the freezing level as in the vertical reflectivity distributions associated with stroke density. However, the median profiles of reflectivity vary between 28 dBZ (NEPAC) and 33 dBZ (NATL) at 6 km, which is noticeably less than the profiles for stroke density at this altitude. Between 5 km and 12 km in the NEPAC basin, the reflectivity profile for the top 10% of maximum currents is higher than the top 1% of maximum currents by 4 – 5 dBZ. In contrast, in the NWPAC, the reflectivity for the top 1% of maximum currents is about 4 dBZ larger than that for the top 10% of maximum flashes. In the NATL basin, the reflectivity for the median, 90^{th} percentile, and 99^{th} percentile are nearly the same.

CHAPTER IV: CONCLUSIONS, DISCUSSION, AND FUTURE WORK

An advanced long-range lightning detection system, GLD360, was used along with the TRMM PR in order to explore the relationship between lightning and rainfall over three ocean basins in the northern hemisphere over the course of a ten-month period. The lightning stroke density and maximum current observed by GLD360 was compared to the maximum and average reflectivity observed by TRMM PR roughly between 0.5 km and 2 km. Positive correlations between the stroke density and maximum reflectivity were found in each of the three ocean basins studied. The result is robust for oceanic regions since there is little variability in the relationships found for individual basins. The maximum current of lightning strokes was found to be insensitive to changes in both maximum and average reflectivity across each of the basins. The vertical distributions of maximum reflectivity are shown to vary for different categories of stoke density while the vertical distributions of maximum reflectivity are similar for different maximum current observations.

Specifically, the stroke density determined using long-distance lightning observations from GLD360 was found to increase logarithmically with increasing maximum reflectivity. There is high confidence in the models found here as the R^2 the values range from 0.68 to 0.89 for the full time period. If lightning is present in oceanic convection, the results of this study say that the maximum reflectivity will be found between 2 – 3 km and that the maximum reflectivity will be between 45 dBZ and 50 dBZ. The range of the logarithmic regressions is consistently about 5 dBZ. Much of the sensitivity in the model is for the lowest stroke density where the

majority of GLD360's observations are. The range of maximum reflectivity over the domain of maximum current is small (less than 4 dBZ across all basins).

The maximum attenuation-corrected reflectivity shows that increases in GLD360's stroke density are associated with higher maximum reflectivity at an altitude of 2 – 3 km as well as above the freezing level. The maximum attenuation-corrected reflectivity for the median, 90th percentile, and 99th percentile of maximum currents observed by GLD360 are very similar which allows any maximum current observed to be linked to a specific reflectivity at a given height.

The results for the average reflectivity over an individual grid box are less certain. Generally low values for R^2 for the logarithmic fits imply that stroke density and maximum current have little skill in predicting average reflectivity. Increases in the goodness-of-fit measures are noted for seasonal data subsets. The stroke density and maximum current compared to the average reflectivity in the NEPAC and NATL during the ten-month period had R^2 values below 0.2 for example. In contrast, NEPAC winter and NATL summer and winter subsets had R^2 values larger than 0.5. Regardless, the confidence in the relationships between stroke density or maximum current and average reflectivity remains low because these results occur for a narrow part of the domains of stroke density and maximum current.

Perhaps the most glaring difference between the results of this study and those of PB2009 is the 10 dBZ disagreement in the reflectivity across the domain of stroke density. To reiterate, PB2009 used a prototype LLS called PacNet, which was composed of four VLF sensors and they covered a large area in the central North Pacific Ocean. The maximum reflectivity was found to be positively correlated with

stroke density, but the magnitude of reflectivity for a given stroke density was much lower than that found in the current study. Although, sferic propagation over the ocean tends to be better because the surface has higher conductivity, the PacNet had maximum DE of 40 – 60 % and LA of 16 – 40 km, which by the standards of modern LLS, are rather poor. It is likely the smaller reflectivity values reported by PB2009 result from the average performance of their chosen LLS and their choice to report average reflectivity over an individual grid box.

Convective and stratiform precipitation are closely associated and are likely to occur together within a 2500 km^2 area (the unit area used in this study). Accordingly the distinction between convective and stratiform precipitation is difficult to make within an area comparable to the area of an individual grid box used in this study. The CAT isolates those grid boxes that have 25% of the total TRMM PR pixels identified as convective. Therefore, the CAT is an attempt to eliminate observations where poorly located lightning strokes from GLD360 are coupled with low reflectivity values (10 – 20 dBZ). The constraint here is that lightning is associated with heavy rainfall over the ocean. The results of this study confirm that the CAT is effective at eliminating noise as the maximum reflectivity values are higher for a given stroke density or maximum current with the inclusion of the CAT in data processing.

As expected, the stroke density is shown to vary based upon the choice of time and space parameters used in the analysis. However, the stroke density found using the highest resolution grid box studied (0.1° x 0.1°) behaves in like the two larger grid box sizes considered in this study as the time and space parameters are varied. The most reliable data from GLD360 to date (generally available after May 2011)

illustrate relationship between lightning and reflectivity is nearly insensitive to the choice of total observation time in the analysis.

For both stroke density and maximum current, the total average vertical distributions of reflectivity (solid black lines in Figs. 37, 38, and 39) are substantially lower than the median vertical reflectivity distributions through the majority of the depth of the troposphere. This means that a large number of lightning observations over the ocean are associated with low reflectivity, or conversely, that the most intense precipitation occurs in a relatively low number of clouds. This result is in agreement, at least in part, with the intercomparison study between GLD360 and NLDN (Demetriades et al. 2010) over the continental United States and the preliminary analysis for the first year of GLD360 observations.

4.1 Discussion and Future Work

It is important to note that the maximum reflectivity values reported by TRMM PR for stroke density and maximum current are relatively large. A difference of 5 dBZ, from 45 dBZ to 50 dBZ, between the lowest and highest stroke density or maximum current implies exponential increases in rainfall, whereas, the differences in precipitation indentified by reflectivity of 20 dBZ and 25 dBZ respectively may be difficult to discern. The logarithmic relationships between lightning and rainfall defined for each basin studied here (Tables 5,6, 8, and 9) can be used to create proxy data sets to be used for weather analysis and numerical weather prediction.

In the current study, the stroke density and maximum current have been compared to both the maximum reflectivity and the reflectivity averaged over the

entire grid box (consistent with PB2009). Note that a substantial decrease in the magnitude of reflectivity values for various stroke densities and maximum currents was shown here when reflectivity was averaged over an individual grid box. Although convective precipitation within a grid box is required before lightning characteristics are documented, adjacent stratiform precipitation regions included in the areal reflectivity average will lead to an overall decrease in the average reflectivity value. In addition, TRMM PR observations of cloud free air (reflectivity near zero) will also reduce the average value substantially.

The insensitivity of the lightning versus rainfall relationship to increases in temporal and spatial resolution during the analysis suggests that GLD360 may be used to study the relationship between lightning and rainfall at higher temporal and spatial resolution. The proxy data sets that are formulated using results from a higher resolution study are more applicable to individual convective clouds.

There were 1307 observations in the NWPAC, 2009 observations in the NATL, and only 243 observations in the NEPAC for the ten-month period. The smaller data population in the NEPAC domain allows for any outlier data to influence the final relationship between stroke density or maximum current and reflectivity; outliers exist even with the inclusion of the CAT in data analysis given the noise characteristics of both GLD360 and TRMM PR data. It is of interest to continue the analysis as data become available to eliminate skepticism based on the argument of small sample size.

The argument that these results are robust using only the current sample is substantiated by the fact that over the ocean conditions are fairly uniform on large

scales. In other words, extending the data set beyond the ten-month period and three domains studied here would add data points in the similar ranges of stroke density and maximum current.

DE of GLD360 has been validated over land using the NLDN. However, a validation study over the ocean has not occurred yet. We must rely on simplified DE and LA models, which are only estimations based on a number of environmental considerations and assumptions discussed above. Advances are being made as evidenced by the growth of the GLD360 network (Vaisala) and the improvements in estimated DE south of the equator.

The area surrounding Hawaii has the lowest estimated DE at present, which may be a concern to local power authorities since lightning impacts electrical infrastructure; knowledge of approaching lightning-producing storms would allow for necessary precautions to be taken to prevent damage to electrical systems. An independent network of MDF/TOA sensors partially funded by local power authorities could be installed throughout the Hawaiian Islands. The installation would allow an accurate assessment of GLD360 performance across the Central Pacific Ocean. Then GLD360's DE model space constants could be adjusted so that the model estimation of DE matches the observations for the region surrounding Hawaii.

A pre-requisite for the development of precipitation-sized ice particles, such as graupel, is an updraft of at least 1 m s^{-1}**.** Smaller ice crystals are lofted to higher altitudes in the cloud. Updrafts in stronger oceanic convection may be 3 – 5 m s^{-1} and therefore could support graupel formation and potential charge separation. Negative

strokes in the GLD360 dataset indicate strokes more closely associated with the core of deep convection (as indicated by TRMM PR) since they emanate from the negatively charged graupel region in the middle to lower levels of the cloud. Therefore, it may also be important to investigate the influence of polarity on the relationships found here.

The current study employs a land mask to determine if the grid box is over ocean or land within the domain of interest. Dust, soot, and other biological or anthropogenic pollutants have been shown to exist great distances far from land out over the ocean (Porter 1988) and in this case continental aerosols may be advected over warm ocean boundary currents from adjacent continents. The introduction of continental aerosols into oceanic parcels leads to a substantial increase in the CCN number concentration (Clark et al. 1997) and therefore cloud electrification processes over the ocean may be affected. The GLD360 lightning stroke density climatology shows maxima in the vicinity of western boundary currents (e.g. Kuroshio current and the Gulf Stream current). These regions may be isolated from the rest of the basin during data analysis to assess their influence on the final results. An analysis without the inclusion of the western boundary current regions would be a step towards analyzing the true oceanic relationship.

McMurdie and Mass (2004) illustrate the difficulty that models often have in representing and forecasting extratropical cyclones moving toward the west coast of North America from the open ocean, where observations are sparse. They find that the average 48-hour forecast location error is 453 km with an absolute average error in central pressure of 7.5 hPa. It was mentioned that efforts to assimilate lightning

data into numerical weather prediction models have produced promising results where oceanic systems are better initialized and more accurately forecasted with the inclusion of lightning data (Alexander et al. 1999; Chang et al. 2001; Pessi and Businger 2009b). Pessi and Businger (2009b) formulated a practical assimilation scheme for operational use. Lightning observations are compiled during an initial observation period (~8 hrs) and then the "observed" rainfall (what is indicated by the lightning from the empirical formula) is used to modify the convective parameterization scheme – specifically the temperature tendency – in the model. An extratropical cyclone was examined in their modeling study. Without lightning data, the cyclone did not rapidly intensify as was observed. The 24-hour forecast with lightning data included in the initialization reproduced the observed central pressure within the system. In addition, the distribution and intensity of convection along the storm's cold front was more accurately represented. This is one example of the improvement in numerical weather prediction using long-range observations of lightning.

GLD360's ability to detect lightning continuously at great distances over the ocean is superior to any other LLS currently operating. The high degree of accuracy of GLD360's observations will allow for improved weather analysis and numerical prediction. The empirical least-squares fits found here can be used in a similar manner to initialize systems in numerical weather models given GLD360's observations of stroke density or maximum current. In addition, the intensity of convective rainfall in electrically-active oceanic storms far beyond the range of land-based Doppler radars can be estimated in near real time using GLD360's continuous

lightning observations and the models developed in this study.

TABLES

Table 1. A summary of available performance statistics for LLS that are currently operating. * indicates WSI GLN peak performance.

Network	DE (%)	LA (km)
GLD360	60–70	2–5
ENTLN	–	–
WWLLN	2–35	~4
WSI GLN	80*	1*

Table 2. A summary of the goodness-of-fit statistics for least-square logarithmic fits for stroke density versus maximum reflectivity. The results from Pessi and Businger 2009a (PB2009) are included for comparison.

Data	Num. of Obs.	R^2	RMSE	Range of Z	Error (dBZ)
NWPAC Full	1307	0.8945	0.4695	46 – 51	0.5816
NEPAC Full	243	0.6825	0.7722	47 – 53	0.5615
NATL Full	2009	0.7602	0.7069	46 – 51	0.5241
PB2009	–	–	–	33 – 39	–
NWPAC Sum.	750	0.8190	0.5754	45 – 51	0.5803
NWPAC Win.	110	0.5815	1.378	47 – 55	0.5641
NEPAC Sum.	51	0.3349	2.267	44 – 52	0.5225
NEPAC Win.	134	0.7201	0.8536	46 – 55	0.5745
NATL Sum.	855	0.6629	0.7746	45 – 50	0.5173
NATL Win.	209	0.6214	1.407	46 – 54	0.5331

Table 3. A summary of the goodness-of-fit statistics for least-square logarithmic fits for maximum current versus maximum near-surface reflectivity. The results from Pessi and Businger 2009a (PB2009) are included for comparison.

Data	Num. of Obs.	R^2	RMSE	Range of Z	Error (dBZ)
NWPAC Full	1307	0.1663	1.103	46 – 51	0.5816
NEPAC Full	243	0.0078	1.127	47 – 48	0.5615
NATL Full	2009	0.0045	0.6143	47	0.5241
NWPAC Sum.	750	0.1452	1.204	46 – 48	0.5803
NWPAC Win.	110	0.00826	1.425	48	0.5641
NEPAC Sum.	51	0.01762	2.265	46 – 48	0.5225
NEPAC Win.	134	0.02988	1.521	48 – 49	0.5745
NATL Sum.	855	0.6649	0.5608	44 – 47	0.5173
NATL Win.	209	0.6146	1.045	46 – 53	0.5331

Table 4. Logarithmic fits of the form $y = A \cdot \ln(x) + C$ by basin for stroke density versus maximum reflectivity. Values in parentheses are 95% confidence bounds. The results of PB2009 are included for comparison.

Data	A	C
NWPAC Full	0.7230 (0.5480, 0.8979)	45.41 (44.77, 46.06)
NEPAC Full	0.5988 (0.3110, 0.8865)	46.68 (45.62, 47.74)
NATL Full	0.6656 (0.4022, 0.9290)	45.25 (44.28, 46.22)
PB2009	0.7200	33.3

Table 5. Logarithmic fits of the form $y = A \cdot \ln(x) + C$ by basin for maximum current versus maximum reflectivity. Values in parentheses are 95% confidence bounds.

Data	A	C
NWPAC Full	0.2818 (-0.1939, 0.7574)	46.15 (44.11, 48.20)
NEPAC Full	0.0758 (-0.7513, 0.9269)	47.83 (44.09, 51.58)
NATL Full	0.0230 (-0.2412, 0.2898)	47.38 (46.24, 48.52)

Table 6. As in Table 2, for stroke density versus average reflectivity.

Data	Num. of Obs.	R^2	RMSE	Range of Z	Error (dBZ)
NWPAC Full	1307	0.6732	1.521	17 – 25	0.5775
NEPAC Full	243	0.0723	2.973	20 – 23	0.5644
NATL Full	2009	0.3317	1.011	17 – 20	0.5250
PB2009	-	-	-	33 – 39	-
NWPAC Sum.	750	0.0035	3.006	19 – 20	0.5740
NWPAC Win.	110	0.3151	3.501	20 – 27	0.5623
NEPAC Sum.	51	0.0087	2.796	18 – 22	0.5305
NEPAC Win.	134	0.6086	2.523	19 – 38	0.5777
NATL Sum.	855	0.1659	1.618	16 – 20	0.5180
NATL Win.	209	0.6893	2.145	18 – 27	0.5367

Table 7. As in Table 3, for maximum current versus average reflectivity.

