# Tracking and prediction of convective cells based on lightning data

Stéphane Pédeboy, Météorage
sp@meteorage.com
--------------------------------
co-authors: Paul Barnéoud
--------------------------------
speaker: Stéphane Pédeboy

Météorage, the French National Lightning Locating System (LLS) operator, has developed the "**S**evere **T**hunderstorm **O**bservation and **R**eporting **M**ethod" (STORM) aiming at detecting active thunderstorms and preventing severe weather based on VLF/LF total lightning data made of Cloud-To-Ground (CG) flashes and Cloud-To-Cloud discharges (CC).

In this study, the capability of STORM to first, identify and track lightning cells and second, to predict severe weather is estimated based on hail ground truth data collected in 2014 across France by the ANELFA [Dessens et al 2006].

## The STORM algorithm

The STORM algorithm relies on two main functions which are the identification of the lightning core cells and the monitoring of lightning jumps in near real time.

### A. Lightning cells identification

Lightning cells identification based on lightning data consists of grouping lightning flashes detected by the LLS in an area representing the electrical active core of a thunderstorm. Because, most of lightning flashes occur close to the updraft region in the thunderstorm it is possible to group them in consistent groups representing the convective core. This can be achieved thanks to data clustering methods.

Out of all the existing clustering methods (DENCLUE, CLIQUE, MAFIA, BIRCH, CURE, GBC, Chamelon…) Météorage has chosen the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (see figure 1) because it is simple to use and is also robust to outliers. It relies on the "Nearby Neighbors Search" technique to group points together according to their separation distance and a given local density of points [Ester et al. 1996].

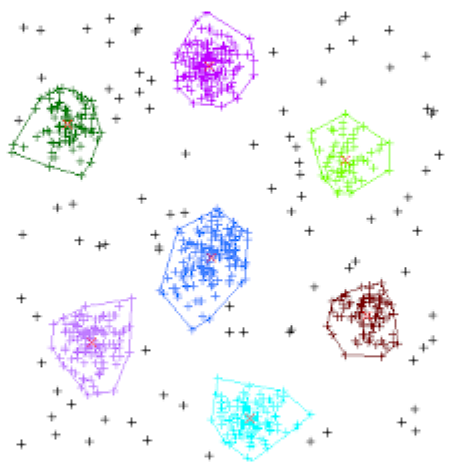


Fig 1. Example of seven clusters determined with the DBSCAN algorithm (in grey are outliers).

A DBSCAN analysis is periodically run on the most recent minutes (5 minutes in this study, but might be more frequent) of lightning data to identify active cores. Results from the current run are then correlated to those obtained from prior runs to either create a new lightning cell or update the status of already existing ones. Then, every individual living cell is tracked and its characteristics (eg. position of the barycentre, direction of propagation, speed, area and number of flashes per minute) are monitored and stored in a dedicated database all long the lifecycle of the cell.

### *B. Cells severity assessment*

The potential severity of every living lightning cell is monitored during every run based on the evolution of their individual lightning rate (see figure 2). The goal is to detect the onset of the lightning jump as soon as it occurs to issue warning messages with the bigger lead time possible. Here again, several algorithms have been developed by different researchers to monitor lightning rate trends and detect the onset of the lightning jump exclusively, according to our knowledge, based on VFH lightning data [Goodman 2005; Gatlin 2006]. Out of these algorithms, the “2σ configuration” has been statically validated on various thunderstorm types and is likely to be the most effective to use for operational early warning usage [Schultz et al. 2009].

