8.3 DISSECTING THE ANOMALY – A CLOSER LOOK AT THE DOCUMENTED ENHANCEMENT IN SUMMERTIME GROUND FLASH DENSITIES IN AND AROUND THE HOUSTON AREA

Michael L. Gauthier* and Walter A. Petersen
The University of Alabama in Huntsville, Huntsville, Alabama

and

Lawrence D. Carey and Richard E. Orville
Texas A&M University, College Station, Texas

1. INTRODUCTION

Over the past two decades, studies of cloud-to-ground (CG) lightning have used data gathered from the National Lightning Detection Network (NLDN) to characterize spatial, and temporal variations in CG lightning activity both regionally and throughout the continental United States (CONUS; e.g. Orville et al., 1983 and Reap and MacGorman, 1986). In addition to variations in positive and negative flash densities, these climatologies have documented spatial and temporal variations in mean peak current, flash multiplicity, and flash polarity (Orville and Huffines, 2001 and Zajac and Rutledge, 2001). On smaller sub-synoptic scales, lightning "hot spots" have been noted in areas of complex and elevated terrain (such as the western United States; Reap, 1986, Gauthier, 1999 and Zajac and Rutlege, 2001), along mesoscale boundaries such as land/sea convergence zones along the Gulf Coast and in Florida, and even over the warm ocean waters of the Gulf Stream (Biswas and Hobbs, 1990; Orville 1990; Petersen and Rutledge, 2004). Not so evident on these maps are more localized lighting enhancements that have been documented to occur in and around several major US cities (e.g. Westcott, 1995; Orville et al., 2001; Steiger et al., 2002).

For example, Westcott (1995) in a four-year study of summertime CG lightning in and around 16 different metropolitan areas, documented enhancements in lightning activity within and downwind of most of the urban areas. Extending Westcott's findings, Orville et al. (2001) used 12-years of CG lightning data (1989-2000) gathered by the NLDN to document a persistent, year-round, enhancement in CG activity downwind of the Houston metropolitan area (a city not studied by Westcott). Steiger et al. (2002), using the same data set (5 km spatial resolution), quantified the enhancement observed by Orville et al., reporting a 45% increase in annual CG lightning flash densities over and downwind of the Houston urban corridor relative to rural surroundings. Their analysis indicated that the enhancement in CG lightning was associated with "large" lightning events, defined as days in which the sum of ground flashes detected within three separate 0.7° latitude by 0.85° longitude boxes (see Figure 1 boxes A, B and C) was greater than, or equal to 100 flashes on any given day. Their findings are generally consistent with Westcott (1995), indicating that observed enhancements in CG flash densities can occur over, and down-wind of urban corridors. Westcott, Orville et al. and Steiger et al., all propose explanations for the observed local enhancements that revolve around "urban" effects, specifically: (1) enhanced convergence, thermodynamic instability, or dynamical influences associated with the urban heat island; (2) altered microphysical processes associated with anthropogenic pollution; and/or (3) mesoscale enhancements in sea breeze convergence.

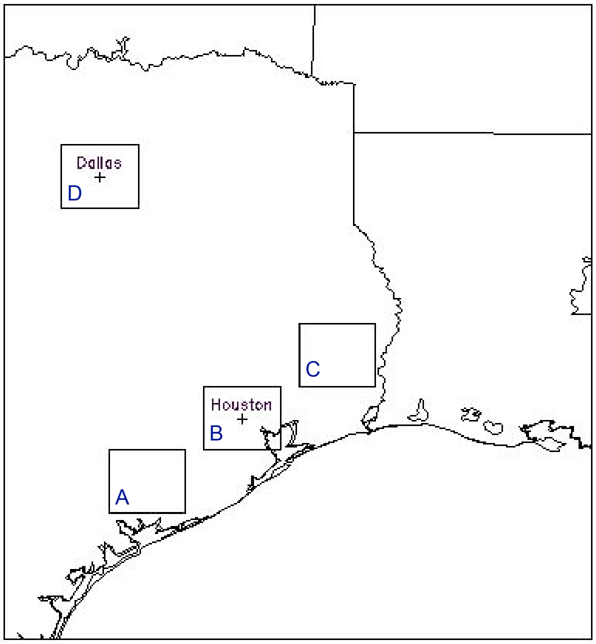


Figure 1. Domain over which cloud-to-ground lightning statistics were generated. Boxes A, B and C (each 0.7^{O} latitude by 0.85^{O} longitude in size) were used in the analysis performed by Steiger et. al. (2002), and are included for reference, with box B approximating the "urban" area associated with the Houston metropolitan area. For comparison, box D approximates the "urban" area associated with the Dallas metropolitan area (same dimensions).