Data	Num. of Obs.	R^2	RMSE	Range of Z	Error (dBZ)
NWPAC Full	1307	0.1663	1.103	18 – 24	0.5775
NEPAC Full	243	0.0959	3.369	18 – 25	0.5644
NATL Full	2009	0.1194	1.223	17 – 20	0.5250
NWPAC Sum.	750	0.4421	1.312	16 – 24	0.5740
NWPAC Win.	110	0.0026	3.280	23 – 17	0.5623
NEPAC Sum.	51	0.0002	5.375	20 – 15	0.5305
NEPAC Win.	134	0.5946	2.363	15 – 31	0.5777
NATL Sum.	855	0.6061	1.499	12 – 22	0.5180
NATL Win.	209	0.7101	1.190	15 – 29	0.5367

Table 8. As in Table 4, for stroke density versus average reflectivity.

Data	A	C
NWPAC Full	1.074 (0.5773, 1.5700)	17.56 (15.59, 19.53)
NEPAC Full	0.4748 (-0.8070, 1.757)	20.66 (16.31, 25.01)
NATL Full	0.3767 (0.0000, 0.7534)	17.19 (15.81, 18.58)
PB2009	0.7200	33.3

Table 9. As in Table 5, for maximum current versus average reflectivity.

Data	A	C
NWPAC Full	0.5031 (-0.2683, 1.274)	18.66 (14.68, 22.64)
NEPAC Full	0.6836 (-1.0270, 2.3950)	18.53 (10.85, 26.20)
NATL Full	0.2803 (-0.3405, 0.9012)	17.15 (14.07, 20.23)

FIGURES

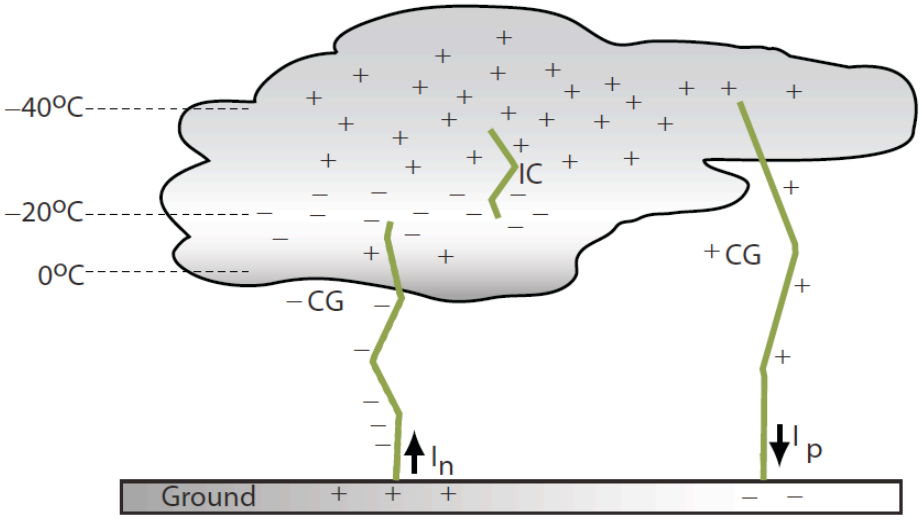


Fig. 1. Schematic of resulting charge centers in a typical thunderstorm after gravitation separation has occurred. Three species of lightning strokes are illustrated: negative cloud-to-ground (CG), positive CG, and intra-cloud (IC).

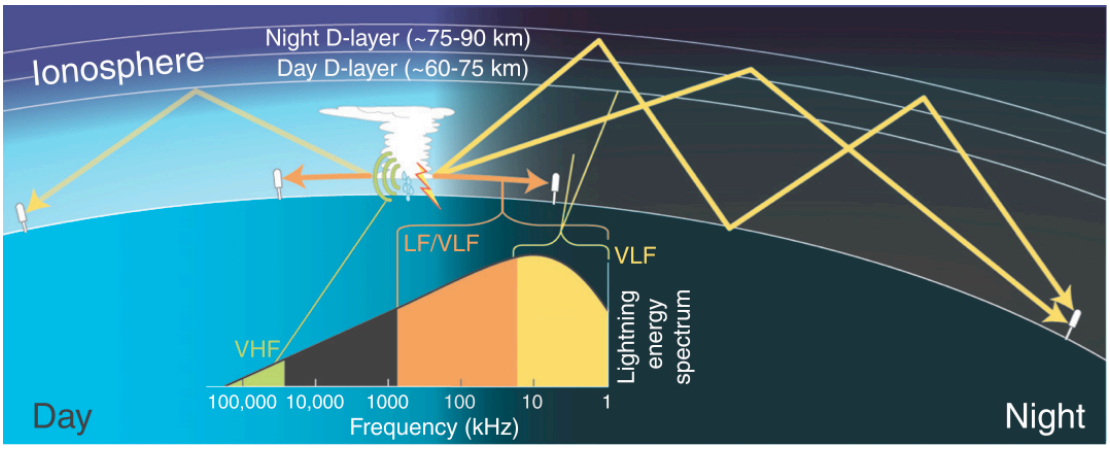


Fig. 2. "Schematic diagram of the earth-ionosphere waveguide, which allows VLF (3-30 kHz) emissions from thunderstorms (sferics) to propagate thousands of kilometers through reflection. The best propagation is observed over the ocean at night" (*Pessi et al., 2009*).

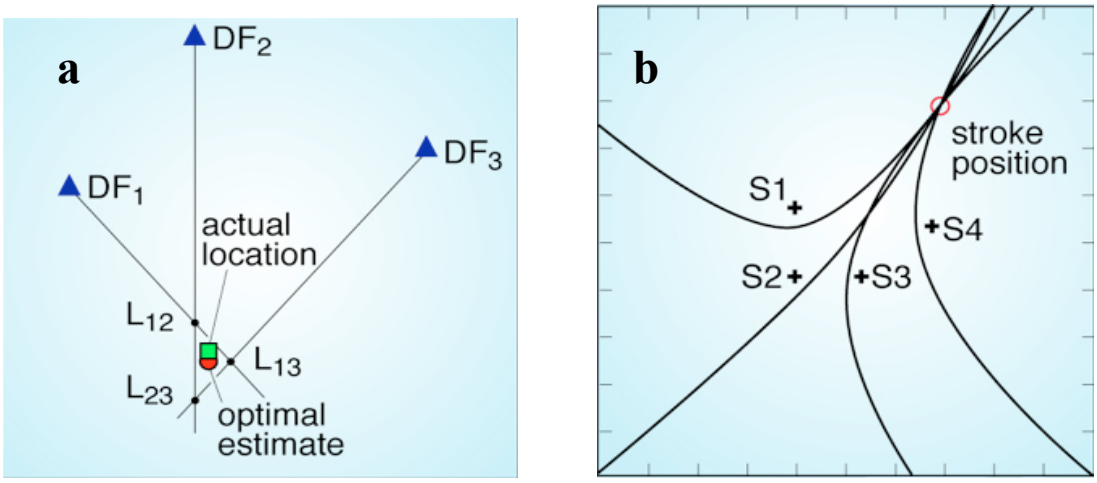


Fig. 3. Schematic diagrams illustrating a)magnetic direction finding, b) time-of-group-arrival methods for lightning stroke location (*Cummins et al. 2000*).

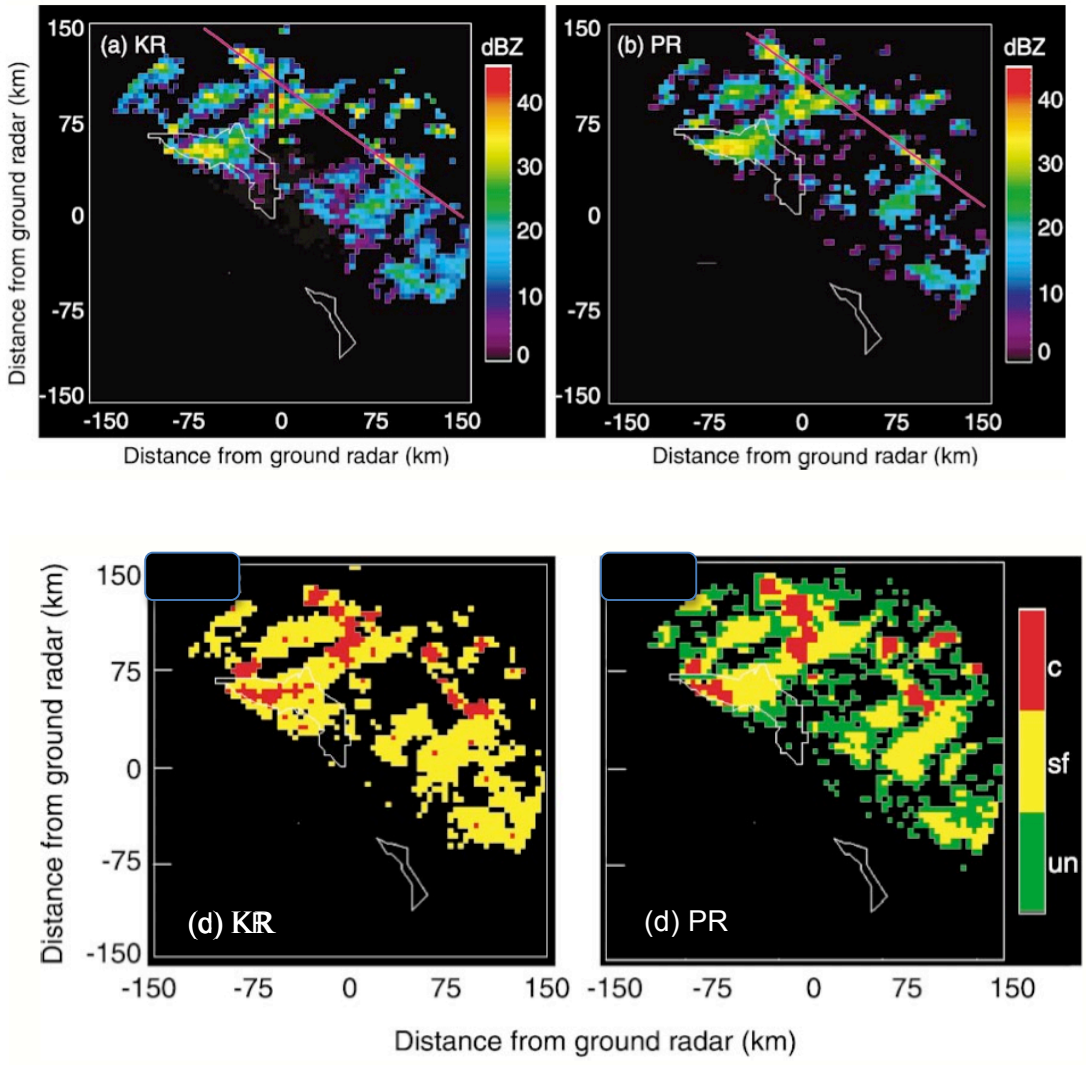


Fig. 4. Reflectivity at a height of 3 km as observed by a) the Kwajalein S-band Doppler radar and b) by TRMM PR. Convective precipitation (red) and stratiform precipitation (yellow) at 3 km are indicated by c) the Kwajalein radar and d) the TRMM PR. Areas where the precipitation type is unknown are shown in green for TRMM PR (*Schumacher and Houze 2000*).

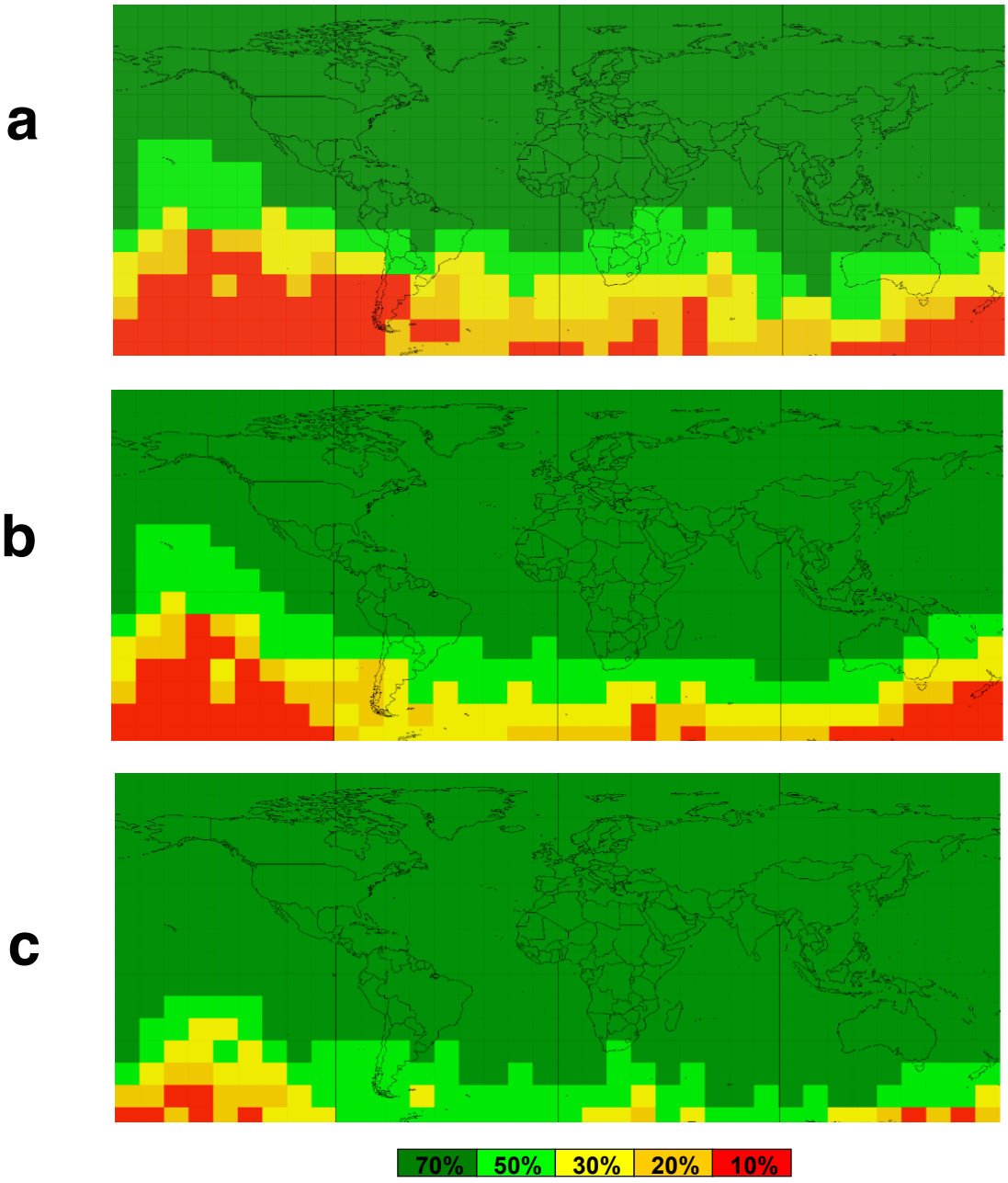


Fig. 5. GLD360 estimated stroke DE across the globe (a) from January 2011 – June 2011, (b) for January 2012, (c) and that projected for September 2012 (*Vaisala*).

