# Spatio-temporal clustering with a temporal sliding window for tracking of convective activity

**Cesar Strauss[1], Stephan Stephany[2]**

[1]Programa de Mestrado ou Doutorado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

[2]Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

cstrauss@cea.inpe.br ,stephan@lac.inpe.br

***Abstract.*** *Meteorological radars allow to detect and track convective activity, but its cover and number are limited. An alternative would be to track convective activity by means of atmospheric electric discharges using data from the RINDAT network. Aiming for this objective, it is needed to study the eventual spatial and temporal delay between precipitation and discharges, proposed in this work by the spatio-temporal clustering of discharges and also by precipitation nuclei with sliding-window.*

**Keywords:** *spatio-temporal clustering, electric atmospheric discharges, meteorological radar, sliding window*

## 1. Introduction

Convective nuclei are cloud formations whose birth, evolution and dissipation are of great interest among meteorologists since they are associated to severe storms. Lightning data can be a way to indirectly track these nuclei. Lima et al. [2006] found a strong correlation between the probability of occurrence of lightnings and the WV-IR index (water vapor minus infrared index) based on GOES satellite images that show cloud tops. Caetano et al. [2009] found a similar correlation comparing the same type of GOES images to the field of density of occurrences of lightning calculated by a kernel estimation technique. The more intense regions of this field correspond to the so-called “nuclei of electrical activity”. Further work [Caetano et al., 2010] suggested a further correlation with precipitation structures in weather radar images. However, the tracking of the evolution of convective activity using lightning data needs to be better analyzed. We propose a new technique for spatio-temporal clustering applied to lightning data as an alternative or a complement to the use of the field of density of occurrences aiming at a deeper investigation of the correlation between convective nuclei and lightnings.

## 2. Proposed spatio-temporal clustering method

The proposed method is based on a temporal sliding window, similarly as the sliding window employed for data flow control in networks. A fixed-width temporal window is slided in time with a constant rate, but for discrete timesteps, in order to process new incoming lightning data. The clusters obtained by this method are also called "nuclei of electrical activity". A major difficulty that we have in the clustering process is how to identify and track a particular nucleous across space and time, while maintaining its identity. After clustering the 2D lightning data defined by a given timestep of the window, the current clusters are a mixture of new clusters and ones that already were defined in a previous timestep of the window.

For instance, in Figure 1, two clusters are found in step T, and are labeled as 1 and 2. In step T+ΔT, four clusters are found. The leftmost cluster has two elements in common with cluster 1, so it receives the same label "1". The second has no elements in common with any cluster in timestep T, so it receives the next available label "3". The two rightmost clusters have each an element in common with cluster 2, so they receive the same label "2".

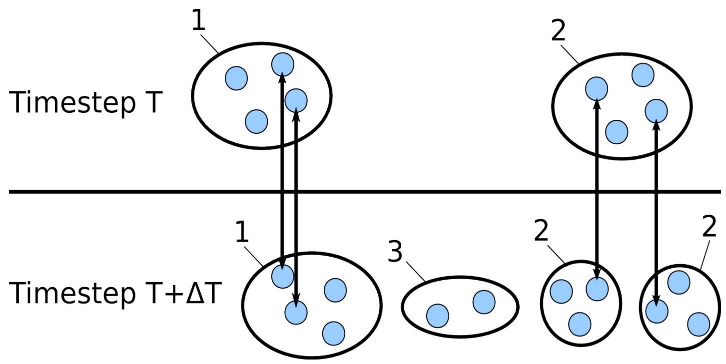


**Figure 1 – Proposed spatio-temporal clustering method, based on temporal a sliding-window.**

## 3. Results

We employed lightning data acquired from the RINDAT network of lightning sensors related to the São Paulo State in the month of January, 2010. In order to select a set of severe convective events, we plotted the total number of lightnings screened by an one-hour sliding window (without clustering) as a function of time, and selected the events given by the highest counts of lightnings. A second selection was related to the choice of events close to the weather radars located at the cities of Bragança Paulista and Presidente Prudente. Finally, a single event was taken as a case study, which occurred on 01/19/2010 in the 10:00 PM to midnight interval. Several nuclei of electrical activity were then found associated to this event by the spatio-temporal clustering.