---

* *The views expressed in this paper are those of the author and do not reflect the official policy or position of the U.S. Air Force, Dept. of Defense or the U.S. Govt.*
Corresponding author address: Major Michael L. Gauthier, Dept. of Atmospheric Science, Univ. of Alabama in Huntsville, Huntsville, AL 35899, e-mail: Michael.Gauthier@msfc.nasa.gov

The intent of this paper is to: (1) extend the findings of Orville et al. (2001) and Steiger et al. (2002) by presenting a statistical analysis of the variance associated with daily summertime (June, July and August) CG flash densities observed throughout eastern Texas and Louisiana, (2) scrutinize the underlying uniqueness of the signal in a regional sense, and (3) examine how flash density characteristics change by selectively excluding relevant subsets of the data (days without lightning, large events, etc.) from subsequent calculations.

2. DATA AND METHOD

Herein we create a regional climatology of total ground flash densities based on 9 years (1995 – 2003) of NLDN CG lightning data for the months of June, July and August for portions of eastern Texas and Louisiana (Figure 1). A complete description of the NLDN can be found in Cummins et al., 1998. We note that following completion of the 1995 NLDN upgrade the detection of a large population of low peak current (<10 kA) positive CG flashes over localized areas began to emerge. In order to compensate for the bias caused by this increase we have chosen to eliminate those flashes from the dataset following the recommendations of Cummins et al. (1998) and Wacker and Orville (1999a,b). Finally, by limiting our analysis to only summer months, we (1) isolate the dominant portion of the annual lightning activity cycle (i.e. summer season) and (2) eliminate the need to account for pre and post-upgrade changes in detection efficiency and/or location accuracy.

2.1 CALCULATION OF FLASH DENSITY

Descriptive statistics such as the variance and cumulative distribution of the daily flash density (flashes/km^2/day) are most easily calculated using an analysis grid of constant spatial resolution. Therefore, ground strikes detected within the period of record for our domain (Fig. 1) were interpolated onto a 143 × 139 Cartesian grid with fixed spatial resolution of 5 km. The total flash density for each of the 19,877 grid points was then calculated for each of the 828 days analyzed, yielding a dataset containing greater than 1.6×10^7 data points from which various descriptive statistics were calculated.

2.2 MEAN AND VARIANCE CALCULATIONS

Daily flash densities for each grid point were used to generate spatial distributions of the nine-year mean summer flash density (flashes/km^2/summer) and the variance of the flash density at each pixel. For these calculations, the temporal sample size (N) is equal to 828 days (92 days/summer × 9 summers).

For further comparison, we partitioned the analysis into "Filtered", "Conditional" and "Conditional-Filtered" mean flash densities, and variances thereof. Filtered mean flash density and variance were calculated in the same manner as the seasonal mean (and variance) except that pixels classified as "large event days" were excluded from the dataset thereby creating a variable sample size from pixel to pixel ($\overline{N}$ = 808 days, $\overline{N}$ = mean number of summer days used in the filtered sample across the entire domain). A "large event day" was defined for a pixel as a day having a flash density that fell within the upper quartile of the positively-skewed flash density distribution of the gridded domain (see Section 2.3 for further discussion). For the case of the conditional mean flash density and variance, rather than eliminating pixels classified as "large event days", we excluded pixels classified as "non-lightning days", or containing no lightning ($\overline{N}$ = 81 days). Finally, the conditional-filtered mean flash density and variance, as their names imply, are a combination of the conditional and filtered partitions, where all pixels classified as "non-lightning" or "large event" were removed from the dataset ($\overline{N}$ = 61 days), in a sense sampling closer to the "inner fence" of the data (Wilkes, 1995).