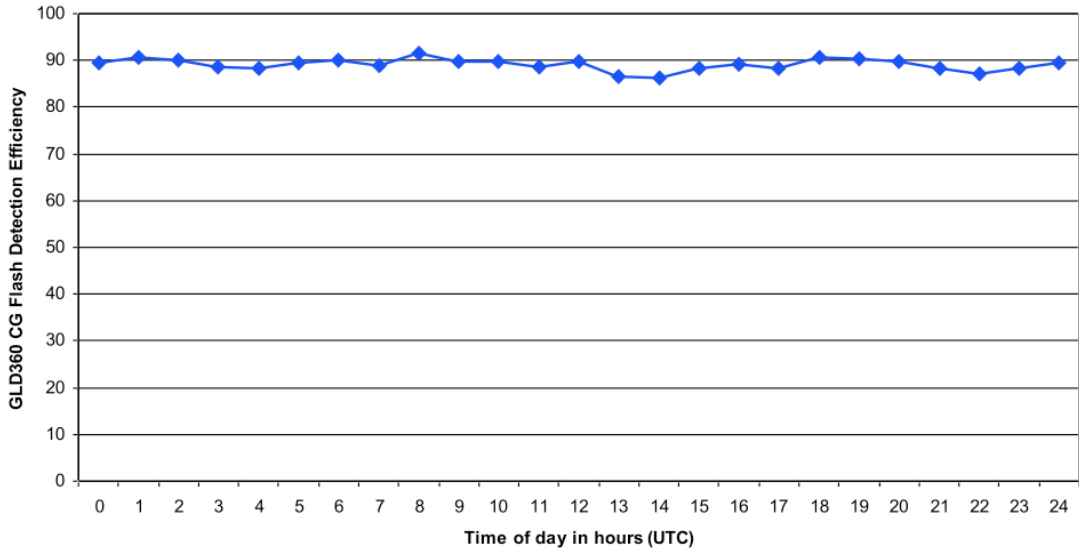


Fig. 6. GLD360's DE as a function of the time of day during reported during a validation study using NLDN as ground truth (*Vaisala*).

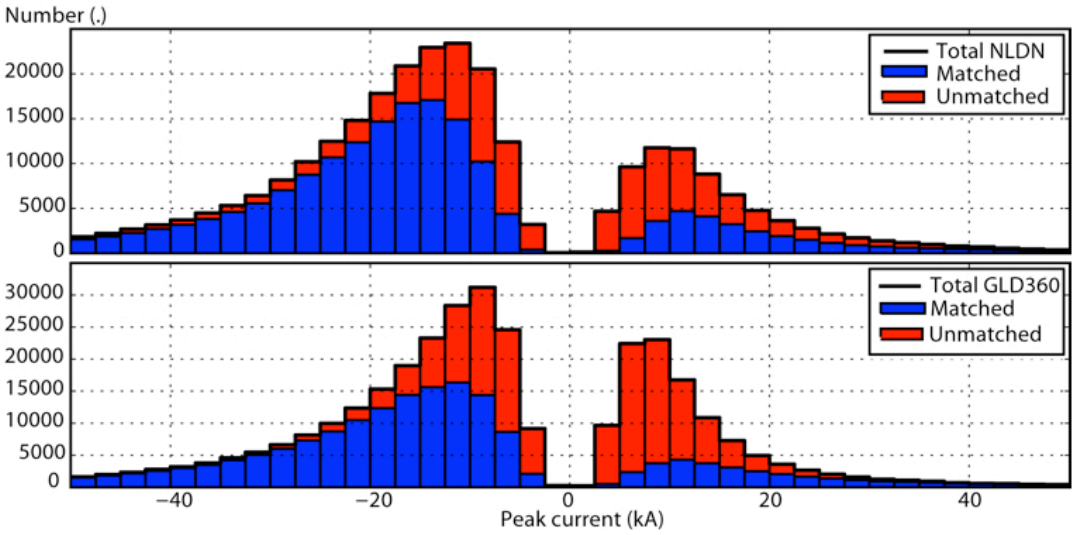
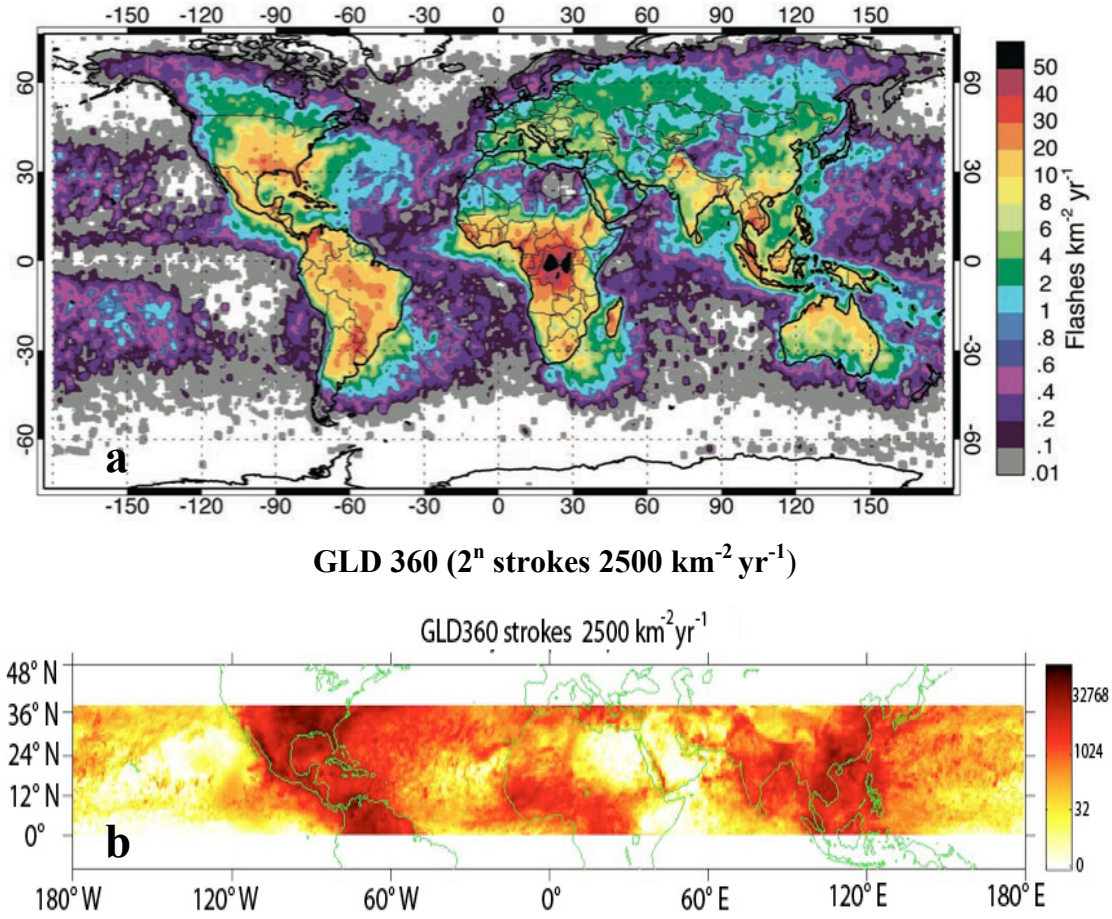


Fig. 7. The numbers of observed (matched/unmatched) flashes as a function of maximum current for both NLDN and GLD360 during a validation study over the continuous United States (*Vaisala*).



GLD 360 (2^n strokes 2500 km^{-2} yr^{-1})

Fig. 8. a) Global climatology of lightning observed by NASA's optical transient detector (*Christian et al. 2003*) and b) a preliminary climatology (January 2010 – December 2010) of CG lightning as observed by GLD360 across the domains of interest.

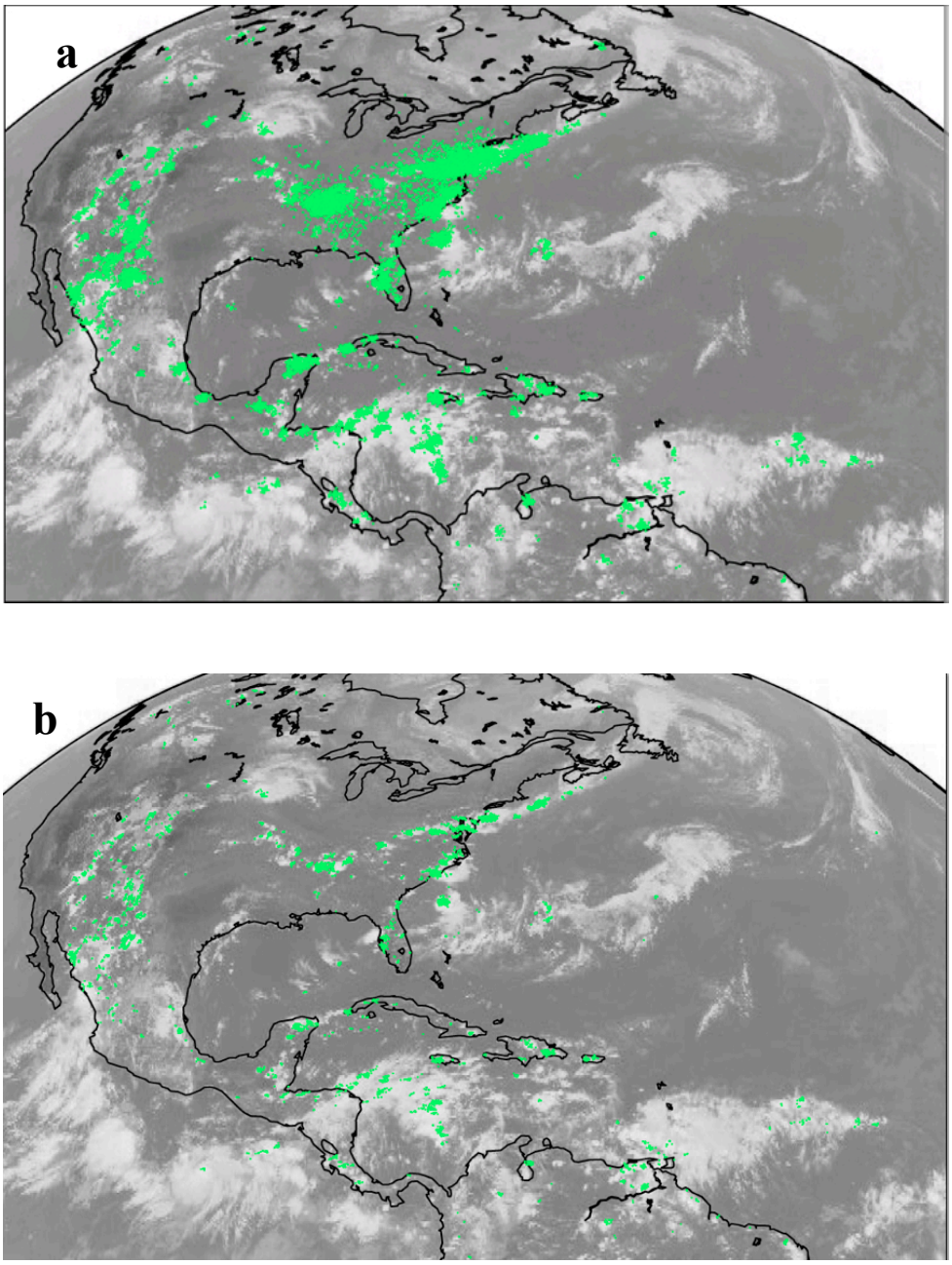


Fig. 9. Geostationary Observation Environmental Satellite infrared imagery with GLD360 flashes overlain (green) a) before reprocessing and b) after reprocessing (*Vaisala*).

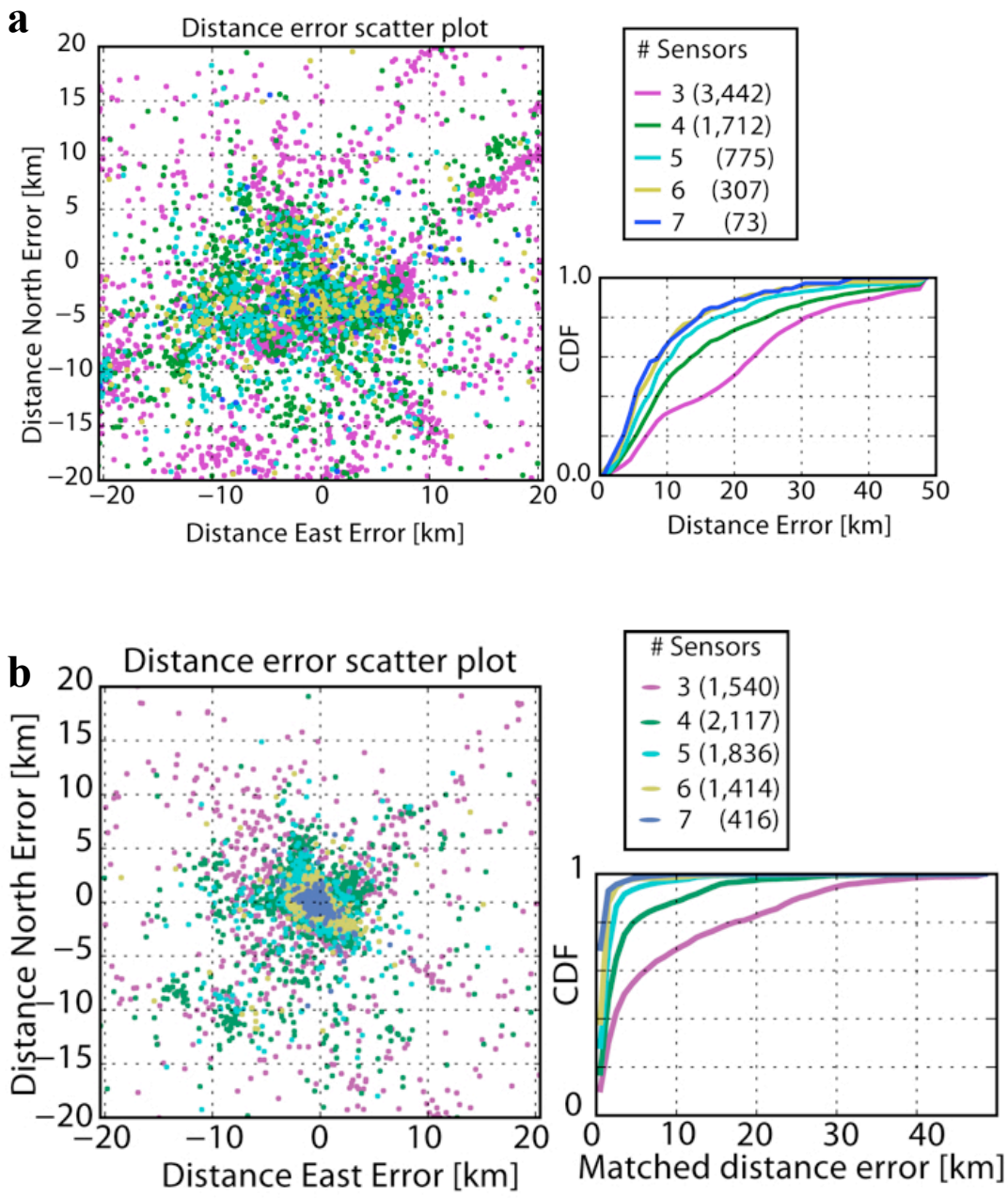


Fig. 10 A visual summary of GLD360's location accuracy a) before reprocessing and b) after reprocessing (*Vaisala*).

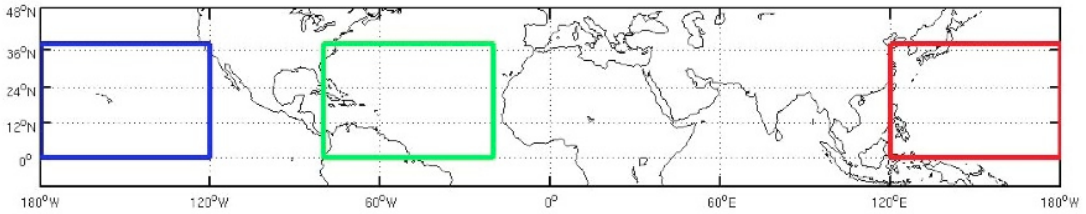


Fig. 11. The three oceanic domains studied. (NWPAC 0° – 38°N / 120°E – 180°E; NEPAC 0° – 38°N / 180°W – 120°W; NATL 0° – 38°N / 80°W – 20°W).