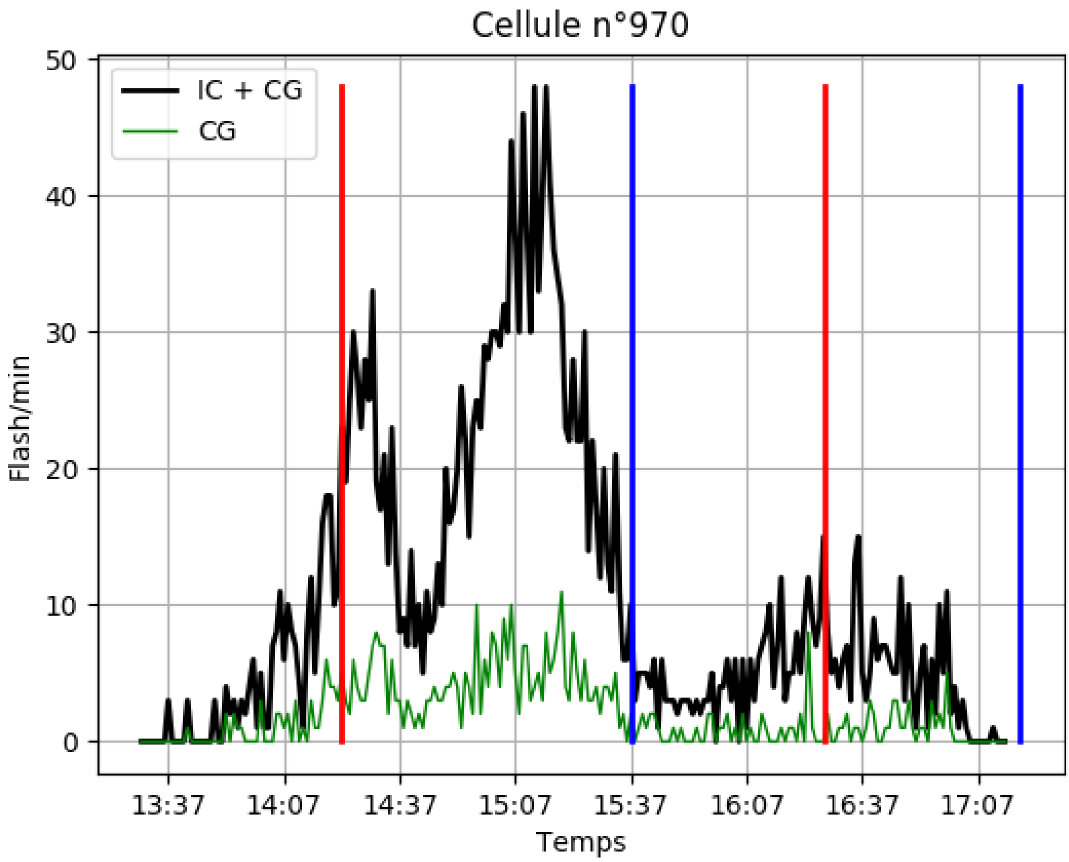



Figure 2- Lightning rate as the number of flashes (IC + CG) per minute. The black curve shows the total lightning dataset, the green only the CG flashes. The red and blue lines indicate the beginning and the end of a severe event alert.

## Data used for the validation of STORM

### A. The lightning dataset

The French National LLS is made of 20 LS7002 sensors by Vaisala dispatched across France to which are added about 60 foreign sensors belonging to neighbouring national LLS partners. Such a system detects the electromagnetic signals generated by the large and intense vertical current discharges occurring either in the cloud or between the cloud and the ground. When several sensors measure the same event, the central computer can locate it using direction finding or difference of time of arrival techniques.

A recent quality control based on high speed video camera records collected during 2015 has shown Météorage's LLS detection efficiency (DE) is 97% for flashes and 94% for strokes. These results are in perfect agreement with those obtained in South-East France and more generally in Europe after similar studies based on the EUropean Cooperation LIghtning Detection (EUCLID) network which uses Météorage's data [Schulz et al, 2014; 2015]. The location accuracy is estimated around 110m based on video analysis of flashes exhibiting multi-strokes ground strike points. The Cloud-to-Cloud detection efficiency (DECC) is estimated to be in a range of 30 to 50% [Pédeboy et al, 2014] depending on the type of thunderstorm (Isolated storm, multi-cellular or supercell).

### B. Hail ground truth dataset

The ANELFA is a French association dedicated to the hail risk prevention. It has deployed several tens of passive hail pads that are disseminated across fifteen departments in France (see figure 3) to track and report hailfalls. Begin date and time of hailfalls are manually determined by local volunteer from the association who send knocked hail pads to the ANELFA scientific research centre who determines from the marks on the sensitive plate the number, the maximum diameter, the cumulative mass and the kinematic energy of hailstones [Farnell et al, 2009].