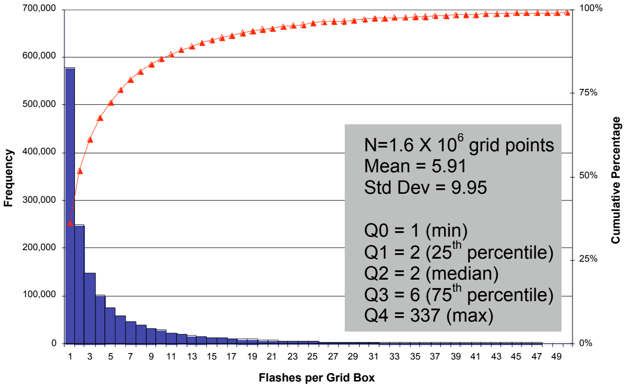


Figure 2. Frequency of occurrence (left ordinate) of daily flash counts for each 5 km x 5 km gridbox located within the entire domain (Fig 1) for June, July and August (JJA) 1995 – 2003. Tabulated statistics are for number of flashes occurring per grid box. Cumulative distribution is indicated on right ordinate.

2.3 "LARGE EVENT" CLASSIFICATIONS

Using daily flash counts for each 5 km × 5 km pixel we used the statistics of the daily flash count distribution to guide classification of "large" event days within our dataset. Due to the skewed, non-Gaussian nature of the flash density distribution (Fig. 2), we were unable to use sample means plus 1 (or 2) standard deviation(s) to identify "large" events. Instead we chose thresholds appropriate to the nature of the flash density histograms based on the cumulative *conditional* flash density distribution (conditioned on the presence of lightning). Specifically, we chose lightning days whose flash counts exceeded the 75th percentile (or fell within the upper quartile) of the cumulative distribution. Using the third quartile (Q3 = 6 flashes) as our threshold allows for an objective classification of the tail of the distribution as "large" events. Therefore, our objective definition of "large" events is any pixel whose flash count is in excess of 6 flashes in a given Julian day (or flash density > 0.24 flashes/km^2), which by definition must constitute 25% of the conditional flash density distribution.

This definition is different than that used by Steiger et al. (2002); in their analysis it was suggested that flash density enhancements over the Houston area were due to those days in which the sum of flashes within three separate geographic boxes (Figure 1, boxes A, B and C) was greater than, or equal to 100. For the purposes of the analysis herein, using 100 flashes as the defining threshold for "large" events is not appropriate. For example, if we reconstruct the Steiger et al. analysis by summing the summer season flashes in each of the three geographic regions depicted in Figure 1 for the period of 1989 – 2003 (Steiger et al. stopped at 2000), we find that during the 1,380 day period there were 1,075 days in which 1 or more flashes occurred in at least one of the three boxes. Of those 1,075 days, 68% (or 733 days) met or exceeded the 100 flash count threshold used in their analysis. This means that only 32% of the summertime lightning days contained less than 100 flashes (total in all 3 boxes). From a statistical standpoint, in this study we did not wish to classify 68% of the lightning distribution as "large" events. Using the quartile method, we find that 1028 flashes/day (Q3 for this distribution) would be comparable to the "large" event classifier specified in the Steiger et al. analysis.

3. RESULTS & DISCUSSION

We begin discussion with a map that shows the number of lightning days occurring for each pixel within our regional domain (Fig. 3). Here we clearly see the influence of mesoscale interactions in shaping the location and frequency of lightning events along the Gulf Coast, with a larger number of lightning days occurring in the Houston area and extending eastward along the coastal region. Comparing the lightning days (Fig. 3) to mean lightning flash density (Fig. 4a), we find flash density enhancements in areas experiencing a greater number of storm days co-located along apparent coastal convergence zones, likely associated with the sea breeze, and evident irregularities in the coastline (e.g. small inlets, bays, coastal lakes etc.; see arrows in Fig. 4a). Consistent with Orville et al. (2001) and Steiger et. al. (2002), we also note the occurrence of a definite localized enhancement in flash densities situated over the Houston metropolitan area (Fig. 4a). The "enhancements" spread eastward from Houston, with similar localized maxima distributed along the coast of southeastern Texas and central portions of Louisiana. Examination of the spatial variance of the mean daily summer flash densities along the Gulf Coast in particular (Fig. 4b), reveals significant variance, often times in excess of the mean, collocated with areas of larger mean flash density, and in particular, in the areas of local "enhancement".

