Long-range spatial correlations and fluctuation statistics
of lightning activity rates in Brazil

H. V. Ribeiro$^{1,2,3}$, F. J. Antonio$^{1,3}$, L. G. A. Alves$^{1,3}$, E. K. Lenzi$^{1,3}$ and R. S. Mendes$^{1,3}$

$^1$ Departamento de Física, Universidade Estadual de Maringá - Maringá, PR 87020-900, Brazil
$^2$ Departamento de Física, Universidade Tecnológica Federal do Paraná - Apucarana, PR 86812-460, Brazil
$^3$ National Institute of Science and Technology for Complex Systems, CNPq - Rio de Janeiro, RJ 22290-180, Brazil

received 5 October 2013; accepted in final form 3 December 2013
published online 8 January 2014

PACS 92.60.Pw – Atmospheric electricity, lightning
PACS 89.75.-k – Complex systems
PACS 89.75.Kd – Patterns

Abstract – We report on a statistical analysis of the lightning activity rates in all Brazilian cities. We find out that the average of lightning activity rates exhibits a dependence on the latitude of the cities, displaying one peak around the Tropic of Capricorn and another one just before the Equator. We verify that the standard deviation of these rates is almost a constant function of the latitude and that the distribution of the fluctuations surrounding the average tendency is quite well described by a Gumbel distribution, which thus connects these rates to extreme processes. We also investigate the behavior of the lightning activity rates vs. the longitude of the cities. For this case, the average rates exhibit an approximate plateau for a wide range of longitude values, the standard deviation is an approximate constant function of longitude, and the fluctuations are described by a Laplace distribution. We further characterize the spatial correlation of the lightning activity rates between pairs of cities, where our results show that the spatial correlation function decays very slowly with the distance between the cities and that for intermediate distances the correlation exhibits an approximate logarithmic decay. Finally, we propose to model this last behavior within the framework of the Edwards-Wilkinson equation.

Introduction. – The search for better understanding of earth-related systems is a stimulating challenge often present in the physicists agenda. Examples of such investigations include study of earthquakes [1–4], geomagnetic activities [5–7], climate [8,9] and weather-related systems [10,11]. Usually, these investigations examine data aiming to uncover patterns or laws that rule the spatiotemporal dynamics of the systems. These empirical-based approaches are also very useful to provide basis for comparing and testing the increasing number of models studied in these fields.

In the previous context, a well-known phenomenon is the lightning strike, which displays an intrinsic complex behavior on its own. Physically speaking, lightning is a massive electric discharge between electrically charged regions in one or more clouds, or between clouds and an earth-bound object (cloud-to-ground lightning). Usually this complex process involves more than one current flow and lasts about half a second, while each one of these flows can take only milliseconds to happen [12]. Despite lightnings having short duration, they are well-known by causing severe injuries or even death when striking humans or animals, and for the damage they usually provoke to buildings and power networks [13,14]. In this sense, monitoring and probing spatiotemporal patterns of lightning occurrences may not only prevent electrical discharge hazard but also provide useful markers for thunderstorms, which are among the the major causes of weather-related damages and economic losses in tropical and middle latitude regions [15].

Around 78% of all lightning events occur in tropical and middle latitude areas, that is, the majority of these events are concentrated between 30° N and 30° S latitudes [16,17]. Thus, understanding the lightning activity in these regions of the world is especially important. We devote this article to investigate the spatial dynamics of lightning activity rates in Brazil, the largest tropical country in the world, which is hit (on average) by more than

$^{(a)}$E-mail: hvr@dfi.uem.br
50 million lightning strikes each year. Furthermore, due its continental dimensions and the existence of reliable data, Brazil can be considered as an ideal place for probing spatial patterns of lightning activity. Here we shall study the lightning activity rates in connection with the geographic localization of all Brazilian cities aiming to uncover statistical patterns and to model these findings within a reductionist approach. In following, we present the data we have analyzed, the procedures we have employed to analyze the data as well as the empirical findings, and finally, we end with a summary and some concluding remarks.

