



## RESEARCH ARTICLE

10.1002/2012JD018819

## Key Points:

- Relationship between lightning rate and rainfall rate over the ocean
- This relationship can be used as a proxy in NWP
- This relationship can be used in nowcasting of thunderstorms over the ocean

## Correspondence to:

S. Businger,  
businger@hawaii.edu

## Citation:

Stolz, D. C., S. Businger, and A. Terpstra (2014), Refining the relationship between lightning and convective rainfall over the ocean, *J. Geophys. Res. Atmos.*, 119, 964–981, doi:10.1002/2012JD018819.

Received 8 SEP 2012

Accepted 16 DEC 2013

Accepted article online 19 DEC 2013

Published online 30 JAN 2014

## Refining the relationship between lightning and convective rainfall over the ocean

Douglas C. Stolz<sup>1</sup>, Steven Businger<sup>1</sup>, and Annick Terpstra<sup>2</sup>

<sup>1</sup>Departement of Meteorology, University of Hawaii, Honolulu, Hawaii, USA, <sup>2</sup>Wageningen University, Wageningen, Netherlands

**Abstract** The lightning stroke density observed by Vaisala's Global Lightning Dataset (GLD360) was compared to maximum reflectivity in the upper and lower troposphere observed by NASA's Tropical Rainfall Measuring Mission Precipitation Radar over remote oceanic regions. We found that GLD360 stroke density is strongly correlated with maximum reflectivity above 0°C and the height of the 30 dBZ isopleth in two ocean basins (best represented by a logarithmic least squares regression). The maximum reflectivity above 0°C increases ~8–10 dBZ and the maximum height of the 30 dBZ isopleth increases ~2.8–3.8 km across the domain of stroke density. Stroke density and the maximum near-surface reflectivity are also correlated, but the range of the resulting logarithmic relationship is less than that found for maximum reflectivity above 0°C. There is high confidence in the logarithmic regressions as  $R^2$  values are generally above 0.9 during the study. A bootstrap resampling approach confirms that the lightning versus rainfall relationship is statistically significant. Vertical profiles of radar reflectivity show that progressively larger values of stroke density are associated with higher reflectivity at most altitudes (most pronounced near 6.5 km mean sea level) and less rapid decreases in reflectivity with increasing height above 0°C. The logarithmic relationships between lightning and rainfall defined in this study have significant utility in creating proxy data sets, such as pseudoreflectivity, for weather analysis and assimilation in numerical weather prediction.

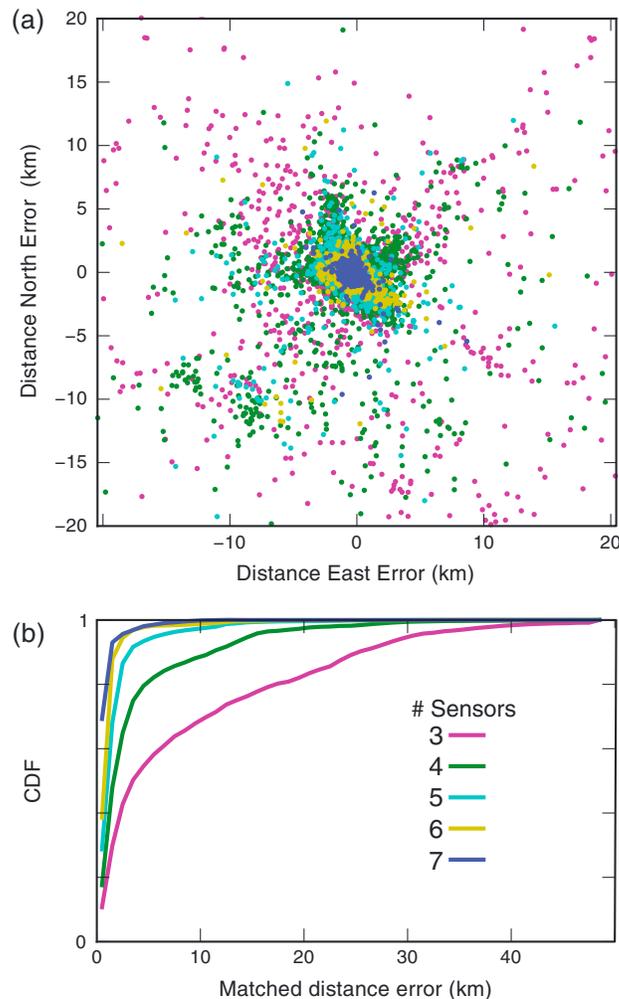
### 1. Introduction

The study of electromagnetic signals emitted by electrical discharges in clouds dates back to the end of the seventeenth century [Norinder, 1953]; however, the first quantitative analysis and visualization of waveforms associated with electric field changes produced by lightning occurred in 1920 [Cummins and Murphy, 2009]. In the 1980s, researchers 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]. The advent of ground-based systems of low frequency/very low frequency (VLF) sensors allowed lightning radio atmospheric, or "sferics," to be detected on much broader scales and this development led researchers to examine relationships between lightning and various indicators of convective intensity in many different regions of the world [Goodman *et al.*, 1988; Zipser and Lutz, 1994; Petersen *et al.*, 1996, 2005; Chèze and Sauvageot, 1997; Tapia *et al.*, 1998; Petersen and Rutledge, 1998, 2001; Toracinta *et al.*, 2002; Cecil *et al.*, 2005]. Lightning and rainfall have been studied extensively over land, and numerous studies document a marked 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; Fuytan 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 the ice water path [Carey and Rutledge, 2000; Cecil *et al.*, 2005; Pessi and Businger, 2009a, hereafter PB2009]. Leveraging these types of relationships, modeling studies have demonstrated subsequent improvement in the prediction of extratropical cyclones [Alexander *et al.*, 1999; Chang *et al.*, 2001; Papadopoulos *et al.*, 2005; Pessi and Businger, 2009b]. PB2009 used the Pacific Lightning Detection Network (PacNet), which consisted of four sensors covering a limited domain in the central North Pacific Ocean. The detection efficiency (DE) and location accuracy (LA) of PacNet were generally below 60% and 16–40 km, respectively. Since that time, long-range lightning detection technology has significantly improved.

In 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

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.



**Figure 1.** (a) Scatterplot of location error for GLD360 events matched to NLDN data for the 20 min period 19:35 to 19:55 UTC 29 July 2010 tracking an isolated thunderstorm. The match criteria used here are a simultaneous spatial coincidence of 50 km and a time coincidence of 150  $\mu$ s between the GLD360 event and the NLDN-reported lightning discharge. The scatterplot shows a window  $\pm$  20 km distance error N-S and E-W, and the points are color coded by the number of contributing sensors. (b) Plot of the corresponding cumulative distribution function (CDF) of the distance error relative to the NLDN events, again color coded by the number of sensors (adapted from Said et al. 2010b).

Lightning Dataset (GLD360). GLD360's performance in long-range lightning detection is unmatched by any other LLS currently operating. In contrast to predecessor LLSs, GLD360 monitors cloud-to-ground (CG) lightning strokes across the Northern Hemisphere with generally high DE and LA, 57% and 2.5 km, respectively [Said et al., 2013]. Thus, better quantification of the relationship between lightning and rainfall over oceanic regions using GLD360's superior lightning detection capability over remote, historically data-sparse regions holds promise for improvement of operational weather analysis and forecasting.

When the locations of GLD360 lightning strokes are compared over the US with the accurate location solutions for National Lightning Detection Network (NLDN) ground strokes and cloud pulses, the GLD360 lightning locations exhibit a tight clustering around the location of isolated thunderstorms (Figure 1). As the number of sensors increases, the added arrival time redundancy reduces the distance error in the geo-location solution. While the LA for three-sensor solutions is worse compared to events with four or more contributing sensors, admitting three-sensor solutions greatly improves the overall detection efficiency for the long-range network.

The low Earth orbits of current space-borne lightning imagers, such as the Tropical Rainfall Measuring Mission's (TRMM) Lightning Imaging Sensor (LIS) and NASA's Optical Transient Detector (OTD), result in observations over any one location on Earth's surface for  $\sim$ 90 s or less. Therefore, these instruments provide only

“snapshot” estimates of lightning flash rate. In contrast to these satellite-borne instruments, lightning observations from GLD360 are spatially and temporally *continuous*. In this work, 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.

The goal of this study is to develop proxy relationships for various rainfall characteristics observed by TRMM PR from GLD360 lightning data for use in nowcasting and numerical weather prediction (NWP). Near-surface and upper-level reflectivity, as well as reflectivity profiles, are analyzed in relation to changes in lightning rate. Low-level reflectivity is useful for both nowcasting and NWP. Forecasters rely heavily on base-level radar reflectivity in their short-term forecasting of convection so a pseudo reflectivity product that targets low-level reflectivity will be easy for forecasters utilize. Low-level reflectivity is a measure of the total latent heat left in the atmospheric column after the rain falls to the surface, which is the relevant thermodynamic quantity for NWP. Reflectivity profiles provide a vertical distribution of latent heating, important in prediction, and investigations of stability. Additionally, the profiles can inform the link between lightning production and elevated reflectivity above the freezing level [Zipser, 1994; Zipser and Lutz, 1994].

## 2. Data and Methodology

### 2.1. The Tropical Rainfall Measuring Mission Precipitation Radar

The TRMM PR aboard a NASA satellite provides coverage across all longitudes from 38°S to 38°N approximately 36°S to 36°N. The inclination angle of 35° allows the satellite to orbit the Earth 16 times each day and observe any single point 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 5.0 km × 5.0 km and the swath width is about 220 km. The TRMM PR operates at 13.8 GHz (a wavelength of roughly 2 cm) examining rainfall characteristics over both land and ocean.

The level 2 TRMM processing algorithm (version 7) Radar Rainfall Rate and Profile (2A25) is used in this study [Iguchi *et al.*, 2000]. In the upgraded 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 is given by  $h_a = (B_o - B_r) \cdot 250$  m, where  $B_r$  is the reported slant range bin. 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,  $\psi$ , the height of the radar return,  $h_r$ , is  $h_r = h_a \cdot \cos(\psi)$ .