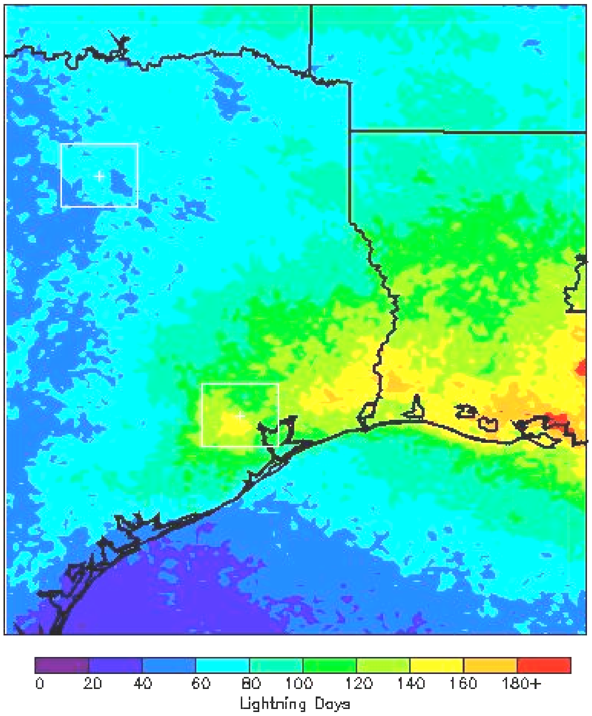


Figure 3. Number of summer season (JJA) lightning days (1995 – 2003); 828 possible days.

Normalizing the mean flash density of each pixel ($\hat{x}$) by the mean flash density of the entire domain ($\bar{x}$), we can define domain flash density "anomalies" (i.e. $\hat{x}/\bar{x}$). For this parameter, values greater then one are referred to as "positive" anomalies and those less than one "negative" anomalies. This approach scales the flash density values across samples (i.e. mean, filtered, conditional and conditional-filtered mean flash density) so that the mean flash density of each sample is equivalent (i.e. mean = 1), thereby enabling a direct comparison between anomaly patterns between the four samples. Similarly, the coefficient of variation (pixel standard deviation normalized by the domain mean) was computed to allow for inter-comparisons of the variance fields for each of these scenarios. Flash density anomalies shown in Figure 5a, reveal that absolute flash densities in the vicinity of the Houston area and along the coast, east of Galveston Bay, are on the order of 2.5 times the domain mean, while flash densities in the Dallas area, as well as throughout north and central Texas, are predominantly centered on the mean. The Dallas area was included for comparison due to the fact that it is another major urban area exiting within a different meteorological regime within our domain. Table 1 provides comparative quantitative statistics for each of three geographic boxes: (1) the Houston area, (2) the entire domain, and (3) the Dallas area.

Referring to Table 1, averaged over the entire domain, the results indicate that 25% of the days in which lightning occurred were classified as "large