Data presentation. – We have accessed data corresponding to the lightning activity rates (only cloud-to-ground lightning discharges) of all Brazilian cities made freely available by the National Institute for Space Research (INPE) [18–21]. These rates were estimated by INPE through data obtained by the sensor Lightning Imaging Sensor [22] that is aboard the Tropical Rainfall Measuring Mission [23]. The rates were calculated using data from the years 1998 to 2013 and are considered a reliable measure for all Brazilian cities. The data consist of the geographic location (latitude $\phi$ and longitude $\lambda$) and the average number of lightnings per km$^2$ per year ($L$). Figure 1 shows the spatial distribution of the lightning activity rates throughout the Brazilian territory.

Data analysis. – We start by investigating the dependence of the lightning activity rates $L$ on the latitudes $\phi$ of the cities. Figure 2(a) shows a scatter plot of $L$ vs. $\phi$ where, despite the considerable noise, we can identify some regularities. We observe the existence of systemically large values for the lightning activity rates around the Tropic of Capricorn and also around the Equator. In order to check whether these visual patterns are statistically meaningful, we have binned the data into 25 windows and, for each window, we calculate the average value (square) of the lightning activity rates as well as the 95% bootstrap confidence interval (error bar). As shown in fig. 2(a), the average value of $L$ is actually statistically significantly larger around the Tropic of Capricorn and the Equator. We note yet the existence of maximum values just before these geographic references.

We have also studied the dependence of the standard deviation of the lightning activity rates on the latitude. Figure 2(b) reveals that the standard deviation $\sigma$ is almost a constant function of the latitude. In particular, we note that practically all values of $\sigma$ are within the 95% confidence interval (light shaded area). A systematic exception occurs around the latitude $-10$, where the standard deviations remain outside of the 99% confidence interval for a constant behavior (dark shaded area).

Aiming to further characterize the lightning activity rates, we have evaluated the residuals surrounding the average tendency of $L$ vs. $\phi$. In order to do so, we have interpolated a spline of third order to the average data and then we calculate the residuals

$$\xi_i = L_i - F(\phi_i),$$

(1)

where $L_i$ is the lightning activity rate in the city $i$, $\phi_i$ is latitude of the city $i$ and $F(\phi_i)$ stands for the value of the interpolating function at the latitude $\phi_i$. Next, we investigate the probability distribution of $\xi_i$ as shown in fig. 2(c). We have found that the Gumbel distribution [24]

$$P(\xi) = \frac{1}{\beta} e^{(-\xi+\alpha)/\beta - e^{(-\xi+\alpha)/\beta}}$$

(2)

with the parameters $\alpha = -1.24$ (location parameter) and $\beta = 2.16$ (scale parameter) describes quite well our empirical distribution. The best fit parameters were obtained by the maximum-likelihood method and the p-value of the Cramér-von Mises test is 0.02, indicating that we cannot reject the Gumbel hypothesis at a confidence level of 99%.