At each pixel, the TRMM PR algorithm uses a climatological sea surface temperature and assumes a lapse rate of 6.0°C km<sup>-1</sup>. Starting at the surface temperature, the temperature is decreased according to the assumed lapse rate and the height at which the temperature is 0°C is the reported freezing height. The 2A25 algorithm requires that the bright band be detected within a layer that extends from 1.5 km below the freezing level to 1.5 km above the freezing level.

The TRMM PR algorithm makes the distinction between convective and stratiform precipitation using horizontal and vertical criteria [Steiner *et al.*, 1995; Awaka *et al.*, 1998; Schumacher and Houze, 2000]. Steiner and Houze [1993] showed that early partition schemes [e.g., Churchill and Houze, 1984] were sensitive to the horizontal resolution of the analysis. Bearing this in mind, the developers of the TRMM algorithm compared convective/stratiform partition results from TRMM PR (~5 km resolution) to a ground validation data set (2 km resolution). The parameters for the TRMM PR convective/stratiform partition were tuned so that the results from each study were approximately equal despite the differing horizontal resolutions. Ground validation indicates that bright band detection (which has implications for rain type classification) is approximately 80% within +/- 7° of the nadir and decreases toward the scan edges to values as low as 20%.

Near-surface reflectivity, the attenuation-corrected reflectivity through the lowest 20 km, and the height of the freezing level are reported for each of the 49 pixels in a scan. Over the ocean, TRMM PR has issues with ground clutter (i.e., ocean surface roughness and sea spray), and accordingly, the algorithm flags potentially contaminated data. 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 0.5–2.0 km.

In this study, the near-surface reflectivity, the reflectivity profile, and the reflectivity in the mixed phase region above the freezing level are compared with lightning density of CG strokes from GLD360. The near-surface reflectivity gives the best measure of the total latent heat that is released in the column above the measurement. Therefore, this observation is the one that is most important for data assimilation into NWP models, as will be discussed further in section 5. The reflectivity profile provides some information on where in the vertical the latent heating may be distributed. Finally, the reflectivity in the mixed phase region is studied, because the microphysics in this region is most directly involved in the production of charge separation required for lightning occurrence, as discussed in section 1. The time intervals chosen for the comparisons are long enough to encompass the evolution of the rainfall following lightning activity in thunderstorms.

## 2.2. 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 specifically tuned to pick up the relatively weak sferics that result from multiple reflections in the waveguide between Earth and the ionosphere. The minimum detection threshold is roughly 5–7 kA, a level at which intracloud strokes can be detected. Comparisons with NLDN data, however, show that LLSs predominantly sample CG flashes (roughly 80%) [Pessi *et al.*, 2009; Said *et al.*, 2010, 2013].

The DE of an individual GLD360 sensor is strongly dependent upon the attenuation characteristics of the environment. The VLF signal strength drops off exponentially with increasing distance; therefore, the attenuation scaling factor,  $A$ , is modeled as:

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

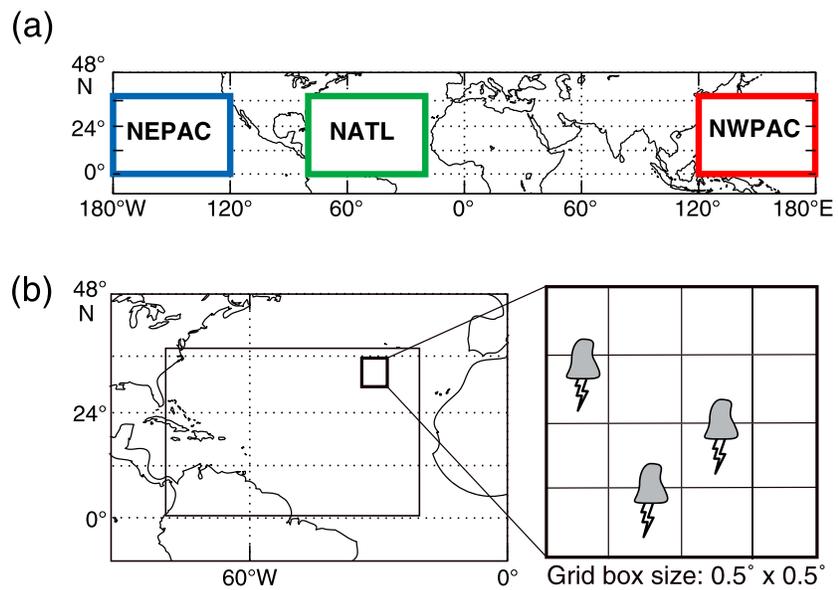
where  $\alpha_f$  is a scaling constant,  $R$  is the distance between the flash and the sensor location,  $\theta = R/R_e$ , and  $R_e$  is the Earth's radius. The scaling constant,  $\alpha_f$ , is adjusted so that the error between the fit and the actual data is minimized. The space constant,  $\lambda_f$ , depends upon surface characteristics and its empirical determination yields a complete description of the attenuation, which in turn provides an accurate model for sferic attenuation [Pessi *et al.*, 2009].

GLD360's CG stroke DE during the time window of this study is estimated to be 60–70% across the northern hemisphere [Said *et al.*, 2010]. For individual strokes, GLD360 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 [e.g., Cummins *et al.*, 1998, 2000]. Furthermore, each sensor uses recognition algorithms to diagnose sferic waveform variations resulting from propagation over great distances through the Earth-ionosphere waveguide, which provides a more accurate geo-location [Said *et al.*, 2010]. Observations are transmitted to Vaisala's Network Control Center in Tucson, AZ for processing and then made available to the data user in near real time [Grogan and Demetriades, 2009].

GLD360 began reporting in late 2009, so preliminary data analysis started from January 2010 and continued through May 2011. In early May 2011 and again in November 2011, Vaisala implemented improved reprocessing algorithms that, coupled with the growing network and improved site calibration, dramatically increased DE and LA across all domains [Said *et al.*, 2010]. DE was generally greater than 50% across the northern hemisphere at the start of 2011 and the region surrounding the Hawaiian Islands was the only area of deficiency across the northern hemisphere. By September 2012, the DE was approximately 70% across the whole northern hemisphere and portions of the southern hemisphere with general decrease in performance poleward of 45°S. At present, the average LA of GLD360 is ~5 km across the Northern Hemisphere.

## 2.3. Domains of Interest and Procedure

The global coverage of GLD360 allows for an expanded analysis over several oceanic domains in the present study (Figure 2a). The domains of interest are the North Atlantic Ocean (NATL), Northeast Pacific Ocean (NEPAC), and Northwest Pacific Ocean (NWPAC). Although analysis of the data began when GLD360 began reporting in January 2010, one year of data from GLD360 and the TRMM PR (May 2011 to April 2012)



**Figure 2.** (a) Three oceanic domains studied (NWPAC 0°N–38°N/120°E–180°E; NEPAC 0°N–38°N/180°W–120°W; NATL 0°N–38°N/80°W–20°W). (b) A schematic diagram in which the CAT is met, i.e., the number of convective pixels equals or exceeds 2% of the pixels in the box.

constitute the final data set in this study following improvements in GLD360's performance and data processing as well as an upgrade to TRMM PR data processing.

Following PB2009, a  $0.5^\circ \times 0.5^\circ$  grid is applied across the domains of interest (Figure 2b), taking into account the latitudinal dependence of grid box area, and the time interval is set to  $\pm 900$  s from TRMM PR observation time. An analysis over the NWPAC using one year of data (January 2010–December 2010) indicates that the sensitivity of the relationship between lightning and rainfall is maximized for a larger grid box size and smaller time window. However, the correlation coefficient is greatest when using a  $0.5^\circ \times 0.5^\circ$  grid box and  $\pm 900$  s time interval. The choice of grid box size and observation time following PB2009 allows for comparison between the current results and those found previously; however, a number of considerations can come into play in the choice of grid box size.

First, the time interval that a convective feature may spend within a grid box should be considered. This interval may vary some with convective system, season, and latitude. Second, the mature, active phase in the typical life cycle of an air mass thunderstorm is on the order of 30–60 min. Given that the cloud and precipitation fields evolve during the life cycle of a thunderstorm, and the representative quality of the TRMM PR observation is taken in a matter of seconds, the time interval for an analysis of this type should encompass the majority of the active life cycle of convective clouds to allow for the observation of lightning [MacGorman and Rust, 1998].

Third, the application of the results can be considered in the choice of time and space parameters. In PB2009, the objective was to improve the performance of a global numerical weather prediction model. Global models traditionally have coarse resolution, generally  $> 0.5^\circ$ , and do not explicitly resolve individual convective clouds. Therefore, a larger area and observation time may be used to determine an appropriate grid-scale relationship between lightning and rainfall.

Finally, the choice of time and space parameters may be dictated by the instrumentation that comprises the lightning network used in the study. PB2009 used PacNet, which had an average DE of  $\sim 40\%$  and LA of 16–40 km. The low DE of PacNet (relative to GLD360's DE) means that a longer time period is needed in order to adequately sample the lightning. In addition, if the horizontal dimension of the grid box chosen is 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 occurring in adjacent grid boxes being counted, which introduces error in the final data set.