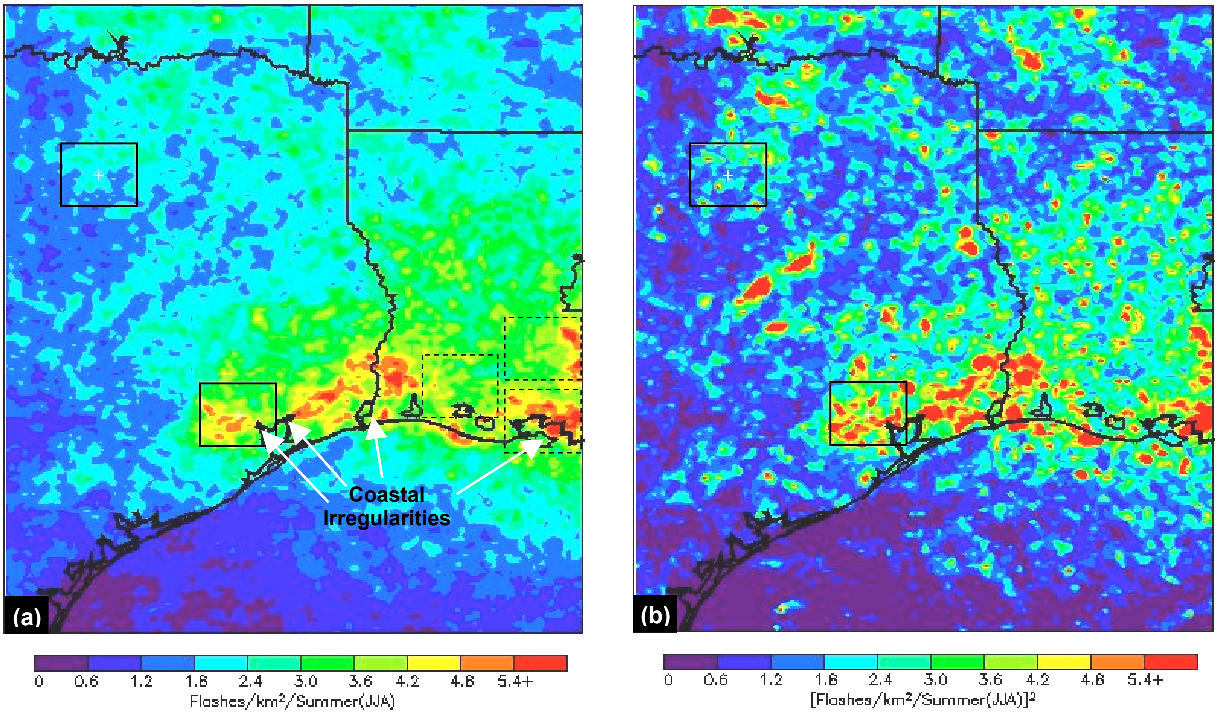


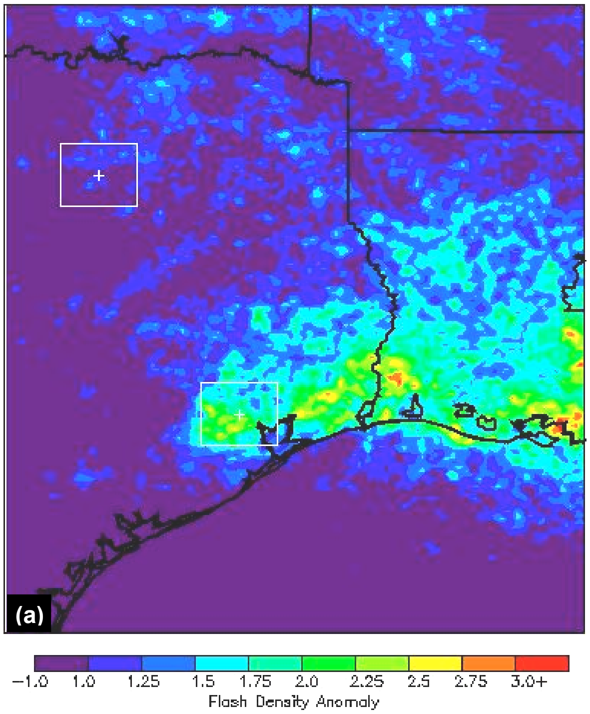
Figure 4. Spatial variations of the 9 year (a) mean summer season (JJA) ground flash density and (b) variance of the daily mean. Arrows in Fig 4a highlight referenced coastal irregularities, while dashed boxes are additional locations in which a student's t-test was applied.

events" (i.e. those days in which flash densities fell within the upper quartile of the cumulative flash density distribution), and that these large events contributed over 70% of the mean flash density within the domain. In order to evaluate the influence of large events on the flash density patterns, we created a spatial distribution of the filtered mean flash density by removing large events from the sample (not shown). As expected, the magnitude of the maximum flash densities decreased significantly across the domain, with the domain mean decreasing to 0.61 flashes/km^2/summer. Normalizing each pixel by the domain filtered mean we are able to examine the spatial distribution of the filtered mean flash density anomaly (Figure 5b). Here we see that flash densities in the vicinity of the Houston area, and immediately east of Galveston Bay are 1.5 to 2.5 times that of the domain mean, with flash densities increasing eastward along the coast reaching peaks in southeastern Louisiana. While there is some suppression of the local anomaly over Houston (see Table 1, line 2), it is important to note that *the enhancement in CG lightning over the Houston area persists even with the removal of these large events*. Further inland, flash densities in the central Texas region have decreased to 0.5 to 1 times the mean. In particular, the mean flash density over the Dallas area is almost half that observed over the Houston area (Table 1). These results indicate that the persistently enhanced flash densities present in the vicinity of the Houston area, as well as along the Gulf Coast convergence zone (as in Figure 5b) are not solely due to the occurrence of large events, as proposed by