In order to overcome possible bias related to the inhomogeneous spatial distribution of the Brazilian cities, we have also calculated the distribution of $\xi$ after grouping cities with close latitudes values, considering the same 25 windows employed for the average and standard deviation analyses (figs. 2(a) and (b)). The results show that the distributions of the residuals by latitude obey the same Gumbel form with parameters $\alpha$ and $\beta$ very close to those obtained for the whole dataset, as shown in fig. 2(d). It is worth noting that the Gumbel distribution is related to the distribution of the maximum of a set of $n \to \infty$ random numbers drawn from a distribution that asymptotically decays faster than any power law. This result thus suggests that the deviations surrounding the average tendency can be also understood as an extreme process. While it is not trivial to connect these results with the mechanisms underlying the physics of lightning process [25], we can consider a cloud-to-ground lightning as an
Fig. 2: (Colour on-line) Dependence of the lightning activity rates on the latitude of the cities. (a) Each red dot represents the lightning activity rate $L$ vs. the latitude of a city. The blue squares are window average values (25 equally spaced windows) and the error bars are 95% bootstrap confidence intervals for these means. The continuous line is a spline interpolation of third order to the average data. We note that the average value displays a maximum value close to the Tropic of Capricorn (horizontal dashed line) and also a local maximum just before the Equator (zero latitude). (b) Standard deviation of the lightning activity rates $\sigma$ vs. the latitude evaluated over same windows utilized for the average values. The blue squares are the standard deviations and the error bars are 95% bootstrap confidence intervals. The horizontal dashed line is the average value of standard deviations and the dark (light) shaded area represents the 95% (99%) bootstrap confidence intervals. We note that except for a few points, the standard deviations do not display statistically significant deviations from a constant behavior. (c) Probability distribution of the residuals of the lightning activity rates around the average tendency described by the spline interpolation. The bins were chosen through the Wand’s procedure and the inset shows the empirical distribution in log-log scale. The continuous line is the Gumbel distribution (eq. (2)) with location parameter $\alpha = -1.24$ and scale parameter $\beta = 2.16$, which were obtained via the maximum-likelihood fit. The $p$-value of Cramér-von Mises test is 0.02. (d) Values of $\alpha$ and $\beta$ of the Gumbel distribution when grouping the residuals by using the same 25 windows employed for the calculation the averages and standard deviations. These values were obtained via the maximum-likelihood fit considering only the windows with more than 250 points within and the asterisks indicate the windows where the Gumbel hypothesis cannot be rejected at a confidence level of 99%. The error bars are 95% bootstrap confidence intervals and the horizontal lines represent the $\alpha$ and $\beta$ obtained for the whole dataset.

We now focus on the dependence of the lightning activity rates $L$ over the longitude $\lambda$ of the cities. We have proceeded as previously, that is, we start with a scatter plot of $L$ vs. $\lambda$, we bin the data into 25 windows and for each one we evaluate the average of $L$, and we interpolate a spline of third order to the average data, as it is shown in fig. 3(a). For this case, we do not observe a clear dependence of the average lightning activity rates over the longitudes within the interval $-78 \lesssim \lambda \lesssim -43$. In this interval, apart from a peak around the latitude $-55$, $L$ displays an approximate constant behavior. However, we do note a systematic decrease in the average values of $L$ for longitudes higher than $\approx -43$. This fact is also visible in fig. 1, where the blue region points out the lower lightning rates in the northeast of Brazil. Analogously to the previous case, we have also studied the standard deviations $\sigma$ within windows as a function of the longitude (fig. 3(b)). Similarly, we have found an almost constant extreme event of the several processes that occur in a thundercloud (in addition to large values of the electric fields necessary for the lightning initiation). Actually, cloud-to-ground lightnings are about 25% of the global lightning activity, all remaining activity does not involve the ground and it is mainly composed by intra-cloud discharges and by (on a much smaller fraction) cloud-to-cloud and cloud-to-air discharges. Therefore, when analyzing only cloud-to-ground lightnings, we are selecting a group of extreme events and thus approaching somehow the mathematical conditions of the extreme value theory where Gumbel distributions naturally emerge.
behavior and that most of the values of $\sigma$ are inside the 99% confidence intervals of the average of the standard deviation over the longitude (though systematic deviations are observed for longitudes larger than $-43$).