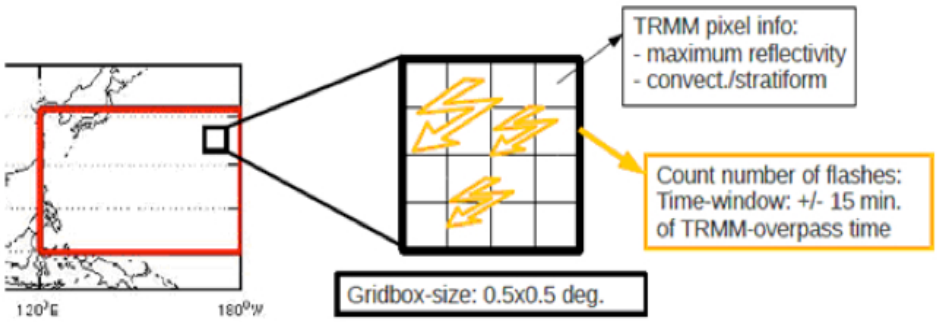


Fig. 12. A schematic diagram of the procedure for this study.

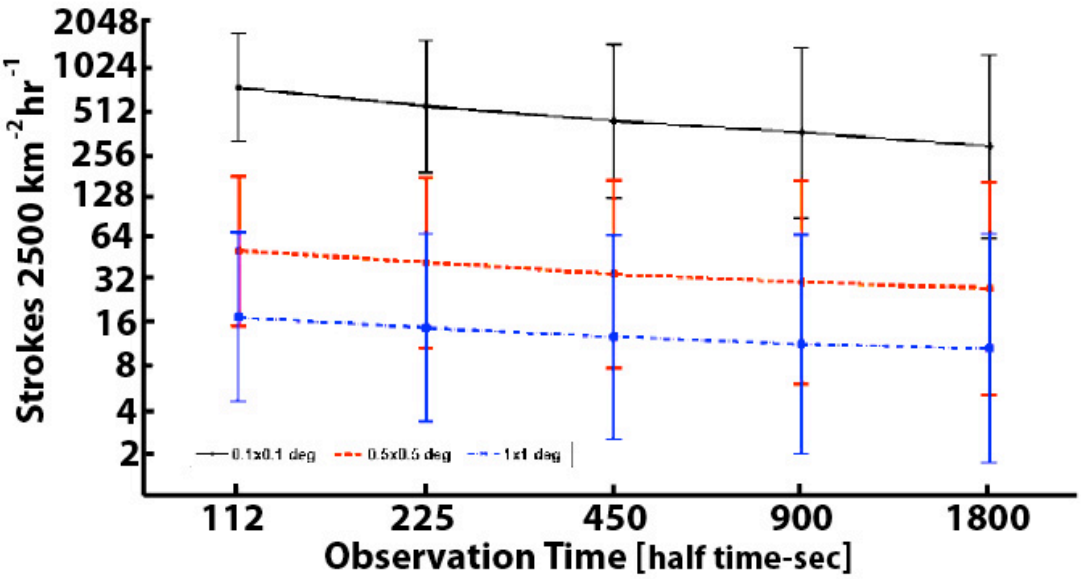


Fig. 13. Mean GLD360 stroke density (+/- σ) as a function of observation time interval for three different grid sizes (0.1°, 0.5°, 1.0°) for January – December 2010.

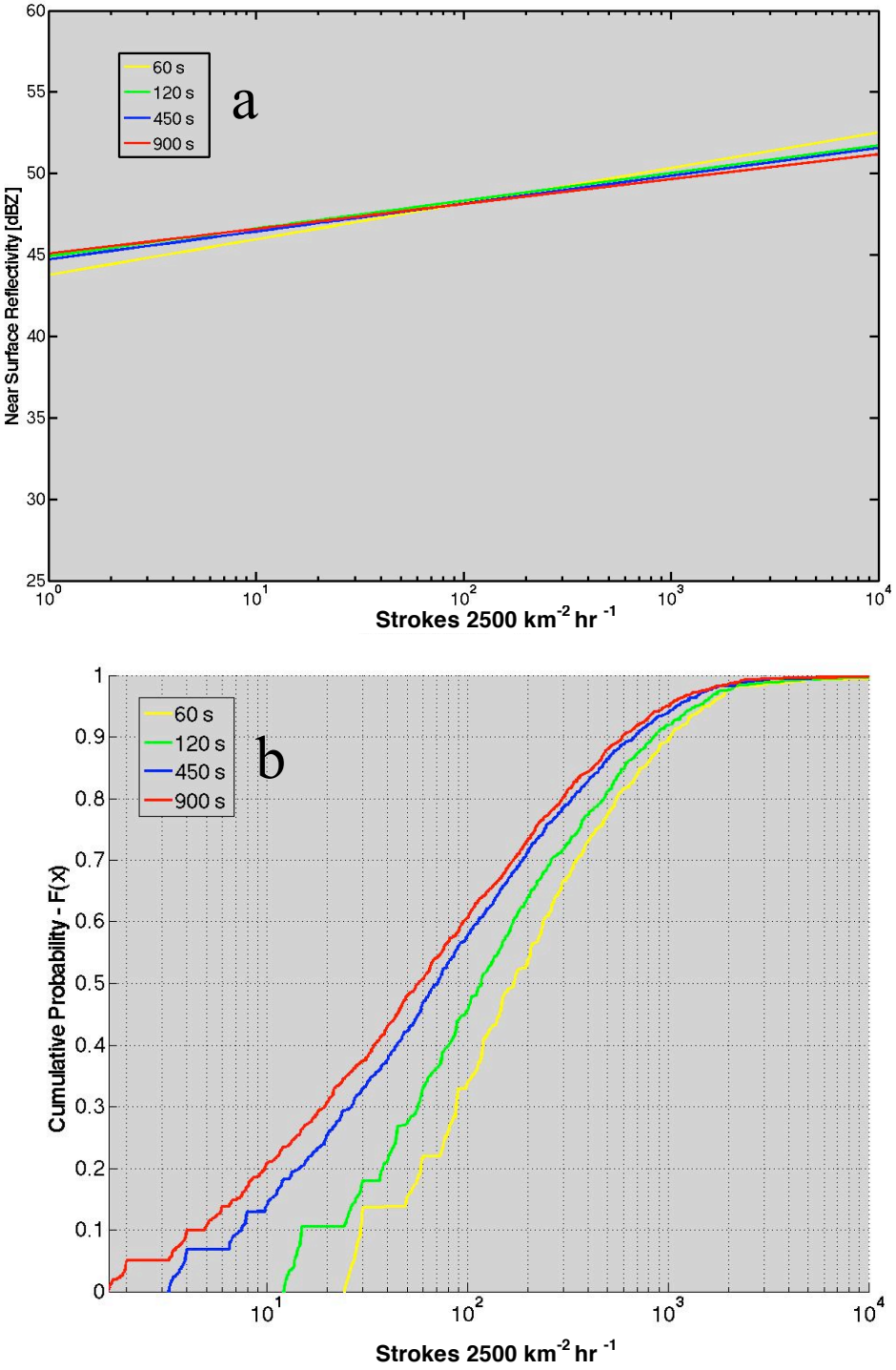


Fig. 14. a) GLD360 (May 2011 – August 2011) CG stroke density versus near-surface reflectivity and b) cumulative distribution functions for four different time parameter choices: 60 s, 120 s, 450 s, and 900 s. The time values indicate half of the total observation time (2 minutes, 4 minutes, 15 minutes, and 30 minutes respectively).

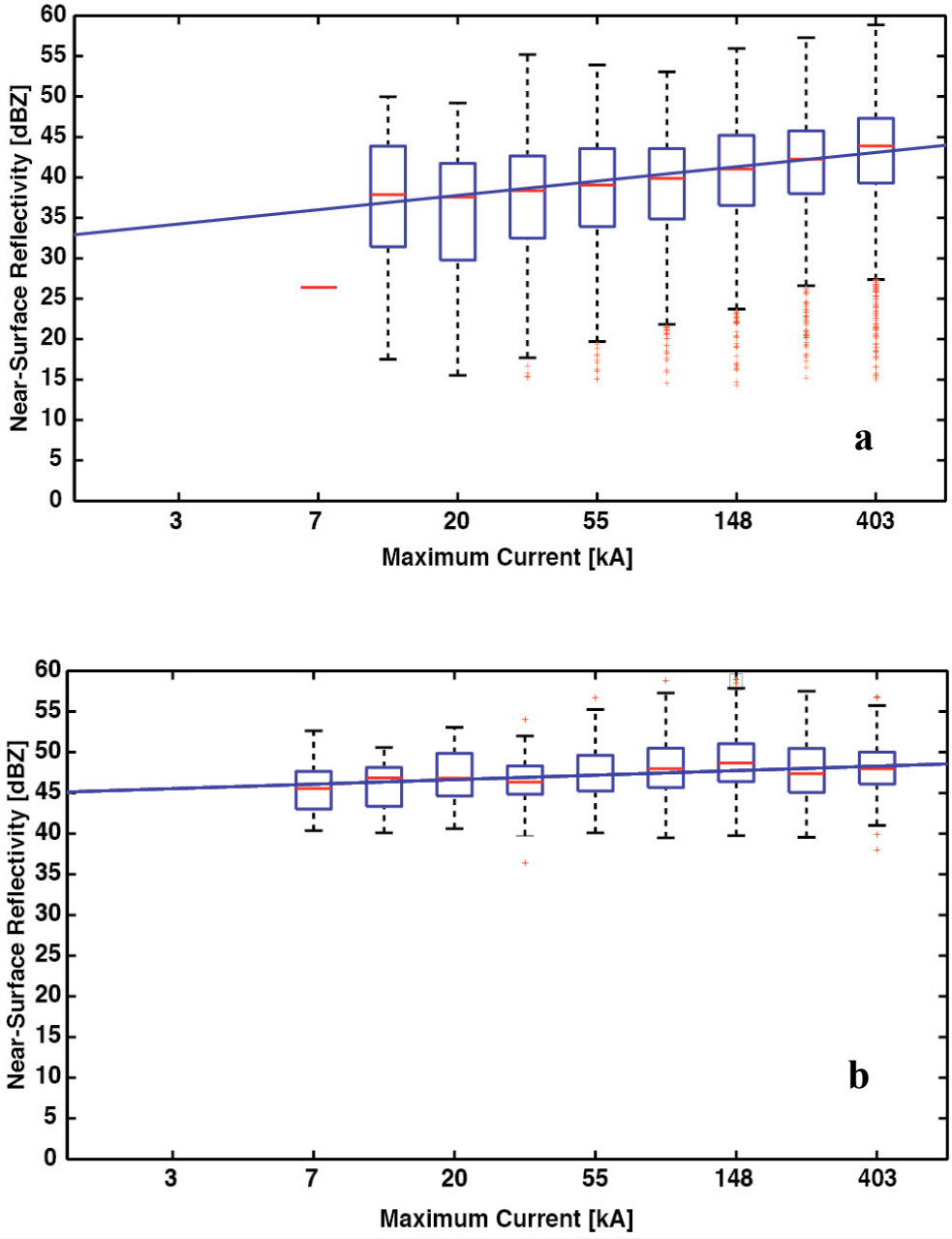


Fig. 15. The maximum current amplitude observed by GLD360 versus near-surface reflectivity a) before b) and after a convective area threshold has been applied (25% of the TRMM PR observations in a grid box area must be convective).

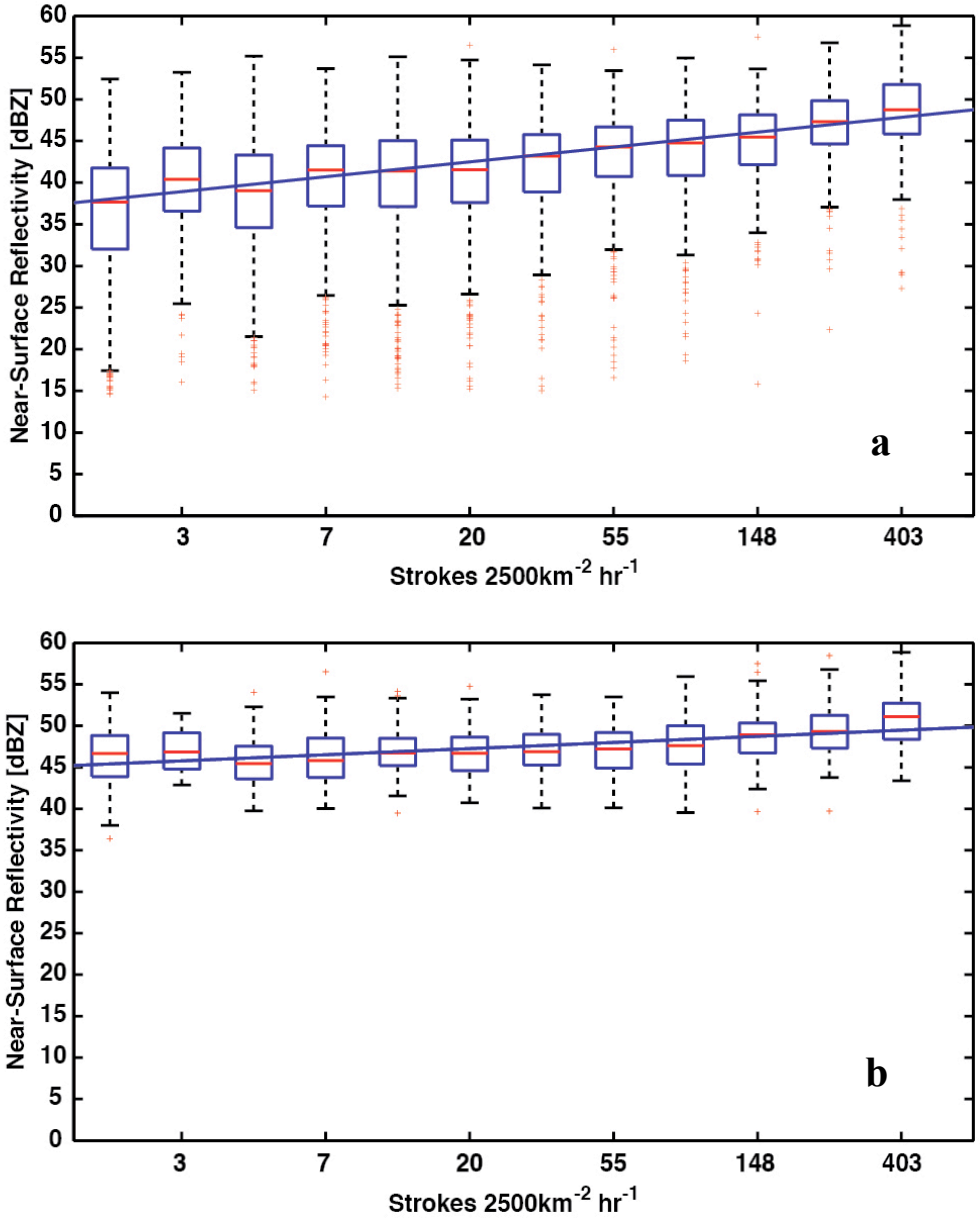


Fig. 16. The stroke density observed by GLD360 versus near-surface reflectivity a) without and b) with a convective area threshold (25% of the TRMM PR observations in a grid box area must be convective).