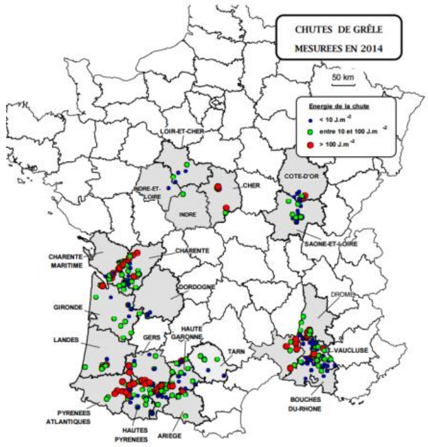
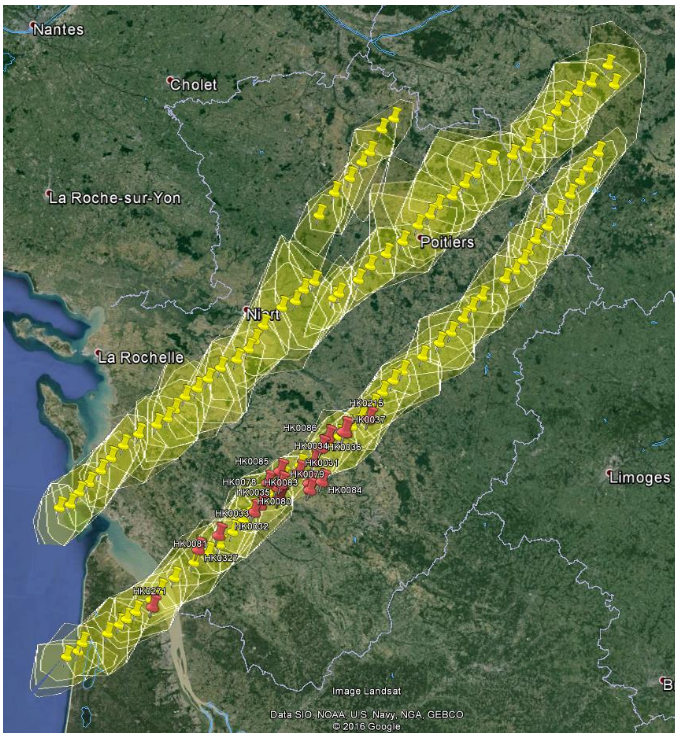


Fig 3. Distribution of hail reports in 2014

The total initial hail dataset consists of a total of 248 reports registered during 2014. However, as the goal of this study is to validate the lightning cell identification, tracking and prediction of severity, only the hail reports that are positively correlated with lightning data are selected. This allows also to filter out inaccurate hail observations.

## Results and comments

A total of 82 different lightning cells were correlated with 145-validated hail reports out of which 19 (22%) exhibit more than one single hail observation. Interesting to note the time correlation between cells and hail reports fit very well as 86% of the observation dataset is correlated within




-/+5 minutes and nearly 70% hail and cell dating match perfectly with no time difference. In addition, the median separation distance between cells barycentre and hail pad locations is about 10 km that is in the order of magnitude of a typical storm cell.

As an example, figure 3 shows an example of two lightning cells identified by STORM on the 9th of June 2014 at night in the South-West France. Both cells are moving in parallel at a similar speed and at distance of 40 km. The total path length corresponding to 3 hours of observation is about 300 km. observed during One can see the very nice match between cell data and hailfalls recorded by the ANELFA. Interesting to note that STORM can detect the moment when the northern cell splits from its mother cell between Niort and Poitiers and track different individual cells from that time.

Out of the 82 hail correlated cells, 19 exhibit no lightning jump so it is impossible to determine the efficiency of the cell severity prediction with these cases. This is consistent with observations that reported some thunderstorms exhibited no lightning jump despite severe events produced [Schultz et al. 2009].

Finally, this is an overall dataset of 63 severe cells associated with 120 hail reports that is available to assess performance of the severe weather prediction algorithm. Assuming, that all 145 hail reports are related to a severe cell that should have produced a lightning jump, STORM reaches a Probability of Detection (POD) is 82%. Interesting to note, the same parameter increases to 89% and 100% for cells exhibiting hailstones respectively larger than 20 mm and 25 mm in diameter. Note this computation is done on a reduced dataset that do not consider cases where STORM failed to identify a lightning cell with hail reports. If these cases are considered, and again the assumption that all 248 hail reports should have led to the identification of a cell, then the POD drops down to 48%, then 60% and 80% respectively for cells producing hailstones larger than 20 mm and 25mm.

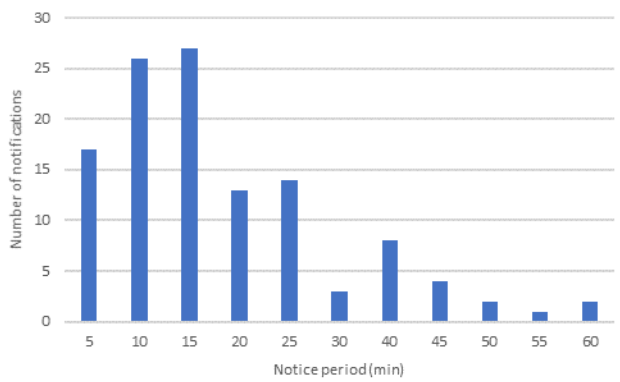



Fig 5. Distribution of the notice period

The warning lead time is defined as the period between the time an alarm is triggered by the lightning jump algorithm and the time of hail observation. Figure 4 shows a quite large distribution of warning lead times ranging from 0 to 60 min. This results from the wide varieties of thunderstorm types being considered in this study occurring in different seasons and terrain conditions. Nevertheless, the mean lead time computed is about 15 min increasing to 18 min when only cells producing hailstones of a diameter greater or equal to 25 mm are considered.