The most striking difference between the relationships of $L$ vs. latitude and $L$ vs. longitude occurs for the distribution of the residuals $\xi$ around the average tendency. While the Gumbel distribution is a quite good description of the dependence of $L$ on the latitude, it cannot describe the dependence of $L$ on the longitude. In fact, the tent-shaped distribution shown in fig. 3(c) is in good agreement with the Laplace distribution

$$P(\xi) = \frac{1}{2\gamma} e^{-|\xi-\mu|/\gamma},$$

where the parameters $\mu = -0.12$ (mean) and $\gamma = 1.40$ (scale factor) were obtained via maximum-likelihood fit and the p-value of Cramér-von Mises test is 0.02, pointing out that the Laplace hypothesis cannot be rejected at a confidence level of 99%. As in the latitude case, we also calculated the distribution of $\xi$ after grouping cities with close longitudes values. Intriguingly, the Laplace form also describes the residuals distributions by longitude with parameters $\mu = -0.12$ and scale parameter $\gamma = 1.40$ which were obtained via the maximum-likelihood fit. The p-value of Cramér-von Mises test is 0.02, pointing out that the Laplace hypothesis cannot be rejected at a confidence level of 99%. As in the latitude case, we also calculated the distribution of $\xi$ after grouping cities with close longitudes values. Intriguingly, the Laplace form also describes the residuals distributions by longitude with parameters $\mu$ and $\gamma$ very close to those obtained for the whole dataset, as shown in fig. 3(d). In fact, since the spatial lightning distribution is closely linked to climate phenomena on Earth (such as the general circulation of the atmosphere between the equator and the middle latitude [26]) and therefore it seems unrelated to longitudes,
where small distances ($r \lesssim 40$ km) the correlation remains almost constant. We next observe an approximate logarithmic decay, where $C(r) \approx -C \ln (r/R)$ describes the correlation function for $60 \text{ km} \lesssim r \lesssim 600 \text{ km}$ (shaded area) with $C = 0.19$ and $R = 8311.32$. For $r \gtrsim 600 \text{ km}$, we note a faster-than-logarithmic decay.

Another intriguing question is whether there is long-range memory in the lightning activity rates $L$. To investigate this hypothesis, we evaluate the spatial correlation function of the lighting activity rates between pairs of cities that are $r$ kilometers distant. Specifically, we have computed

$$C(r) = \frac{\langle (L_i - \bar{m}(r)) (L_j - \bar{m}(r)) \rangle_{r_{ij}=r}}{s(r)^2},$$  

where $\bar{m}(r)$ is the mean value and $s(r)$ is the standard deviation of the lightning activity rates of cities separated by $r$ kilometers, $L_i$ is the lightning activity rate in the city $i$, and $\langle \ldots \rangle_{r_{ij}=r}$ stands for the average value over cities where the distance $r_{ij}$ that separates them are equal to $r$. Due to the discrete nature of our spatial data, we have actually considered logarithmically spaced intervals of $r$ for evaluating eq. (4). Figure 4 shows the spatial correlation $C(r)$ in a lin-log scale where it is remarkable a very slow decay of the correlation function.

In particular, we have identified three regimes: i) an initial plateau in the interval $10 \text{ km} \lesssim r \lesssim 40 \text{ km}$ where $C(r) \approx 1$; ii) an approximate logarithmic decay in the interval $60 \text{ km} \lesssim r \lesssim 600 \text{ km}$ (shaded area) where $C(r) \approx -C \ln (r/R)$; and iii) a faster-than-logarithmic decay for distances larger than $\approx 600 \text{ km}$. We thus conclude that the lightning activity rates are long-range correlated in space, since the correlation function $C(r)$ decays slower than any power law function for distances smaller than $\approx 600 \text{ km}$. Another interesting point is concerning the relationships between these long-range correlations and the non-Gaussian residual distributions previously presented. In several contexts, long-range correlations emerge in coupled-like manner with non-Gaussian distributions (see ref. [28] for a specific example) and here the same happens. Furthermore, if we shuffle the lightning rates among the cities, besides destroying the spatial correlations this process drastically changes the profile of the residual distributions. Specifically, both residual distributions for latitude and longitude become equal and they pass simply to reflect the spatial distribution of the Brazilian cities.