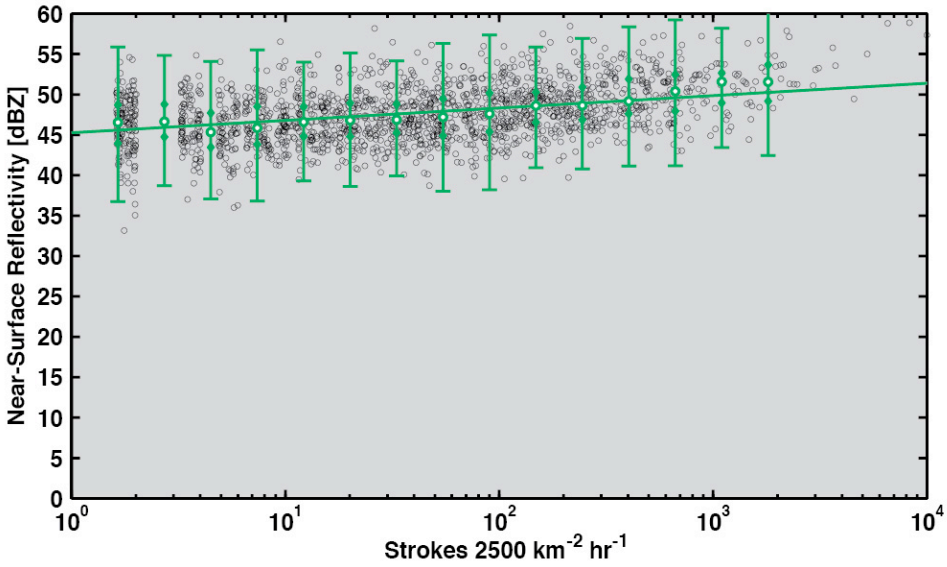


Fig. 17. GLD360 stroke density versus TRMM PR maximum reflectivity over the NATL domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in green. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

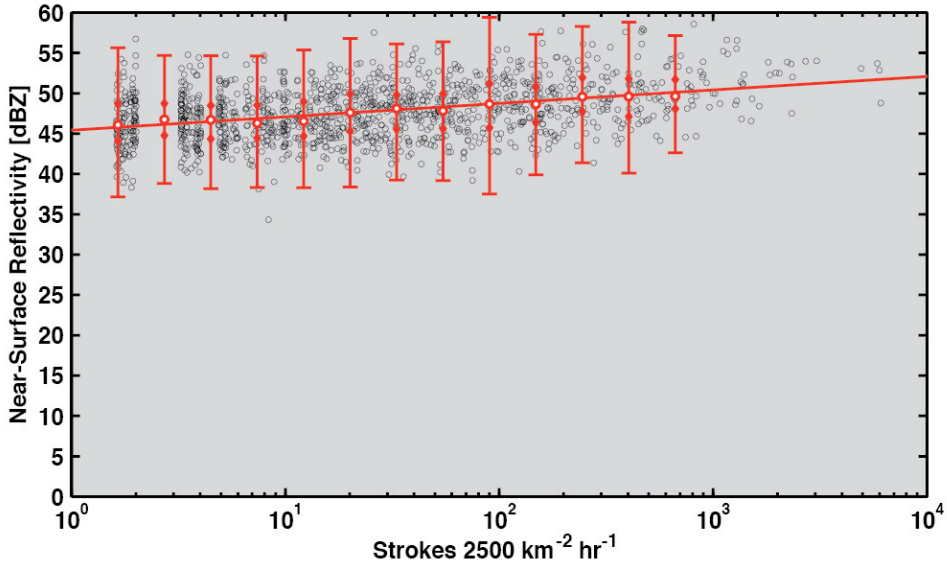


Fig. 18. GLD360 stroke density versus TRMM PR maximum reflectivity over the NWPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in red. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

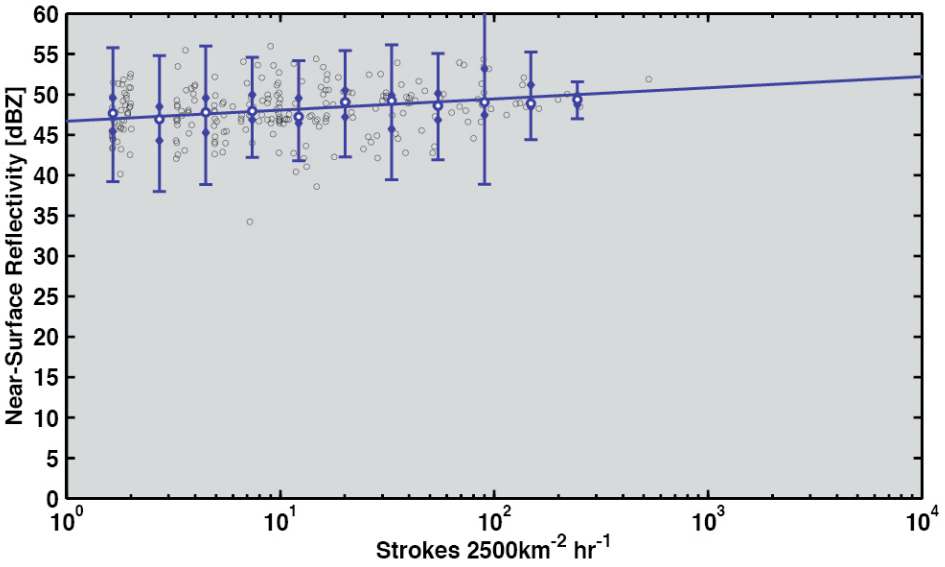


Fig. 19. GLD360 stroke density versus TRMM PR maximum reflectivity over the NEPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in blue. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

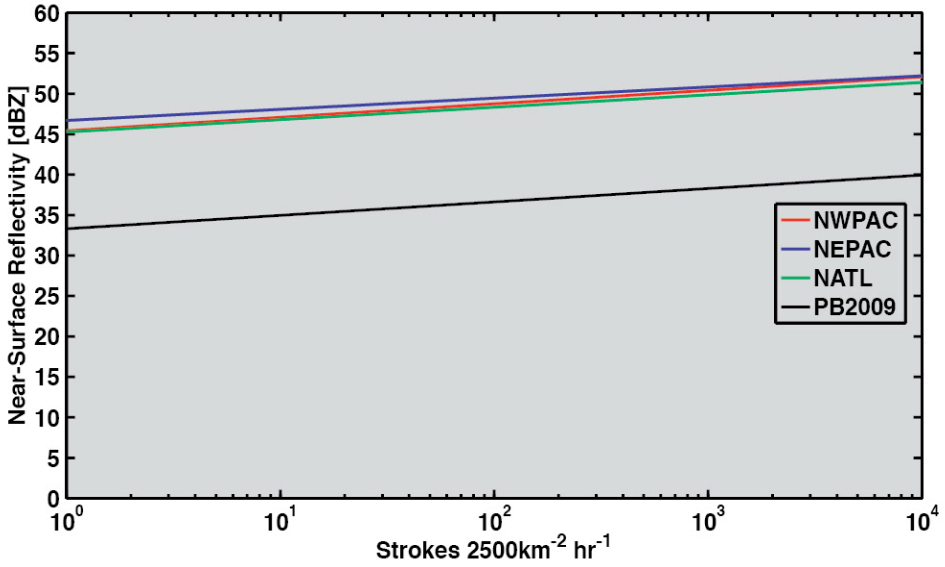


Fig. 20. The resultant logarithmic least-squares fit to bin medians of NWPAC, NEPAC, and NATL of maximum reflectivity versus GLD360 stroke density for the full time period, May 2011 – February 2012. The results of PB2009 (solid black line) are included for comparison.

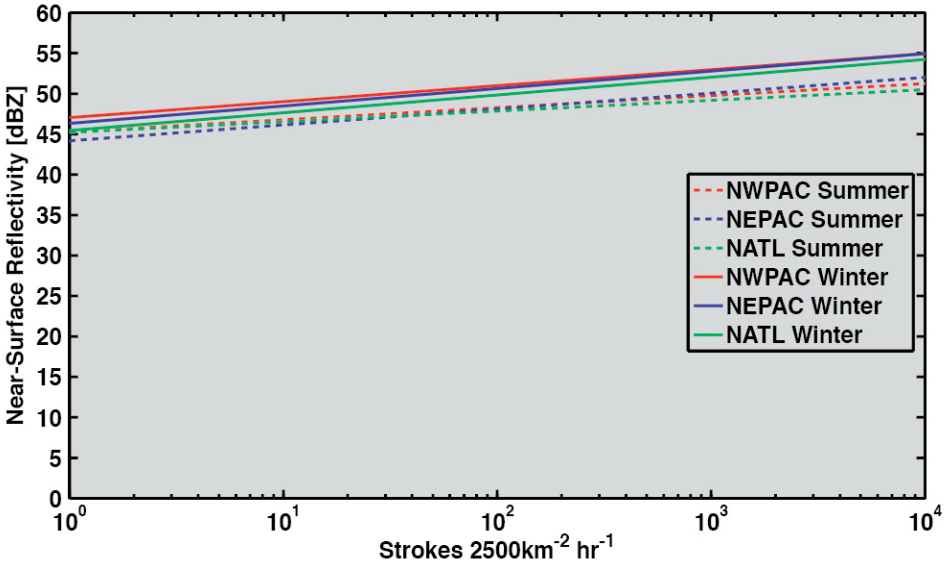


Fig. 21. As in figure 20, but the results separated by season. The summer subset consists of data from June, July and August of 2011. Data from December 2011, January 2012, and February 2012 constitute the winter subset.

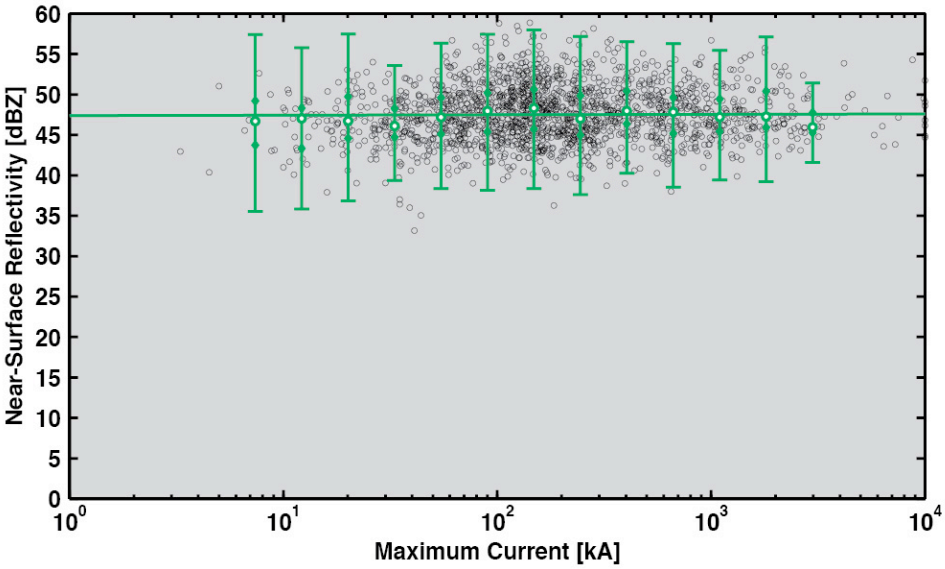


Fig. 22. GLD360 maximum current versus TRMM PR maximum reflectivity over the NATL domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in green. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

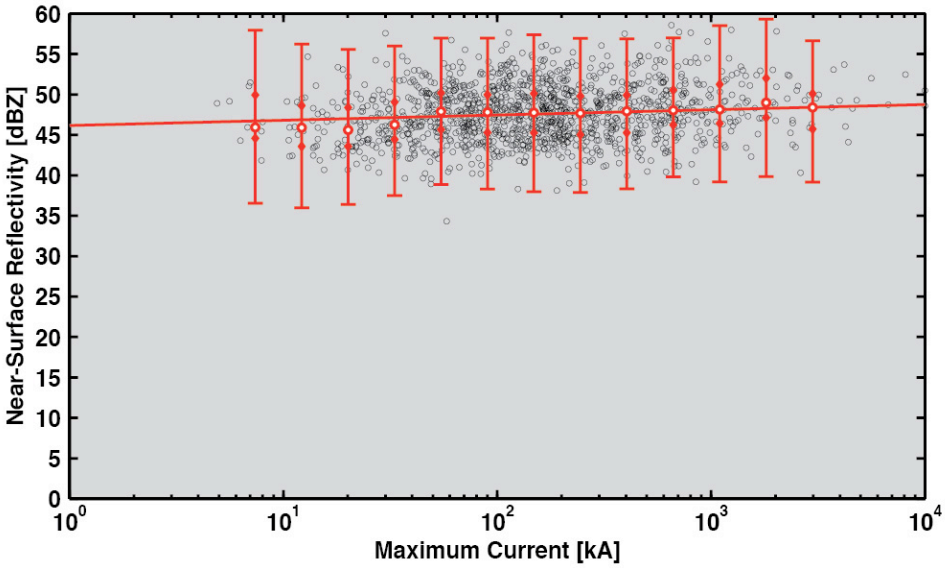


Fig. 23. GLD360 maximum current versus TRMM PR maximum reflectivity over the NWPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in red. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

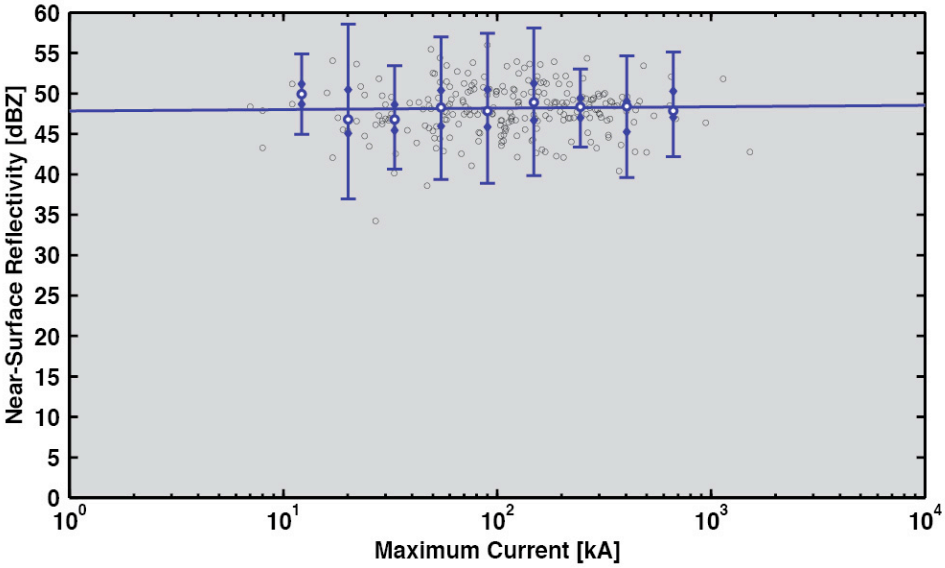


Fig. 24. GLD360 maximum current versus TRMM PR maximum reflectivity over the NEPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in blue. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

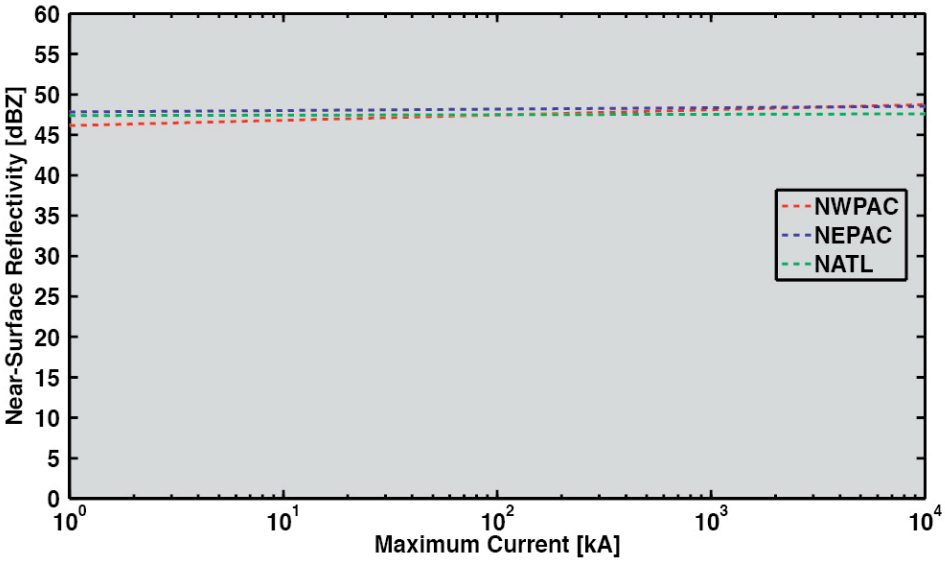


Fig. 25. The resultant logarithmic least-squares fit to bin medians of NWPAC, NEPAC, and NATL near-surface versus GLD360 maximum current for the full time period, May 2011 – February 2012.