This section describes the method used to obtain data for this study. If TRMM PR determines that rain is present within a grid box in the domain, then the "rainType" parameter of 2A25 identifies whether TRMM PR pixels in

**Table 1.** A Summary of Statistics for the Least Square Logarithmic Fits for Stroke Density Versus Maximum Reflectivity Above the Freezing Height

Data	Num. of Obs.	R <sup>2</sup>	RMSE	Range of Z
NWPAC Full	18,856	0.9667	0.3703	7.97 (35.5 – 43.472)
NATL Full	24,907	0.9535	0.5420	9.81 (35.24 – 45.06)
NWPAC Sum.	6612	0.9325	0.5294	8.78 (35.25 – 44.03)
NWPAC Win.	2722	0.6034	0.6361	4.93 (36.95 – 41.89)
NATL Sum.	6793	0.9205	0.8370	11.34 (34.55 – 45.93)
NATL Win.	4347	0.8815	0.7207	10.63 (35.25 – 44.03)

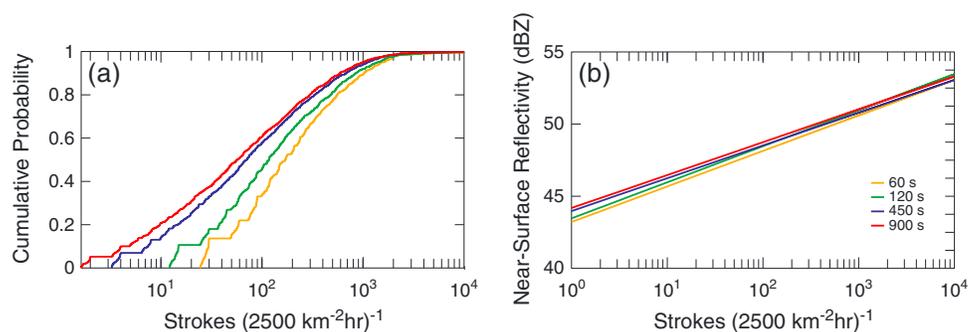
the grid box are stratiform or convective based on horizontal and vertical sampling methods mentioned previously. If convective rain pixels are found within the grid box (the number of convective pixels in the grid box is tabulated), a window of 15 min before and 15 min after ( $\pm 900$  s) the TRMM PR overpass time is allotted during which the number of lightning strokes observed by GLD360 is 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 ( $\text{km}^2$ ) depending on the size of the grid box used—here the normalization is  $2500 \text{ km}^2$  following PB2009.

The corresponding recorded TRMM PR indices include: (i) the maximum reflectivity above the freezing height (MAXZFH), (ii) the maximum near-surface reflectivity (MAXNSZ), (iii) the maximum height of the 30 dBZ isopleth (MAXH30), and (iv) the profile of attenuation-corrected reflectivity in the grid box. It is widely accepted that the noninductive charge separation process, i.e., collisions between graupel and small ice in the presence of supercooled liquid within a cloud’s mixed phase, is the most plausible mechanism to explain the observed charge structure in a thunderstorm [Takahashi, 1978, 1984; Saunders, 1993; Pruppacher and Klett, 1998]. Thus, it is expected that radar observations of the cloud’s mixed phase characteristics will exhibit strong correlation with lightning stroke density [Williams et al., 1992; Rutledge et al., 1992; Zipser and Lutz, 1994; and Petersen et al., 1996].

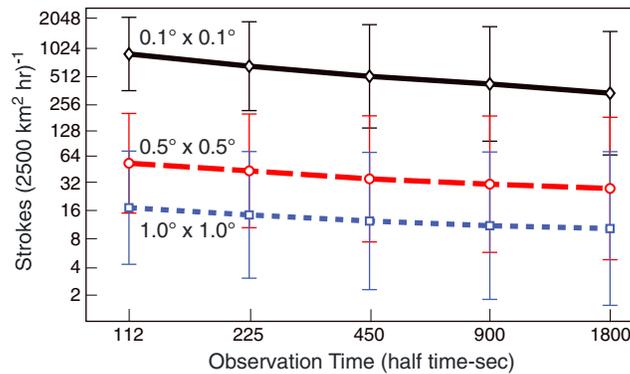
It is important to note the changing character of the diverse weather phenomena that occur within our domains of study as a function of the time of year. Accordingly, a winter subset (December 2011, January 2012, and February 2012) and a summer subset (June 2011, July 2011, and August 2011) were formed from the full data set to investigate potential seasonal variability in the relationship between lightning and rainfall.

The frequency distribution for the initial GLD360 data sample (January 2010–mid 2011) showed that there is a higher probability of occurrence of relatively low stroke density,  $O[10^0\text{--}10^2 \text{ strokes } (2500 \text{ km}^2 \text{ h})^{-1}]$ . The maximum observed stroke density was near  $10^4\text{--}10^5 \text{ strokes } (2500 \text{ km}^2 \text{ h})^{-1}$ . The data were binned according to the natural logarithm of the stroke density to better visualize the spread of observed stroke density across multiple orders of magnitude. A least squares fit was then applied to the bin medians of TRMM PR parameters for each stroke density. The least squares fit follows the general form:

$$f(x) = A \cdot \ln(x) + C.$$



**Figure 3.** (a) 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, 4, 15, and 30 min, respectively). (b) GLD360 (May 2011–August 2011) CG stroke density versus maximum near-surface reflectivity.



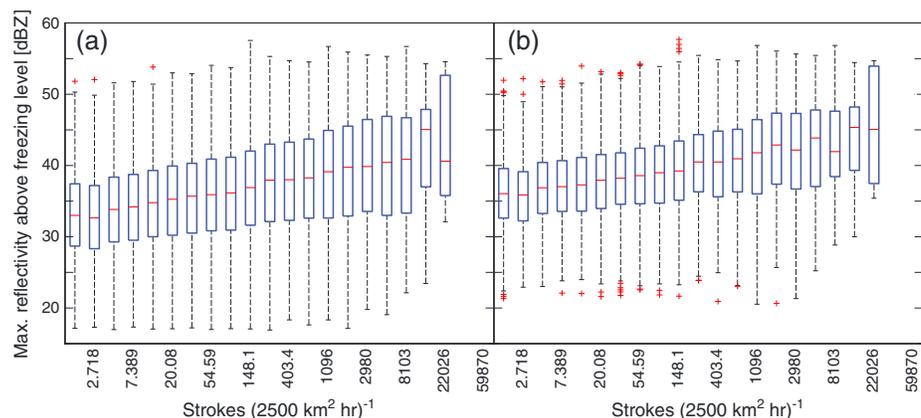
**Figure 4.** Mean GLD360 stroke density (+/− σ) as a function of observation time interval for three different grid sizes (0.1, 0.5, and 1.0°) for January–December 2010.

The vertical distribution of the attenuation-corrected reflectivity is determined for various categories of lightning stroke density, including for the total data set, for stroke density greater than or equal to the 50<sup>th</sup> percentile, for stroke density greater than or equal to the 90<sup>th</sup> percentile, and for stroke density greater than or equal to the 99<sup>th</sup> percentile. The data are analyzed for the full one-year period in each 0.5° grid box across each basin. The vertical profile of attenuation-corrected reflectivity corresponding to the observation of maximum reflectivity above the freezing level (values between 0 dBZ and 100 dBZ) is recorded in a grid box. Then the vertical distributions within each grid box in the domain are averaged to obtain the final vertical distributions of reflectivity, which are presented in section 3.

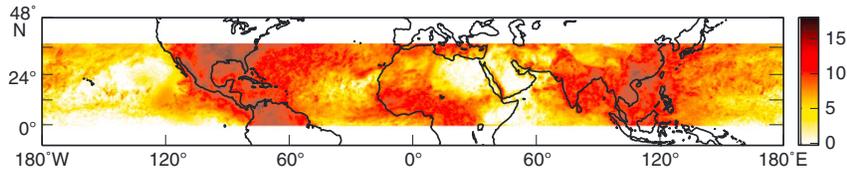
**2.4. Sensitivity Tests**

Data subsets were used to conduct 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° × 0.1°, 0.5° × 0.5°, or 1.0° × 1.0° areas and various time intervals. Data from May 2011 to August 2011 over the NATL domain were used to test for sensitivity in the lightning versus rainfall relationship using four different observation time intervals (±60, ±120, ±450, and ±900 s) on a 0.5° × 0.5° grid.

In *LeMone and Zipser* [1980], the upper 10% of updraft core diameters observed in convective clouds during the Global Atmospheric Research Program’s Atlantic Tropical Experiment (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. Using the method described above, there is 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



**Figure 5.** The stroke density observed by GLD360 versus maximum near-surface reflectivity using a 0.5° × 0.5° grid box and ±900 s time interval, (a) with and (b) without a convective area threshold (CAT) of 2%.



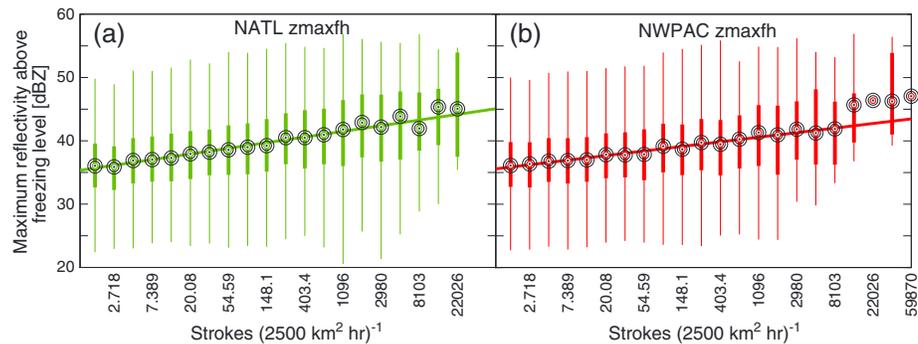
**Figure 6.** Climatology of CG lightning as observed by GLD360 across the domains of interest for the year 2010. The values on the color scale represent  $2^n$  flashes per  $(2500 \text{ km}^2 \text{ yr})^{-1}$ , where  $n$  is 0, 5, 10, and 15.

observation set. However, we note that a  $\sim 100 \text{ km}^2$  area may not be representative of the conditions within the larger grid box ( $2500 \text{ km}^2$  in this case) as a whole.

Despite the general high accuracy and precision of GLD360 observations, the detection and location of individual lightning strokes are not perfect. Poor location of lightning strokes can result from lightning occurring where the distance between sensors is large or overall sensor density is low (Figure 1). Additional sources of error are radio signal interference from power lines and site-specific errors. Occasionally, a single, isolated lightning stroke can be identified with cloud tops that are below the freezing level. Because lightning requires high concentrations of frozen hydrometeors above the freezing level [Takahashi, 1981], lightning observations near warm cloud tops are likely spurious and contribute to outliers in the comparison of lightning rate and rainfall.