The approximate logarithmic decay is very intriguing and it can be modeled on the theoretical canvas of the Edwards-Wilkinson equation [29]

$$\frac{\partial}{\partial t} L(r,t) = D \nabla^2 L(r,t) + \eta(r,t),$$  

where $L(r,t)$ represents the lighting rate in a point localized by the vector $r$, $D$ is a constant, $\nabla^2$ is the Laplacian, and $\eta(r,t)$ is an uncorrelated noise (in space and time) with zero mean and finite variance. Because this equation is mainly used to describe stochastic kinetics of a growing interface, it may appear as an ad hoc hypothesis for modeling our data. However, the Edwards-Wilkinson equation contains two important phenomenological ingredients related to the spatial distribution of the lighting activity: diffusion and randomness. The former (Laplacian term) can be related to the diffusive aspects of the thunderstorms and the second (noise term) reflects the natural complexity of the lightning phenomenon (as well as other sources of randomness). Nonetheless, modeling the spatial distribution of lighting rates only with this equation represents a quite crude approximation where several important processes are not considered. For instance, we can not expect this simple model to describe the residual distributions. On the other hand, the autocorrelation function related to eq. (5) can be obtained in $d$ dimensions by taking Fourier transforms in space and time, evaluating the correlation in this double-transformed Fourier space, and next returning to the usual space via inverse Fourier transforms (see refs. [29–31] for more details). For equal-times and in the equilibrium ($t \gg R^2/D$, see below), the correlation is written as $C(r) \sim r^{2-d}$ when $d \neq 2$, and for the two-dimensional case (that is, for our case) it is

$$C(r) \approx -C \ln \left( \frac{r}{R} \right) \quad \text{for} \quad r_{\text{min}} < r < R,$$  

where $C > 0$ is constant, $r_{\text{min}}$ is a short-distance cut-off and $R$ is the length of the system. We thus observe that the logarithmic decay is intimately connected with the two-dimensional nature of the problem. Furthermore, it is equally intriguing that similar logarithmic decays were reported in very different contexts related to turnoff rates in elections [32,33] and spatial spreading of obesity and other diseases [34], which can be somehow related to remnant effects of a universal mechanism underlying these two-dimensional processes.
Summary. — We characterized the lightning activity rates over all the Brazilian cities by taking their geographic location into account. We firstly studied the dependence of the lightning activity rates on the latitude of the cities, where we observed two peaks in the average value of these rates: one just before the Tropic of Capricorn and another one just before the Equator. We also investigated the fluctuations surrounding the average tendency, where we found that the standard deviation is almost a constant function of the latitude and that the distribution of these fluctuations (residuals) is in quite good agreement with a Gumbel distribution. The dependence of the lightning activity rates on the longitude of the cities was also characterized, where we reported that the average value exhibits an approximate plateau for a wide range of longitude values. For the longitude case, the standard deviation of the residuals also presents an almost constant behavior; however, the distribution of the residuals was well described by a Laplace distribution. We further investigated the spatial correlations of the lightning activity rates between a pair of cities. Our results show that the correlation function decays very slowly with the distance between the cities and that for intermediate distances (60 km ≲ r ≲ 600 km) the correlation exhibits an approximate logarithmic decay. Finally, we proposed to model the logarithmic decay via the Edwards-Wilkinson equation, where we observed that this behavior is closely related to the two-dimensional nature of the system and also that the same behavior was reported in very different contexts, which somehow points to a universal mechanism.

***

We thank Capes, CNPq and Fundação Araucária for financial support. HVR is especially grateful to Capes/Fundação Araucária for financial support under grant number 113/2013.