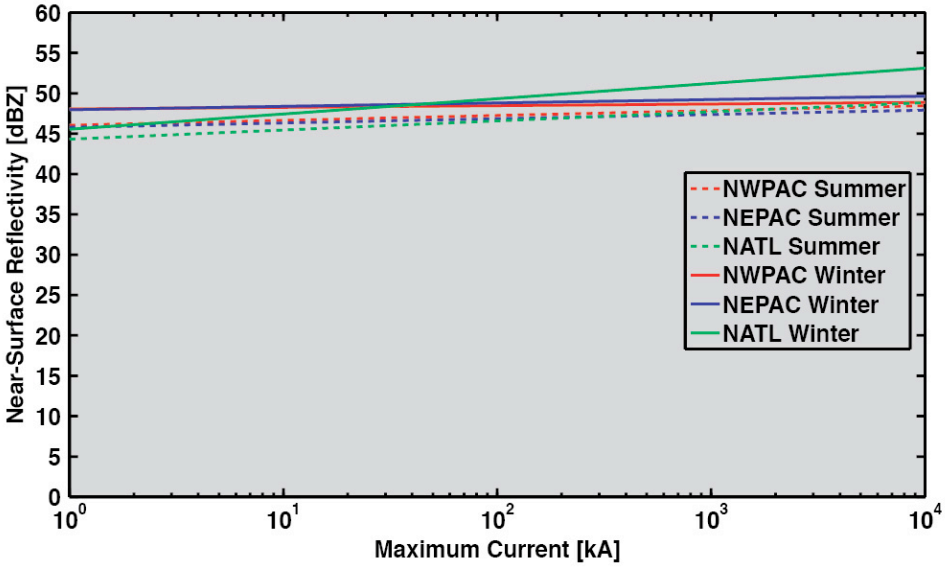


Fig. 26. As in figure 26, but the results separated by season. The summer subset consists of data from June, July and August of 2011. Data from December 2011, January 2012, and February 2012 constitute the winter subset.

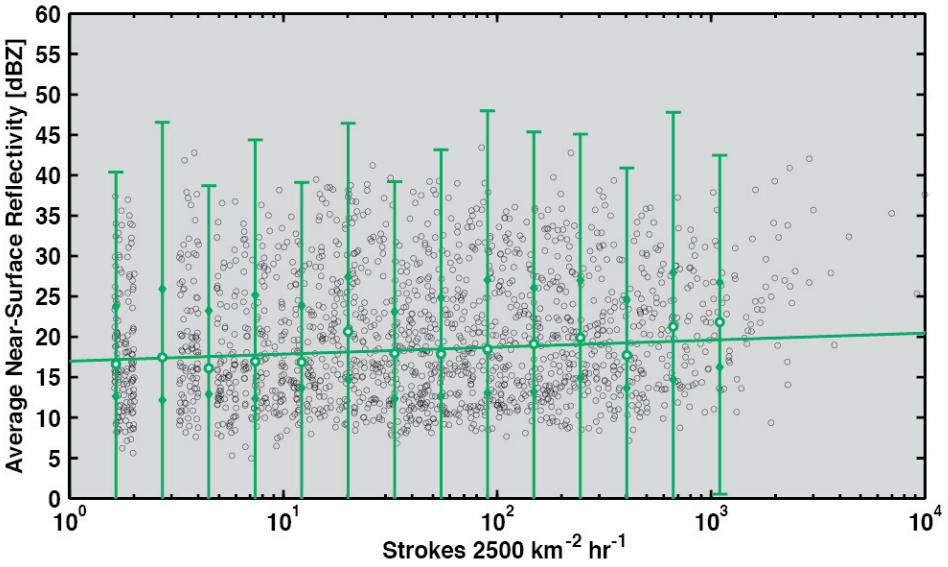


Fig. 27. GLD360 stroke density versus TRMM PR average reflectivity over the NATL domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in green. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

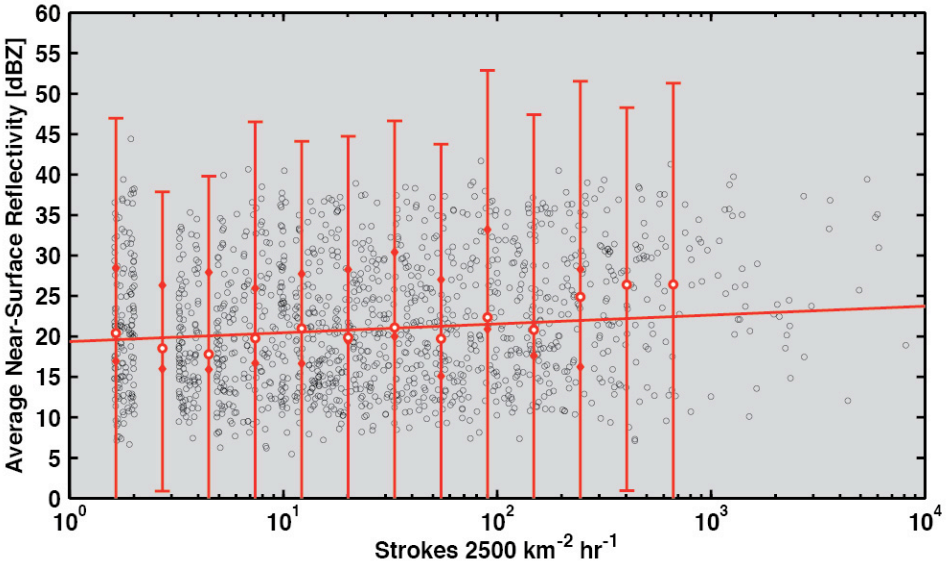


Fig. 28. GLD360 stroke density versus TRMM PR average reflectivity over the NWPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in red. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

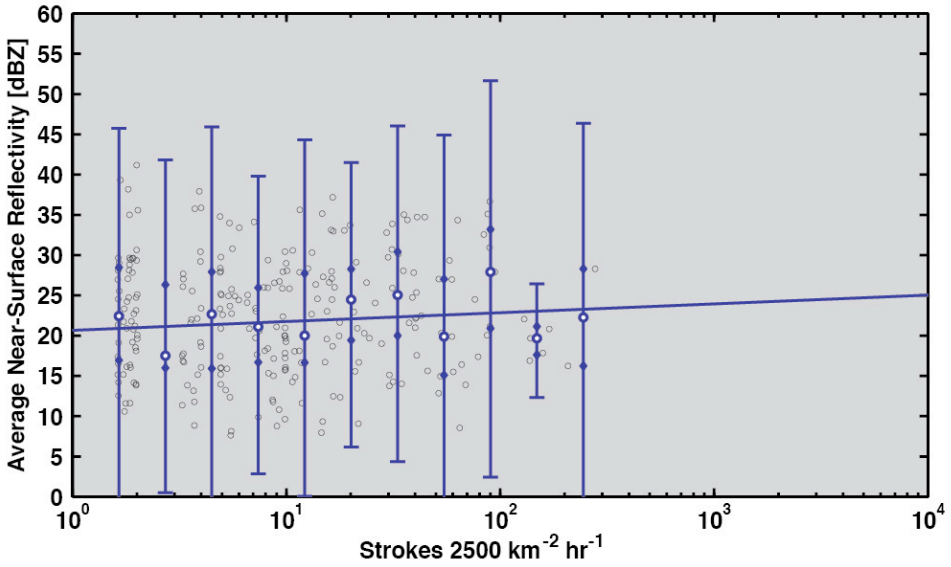


Fig. 29. GLD360 stroke density versus TRMM PR maximum reflectivity over the NEPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in blue. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

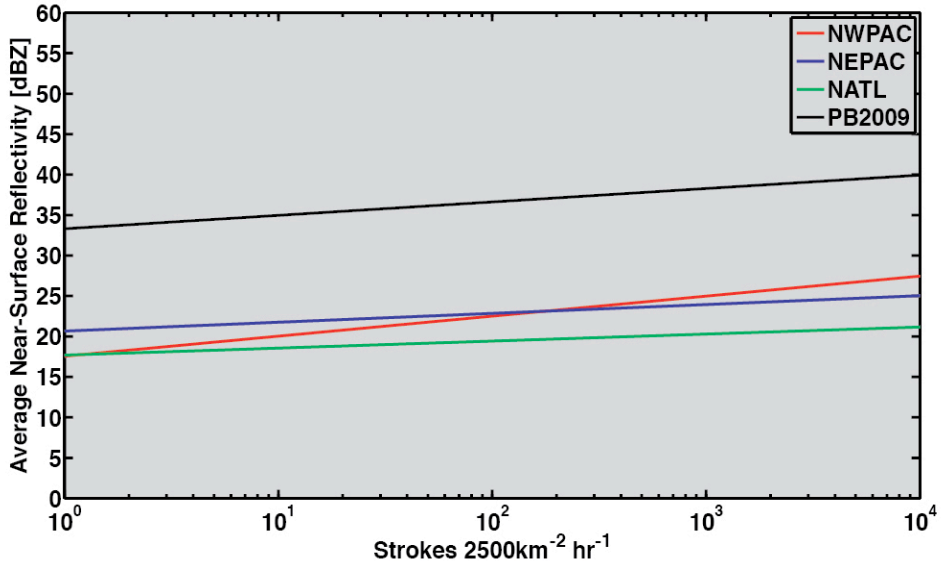


Fig. 30. The resultant logarithmic least-squares fit to bin medians of NWPAC, NEPAC, and NATL average near-surface versus GLD360 stroke density for the full time period, May 2011 – February 2012.

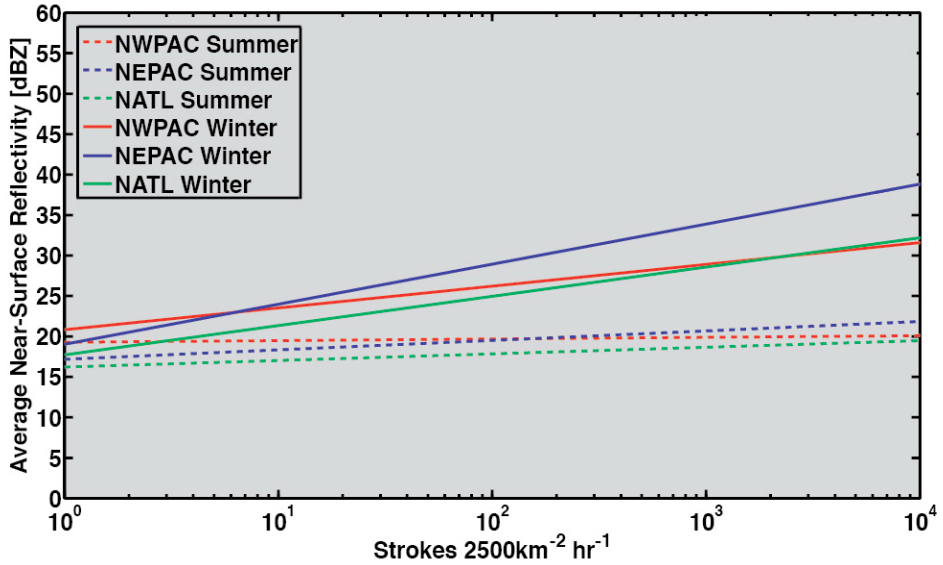


Fig. 31. As in figure 31, but the results separated by season. The summer subset consists of data from June, July and August of 2011. Data from December 2011, January 2012, and February 2012 constitute the winter subset.

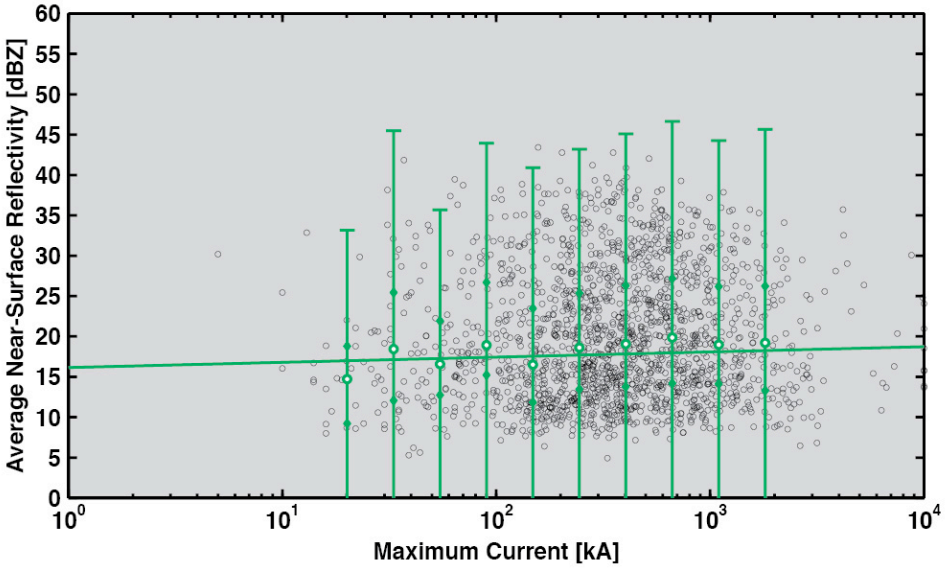


Fig. 32. GLD360 maximum versus TRMM PR average reflectivity over the NATL domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in green. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

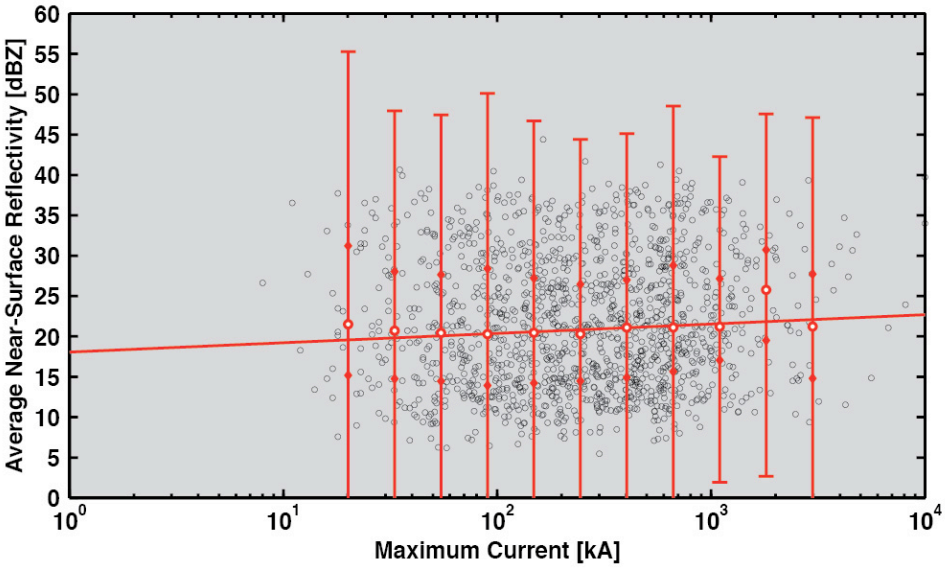


Fig. 33. GLD360 maximum current versus TRMM PR average reflectivity over the NWPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in red. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

