2010 21st International Lightning Detection Conference
19 - 20 April • Orlando, Florida, USA
3rd International Lightning Meteorology Conference
21 - 22 April • Orlando, Florida, USA

USING THE GAUSSIAN-BASED CONFIDENCE ELLIPSE PROVIDED BY LIGHTNING LOCATION SYSTEMS TO CREATE MORE RELIABLE LIGHTNING DENSITY MAPS

V. Bourscheidt, O. Pinto Jr., K. P. Naccarato

Brazilian National Institute for Space Research
São José dos Campos – Brazil

1. INTRODUCTION

Nowadays, one of the most common applications of data obtained by Lightning Location Systems (LLS's) is the creation of lightning density maps for spatial distribution analysis. These data has brought possible to obtain the distribution of the lightning over large areas around the world and the density estimative has become essential in many applications. And for a good application of this data, an important point to include in the analysis are the errors resulting from the system solutions.

The errors regarding the lightning data might be divided into two groups: the temporal error and the spatial error. The temporal error would depend basically on the amount of lightning data and on the network efficiency. On the other hand, the spatial error would be more related to the system location accuracy, the grid resolution and to the different existent methods for the data interpolation. All this aspects lead to errors in the density maps, which can be quantified and included, for example, in lightning strike risk analysis.

The amount of data is important to reduce the (temporal) error on the estimative. Diendorfer et al (2008), based on the Poisson probability distribution and assuming the lightning as a stochastic phenomenon stated that a value of 80 discharges in each grid cell, for example, would be associated to a 20% achievable uncertainty for random lightning discharges.

It is also important, when considering the amount of data, analyze the grid resolution. In our case, the resolution has a particular importance due its dependence on the confidence ellipses size. The confidence ellipses are a product of the LLS's to quantify random and systematic errors. It results from a 2D Gaussian model related to the location accuracy and it size depends basically on the on the network geometry (Cummins *et al.*, 1998). To reduce the spatial errors and uncertainties, the map (or grid) resolution should not be higher (or finer) than the size of the ellipse. But in some cases, like Brazil, where the VLF/LF network is hybrid (TOA and IMPACT technology) and some sensors have a large downtime or gaps (Naccarato *et al.*, 2008), it is reasonable to find spatial variation in the ellipses size, and this variation become very important on the resolution of the density maps for the different regions.

A way to overcome this point might be the use of the ellipses as a reference to determine whenever a discharge is inside a grid cell or not. On this way, an early method was already used, which used the confidence ellipse (50%) area to represent the discharge and the proportional area inside each grid cell (Campos and Pinto Jr., 2007). However, this approximation gives to every place inside the ellipse region the same probability (i.e., depends only on the area of the ellipse).

Our idea here is use the confidence ellipses by reconstructing the Gaussian distribution for each discharge and evaluate the probability (or proportional probability) of the discharge for each map grid cell. This approximation is more realistic and can include more appropriately the spatial error on the maps.

2. DATA AND METHODOLOGY

To evaluate the proposed method, we have used as reference a part of the Southeast region (especially over the São Paulo State, as showed in the Figure 2), which is on the center of the network and at the expected higher detection efficiency region. The resolution of the grids is 5 kilometers. Cloud-to-ground (CG) lightning data were obtained from the Brazilian National Lightning Detection network (BrasilDAT) for the period from Jan 2006 to May 2008.

Three different methodologies were evaluated in the map creation (as showed in Figure 2):

- Simple data count, were each flash is represented by a defined Lat/Lon and included in the grid cell;
- The weighted count, were every flash is represented by an area (obtained through the confidence ellipse), which is proportionally calculated for each grid cell.

- The so called "Gaussian count", where the ellipses are used in intervals and each interval is proportionally calculated for each grid cell.