REFERENCES

[1] Mendes R. S., Malacarne L. C., Santos R. P. B., Ribeiro H. V. and Picoli S., EPL, 92 (2010) 29001.
[2] Kawamura H., Hatano T., Kato N., Biswas S. and Chakrabarti B. K., Rev. Mod. Phys., 84 (2012) 839.
[3] Daviden J. and Kwiatek G., Phys. Rev. Lett., 110 (2013) 068501.
[4] Lippiello E., Godano C. and de Arcangelis L., EPL, 102 (2013) 59002.
[5] Zheng H., Fu S. Y., Zong Q. G., Pu Z. Y., Wang Y. F. and Parks G. K., Phys. Rev. Lett., 109 (2012) 205001.
[6] Reeves G. D. et al., Science, 341 (2013) 991.
[7] Turner D. L., Shprits Y., Hartinger M. and Angelopoulos V., Nat. Phys., 8 (2012) 208.
[8] Rybski D., Bunde A., Havlin S. and von Storch H., Geophys. Res. Lett., 33 (2006) L06718.
[9] Boettle M., Rybski D. and Kropp J. P., Water Resour. Res., 49 (2013) 1199.
[10] Rybski D., Holsten A. and Kropp J. P., Water Resour. Res., 390 (2011) 680.
[11] Prahl B. F., Rybski D., Kropp J. P., Burghoff O. and Held H., Geophys. Res. Lett., 39 (2013) L06806.
[12] Telese L., Bernardi M. and Rovelli C., Commun. Nonlinear Sci., 13 (2008) 1384.
[13] Elsom D. M., Atmos. Res., 56 (2001) 325.
[14] Cooray V., Cooray C. and Andrews C. J., J. Electrostat., 65 (2007) 386.
[15] Enno S. E., Briede A. and Valiukas D., Theor. Appl. Climatol., 111 (2013) 309.
[16] Christian H. J. et al., J. Geophys. Res.: Atmos., 108 (2003) ACL 4-1.
[17] Williams E. R., Chan T. and Bocchicchio D., J. Geophys. Res.: Atmos., 109 (2004) D16206.
[18] Brazilian National Institute For Space Research (INPE) and Research Group on Atmospheric Electricity (ELAT), http://www.inpe.br/webelat/homepage/ (accessed on June 2013).
[19] Ballarotti M. G., Medeiros C., Save M. M. F., Schulz W. and Pinto O. jr., J. Geophys. Res., 117 (2012) D06112.
[20] Bourcheidt V., Cummins K. L., Pinto O. jr. and Naccarato K. P., J. Atmos. Ocean. Technol., 29 (2012) 1304.
[21] Pinto O. jr., Pinto I. R. C. A. and Ferro M. A. S., J. Geophys. Res., 118 (2013) 5231.
[22] Lightning Imaging Sensor (LIS), http://pmm.nasa.gov/node/163 (accessed on June 2013).
[23] Tropical Rainfall Measuring Mission (TRMM), http://trmm.gsfc.nasa.gov (accessed on June 2013).
[24] Gumbel E. J., Statistics of Extremes (Columbia University Press, New York) 1958.
[25] Rakov V. A., Surv. Geophys., 34 (2013) 701.
[26] Price C., Global thunderstorm activity, in Sprites, Elves and Intense Lightning Discharges, edited by Fullekrug M. et al. (Springer, Amsterdam) 2006.
[27] Price C., Surv. Geophys., 34 (2013) 755.
[28] Mendes R. S., Ribeiro H. V., Freire F. C. M., Tateishi A. A. and Lenzi E. K., Phys. Rev. E, 83 (2011) 017101.
[29] Edwards S. F. and Wilkinson D. R., Proc. R. Soc. London, Ser. A, 381 (1982) 17.
[30] Nattermann T. and Tang L.-H., Phys. Rev. A, 45 (1991) 7156.
[31] Barabási A.-L. and Stanley H. E., Fractal Concepts in Surface Growth (Cambridge University Press, Cambridge) 1995.
[32] Borgesi C. and Bouchaud J.-P., Eur. Phys. J. B, 75 (2010) 395.
[33] Borgesi C., Raynal J.-C. and Bouchaud J.-P., PLoS ONE, 7 (2012) e36289.
[34] Gallos L. K., Barttfeld P., Havlin S., Sigman M. and Makse H. A., Sci. Rep., 2 (2012) 454.