In order to compare these nuclei to the nuclei of precipitation we employed the 3.1 km CAPPI images from the radars mentioned above. These CAPPI data was interpolated on a 0.02 degree regular grid and a 10 minutes temporal resolution. The threshold employed for segmentation was 30 dBZ, which provided a good trade-off between size and number of the nuclei of precipitation. In the sliding window spatio-temporal clustering of lightnings, we selected a window width of 10 minutes and a timestep of 5 minutes, in order to have compatibility to radar data.

The strongest 4 precipitation nuclei were selected and then we found 4 nuclei of electrical activity that would be corresponding since they were in similar locations at similar times, as demonstrated by the trajectories of the centroids of both nuclei. These trajectories were plotted as function of time, as shown in Figure 3. The temporal plot of the number of lightnings (total, negative, positive, cloud-to-ground and cloud-to-cloud) appears in Figure 4 for one of the precipitation nuclei, showing a delay between the peaks of electrical activity and precipitation.

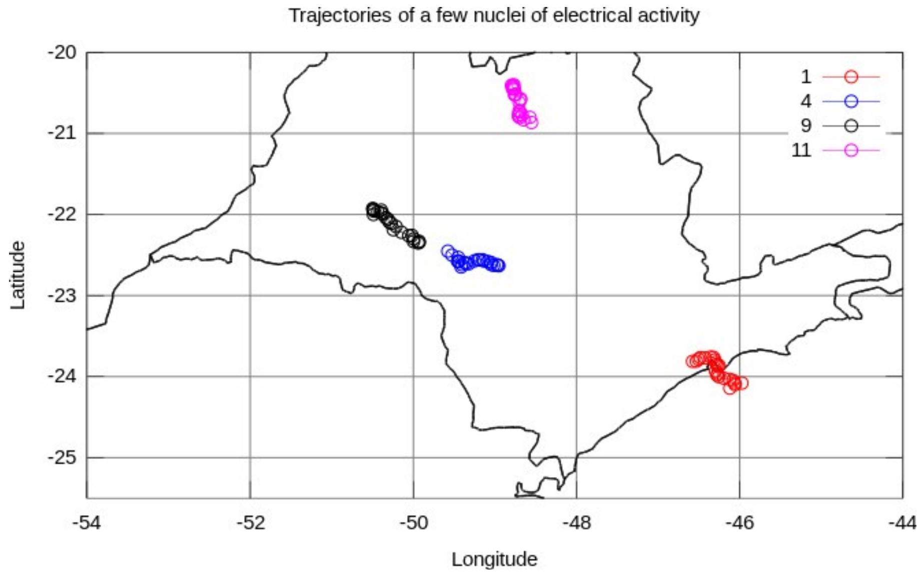


**Figure 2 – Trajectories of the four strongest nuclei of electrical activity, which correspond to the four strongest nuclei of precipitation.**

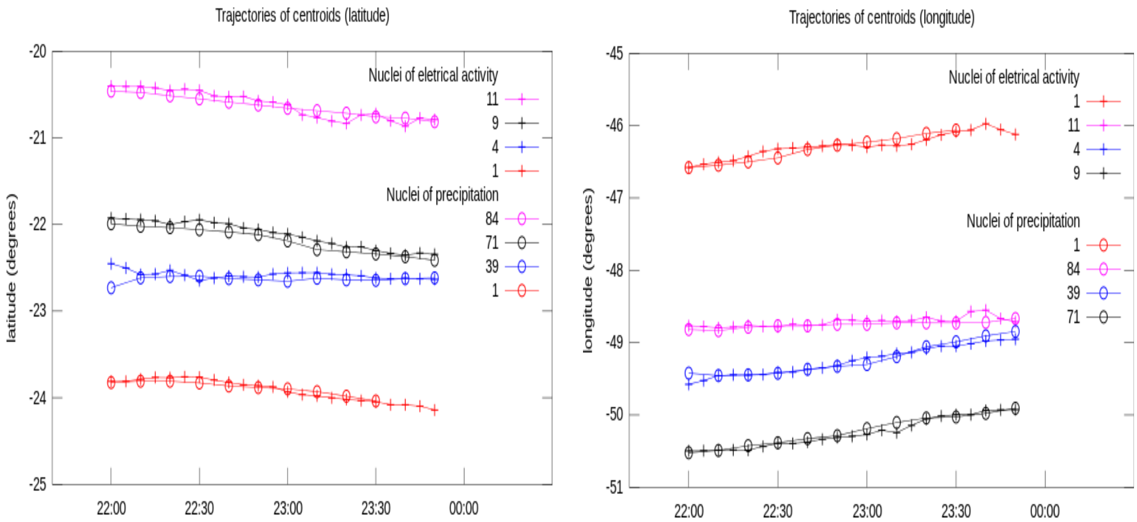


**Figure 3 – Trajectories of the centroids for 4 precipitation nuclei, identified by the numbers 1, 39, 71 and 84, and the corresponding nuclei of electrical activity, identified by the numbers 1, 4, 9, 11 (left, in latitude and, right, in longitude).**