To mitigate these problems, the data were subjected to a convective area threshold, CAT(%). First, the total number of convective pixels within a grid box was multiplied by the area of the TRMM PR footprint (assumed to be  $25 \text{ km}^2$ ). Then the fractional convective area in the grid box was computed by finding the ratio of the total “convective” area to the total grid box area (taking the latitudinal dependence of total grid box area into account). Finally, the CAT was varied between 0.01 and 0.2 (e.g., 1% and 20%) and in each case, only those pairs of TRMM PR observables/GLD360 lightning observations with fractional convective area larger than the CAT were admitted into the data population prior to statistical analysis. Data from the NATL basin exhibited the highest number of observations within each bin of stroke density and were therefore used to define the CAT. Least squares logarithmic regression coefficients and goodness-of-fit parameters were calculated as the CAT was varied over the experimental range.

It was found that the root-mean-square (RMS) error of the fit was minimized for a CAT of 2%. Furthermore, with the CAT set to 2%, data points with high lightning density coupled to low TRMM PR observables (e.g., MAXZFH, MAXNSZ, and MAXH30) were effectively eliminated. Thus, using the CAT increases the probability that GLD360’s lightning observations are more closely coupled to convective cores rather than to shallow clouds lacking mixed phase development.



**Figure 7.** GLD360 stroke density versus TRMM PR maximum reflectivity above the freezing height from May 2011 to April 2012. The least squares logarithmic fit to the bin medians (outlined circles) is overlaid. The interquartile range is denoted by the thick vertical solid lines while the error bars (thin solid lines) cover approximately 99% of the data within a given stroke density bin. (a) Results for the NATL domain and (b) results for NWPAC domain.

**Table 2.** Logarithmic Fits of the Form  $y = A \cdot \ln(x) + C$  by Basin for Stroke Density Versus Maximum Reflectivity Above the Freezing Height<sup>a</sup>

Data	A	C
NWPAC Full	0.7248 (0.6535, 0.7961)	35.50 (35.12, 35.89)
NATL Full	0.8926 (0.7822, 0.9970)	35.24 (34.67, 35.80)
PB2009	0.7200	33.3

<sup>a</sup>Values in parentheses are 95% confidence bounds.

### 3. Results

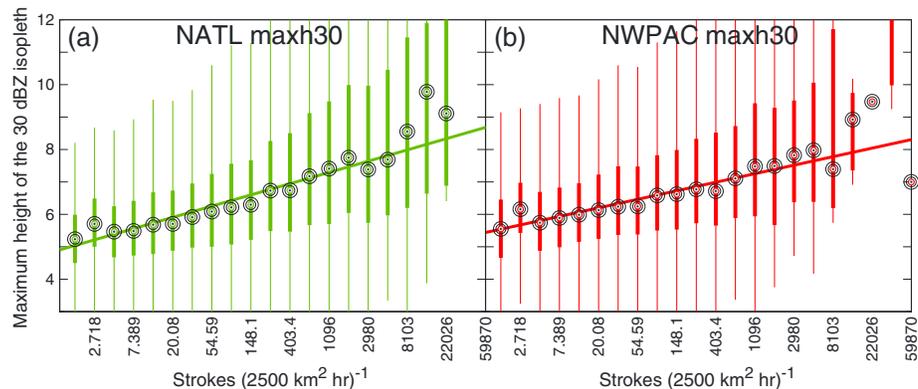
Observations that satisfied the criteria defined in section 2 resulted in sample sizes of 18,856, 24,907, and 4713 for the NWPAC, NATL, and NEPAC, respectively, for a one-year period (Table 1). Note that an appreciable portion of the NEPAC is subject to relatively low GLD360 CG DE and a comparison of sample sizes by basin demonstrates the general lack of lightning in the NEPAC. It is suggested that the prevailing subtropical high and subsidence over much of the NEPAC inhibits the development of deep convection in that domain. In addition, over the ocean the reflectivity tends to be concentrated below the freezing level, even in deeper clouds. Therefore, these clouds lack a healthy mixed-phase region that is a requirement for lightning production. For these reasons, the NEPAC is excluded from the current statistical analysis. From this point forward, results from the NWPAC and NATL will be presented.

The analysis shows that the stroke density is concentrated between magnitudes of  $10^0$  and  $10^3$  strokes  $(2500 \text{ km}^2 \text{ h})^{-1}$  (Figure 3a). An exponential decrease in the number of observations with increasing stroke density is apparent for all time intervals studied. The climatology of strokes shows that the maximum stroke density occurs in the boreal summer in both the NWPAC and NATL, yet stroke density of this magnitude constitutes less than ~1% of the total data population. The reflectivity observations paired with these rare, high stroke densities introduce an additional source of noise in the equal-weighting least squares regression process. Therefore, these data pairs are excluded from the population during the logarithmic regression process.

The lightning versus maximum near-surface reflectivity relationship appears to be insensitive to the choice of observation time (Figure 3b). The range of reflectivity values across the domain of stroke density is from ~45 dBZ to ~53 dBZ (~6–8 dBZ). For a  $\pm 60$  s observation time, the slope is very slightly larger. The agreement between the logarithmic fits resulting from different time windows has important implications and these will be discussed briefly in section 4.

The lightning versus maximum near-surface reflectivity relationship appears to be insensitive to the choice of observation time (Figure 3b). The range of reflectivity values across the domain of stroke density is from ~45 dBZ to ~53 dBZ (~6–8 dBZ). For a  $\pm 60$  s observation time, the slope is very slightly larger. The agreement between the logarithmic fits resulting from different time windows has important implications and these will be discussed briefly in section 4.

In general, the stroke density decreases as the grid box size and observation interval increase (Figure 4). The range of observed mean stroke density for a  $1^\circ \times 1^\circ$  grid box resolution is about 8–16 strokes  $(2500 \text{ km}^2 \text{ h})^{-1}$ . This range increases approximately exponentially as the grid box size decreases; for the  $0.1^\circ \times 0.1^\circ$  grid box, the range is roughly 350–725 strokes  $(2500 \text{ km}^2 \text{ h})^{-1}$ . Note that the vertical axis in Figure 4 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 is about an order of magnitude larger than that for the largest grid box for the same temporal window. Longer observation times generally allow more observations of lower stroke density (Figure 4). Additionally, the lowest observable stroke density is a function of the observation time. A  $\pm 900$  s observation time resolves stroke density near 0  $(2500 \text{ km}^2 \text{ h})^{-1}$  while a total observation time of  $\pm 60$  s shows the data to truncate at about 25 strokes  $(2500 \text{ km}^2 \text{ h})^{-1}$ .



**Figure 8.** As in Figure 7, for GLD360 stroke density versus maximum height of the 30 dBZ isopleth.

**Table 3.** As in Table 1, for Stroke Density Versus Maximum Height of the 30 dBZ Isopleth

Data	Num. of Obs.	R <sup>2</sup>	RMSE	Range (m)
NWPAC Full	18,856	0.9112	226	2894 (5410 – 8304)
NATL Full	24,907	0.9316	258	3813 (4865 – 8678)
NWPAC Sum.	6612	0.9165	199	2952 (5392 – 8344)
NWPAC Win.	2722	0.3877	465	2333 (4820 – 7153)
NATL Sum.	6793	0.7818	567	4291 (5030 – 9321)
NATL Win.	4347	0.6633	280	2225 (4715 – 6840)

To illustrate the utility of the CAT in the analysis of stroke density versus maximum near-surface radar reflectivity, a comparison of results with and without a conservative 2% CAT is given in Figure 5. Since the outliers tend to fall below the median, a reduction of outlier data implies an increase in the medians of MAXZFH across all stroke density bins. In addition, the interquartile ranges (IQR) in each of the stroke density bins become smaller and the IQR is generally more symmetric about the bin medians when a CAT of 2% is implemented (Figure 5a) versus when no CAT is applied (Figure 5b). To reiterate, the outliers that are thus eliminated are likely associated with scatter due to the location error growth at greater distances from sensors over the open ocean (Figure 1). With the inclusion of the CAT, the number of outliers is significantly reduced across the domain (Figure 5a). It is also important to emphasize that the CAT of 2% minimizes the RMS error of the log fit.

*Petersen and Rutledge* [2001] present relative frequency distributions of TRMM PR radar reflectivity as a function of height for a variety of locations in the Tropics (cf., Figures 4–6 of that study). They find a larger relative percentage of reflectivity  $\geq 30$  dBZ above 5 km in regions where the climatological distribution (Figure 6) shows the greatest preference for lightning to occur. Furthermore, studies of deep convection in the Tropics generally show that the height of the 30 dBZ echo can serve as a specific measure of convective intensity or the likelihood for a convective cloud to produce lightning [e.g., *Petersen et al.*, 1996; *DeMott and Rutledge*, 1998]. Because lightning activity tends to track the strongest cells in a convective system, it is logical to look at the relationship between lightning density and maximum reflectivity, specifically that which is associated with a convective cloud’s mixed phase [*Petersen et al.*, 1996].

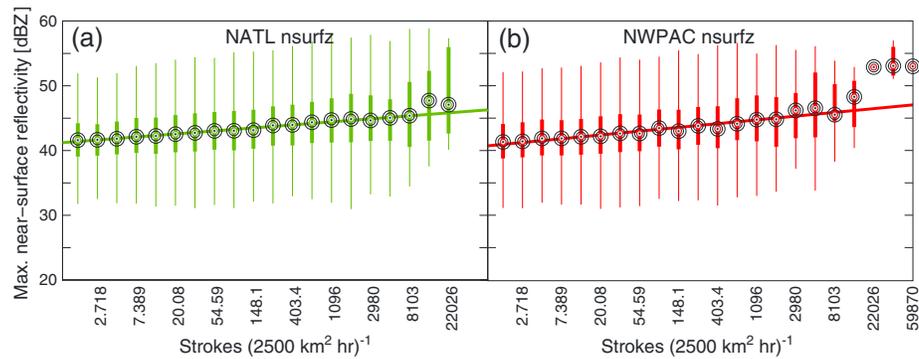
Figure 7 shows that the MAXZFH increases logarithmically with increasing stroke density in the NWPAC and NATL. If the logarithmic fit for each basin is evaluated across the domain of stroke density, then the range of MAXZFH is 35.5 to 43.47 (7.97) dBZ in the NWPAC and 35.24 to 45.06 (9.81) dBZ in the NATL. Tables 1, 2 provide statistics and coefficients for various logarithmic regressions computed. In both basins, the range of the logarithmic fit is found to be larger in boreal summer than in boreal winter. However, the seasonal variability in the logarithmic relationship is smaller in the NATL than in the NWPAC. The range of one-year and seasonal data in the NATL shows an excursion that is consistently near  $\sim 10$  dBZ while the excursions of the one-year and seasonal data in the NWPAC span from  $\sim 5$  dBZ (winter) to  $\sim 9$  dBZ (summer). There is high confidence in these results because the  $R^2$  values are generally range from about 0.8 to 0.97 (although a minimum  $R^2 \sim 0.6$  is noted in NWPAC winter and attributed to smallness of the data population in that basin and season).