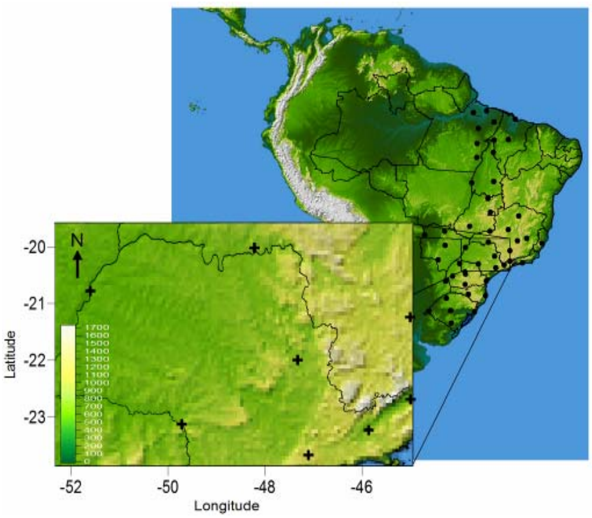

Figure 1. Region used for evaluation: sensors are represented by black dots. Green to yellow indicates increasing altitudes.

To obtain the intervals the equation available from Vaisala, Inc. was used:

$$Scaling\ Const = \frac{\sqrt{-2 \cdot \ln(1 - probability)}}{1.177}$$

Each probability step result in a scale constant that is used to build the ellipses set, given the Gaussian aspect for each flash. The final percentage for each flash inside each grid cell could be represented by the following equation (see Figure 2):

$$P_{flash/grid\ cell} = \frac{\sum_{i=1}^{n} \left(\frac{A_{Gridcell}}{A_{Ring(i)}}\right) \cdot P_{(ring)}}{P_{max}}$$

To calculate this, pairs of ellipses were used as rings with a certain probability (P_{ring}). For each of these rings, the proportional area regarding each grid cell was calculated and then a final percentage was obtained. This percentage might also denotes the probability of a flash be inside a grid cell. The P_{max} represent the maximum probability of the ellipse and is used to assume this maximum to represent one flash.

To evaluate each method, three ellipse size (probability limits) were used: 50, 90 and 99% of confidence with steps ranging from 3 to about 5%. The difference among the simple count and the other methods were also calculated. The results were analyzed through maps.

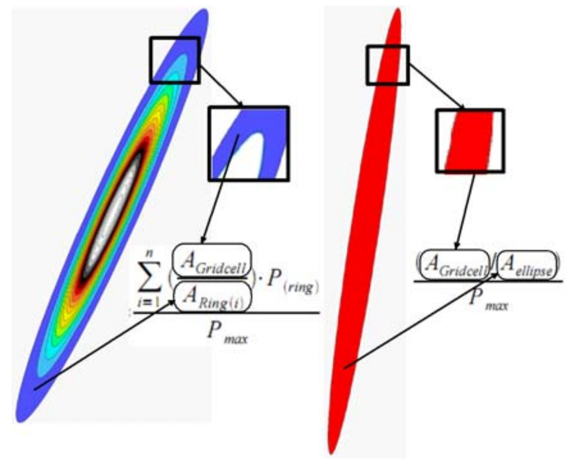

Figure 2. Illustration of the weighted count and Gaussian count methods for each map grid cell.

3. RESULTS AND DISCUSSION

3.1 Effects of the weighted count method

To make the comparisons, we used the simple data count as reference, due its common use in the lightning maps. The maps were made using ranges colors, to maintain the values unchanged and better see the impact of each method and ellipse size.

The results show very small variations for the entire region. The more significant effect that occurs when using the weighted count (with 50% confidence ellipse, Figure 4) is the smoothness of the values, especially if we look at the highlighted region. The same occur for the other (not showed) ellipse conditions (90% and 99% probability), with small differences related to the first case.

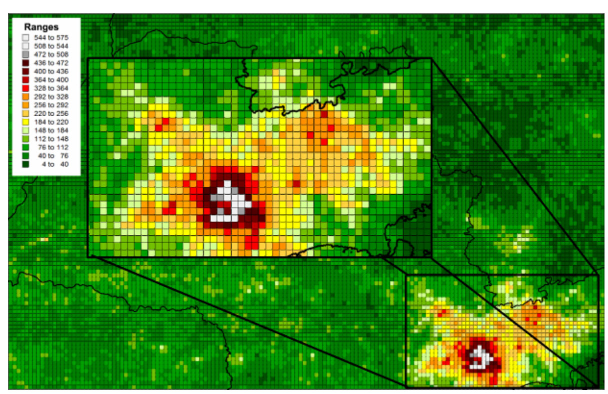

Figure 3.CG Lightning count using the simple count method. Zoom over the São Paulo metropolitan region.