Assuming a 10-min delay is enough to deliver an efficient severe weather warning for most of operational applications, then STORM is successful to release relevant warning messages in 63% of all cases.

## Discussion

This study has demonstrated the capability of STORM to reproduce realistic active core cells as they produce in nature and predict the production of severe weather, in this case hailfalls. Surprisingly, the results in terms of POD and warning lead time are consistent with those that can be found in literature [Schultz et al, 2014] and that are based on a high-resolution Lightning Mapping Array lightning data. Indeed, these authors who have extensively studied several lightning jump algorithms claim the POD for a "2σ configuration" is 89% and the mean lead time is 20 minutes to be respectively compared to our 80% and 15 minutes. Thus, the first result obtained in this study demonstrate the relevancy of using VLF/LF lightning in lightning cell identification and lightning jump detection.

The capability of STORM to detect severe weather is not only depending on the lightning jump algorithm settings but relies also on the lightning cell identification. Of course, thunderstorms producing few lightning are not likely to be detected that is somehow not a limitation of STORM but clearly a physical characteristic of thunderstorms. However, putting apart these special cases, the tuning of the clustering algorithm is crucial, and a special attention shall be applied to the settings in order it do not to split or merge a real storm preventing either the detection of a lightning jump or issue false alarms. It must be noted, the False Alarm Rate (FAR) parameter has not been assessed in this study but it remains an objective in a future work.

The good performances of STORM are interesting since this opens doors for new applications needing to monitor in real time the lightning activity over large areas. Out of them, applications related to air traffic control, ground operation and on-board information for pilots. The usage of reconstructed lightning cell data based on individual lightning flash location provides a synthetic way to represent thunderstorms, giving a clear view to end users even during complex episodes. As a result, the decision-making process can be very fast and efficient because pilots or air traffic controllers relies on a simple representation of the contours of active cells associated with the past and the future trajectories. They immediately know where the convective systems are and they can check in advance their route is not in conflict with the progression of a thunderstorm, and eventually modify their short-term flight plan. On another hand, the information regarding the potential severity of a cell can be interesting for take-off, landing and ground operations since lightning jump signature are not only related to hailfall but to all dangerous event like strong winds and heavy precipitations that may disturb the approach phase of an aircraft.

## References

Dessens, J., Berthet, C., and Sanchez, J. L., 2007: ´ A point hailfall classification based on hailpad measurements: The ANELFA scale, Atmos Res, 83, 132–139.

Ester M., H-P Kriegel, J. Sander, X. Xu, 1996, Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)

Farnell, C., Aran, M., Andres, A., Busto, M., Pineda, N., and Tor ´ a,` M., 2009: Study of the September 17th 2007 severe hailstorm in Pla d'Urgell. Part I: fieldwork and analysis of the hailpads, Tethys, 6, 69–81, doi: 10.3369/tethys.2009.6.05.

Gatlin, P., 2006: Severe Weather Precursors in the Lightning Activity of Tennessee Valley Thunderstorms. M.S. Thesis, University of Alabama-Huntsville, 87 pp.

Goodman, S. J. and Coauthors, 2005: The North Alabama Lightning Mapping Array: Recent severe storm observations and future prospects. Atmos. Res., 76, 423–437.

Pédeboy, S., Defer, E., and Schulz, W.: Performance of the EUCLID network in cloud lightning detection in the South-East France, in: 8th HyMeX Workshop, Valletta, Malta, 2014.

Rison, W., R.J. Thomas, P.R. Krehbiel, T. Hamlin, and J. Harlin, A GPS-based Three-Dimensional Lightning Mapping System: Initial Observations in Central New Mexico Geophysical Research Letters, 26, 3573-3576, 1999

Schulz, W., Pedeboy, S., Vergeiner, C., Defer, E., and Rison, W.: Validation of the EUCLID LLS during HyMeX SOP1, in: International Lightning Detection Conference and International 20 Lightning Meteorology Conference (ILDC/ILMC), Tucson, Arizona, USA, 2–5, 2014.

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European lightning location system EUCLID – Part 1: Performance validation, Nat. Hazards Earth Syst. Sci. Discuss., 3, 5325-5355, doi:10.5194/nhessd-3-5325-2015, 2015.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Pre-liminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather.J.Appl. Meteor. Climatol.,48,2543–2563, doi:10.1175/2009JAMC2237.1.

Schultz,E. V, Themis Chronis1 , Kiel L. Ortega, Lawrence D. Carey1, Kristin M. Calhoun, Greg Stumpf, Christopher J. Schultz1, Geoffrey T. Stano, Steven J. Goodman, Enhanced Verification of the Lightning Jump Algorithm, XV International Conference on Atmospheric Electricity, 15-20 June 2014, Norman, Oklahoma, U.S.A.