Given the robust relationship found between stroke density and MAXZFH, it follows that lightning density should also be well correlated with the height of the 30 dBZ isopleth. The height of the 30 dBZ echo is found to increase by 2.8 km and 3.8 km in the NWPAC and NATL, respectively, across the domain of stroke density (Figure 8). Again, the seasonal variability of the relationship is apparent (Tables 3, 4). The range of 30 dBZ echo heights is larger in the warm season than in the cold season in both basins, but the magnitude of variability is larger in the NATL domain (see Table 3). The seasonal difference between the ranges of the logarithmic fits is approximately 2 km in the NATL ( $\sim 4.2$  km in summer and  $\sim 2.2$  km in winter) and only about 0.6 km in the NWPAC ( $\sim 2.9$  km in summer and  $\sim 2.3$  km in winter). The  $R^2$  values for one-year data and seasonal data subsets in both the NWPAC and NATL generally range from 0.66 to 0.93 (with the minimum  $R^2 = 0.38$  again found for the NWPAC winter data subset).

**Table 4.** As in Table 2, for Stroke Density Versus Maximum Height of the 30 dBZ Isopleth

Data	A	C
NWPAC Full	263.1 (219.6, 306.7)	5410 (5175, 5646)
NATL Full	346.7 (296.9, 396.5)	4865 (4595, 5134)

To this point, we have shown that GLD360 stroke density is well correlated with mixed-phase microphysical development. It is of interest to understand how lightning may relate to near-surface reflectivity because it provides the best indication of the total



**Figure 9.** As in Figure 7, for GLD360 stroke density versus maximum near-surface reflectivity.

latent-heat released through the depth of the raining column—a quantity that is of chief importance for atmospheric modeling applications and also for nowcasting of flood potential. Near-surface rainfall in deep convection does rely on the microphysical processes that take place in the mixed-phase region, including enhanced deposition on ice crystals, riming, and aggregation. A full description of the mixed-phase processes and the resulting enhancement of surface rainfall are beyond the scope of this paper. It is not surprising to find that MAXNSZ exhibits a correlation with MAXZFH in both the NATL ( $R^2 = 0.55$ , full year;  $R^2 = 0.43$ , summer;  $R^2 = 0.50$ , winter) and NWPAC ( $R^2 = 0.53$ , full year;  $R^2 = 0.43$ , summer;  $R^2 = 0.46$ , winter). The fact that MAXNSZ and MAXZFH are correlated suggests that the mixed-phase precipitation processes translate downward to generate discernable changes in the character of near-surface rainfall [Petersen and Rutledge 2001].

Next, we show that near-surface reflectivity increases with increasing GLD360 stroke density (Figure 9). However, the slope or sensitivity of the logarithmic relationship is less than what was found for relationships between lightning and mixed-phase radar observations. Across the domain of stroke density, the near-surface reflectivity increases from 40.65 to 47.04 (6.39) dBZ in the NWPAC and from 41.16 to 46.28 (5.12) dBZ in the NATL for the full year data set. Seasonal variability in the NATL domain is negligible; however, the ranges of both seasonal regressions surpass that for the full year of data. In the NWPAC domain, seasonal variability is more pronounced with the range of the warm season regression being larger than that for the cold season by ~2 dBZ. In all cases, the least squares  $R^2$  values fall between 0.81 and 0.98, which lends high confidence to the logarithmic relationships found. The details of these relationships and coefficients of the regressions are presented in Tables 5 and 6. Note that the results of PB2009 are also shown in these tables and that the range of the relationships found here is comparable to those found in the previous study. The values of near-surface reflectivity found in this study are systematically larger than those found in PB2009, a finding that will be discussed further in section 4.

To verify the significance of the least squares regressions, a bootstrap approach was employed to resample the data. Indices derived from a uniform random number generator were used to select a subset of the reflectivity bin medians. Then the coefficients of the logarithmic fit on the data subset were determined according to the same least squares methodology described above. The process was repeated 1000 times and after each iteration the coefficients (slope and intercept parameters) were recorded. Figure 10 shows each of the 1000 resampled regressions with the bin medians overlaid, which graphically shows that the relationships between lightning and MAXZFH, MAXH30, and MAXNSZ are decidedly positive in each basin. To quantify this result, the slopes of the logarithmic regressions were sorted and then binned (Figure 11) with the 5<sup>th</sup> and 95<sup>th</sup> percentile slopes indicated by vertical (black) solid lines. The results of the resampling

**Table 5.** As in Table 1, for Stroke Density Versus Maximum Near-surface Reflectivity

Data	Num. of Obs.	$R^2$	RMSE	Range of Z
NWPAC Full	18,856	0.9220	0.4649	6.389 (40.65 – 47.04)
NATL Full	24,907	0.9823	0.1719	5.124 (41.16 – 46.28)
PB2009	–	–	–	~33 – 41
NWPAC Sum.	6612	0.8670	0.7684	7.455 (40.21 – 47.66)
NWPAC Win.	2722	0.8164	0.4052	5.375 (41.14 – 46.51)
NATL Sum.	6793	0.8810	0.6296	6.850 (40.15 – 46.99)
NATL Win.	4347	0.9169	0.3584	6.440 (41.55 – 47.99)

**Table 6.** As in Table 2, for Stroke Density Versus Maximum Near-surface Reflectivity<sup>a</sup>

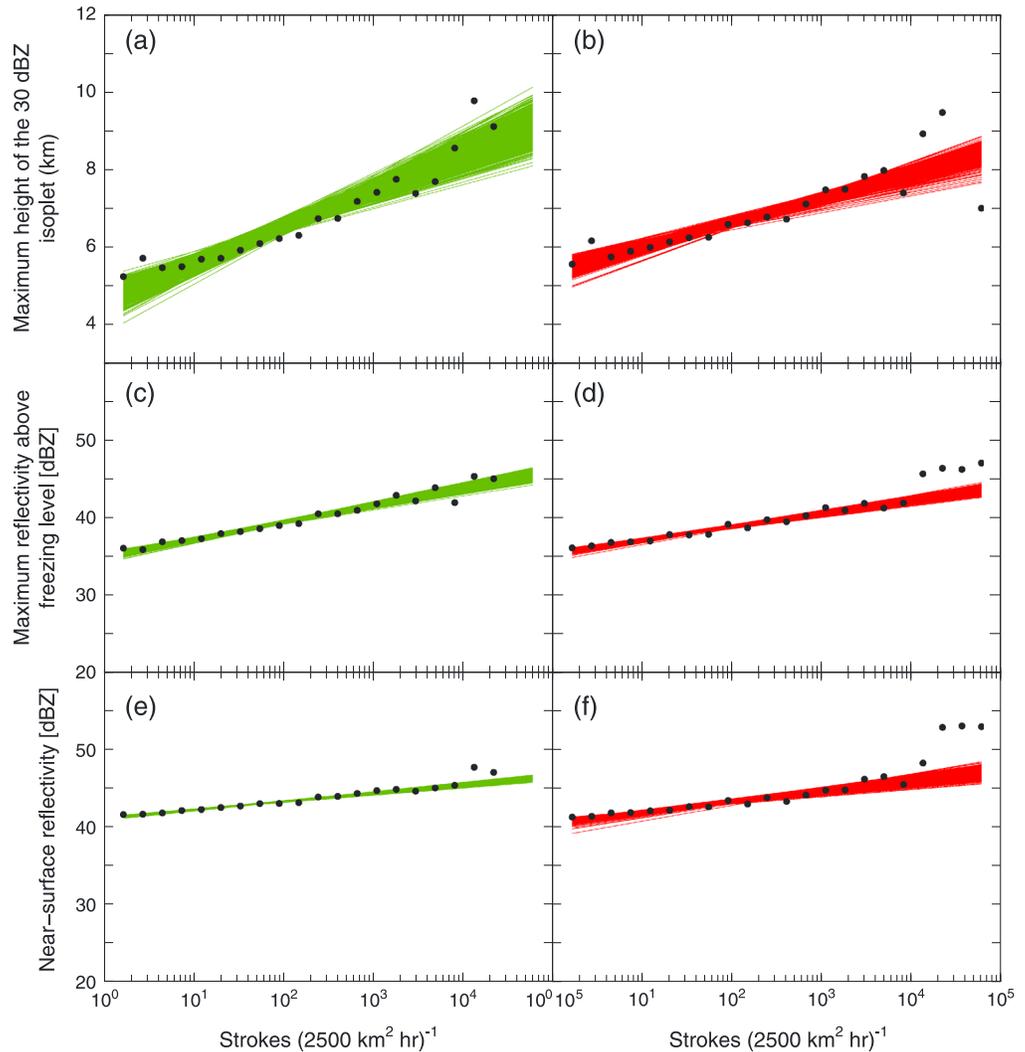
Data	A	C
NWPAC Full	0.5808 (0.4912, 0.6703)	40.65 (40.17, 41.14)
NATL Full	0.4887 (0.4327, 0.4989)	41.16 (40.98, 41.34)
PB2009	0.7200	33.3

<sup>a</sup>The results of Pessi and Businger [2009a] are included for comparison.