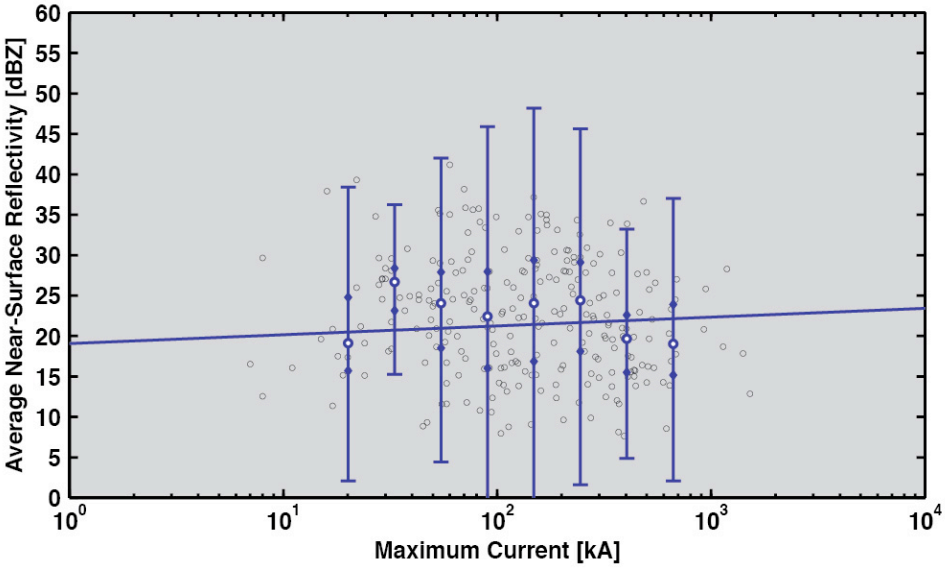


Fig. 34. GLD360 maximum current versus TRMM PR average reflectivity over the NEPAC domain from May 2011 – February 2012. The least-squares logarithmic fit to the bin medians (outlined circles) is overlaid in blue. The upper and lower quartiles are denoted by diamonds while the error bars cover approximately 99% of the data within a given stroke density bin.

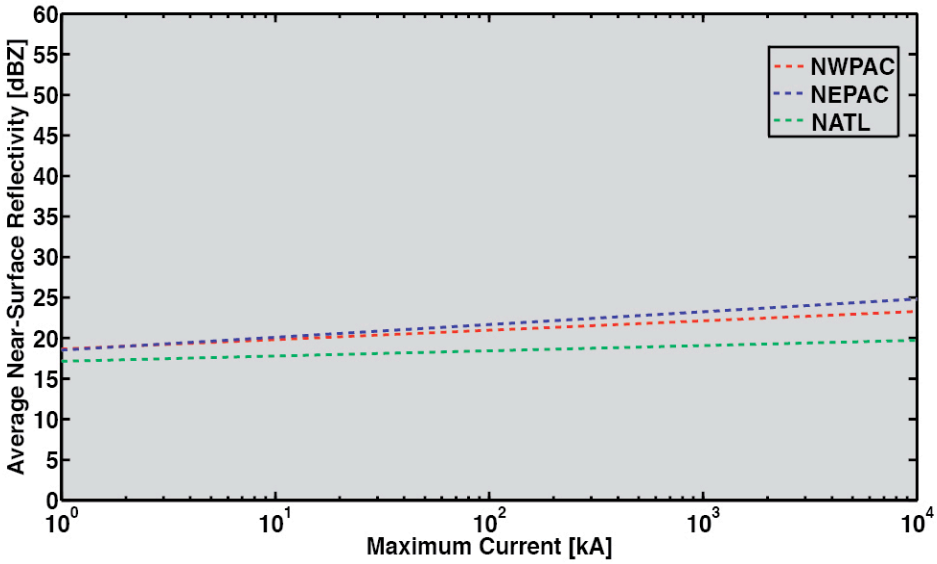


Fig. 35. The resultant logarithmic least-squares fit to bin medians of NWPAC, NEPAC, and NATL average reflectivity versus GLD360 maximum current for the full time period, May 2011 – February 2012.

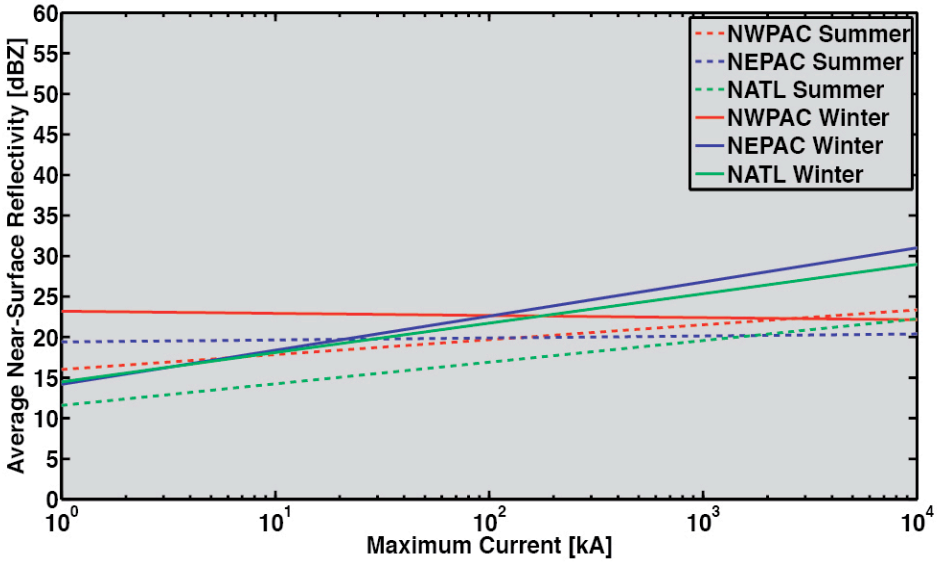


Fig. 36. As in figure 35, but the results separated by season. The summer subset consists of data from June, July and August of 2011. Data from December 2011, January 2012, and Februrary 2012 constitute the winter subset.

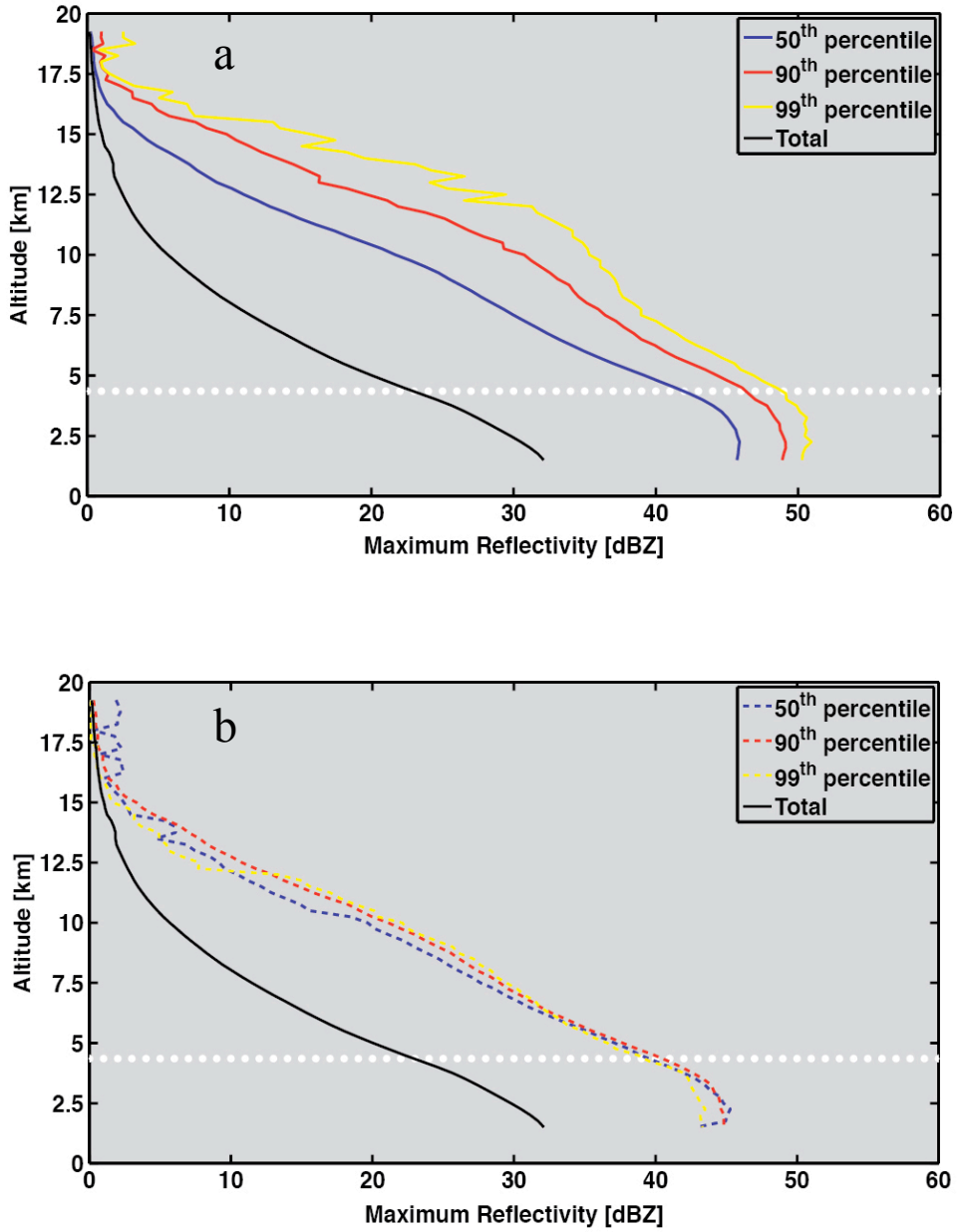


Fig. 37. The maximum attenuation-corrected reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NATL. The horizontal dotted white line indicates the mean height of the freezing level in the NATL (3772 m). The 50th, 90th, and 99th percentiles for stroke density are 35 strokes 2500 km^{-2} hr^{-1}, 450 strokes 2500 km^{-2} hr^{-1}, and 2008 strokes 2500 km^{-2} hr^{-1} respectively. The 50th, 90th, and 99th percentiles for maximum current are 302 kA, 1155 kA, and 3984 kA respectively.

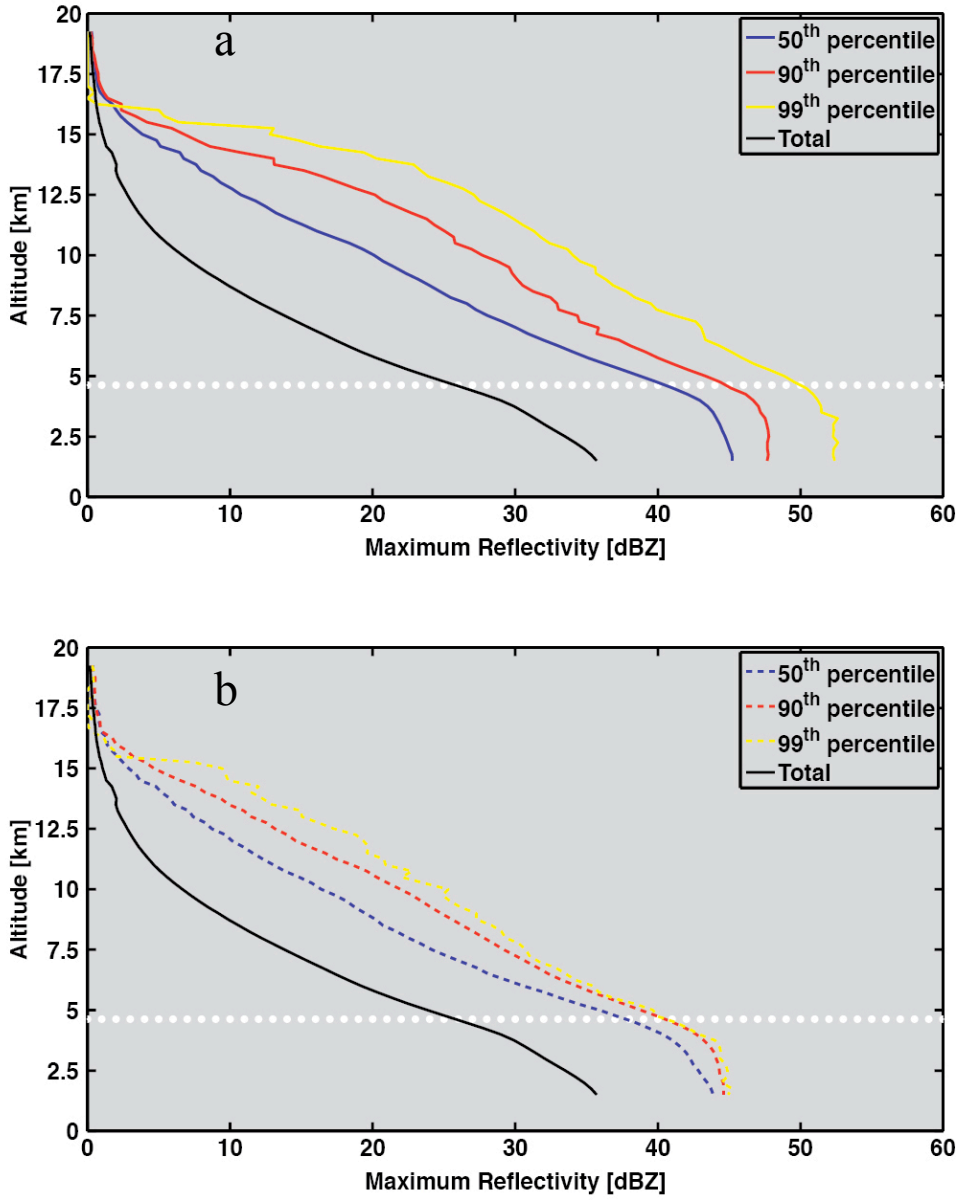


Fig. 38. The maximum attenuation-corrected reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NWPAC. The horizontal dotted white line indicates the mean height of the freezing level in the NWPAC (4625 m). The 50th, 90th, and 99th percentiles for stroke density are 22 strokes 2500 km^{-2} hr $^{-1}$, 293 strokes 2500 km^{-2} hr $^{-1}$, and 1825 strokes 2500 km^{-2} hr $^{-1}$ respectively. The 50th, 90th, and 99th percentiles for maximum current are 150 kA, 747 kA, and 3391 kA respectively.