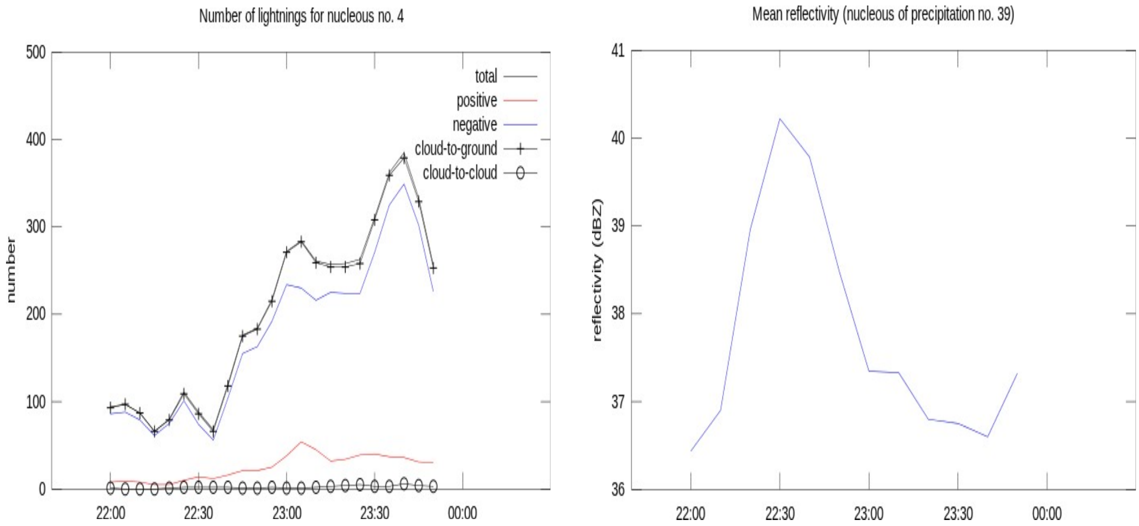


**Figure 4 – Number of lightnings (total, negative, positive, cloud-to-ground and cloud-to-cloud) as function of time for a selected nucleous of electrical activity (identified as number 4, eft) and mean reflectivity of the corresponding precipitation nucleous (identified as number 39, right).**

## 4. Conclusions

A suitable spatio-temporal clustering of lightning data was performed employing an innovative sliding-window scheme. It was possible to show correlation between the trajectories of the centroids of the nuclei of electrical activity resulting from the clustering and the centroids of the precipitation nuclei derived from the CAPPI radar images. However, further studies are required to better investigate the temporal correlation between the intensity of the electrical activity and intensity of the precipitation.

## References

Caetano, M.; Escobar, G.C.J.; Stephany, S.; Menconi, V.E.; Ferreira, N.J.; Domingues, M.O.; Mender Junior, O. Visualização de campo de densidade de ocorrências de descargas elétricas atmosféricas como ferramenta auxiliar no nowcasting. In: XIII LATIN AMERICAN AND IBERIAN CONGRESS ON METEOROLOGY (CLIMET XIII) AND X ARGENTINE CONGRESS ON METEOROLOGY (CONGREMET X). **Proceedings**... Buenos Aires: 2009.

Caetano M.; D'Oliveira, A.B.; Strauss, C.; Stephany, S.;Ferreira, N.J.; Análise comparativa de imagens de radar meteorológico e de campos de densidade de descargas elétricas atmosféricas para eventos convectivos severos. In: XVI Congresso Brasileiro de Meteorologia, **Proceedings**, Belém, 2010.

Lima, W.; Machado, L.; Morales, C.; Pinto Jr., O. Estimativa de probabilidade de ocorrência de relâmpagos. In: XIV Congresso Brasileiro de Meteorologia, **Proceedings...**, 2006.