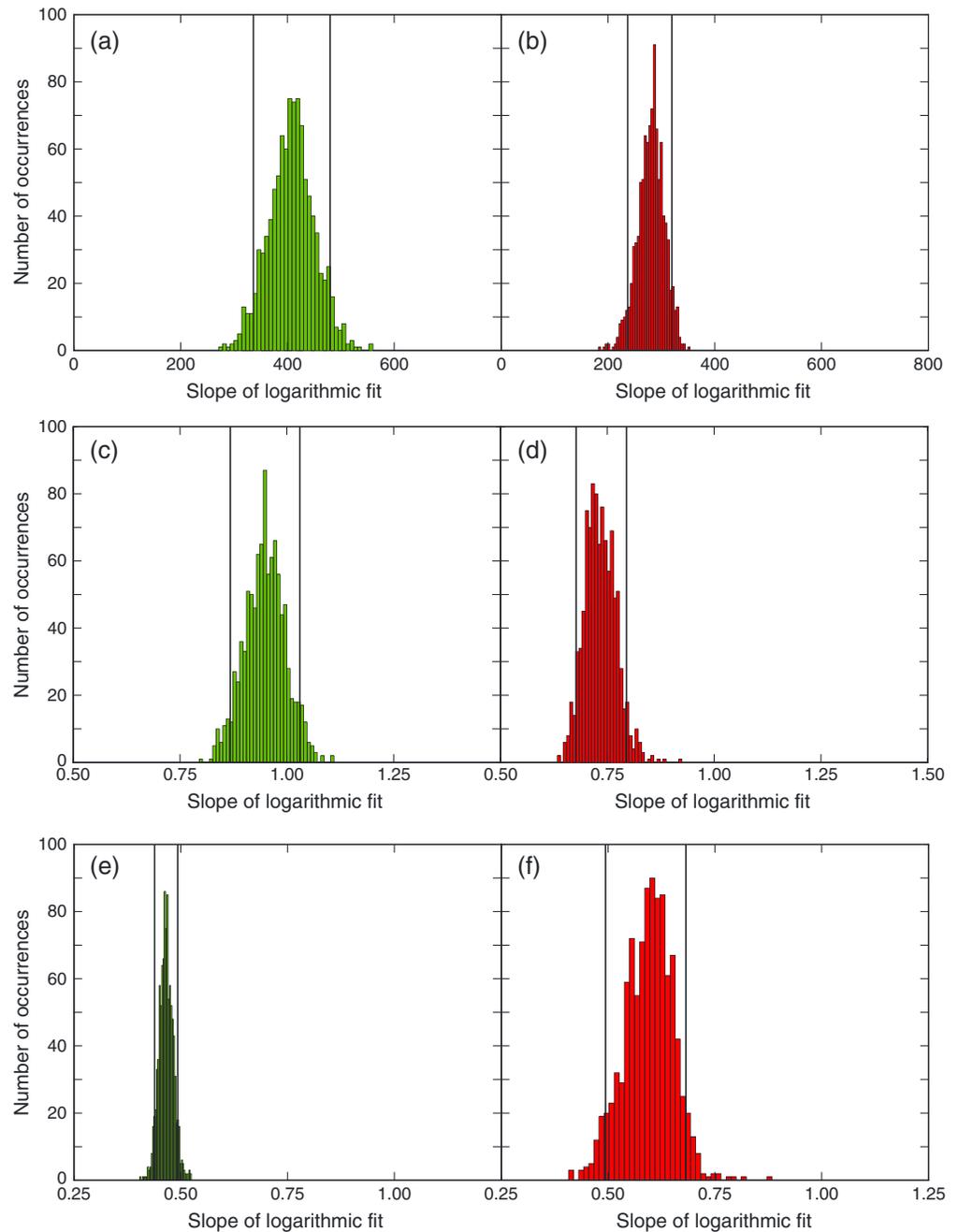
indicate that the 95% confidence intervals for regression slopes were consistently far from zero in all basins for all parameters, confirming that the lightning versus rainfall relationships are statistically significant.

As a final point, we present domain averages of the vertical profiles of attenuation-corrected reflectivity corresponding to the maximum reflectivity above the freezing height in each grid box. By definition, these profiles are most closely representative of convective clouds with well-developed mixed-phase microphysics conducive to lightning. The profiles for the domain average (black), median (blue), 90<sup>th</sup> percentile (red), and 99<sup>th</sup> percentile (yellow) of GLD360 stroke density are plotted together for comparison for the NATL (Figure 12a) and NWPAC (Figure 12b).

Note that the total average profile (black) is derived for grid boxes with lightning, as opposed to every profile observed in the domain. In the latter case, a domain average profile would likely exhibit considerably lower mixed phase reflectivity. The average height of the near-surface reflectivity bin (i.e., the lowest bin that is free



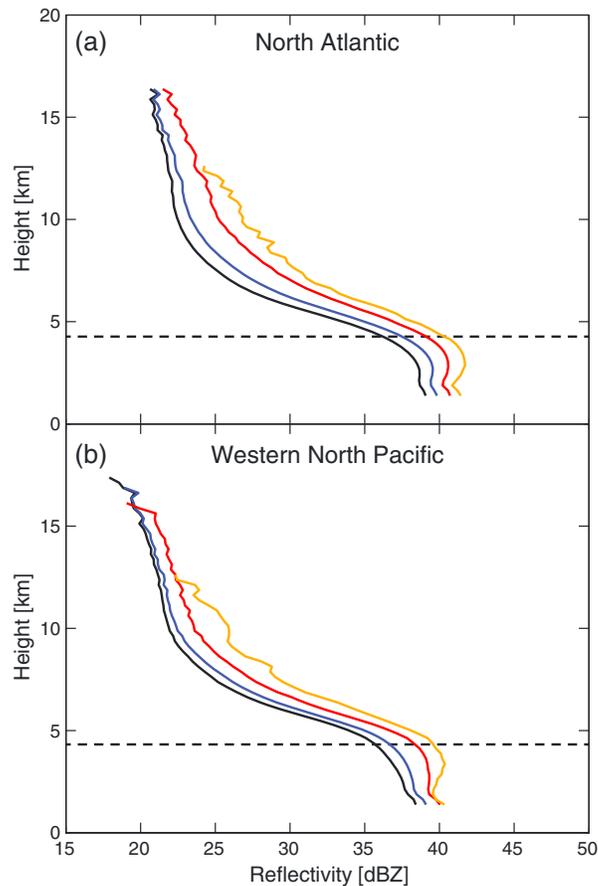
**Figure 10.** A 1000 member logarithmic regression ensemble of GLD360 stroke density versus (a and b) maximum height of the 30 dBZ isopleth, (c and d) maximum reflectivity above the freezing level, and (e and f) maximum near-surface reflectivity derived from a bootstrap resampling approach for the NATL (green) and NWPAC (red). Bin medians of stroke density are overlaid in black.



**Figure 11.** Binned results for the slope of the logarithmic fit derived from a bootstrap resampling approach for GLD360 stroke density versus (a and b) maximum height of the 30 dBZ isopleth, (c and d) maximum reflectivity above the freezing level, and (e and f) maximum near-surface reflectivity for the NATL (green) and NWPAC (red). The 95% confidence intervals are overlaid in each plot (vertical solid black lines) and values are as follows: (287.88, 367.67); (238.18, 316.36); (0.8723, 1.0318); (0.6752, 0.7984); (0.4388, 0.4929); and (0.4940, 0.6823).

of ground-clutter contamination) is ~1.5 km in both basins. Accordingly, only data above that altitude are shown in Figure 12.

All profiles in both basins exhibit distinct maxima in reflectivity immediately below the freezing height (denoted by the horizontal white dotted line), which is consistent with the dielectric increase in the phase change from ice to liquid hydrometeors. Above this altitude, reflectivity drops off rapidly (i.e., between 5 and 9 km in the mixed-phase region). In line with previous studies [e.g., Zipsper and Lutz, 1994; PB2009], the



**Figure 12.** The maximum attenuation-corrected reflectivity associated with the total (black), 50<sup>th</sup> (blue), 90<sup>th</sup> (red), and 99<sup>th</sup> (yellow) percentiles of GLD360 stroke density in the (a) NATL and (b) in the NWPAC. The horizontal dotted white line indicates the mean height of the freezing level in each basin (NATL, 4.26 km; NWPAC, 4.32 km).

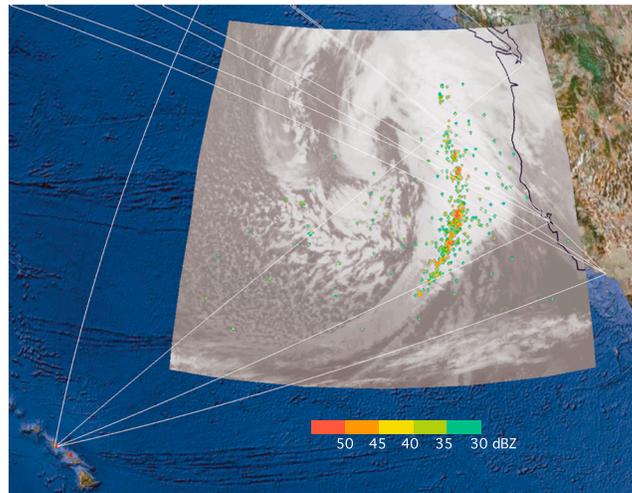
maximum attenuation-corrected reflectivity distributions generally illustrate greater reflectivity for larger stroke density at all altitudes. The profiles in Figure 12 closely resemble the tropical oceanic median profile from *Zipser and Lutz* [1994; c.f. Figure 3]. However, the 99<sup>th</sup> percentile profile in both basins exhibits characteristics more similar to the tropical continental median profile found in that study.