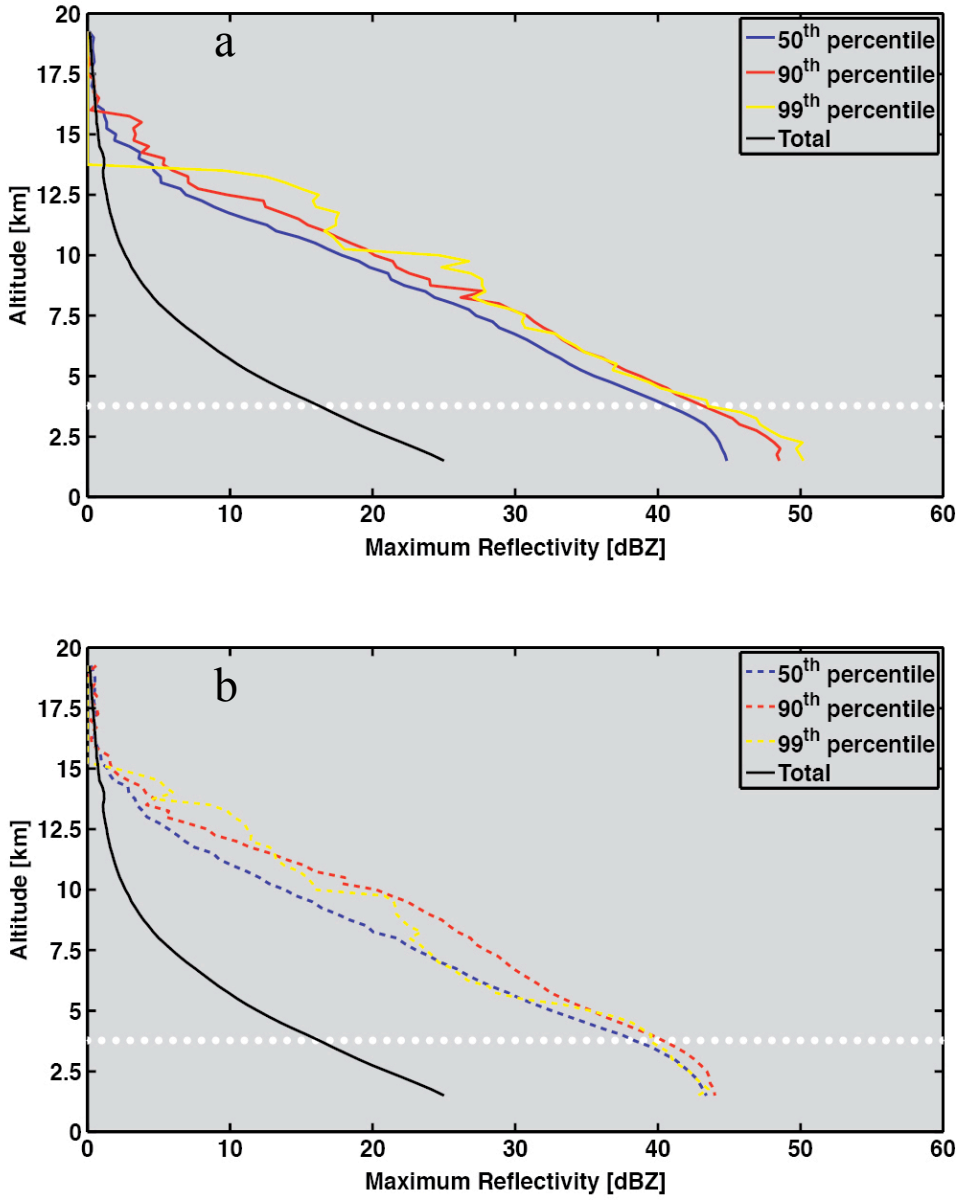


Fig. 39. The maximum attenuation-corrected reflectivity associated with the total, 50th, 90th, and 99th percentiles of a) GLD360 stroke density and b) GLD360 maximum current in the NEPAC. The horizontal dotted white line indicates the mean height of the freezing level in the NEPAC (3772 m). The 50th, 90th, and 99th percentiles for stroke density are 9 strokes 2500 km^{-2} hr $^{-1}$, 69 strokes 2500 km^{-2} hr $^{-1}$, and 222 strokes 2500 km^{-2} hr $^{-1}$ respectively. The 50th, 90th, and 99th percentiles for maximum current are 115 kA, 337 kA, and 938 kA respectively.

REFERENCES

Abarca, S. F., K. L. Corbosiero, and T. J. Galarneau Jr, 2010: An evaluation of the Worldwide Lightning Location Network (WWLLN) using the National Lightning Detection Network (NLDN) as ground truth. *J. Geophys. Res.*, 115, doi:10.1029/2009JD013411.

Albrecht, B. A., 1989: Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science*, 245 (4923), 1227-1230.

Alexander, G. D., J. A. Weinman, V. M. Karyampudi, W. S. Olson, and A. C. L. Lee, 1999: The effect of assimilating rain rates derived from satellites and lightning on forecasts of the 1993 superstorm. *Mon. Wea. Rev.*, 127, 1433–1457.

Carey, Lawrence D., Steven A. Rutledge, 2000: The Relationship between Precipitation and Lightning in Tropical Island Convection: A C-Band Polarimetric Radar Study. *Mon. Wea. Rev.*, 128, 2687–2710.

Cecil, Daniel J., S. J. Goodman, D. J. Boccippio, E. J. Zipser, S. W. Nesbitt, 2005: Three Years of TRMM PRecipitation Features. Part I: Radar, Radiometric, and Lightning Characteristics. *Mon. Wea. Rev.*, 133, 543–566.

Chang, D. E., J. A. Weinman, C. A. Morales, and W. S. Olson, 2001: The effect of spaceborne microwave and ground-based continuous lightning measurements on forecasts of the 1998 Groundhog Day storm. *Mon. Wea. Rev*., 129, 1809–1833.

Christian, H. J., R. J. Blakeslee, D. J. Boccippio, W. L. Boeck, D. E. Buechler, K. T. Driscoll, S. J. Goodman, J. M. Hall, W. J. Koshak, D. M. Mach, and M. F. Stewart, 2003: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector. *J. Geophys. Res.*, 108(D1), doi:10.1029/2002JD002347.

Clarke, A.D., T. Uehara, and J. N. Porter, 1997. Atmospheric nuclei and related aerosol fields over the Atlantic: Clean subsiding air and continental pollution during ASTEX, *J. Geophys. Res.*, 102, 25281-25292.

Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer, 1998: A combined TOA/MDF technology upgrade of the U.S. national lightning detection network, *J. Geophys. Res.*, 103 (D8), 9035–9044, doi:10.1029/98JD00153.

Cummins, K. L., M. J. Murphy, J. V. Tuel, 2000: Lightning Detection Methods and Meteolorogical Applications. *IV International Symposium on Military Meteorology*, September 26-28, Malbork, Poland, 13 pp.

Demetriades, N.W .S., M.J. Murphy, and J.A. Cramer, 2010: Validation of Vaisala's Global Lightning Dataset (GLD360) over the continental United States. *Preprints, 29th Conf. Hurricanes and Tropical Meteorology*, May 10-14, Tucson, AZ, 6 pp.

Futyan, J. M., and A. D. Del Genio, 2007a: Deep convective system evolution over Africa and the tropical Atlantic. *J. Climate*, 20, 5041–5060.

Global Lightning Network. Weather Services International. 20 Jan. 2012 http://www.uspln.com/documentation/GLN%20Data%20Sheet%202011.pdf

Grogan, M. and Demetriades, N.W.S., 2009: Vaisala Global Lightning Dataset GLD360: Technology, Operations, and Applications Overview. Vaisala, Inc.

Hitschfeld, W. and J. Bordan, 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths, *J. Meteor.*, 11, 58-67.
Iguchi, T., R. Meneghini, 1994: Intercomparison of Single-Frequency Methods for Retrieving a Vertical Rain Profile from Airborne or Spaceborne Radar Data. *J. Atmos. Oceanic Technol.*, 11, 1507–1516.
Illingworth, A. J., and J. M. Caranti, 1985: Ice Conductivity Restraints on the Inductive Theory of Thunderstorm Electrification, *J. Geophys. Res.*, 90, 6033–6039.
Jayaratne, E.R., C.P.R. Saunders, and J. Hallet, 1983: Laboratory studies of the charging of soft-hail during ice crystal interactions. *Quart. J. Roy. Meteor. Soc.*, 101,
227-234.
Krider, E. P., R. C. Noggle, A. E. Pifer, D. L. Vance, 1980: Lightning Direction-Finding Systems for Forest Fire Detection. *Bull. Amer. Meteor. Soc.*, 61, 980–986.
Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM PR) sensor package. *J. Atmos. Oceanic Technol*., 15, 809-817.
Latham, J., 1981: The electrification of thunderstorms. *Quart. J. Roy. Meteor. Soc.*, 107, 277-298.
LeMone, Margaret A., E. J. Zipser, 1980: Cumulonimbus Vertical Velocity Events in GATE. Part I: Diameter, Intensity and Mass Flux. *J. Atmos. Sci.*, 37, 2444–2457.
Lewis, E. A., R. B. Harvey, and J. E. Rasmussen, 1960: Hyperbolic direction finding with sferics of transatlantic origin, *J. Geophys. Res*., 65, 1879–1905.
MacGorman, D. R., and W. D. Rust, 1998: *The Electrical Nature of Storms*. Oxford University Press, 422 pp.
Marshall, T. C., M. Stolzenburg, C. R. Maggio, L. M. Coleman, P. R. Krehbiel, T. Hamlin, R. J. Thomas, and W. Rison (2005), Observed electric fields associated with lightning initiation. *Geophys. Res. Lett*., 32, L03813, doi:10.1029/2004GL021802.
McMurdie, L., and C. Mass, 2004: Major numerical forecast failures over the Northeast Pacific. *Wea. Forecasting*, 19, 338-356.
Nelson, J. and Baker, M., 2003: Charging of ice-vapor interfaces: applications to thunderstorms, *Atmos. Chem. Phys*., 3, 1237-1252.
Nesbitt, S. W., E. J. Zipser, and D. J. Cecil, 2000: A census of precipitation features in the tropics using TRMM PR: Radar, ice scattering, and lightning observations. *J. Climate*, 13, 4087–4106.
Orville, Richard E., 2008: Development of the National Lightning Detection Network. *Bull. Amer. Meteor*. Soc., 89, 180–190.
Orville, Richard E., Ronald W. Henderson, 1986: Global Distribution of Midnight Lightning: September 1977 to August 1978. *Mon. Wea. Rev*., 114, 2640–2653.
Orville, Richard E., Ronald W. Henderson, Lance F. Bosart, 1983: An East Coast Lightning Detection Network. *Bull. Amer. Meteor. Soc*., 64, 1029–1037.
Papadopoulos, A., Themis G. C., Emmanouil N. A., 2005: Improving Convective

Precipitation Forecasting through Assimilation of Regional Lightning Measurements in a Mesoscale Model. *Mon. Wea. Rev*., 133, 1961–1977.

Pessi, A., and S. Businger, 2009a: Relationships Among Lightning, Precipitation, and Hydrometeor Characteristics over the North Pacific Ocean. *J. Appl. Meteor*. Clim., 48, 833-848.

Pessi, A., and S. Businger, 2009b: The Impact of Lightning Data Assimilation on a Winter Storm Simulation over the North Pacific Ocean. *Mon. Wea. Rev*., 137, 3177-3195.

Pessi, A., S. Businger, K. L. Cummins, N. W. S. Demetriades, M. Murphy, and B. Pifer, 2009: Development of a Long-Range Lightning Detection Network for the Pacific: Construction, Calibration, and Performance. *J. Atmos. Oceanic Technol*., 26, 145-166.

Petersen, W. A., Steven A. Rutledge, Richard E. Orville, 1996: Cloud-to-Ground Lightning Observations from TOGA COARE: Selected Results and Lightning Location Algorithms. *Mon. Wea. Rev*., 124, 602–620.

Porter, J., and Clarke, A., Aerosol Size Distribution Models Based on In-Situ Measurements. *J. Geophy. Res*., 102, 6035-6045, 1997.

Porter, J. N. Marine Aerosol: "Their Measurement and Influence on Cloud Base Properties." MS thesis. University of Hawaii - Manoa, 1988. Print.

Pruppacher, H.R. and Klett, J. D., 1998: Microphysics of Clouds and Precipitation. D. Reidel Publishing Company, Boston, USA., 71-826.

Reynolds, S. E., M. Brook, Mary Foulks Gourley, 1957: Thunderstorm Charge Separation. *J. Meteor*., 14, 426–436.

Rosenfeld, Daniel. "Modulation of Hail and Tornadic Storms By Aerosols." Fourth Symposium on Aerosol-Cloud-Climate Interactions. New Orleans, LA. 25 Jan 2012. Keynote Address.

Said, R. K., U. S. Inan, and K. L. Cummins, 2010: Long‐range lightning geolocation using a VLF radio atmospheric waveform bank, *J. Geophys. Res*., 115, D23108.

Saunders, C. P. R., 1993: A Review of Thunderstorm Electrification Processes. *J. Appl. Meteor.,* 32, 642-655.

Schumacher, Courtney, Robert A. Houze, 2000: Comparison of Radar Data from the TRMM PR Satellite and Kwajalein Oceanic Validation Site. *J. Appl. Meteor*., 39, 2151–2164.

Steiner, M., R. A. Houze, Jr., and S. E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteor.*, 34, 1978-2007.

Takahashi, Tsutomu, 1978: Riming Electrification as a Charge Generation Mechanism in Thunderstorms. *J. Atmos. Sci*., 35, 1536–1548.

Takahashi, T., 1984: Thunderstorm electrification—A numerical study. *J. Atmos. Sci*., 41, 2541–2558.

Takahashi, Tsutomu, 1981: Warm Rain Study in Hawaii—Rain Initiation. *J. Atmos. Sci.*, 38, 347–369.

Takahashi, T., Kuniko Miyawaki, 2002: Reexamination of Riming Electrification in a Wind Tunnel. *J. Atmos. Sci*., 59, 1018–1025.

Thomson, N. R., 1993: Experimental daytime VLF ionospheric parameters. *J. Atmos. Terr. Phys*., 55, 173–184.
Thomson, N. R., M. A. Clilverd, and W. M. McRae, 2007: Nighttime ionospheric D region parameters from VLF phase and amplitude. *J. Geophys. Res*., 112, A07304, doi:10.1029/2007JA012271
Toracinta, E.R., D. J. Cecil, E. J. Zipser, and S. W. Nesbitt, 2002: Radar, passive microwave, and lightning characteristics of precipitating systems in the tropics. *Mon. Wea. Rev*., 130, 802– 824.
Total Lightning Network. Earth Networks, 2011. 20 Jan. 2012 http://earthnetworks.com.dnnmax.com/Products/TotalLightningNetwork.aspx
Vonnegut, B., 1982: The Physics of Thunderclouds. CRC Handbook of Atmospherics, Vol. 1, CRC Press, 1-22.
Williams, E. R., S. A. Rutledge, et al., 1992: A Radar and Electrical Study of Tropical "Hot Towers". *J. Atmos. Sci*., 49, 1386-1395.
Williams, E., K. Rothkin, D. Stevenson, D. Boccippio, 2000: Global Lightning Variations Caused by Changes in Thunderstorm Stroke and by Changes in the Number of Thunderstorms. *J. Appl. Meteor*., 39, 2223–2230.
Wormell, T.W., 1983: Atmospheric Electricity: Some recent trends and problems. *Quart. J. Roy. Meteor. Soc*., 79, 3.
Zipser, E.J., and K. R. Lutz, 1994: The vertical profile of radar reflectivity of convective cells: A strong indicator of storm intensity and lightning probability? *Mon. Wea. Rev.*, 122, 1751–1759.
Zipser, Edward J., 1994: Deep Cumulonimbus Cloud Systems in the Tropics with and without Lightning. *Mon. Wea. Rev*., 122, 1837–1851.