		Houston						Domain						Dallas					
		Mean Number of Days	Contribution by "Large Events"	Flash Density (fl/km^2/sumr)	Flash Density Anomaly	Variance (fl/km^2/sumr)^2	Coefficient of Variation	Mean Number of Days	Contribution by "Large Events"	Flash Density (fl/km^2/sumr)	Flash Density Anomaly	Variance (fl/km^2/sumr)^2	Coefficient of Variation	Mean Number of Days	Contribution by "Large Events"	Flash Density (fl/km^2/sumr)	Flash Density Anomaly	Variance (fl/km^2/sumr)^2	Coefficient of Variation
1.	**Mean**	828	76%	3.79	1.78	3.91	0.91	828	71%	2.13	1.0	1.88	0.60	828	77%	2.01	0.94	2.06	0.65
2.	**Filtered Mean**	794	N/A	0.92	1.51	0.11	0.54	808	N/A	0.61	1.0	0.07	0.44	809	N/A	0.47	0.77	0.06	0.40
3.	**Conditional Mean**	124	68%	25.3	1.18	20.2	0.21	81	62%	21.3	1.0	13.9	0.16	64	68%	25.9	1.21	19.9	0.20
4.	**Cond.-Filtered Mean**	90	N/A	8.11	1.01	0.33	0.07	61	N/A	8.01	1.0	0.33	0.07	45	N/A	8.35	1.04	0.35	0.07
		Houston						**Domain**						**Dallas**					

Table 1. Comparative quantitative statistics for the Houston and Dallas metropolitan areas as well as for the entire domain.

Steiger et al. (2002); we speculate that *the enhancements are due, in part, to a much greater frequency of occurrence* (see Figure 3), likely associated with the frequent forcing along coastal mesoscale boundaries.



Although the flash density and variance calculations presented thus far include non-lightning producing periods, additional insight as to the relative importance of the anomalies and intensity of storms throughout the domain can be gained by examining only days with lightning producing storms. If only lightning days (i.e. one or more CG flashes in a grid box) are considered in our analysis, a marked change in the flash density anomaly pattern becomes evident, relative to patterns in the absolute mean anomalies presented in Figure 5a. As expected, conditional mean flash densities increase significantly by removing the non-events from our analysis (Table 1, line 3), but more intriguing is the fact that the flash densities become more uniform across the domain (smaller mean coefficient of variation), with the positive anomaly previously evident in the Houston area being significantly diminished. In fact, peak conditional mean flash densities within the domain are now located in, and north of, the Dallas area. Further examination of Fig. 5c reveals that the mean flash density in the Houston area is only *slightly* greater than the domain mean, and more similar to that observed in central Texas, south of Dallas. By removing all data points classified as either non-lightning days or “large event” days, we find that the conditional-filtered mean flash density has all but converged on the domain mean with only slight spatial variation about this value (not shown), with no sign of an “urban” anomaly over Houston.