Specifically, there is a less rapid decrease in reflectivity with increasing height for higher stroke density. For example, in the 99<sup>th</sup> percentile profiles reflectivity  $\geq 30$  dBZ extends up to  $\sim 7.6$  km ( $\sim 7.1$  km) in the NATL (NWPAC) domain while in the total profiles reflectivity  $\geq 30$  dBZ only reaches  $\sim 5.6$  km ( $\sim 6.1$  km) in the NATL (NWPAC). In addition, the maximum deviation between the total and 99<sup>th</sup> percentile profiles in both basins is found between  $\sim 6.3$  and  $6.5$  km, i.e., in the heart of the mixed-phase region. This result again implies that high lightning rates are associated with more robust mixed-phase development.

#### 4. Summary and Conclusions

Data from GLD360, an advanced long-range lightning detection network, were combined with TRMM PR data to explore the relationship between lightning and rainfall over three ocean basins in the northern hemisphere. The lightning stroke density observed by GLD360 was compared to the maximum reflectivity above the freezing height, near the surface, and the vertical profile of reflectivity observed by TRMM PR over a one-year period (May 2011–April 2012). The most important conclusions from the work are summarized below.

GLD360 stroke density is found to be strongly correlated with both the maximum reflectivity above the freezing height as well as the maximum height of the 30 dBZ isopleth (best depicted by logarithmic least squares regressions) in the NATL and NWPAC domains. The range of maximum reflectivity above the freezing level was



**Figure 13.** Example of lightning-derived surface reflectivity product ( $0.5^\circ \times 0.5^\circ$  grid and  $\pm 900$  s time window) with commercial airline flight tracks (white lines) overlaid on an IR satellite image. The satellite image valid at 0330 UTC 19 December 2002 [see Pessi and Businger 2009b for more information on this storm].

$\sim 8$ – $10$  dBZ while the range of the maximum height of the 30 dBZ isopleth was  $\sim 2.8$ – $3.8$  km across the domain of stroke density for the one-year data set. The relationship between stroke density and near-surface reflectivity is well represented by the logarithmic fit but the range of the regression is smaller ( $\sim 5$ – $6$  dBZ) than was found for the relationship between stroke density and maximum reflectivity above the freezing height. The resulting regression statistics and coefficients are presented in Tables 1–6. We also found seasonal variations in each of the logarithmic relationships. The ranges of the regressions are consistently larger in boreal summer as compared to boreal winter in both the NATL and NWPAC.

A convective area threshold (CAT) was applied, which stipulates that a minimum fractional area of convection is present in the TRMM scan sample, prior to being included in the comparison. This minor restriction effectively reduces outliers associated with network errors in the location of lightning strokes and therefore significantly improves the correlations.

There is high confidence in the proxy relationships found here as the  $R^2$  values are consistently greater than 0.90 (reaching as high as 0.98) over the period of the study. Here we emphasize that these relationships are consistent across both the NATL and NWPAC, lending credibility and generality to the results. To verify the significance of the least squares regressions, a bootstrap approach was employed to resample the data. The results of the resampling indicate that the 95% confidence intervals for regression slopes were consistently far from zero in the NATL and NWPAC, confirming that the lightning versus rainfall relationships are statistically significant. The logarithmic relationships between lightning and rainfall defined for each basin studied here can be used to create proxy data sets, such as pseudo reflectivity for weather analysis and numerical weather prediction (Tables 1–6). An example of a lightning-derived reflectivity product using a lightning-reflectivity proxy is shown in Figure 13.

The maximum attenuation-corrected reflectivity shows that increases in GLD360 stroke density are associated with higher reflectivity values, both above and below the freezing level. In line with previous results, the profiles for the highest stroke density feature less rapid decreases in reflectivity with increasing height above the freezing level. The maximum difference between the 99<sup>th</sup> percentile and total profiles of stroke density is found in the middle of the mixed-phase region (near 6.5 km altitude).

As expected, the stroke density is shown to vary with the choice of time and space parameters used in the analysis. The stroke density found using the highest resolution grid box studied ( $0.1^\circ \times 0.1^\circ$ ) behaves like that associated with two larger grid box sizes with varied time intervals. Analysis of the most reliable data from GLD360 shows that the relationship between lightning and reflectivity is relatively insensitive to the choice of total observation time in the analysis. These combined results suggest that the GLD360 can be used to study the relationship between lightning and rainfall at higher temporal and spatial resolution.

## 5. Discussion and Future Work

This study uses data from the advanced GLD360 long-range lightning network to quantify relationships between lightning and TRMM radar products. GLD360 data coverage is global and includes remote oceanic regions where continuous monitoring is not available by traditional lightning detection networks, such as the NLDN, whose sensors operate at higher frequencies. In addition to accurately locating lightning strokes, GLD360 also provides data on current and polarity associated with the strokes. As long-range lightning detection networks continue to expand and improve, the data become increasingly accurate and valuable in nowcasting and forecasting applications [Pessi and Businger, 2009b]. To assimilate this continuous global data stream in numerical weather models, accurate relationships between lightning rate and a model variable need to be established. This study has the benefit of a robust data set over the western Pacific Ocean and Atlantic Ocean and our careful methodology, which includes focus on mixed-phase processes, has shown strong statistical significance for the relationships uncovered.

We showed that lightning rates are linked to relatively high levels of reflectivity in the lower troposphere in regions of deep convection over the ocean. This relationship is shown to be robust and statistically significant. Because the reflectivity at low levels is closely tied to the surface rainfall rate, the pseudo reflectivity derived from GLD360 lightning data can be used in NWP models as a proxy for latent heating in the column, and the vertical distribution of the reflectivity can be used to distribute the latent heating in the column.

Small seasonal and spatial variability of the relationships was found, although no effort was made to investigate the source of this variability as it is beyond the scope of the current study. It is suggested that the variability is driven, at least in part, by variations in aerosol loading and or variations conditional instability. Two separate hypotheses surrounding each one of these factors have been advanced in the literature [Albrecht, 1989; Khain et al., 2005; Rosenfeld et al., 2008; Lee, 2012; Williams et al., 1992; Williams and Satori, 2004; and others]. We explicitly isolate oceanic convection in the current study, but by our method it is not possible to isolate the influence of aerosols that are advected offshore downwind of major continents [e.g., Porter, 1988; Clarke et al., 1997; Porter and Clarke, 1997] and changes in environmental conditional instability. To properly address these issues, a large data set of convective scale observations is needed. It may be possible to address this question with data from hyperspectral instruments currently flying on geostationary satellites and polar orbiters.

In this study, we rely on DE and LA models whose empirical validation is based on previous work [e.g., Pessi et al., 2009]. The DE of GLD360 (and other currently operating long-range lightning detection systems) has been validated over land. To our knowledge, however, a comprehensive validation study of GLD360 performance over the ocean has not been undertaken to date. If an independent network of high-frequency sensors was installed in the Hawaiian Islands, it would allow for an independent assessment of the performance of GLD360 over the Central Pacific Ocean. In particular, DE model space constants for GLD360 could be empirically determined so that the model estimation of DE matches the observed lightning strokes over the water in the near vicinity of Hawaii by high-frequency detectors.

In PB2009, the maximum reflectivity was found to be positively correlated with stroke density but the magnitude of reflectivity for a given stroke density was ~7 dBZ lower than that found in the current study. This difference is likely because PB2009 use the average reflectivity across the area, rather than the maximum or core reflectivity in the grid box. Extensive stratiform cloudiness is nearly always associated with deep convection in the Tropics. An area average approach would tend to better reflect this reality, as the stratiform component of the radar signature would serve to "dilute" the relatively rare but more intense convective radar signatures. To reiterate, our choice to report maximum reflectivity is grounded in our belief that CG lightning, which is most commonly observed by GLD360, tends to track the strongest convective cells. Forecasters who use our product have found great utility in being able to pin point regions of deep convection over the ocean and track them in time (Figure 13).

The absence of lightning does not mean the absence of rainfall, as we know from personal experience. But, this fact does not diminish the importance of deep thunderstorms. Historically, case studies have shown that central sea-level pressure of oceanic storms that exhibit extensive deep convection is often seriously underforecast by NWP models [e.g., Reed and Albright, 1986; Alexander et al., 1999; Chang et al., 2001; McMurdie and Mass, 2004; Pessi and Businger, 2009b]. And from a nowcasting perspective, it is clearly useful to know when a lightning-active squall

line is approaching the coast and that it will be accompanied by heavy (45–50 dBZ) rainfall. The intensity of convective rainfall in electrified 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 relationships derived in this study (e.g., Figure 13). Additional work is being done on the pseudo reflectivity product to include radar observed average gradients in reflectivity surrounding convective echoes, allowing the appearance of the product to more closely mimic that of radar reflectivity of convective storms.

The results presented here show that the maximum reflectivity above the freezing level provides the best correlation to lightning density and that maximum reflectivity aloft is correlated to maximum surface reflectivity. The significance of this work is most important in its application to numerical weather prediction. Operational weather models do not predict lightning; therefore, lightning rate must be converted to a variable that the models can utilize, such as latent heating. Well-tested methods exist for assimilating radar reflectivity data. Therefore, an accurate reflectivity proxy based on the results reported here can easily be implemented operationally into a data assimilation procedure [e.g., Benedetti *et al.*, 2005]. Pessi and Businger [2009b] demonstrated the potential utility of assimilating profiles of latent heating associated with lightning rates in a poorly forecast winter storm off the California coast. The simulated storm with the lightning data assimilation was 10 hPa deeper than the control run without the lightning data assimilation, matching the observed storm's central pressure and location. As global model resolution continues to increase, the value of the data with an accurate reflectivity proxy for data assimilation increases as well. It is hoped that the robust empirical least squares fits derived in this study will provide impetus for additional studies of lightning-data assimilation.