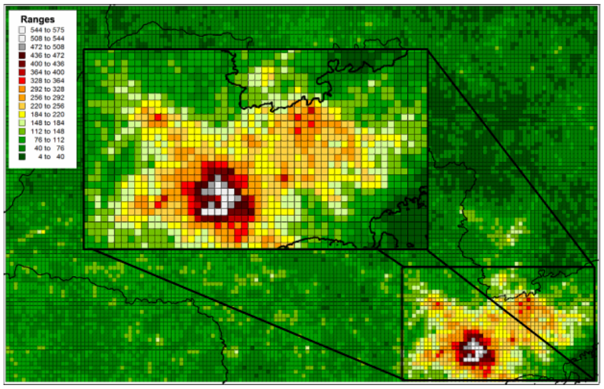

Figure 4. CG Lightning count using the weighted count with the 50% confidence ellipse.

The more significant variations between each early mentioned method are found in the region where the large number of lightning takes place. To show this variation, a "difference map" was done and is showed in Figure 5.

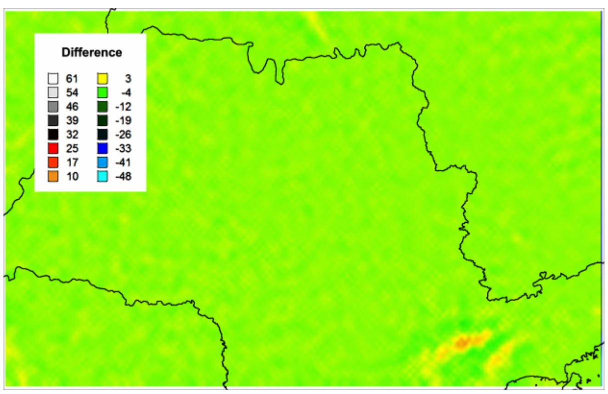

Figure 5. Map of the difference between the simple count and the weighted count for the 50% confidence ellipse.

This variation is also observed in Table 1: with the mean around zero, the variance increase (i.e., more high difference values) as we use large ellipse sizes. This increase indicates a growing smoothness on the maps.

3.2 Effects of the "Gaussian" count method

In the case of the Gaussian method, the values were almost the same as the values for the weighted count, as is possible to see in the Figure 6.

There were not significative spatial variations among the different ellipse limit size, as is possible to see by comparing Figures 6 (50 % ellipse) and 7 (99% ellipse). But this could be changed if other resolution is used to make the maps. The used resolution is larger than the mean ellipse size at this region and this might been affecting the results.

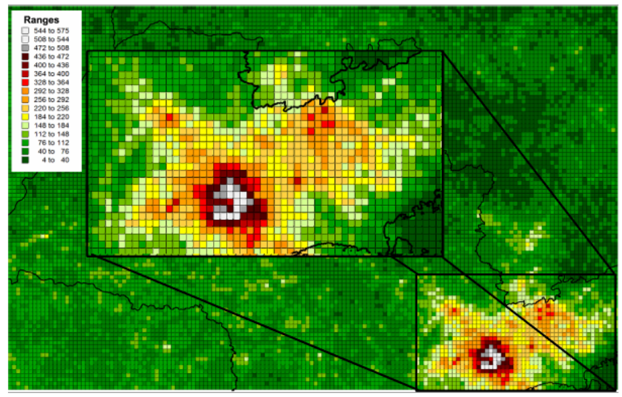

Figure 6. CG Lightning count using the "Gaussian" count with the 50% confidence ellipse.