The disappearance of the Houston flash density

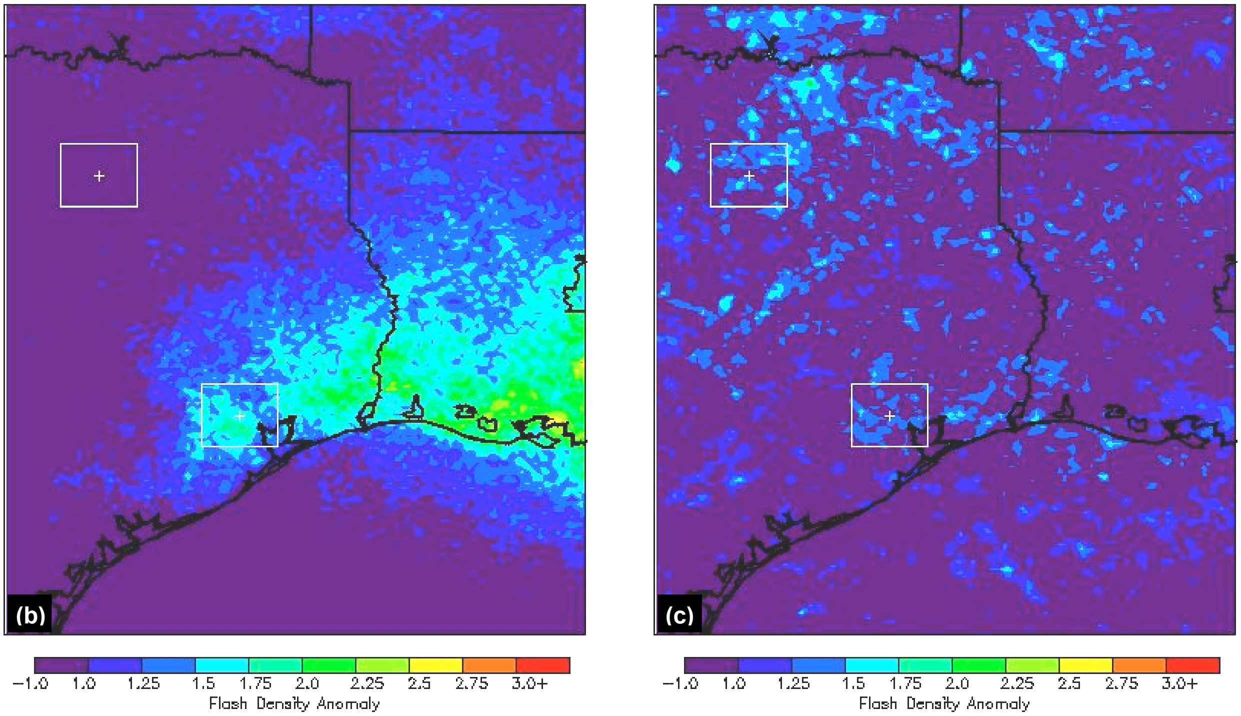


Figure 5. Summer season (JJA, 1995 – 2003) flash density anomalies created by normalization of (a) the mean flash density of each pixel by the domain mean, (b) the filtered pixel mean flash densities by the filtered mean of the domain (i.e. “large” events removed from analysis) and (c) the conditional pixel mean flash densities by the conditional mean of the domain (i.e. only days in which lightning occurred were included in the analysis). In all cases, values > 1 indicate positive anomalies and values < 1 negative anomalies.

anomaly, just described, is the result of eliminating both the non-lightning days as well as the "large" event days from our analysis. This indicates that both increased frequency, as well as the occurrence of large event days contribute to the flash density "enhancements" observed over the Houston area (Figure 4a). Table 1 (lines 3 and 4) and Figure 3, show that on average, there are twice as many lightning days occurring over the Houston area, in comparison to the Dallas area, but the conditional mean flash density in each of these locations are very similar (Houston: 25.3 flashes/km^2/summer; Dallas: 25.9 flashes/km^2/summer). This suggests that, on average, storms occurring over the Dallas area may produce twice as much lightning as those occurring over the Houston area. Further, these findings are not dependent on the number of "large" events in the sample.

4. CONCLUSIONS

The documented enhancement in cloud-to-ground lightning flash densities in and around the Houston metropolitan area was examined using 9 years of NLDN cloud-to-ground lightning data. Collectively, our findings indicate, as in previous studies (i.e. Orville et al., 2000 and Steiger et al., 2002) that the summer season flash density anomaly situated over the Houston area is a robust feature that continues to persist even when large event days associated with the upper quartile of the positively skewed flash density distribution are removed from the analysis. This subset comprises less than 5% of the total daily sample, but produces in excess of 75% of the total lightning in the Houston area.

By comparison, when examining inland regions (e.g., toward Dallas) we found that even fewer large events (relative to the total daily sample) produced roughly the same contribution to the total flash density in the Dallas area. If only lightning-days were averaged to produce a "conditional mean" we found that the anomaly in the Houston area became almost nonexistent, and that the conditional mean flash density was actually larger moving into central and northern Texas. Our combined findings suggest that although the Houston area sees an increased frequency of lightning producing storms (including more large flash density events), storms occurring further inland (e.g., in and north of the Dallas area), actually appear to produce more lightning on an event basis.