## References

- Albrecht, B. A. (1989), Aerosols, cloud microphysics, and fractional cloudiness, *Science*, *245*, 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. Weather Rev.*, *127*, 1433–1457.
- Awaka, J., T. Iguchi, and K. Okamoto (1998), Early results on rain type classification by the Tropical Rainfall Measuring Mission (TRMM PR) precipitation radar, in *Proc. Eighth URSI Commission F Open Symp.*, pp. 134–146, Union Radio-Scientifique Internationale, Avenir, Portugal.
- Benedetti, A., P. Lopez, P. Bauer, and E. Moreau (2005), Experimental use of TRMM precipitation radar observations in 1D + 4D-Var assimilation, *Q. J. R. Meteorol. Soc.*, *131*, 2473–2495, doi:10.1256/qj.04.89.
- Boccippio, D. J., S. J. Goodman, and S. Heckman (2000), Regional Differences in Tropical Lightning Distributions, *J. Appl. Meteorol.*, *39*, 2231–2248.
- Carey, L. D., and S. A. Rutledge (2000), The relationship between precipitation and lightning in tropical island convection: A C-band polarimetric radar study, *Mon. Weather Rev.*, *128*, 2687–2710.
- Cecil, D. J., S. J. Goodman, D. J. Boccippio, E. J. Zipser, and S. W. Nesbitt (2005), Three years of TRMM precipitation features. Part I: Radar, radiometric, and lightning characteristics, *Mon. Weather 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. Weather Rev.*, *129*, 1809–183.
- Chèze, J.-L., and H. Sauvageot (1997), Area-average rainfall and lightning activity, *J. Geophys. Res.*, *102*, 1707–1715.
- Christian, H. J., et al. (2003), Global frequency and distribution of lightning as observed from space by the Optical Transient Detector, *J. Geophys. Res.*, *108*(D1), 4005, doi:10.1029/2002JD002347.
- Churchill, D. D., and R. A. Houze Jr. (1984), Development and structure of winter monsoon cloud clusters on 10 December 1978, *J. Atmos. Sci.*, *41*, 933–960.
- 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*, 25,281–25,292.
- Cummins, K. L., and M. J. Murphy (2009), An overview of lightning locating systems: History, techniques, and data uses, with an in-depth look at the U.S. NLDN, *IEEE Trans. Electromagn. Compat.*, *51*(3), 499.
- 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*, 9035–9044, doi:10.1029/98JD00153.
- Cummins, K. L., M. J. Murphy, and J. V. Tuel (2000), Lightning detection methods and meteorological applications, *IV International Symposium on Military Meteorology*, September 26–28, Malbork, Poland, 13 pp.
- DeMott, C. A., and S. A. Rutledge (1998), The vertical structure of TOGA COARE convection. Part I: Radar echo distributions, *J. Atmos. Sci.*, *55*, 2730–2747.
- Futyan, J. M., and A. D. Del Genio (2007), Deep convective system evolution over Africa and the tropical Atlantic, *J. Clim.*, *20*, 5041–5060.
- Goodman, S. J., D. E. Buechler, P. D. Wright, and W. D. Rust (1988), Lightning and precipitation history of a microburst producing storm, *Geophys. Res. Lett.*, *15*, 1185–1188.
- Grogan, M., and N. W. S. Demetriades (2009), Vaisala Global Lightning Dataset GLD360: Technology, Operations, and Applications Overview, Vaisala, Inc.
- Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto (2000), Rain profiling algorithm for the TRMM precipitation radar, *J. Appl. Meteorol.*, *39*, 2038–2052.
- Jayarathne, E. R., C. P. R. Saunders, and J. Hallet (1983), Laboratory studies of the charging of soft-hail during ice crystal interactions, *Q. J. R. Meteorol. Soc.*, *101*, 227–234.
- Khain, A., D. Rosenfeld, and A. Pokrovsky (2005), Aerosol impacts on the dynamics and microphysics of deep convective clouds, *Q. J. R. Meteorol. Soc.*, *161*, 2639–2663.
- 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.

## Acknowledgments

We are grateful to Nancy Hulbert for her assistance with the graphics. TRMM 2A12 and 2A25 data products were obtained from NASA Goddard Space Flight Center. This work is supported by NOAA under cooperative agreement NA11NMS4320128 and by the Office of Naval Research under grant N000140810450.

- Latham, J. (1981), The electrification of thunderstorms, *Q. J. R. Meteorol. Soc.*, *107*, 277–298.
- Lee, S. S. (2012), Effects of aerosol on circulations and precipitation in deep convective clouds, *J. Atmos. Sci.*, *69*, 1957–1974.
- LeMone, M. A., and E. J. Zipser (1980), Cumulonimbus Vertical Velocity Events in GATE. Part I: Diameter, Intensity and Mass Flux, *J. Atmos. Sci.*, *37*, 2444–2457.
- MacGorman, D. R., and W. D. Rust (1998), *The Electrical Nature of Storms*, pp. 422, Oxford Univ. Press, New York.
- McMurdie, L., and C. Mass (2004), Major numerical forecast failures over the Northeast Pacific, *Weather Forecast.*, *19*, 338–356.
- 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. Clim.*, *13*, 4087–4106.
- Norinder, H. (1953), Long distance location of thunderstorms, in *Thunderstorm Electricity*, edited by H. Y. Byers, pp. 276–327, University Press, Chicago, IL.
- Orville, R. E. (2008), Development of the National Lightning Detection Network, *Bull. Am. Meteorol. Soc.*, *89*, 180–190.
- Orville, R. E., and R. W. Henderson (1986), Global Distribution of Midnight Lightning: September 1977 to August 1978, *Mon. Weather Rev.*, *114*, 2640–2653.
- Orville, R. E., R. W. Henderson, and L. F. Bosart (1983), An East Coast Lightning Detection Network, *Bull. Am. Meteorol. Soc.*, *64*, 1029–1037.
- Papadopoulos, A., G. C. Themis, and N. A. Emmanouil (2005), Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model, *Mon. Weather Rev.*, *133*, 1961–1977.
- Pessi, A., and S. Businger (2009a), Relationships among lightning, precipitation, and hydrometeor characteristics over the North Pacific Ocean, *J. Appl. Meteorol. Climatol.*, *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. Weather 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., and S. A. Rutledge (1998), On the relationship between cloud-to-ground lightning and convective rainfall, *J. Geophys. Res.*, *103*, 14,025–14,040.
- Petersen, W. A., and S. A. Rutledge (2001), Regional variability in tropical convection: Observations from TRMM, *J. Clim.*, *14*, 3566–3586.
- Petersen, W. A., S. A. Rutledge, and R. E. Orville (1996), Cloud-to-ground lightning observations from TOGA COARE: Selected results and lightning location algorithms, *Mon. Weather Rev.*, *124*, 602–620.
- Petersen, W. A., H. J. Christian, and S. A. Rutledge (2005), TRMM observations of the global relationship between ice water content and lightning, *Geophys. Res. Lett.*, *32*, L14819, doi:10.1029/2005GL023236.
- Porter, J. N. (1988), Marine aerosol: Their measurement and influence on cloud base properties, MS thesis, University of Hawaii-Manoa.
- Porter, J., and A. Clarke (1997), Aerosol size distribution models based on in-situ measurements, *J. Geophys. Res.*, *102*, 6035–6045.
- Pruppacher, H. R., and J. D. Klett (1998), *Microphysics of Clouds and Precipitation*, pp. 71–826, D. Reidel Publishing Company, Boston, USA.
- Reed, R. J., and M. D. Albright (1986), A case study of explosive cyclogenesis in the Eastern Pacific, *Mon. Weather Rev.*, *114*, 2297–2319.
- Rosenfeld, D., U. Lohmann, G. B. Raga, C. D. O'Dowd, M. Kulmala, S. Fuzzi, A. Reissell, and M. O. Andreae (2008), Flood or drought: How do aerosols affect precipitation?, *Science*, *321*, 1309–1313.
- Rutledge, S. A., E. R. Williams, and T. D. Keenan (1992), The Down Under Doppler and Electricity Experiment ( DUNDEE ) : Overview and preliminary results, *Bull. Am. Meteorol. Soc.*, *73*, 3–15.
- 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, doi:10.1029/2010JD013863.
- Said, R. K., M. B. Cohen, and U. S. Inan (2013), Highly intense lightning over the oceans: Estimated peak currents from global GLD360 observations, *J. Geophys. Res. Atmos.*, *118*, 6905–6915, doi:10.1002/jgrd.50508.
- Saunders, C. P. R. (1993), A review of thunderstorm electrification processes, *J. Appl. Meteorol.*, *32*, 642–655.
- Schumacher, C., and R. A. Houze (2000), Comparison of radar data from the TRMM PR satellite and Kwajalein oceanic validation site, *J. Appl. Meteorol.*, *39*, 2151–2164.
- Steiner, M., and R. A. Houze Jr. (1993), Three-dimensional validation at TRMM ground truth sites: Some early results from Darwin, Australia, in *Preprints 26<sup>th</sup> Int. Conf. on Radar Meteorology*, pp. 5–8, Amer. Meteor. Soc., Tallahassee, FL.
- 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. Meteorol.*, *34*, 1978–2007.
- Takahashi, T. (1978), Riming Electrification as a Charge Generation Mechanism in Thunderstorms, *J. Atmos. Sci.*, *35*, 1536–1548.
- Takahashi, T. (1981), Warm rain study in Hawaii—Rain initiation, *J. Atmos. Sci.*, *38*, 347–369.
- Takahashi, T. (1984), Thunderstorm electrification—A numerical study, *J. Atmos. Sci.*, *41*, 2541–2558.
- Tapia, A., J. A. Smith, and M. Dixon (1998), Estimation of convective rainfall from lightning observations, *J. Appl. Meteorol.*, *37*, 1497–1509.
- 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. Weather Rev.*, *130*, 802–824.
- Williams, E. R., and G. Satoru (2004), Lightning, thermodynamic and hydrological comparison of the two tropical continental chimneys, *J. Atmos. Sol. Terr. Phys.*, *66*, 1213–1231, doi:10.1016/j.jastp.2004.05.015.
- Williams, E. R., S. A. Rutledge, S. C. Geotis, N. Renno, E. Rasmussen, and T. Rickenbach (1992), A radar and electrical study of tropical hot towers, *J. Atmos. Sci.*, *49*, 1386–1395.
- Zipser, E. J. (1994), Deep cumulonimbus cloud systems in the tropics with and without lightning, *Mon. Weather Rev.*, *122*, 1837–1851.
- 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. Weather Rev.*, *122*, 1751–1759.