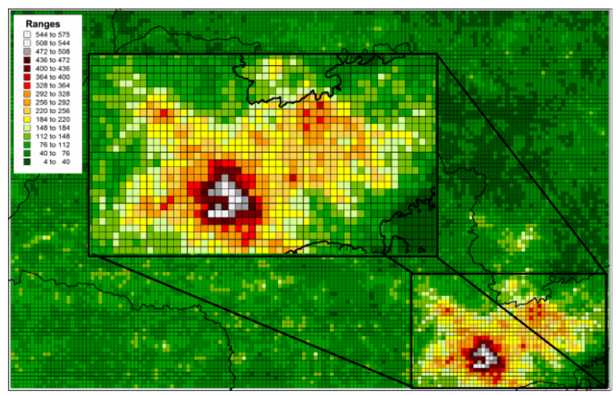

Figure 7. CG Lightning count using the "Gaussian" count with the 99% confidence ellipse.

.

The difference analysis (Table 1 and Figure 8) for this case shows also an increase in the variance for larger ellipse sizes. On the other hand, it shows smaller values when compared to the weighted count method. This might be related to the process used on the calculus: the weighted method assumes the same probability for all the region cover by the ellipse. The Gaussian method assumes that the probability increase as near from the estimated flash location. So, it can be expected that the difference (and also the smoothness) of the Gaussian method related to the simple count will be smaller.

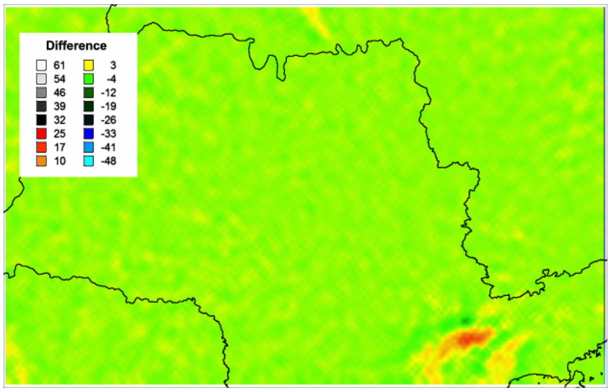

Figure 8. Map of the difference between the simple count and the "Gaussian" count for the 99% confidence ellipse.

Table 1 – Descriptive statistic parameters for the difference among all methods and ellipse size.

Method	***Mean***	***Variance***	***Kurtosis***
Weighted 50%	-0,00250	9,20	5,18
Weighted 90%	-0,00476	16,00	7,94
Weighted 99%	-0,00700	22,29	10,40
Gaussian 50%	-0,00312	8,48	5,04
Gaussian 90%	-0,00337	11,98	6,71
Gaussian 99%	0,22151	12,98	7,97

4. CONCLUSIONS

For each method we do not find large variations on the spatial distribution. The most significant variations occur in the region with large amount of data. The variations increase for increasing ellipse sizes, especially for the weighted count method, which assumes the same probability for all the region cover by the ellipse, inducing the smoothness. For the Gaussian method (where the probability is large nearby the estimated lightning position) the variations increase but less than the early method.

The effects of each method tend to be greater if the resolution of the map is higher than the mean ellipse size. In these cases, the method of the Gaussian count should be the best to evaluate the lightning distribution, since it includes the expected random and systematic errors and gives the appropriated probability distribution for the flash location.

4. ACKNOWLEDGMENTS

The results are part of a PhD Thesis of the first author. The authors would like to acknowledge the National Council for Scientific and Technological Development (CNPq) and the CAPES (Coordenação para Aperfeiçoamento de Pessoal de Nível Superior).

5. REFERENCES

CAMPOS, D. R.; PINTO JR., O. Investigation about the intensity and location of the maximum cloud-to-ground lightning flash densityin the city of São Paulo. *In: International Simposium on Lightning Protection*, Foz do Iguaçu. 2007.

CUMMINS, K. L. *et al.* The US national lightning detection network and applications of cloud-to-ground lightning data by electric power utilities. *IEEE Transactions on Electromagnetic Compatibility* [S.I.], v. 40, n. 4, p. 465-480, 1998.

DIENDORFER, G. Some comments on the achievable accuracy of local ground flash density values. *In: 29th International Conference on Lightning Protection*, Uppsala, Sweden. 2008. p.281-286.

NACCARATO, K. P. *et al.* Performance analysis of the BrasilDAT network. *In: International Conference on Grounding and Earthing 2008 and 3rd International Conference on Lightning Physics and Effects*, Florianópolis, Brazil. 2008. p.329-338.