Finally, our findings highlight the fact that the local Houston cloud-to-ground lightning anomaly, while being a spatially intriguing and persistent feature, is non-unique along the Gulf Coast. There are numerous areas of enhanced mean flash density located along the southeastern coast of Texas and Louisiana. Application of a simple two-sample t-test comparing the means and variances of CG flash density for numerous Gulf Coast locations (see dashed boxes in Fig 4a), including Houston indicate that although the Houston flash density enhancement clearly exists in a spatial sense, the flash density magnitude compared to other coastal locations is not statistically unique (at p-values < .05). Although hypotheses invoking anthropogenic influences have been offered by Orville et al. (2001) and Steiger et al. (2002) to explain the observed increases in flash density over the Houston area (e.g. aerosol influences on storm microphysics, urban heat island, etc.), it seems equally plausible that regular daily mesoscale influences on convective forcing associated with the coastline may also contribute to the observed Houston "anomaly". On the other hand, the mesoscale influence of the coastline and irregularities in the coastline do not explain previously documented lightning enhancements observed over and down wind of other inland cities, and therefore may have little to do with the Houston spatial anomaly in flash density. Clearly, the problem is complex and, to the extent possible, requires a comprehensive set of surface and tropospheric measurements (e.g., aerosol, cloud microphysical and precipitation structure, total lightning, thermodynamic, land surface, regional circulation etc.) that targets each of the most likely forcing mechanisms.

Acknowledgments. Post-processed lightning data were provided by the *NASA Lightning Imaging Sensor (LIS) instrument team and the LIS data center via the Global Hydrology Resource Center (GHRC) through a license agreement with Global Atmospherics, Inc (GAI).*

5. REFERENCES

Biswas, K. R., and P. V. Hobbs, 1990: Lightning over the Gulf Stream. *Geophys. Res. Lett.,* **17**, 941-943.

Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, 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.

Gauthier, M.L., 1999: Mutliparameter investigation of significant lightning producing storms in northeastern Colorado. MS Thesis, Dep. Atmos. Sci., Colorado State University., 164pp.

Orville, R. E., 1990: Winter lightning along the east coast. *Geophys. Res. Lett*., **17**, 713-715.

Orville, R. E., R. W. Henderson, L. F. Bosart, 1983: An East Coast Lightning Detection Network. *Bulletin of the American Meteorological Society*: Vol. 64, No. 9, pp. 1029–1030.

Orville, R. E., G. R. Huffines, 2001: Cloud-to-Ground Lightning in the United States: NLDN Results in the First Decade, 1989–98. *Monthly Weather Review*: Vol. 129, No. 5, pp. 1179–1193.

Orville, R. E., G. R. Huffines, J. Nielsen-Gammon, R. Zhang, B. Ely, S. Steiger, S. Phillips, S. Allen, W. Read, 2001: Enhancement of cloud-to-ground lightning over Houston, Texas, Geophys. Res. Lett., **28**(13), 2597-2600.

Petersen, W. A., and S. A. Rutledge, 2004: Meteorological forcing of lightning over the Gulf Stream. *Transactions*, Fall Meeting of the American Geophysical Union, 13-17 December 2004, San Francisco, California.

Reap, R. M. 1986: Evaluation of Cloud-to-Ground Lightning Data from the Western United States for the 1983–84 Summer Seasons. *J. of Appl. Meteor.,* **25**, 785–799.

Steiger S. M., R. E. Orville, and G. R. Huffines, 2002: Cloud-to-ground lightning characteristics over Houston, Texas: 1989–2000, J. Geophys. Res., **107** (D11).

Wacker, R. S., R. E. Orville, 1999a: Changes in measured lightning flash count and return stroke peak current after the 1994 U.S. National Lightning Detection Network upgrade: 1. Observations, J. Geophys. Res., **104**(D2), 2151-2158.

Wacker, R. S., R. E. Orville, 1999b: Changes in measured lightning flash count and return stroke peak current after the 1994 U.S. National Lightning Detection Network upgrade: 2. Theory, J. Geophys. Res., **104**(D2), 2159-2162.

Westcott, N. E., 1995: Summertime Cloud-to-Ground Lightning Activity around Major Midwestern Urban Areas. *Journal of Applied Meteorology*: Vol. 34, No. 7, pp. 133–1642.

Wilkes, D. S., 1995: Statistical methods in the Atmospheric Sciences. *Academic Press*, 465 pp.

Zajac, B. A., S. A. Rutledge, 2001: Cloud-to-Ground Lightning Activity in the Contiguous United States from 1995 to 1999. *Monthly Weather Review*: Vol. 129, No. 5, pp. 999–1019.