21, rue d'Artois, F-75008 PARIS
https://www.cigre.org

International Colloquium on
Lightning and Power Systems

Delft 2019

Review of the lightning dataset and Lightning Locating Systems performances as recommended by the IEC 62858 standard

Stéphane Pédeboy
Météorage
France

SUMMARY

The lightning data collected by Lightning Locating Systems operating in the Low Frequency range is of great interest in many end user applications. Because this data is derived from remote sensing measurements it suffers from inaccuracies depending on system technology and the effectiveness of the operations. To guarantee the reliability and homogeneity of the various lightning datasets made available on the market, the IEC 62858 standard requires a minimum level of performance for an LLS to be qualified and deliver reliable data. This document reviews the main performance parameters, the minimum requirements that must be achieved It is defined. by an LLS and introduce the lightning data structure and its relations with the natural lightning related electrical processes. The objective of this paper is to give end users that are not knowledgeable on LLS some information about the limitations of LLS and the different data that are made available.

KEYWORDS

Lightning Locating Systems (LLS); Lightning Data; IEC 62858 standard

sp@meteorage.com.

I. INTRODUCTION

Lightning data collected by Lightning Locating Systems (LLS) operating in the Low Frequency (LF) range has become an important dataset for many applications in many fields. Indeed, such systems can locate and estimate in real time different current parameters of the return strokes occurring in Cloud-to-Ground (CG) lightning flashes over wide areas, as large as a country up to a continent. Usually the real time data is stored in a database making available to end users long term series of data for historical analysis.

The raw dataset, so called “stroke data”, is derived from a calculation based on return stroke related signals measurements gathered by remote sensing sensors. Thus, the dataset suffers inevitably from inaccuracies depending on the technology of the sensors, the sensors network design, the data processing algorithm and the operations [1] that might affect the reliability of end user applications. Therefore, the performance of the LLS and the quality of the lightning data are key parameters that must be considered prior and during the usage of the data.

Out of the “stroke data”, some LLS can identify CG flashes using a flash grouping algorithm running on the lightning processor. These algorithms are usually based on spatial-temporal criteria to identify the strokes that are related to the same flash regarding the inter-stroke delay and distance leading to the so-called “flash data” [1][2]. In addition, for high resolution stroke data, it is possible to determine the Ground strike Points (GSP) resulting from individual downward channels in a flash, so called ‘Ground Strike Point data’. As a result, the lightning dataset that can be provided by LLS in LF range is determined by a pyramidal data structure, each level corresponding to a given physical process in a lightning flash. Like for the LLS performances, the end user must be aware of this data structure and therefore the physical process he is dealing before defining its application or service.

Because of the increasing number of LLS in the world, the variety of technologies used in such systems and the different way they are operated it exists on the market various lightning datasets, more of less documented and all being of different quality. Then, it is difficult for an end user that is not knowledgeable about LLS to clearly understand the limits and the applicability of the corresponding lightning data. Consequently, it became necessary to release an international standard giving a common framework to LLS operators for determining the minimum level of performance an LLS must achieve in order its data can be used reliably [3].

This paper intends to introduce the IEC 62858 Edition 2 [4] standard focussing on the first part explaining the different important parameters that must be determined to assess the LLS performances. Then, a detailed overview of the complete lightning data structure is given with the objective to explain the different levels of abstraction characterizing the lightning phenomenon as observed by an LLS. Finally, the reader is made sensitive to a new dataset that can identify and locate individual Ground Strike Point (GSP) that may occur in branched downward flashes. Indeed, this datum might better access the real lightning threat in delivering a more accurate information resulting in an improvement of the lightning protection systems.

II. RECOMMENDED LLS PERFORMANCES BY THE IEC 62858 STANDARD

The main goal of the IEC 62858 standard is to deliver a common framework to the LLS operators for the lightning density (Ng) computation. However, the reliability and the homogeneity of the N_g values are not only depending on the computational method as it is also extremely sensitive to the performance of the LLS. Of course, a system only detecting a fraction of the CG lightning flashes will not be able to deliver a reliable N_g although the method used for the computation is correct. As a result, the first part of the IEC 62858 defines specifications and recommendations for LLS operators to assess the real performances of their system.

Note, the Ng is not the only one application of the lightning data that is sensitive to LLS performances. Indeed, the reliability of all services based on lightning data, like incident or fault correlation, warning systems or statistics depend also very much on overall performances of LLS.

There is no doubt that lightning data collected by high resolution LLS operating in LF are of great interest in many applications and more particularly power line fault correlation. Nowadays, the most modern and up-to-date systems can detect 90% and more of the return strokes with a median location error in a range of 100 meters [5][6][7]. The LLS performance depends on several parameters like the sensor technology, the design of the network of sensors, namely the sensor baseline distances, and the location processing algorithm. However, the quality of the system operation is one of the most important factors as it can dramatically drop down the overall quality of the lightning dataset. Thus, system and sensors failures, power and telecommunication breakdowns, sensors and location algorithm settings or the limited specifications of the hardware architecture (i.e. CPU, RAM, LAN…) will affect the detection efficiency.

Because it exists various lightning dataset worldwide, all exiting a different level of quality it is a complex task for an engineer to simply choose the adequate dataset for his application. This become even more misleading when several datasets of different quality are compared. To prevent such problems, the IEC 62858 defines LLS performance parameters and requires a minimum level of quality on those parameters. Generally, the performance of an LLS is defined by the combination of at least three parameters that are the:

Detection Efficiency (DE), defined as the percentage of number of lightning event (stroke or flash) that are observed by the LLS in respect to the actual number of real events. Determining the detection efficiency is quite challenging since it is really difficult not to say impossible to account for the exact number of events occurring in nature. The minimum flash DE requiered by the standard is 80%.

Location Accuracy (LA), defined in meters as the distance between the located return stroke and its real position on the ground. Again, this parameter is not easy to assess as the ground truth data used as reference must be accurately timestamped for an efficient stroke time correlation. The standard recommends a maximum error of 500 meters on the flash location.

Classification Accuracy (CA), is a new parameter that appeared with the recent capability of LLS to detect Intra-Cloud (IC) discharges. It is defined as the percentage of IC or CG events classified by the LLS against the real class of the event. The issue with a bad IC/CG classification is that the number of CG flashes considered in the N_g calculation can either be overestimated because IC flashes are wrongly considered or on the opposite underestimated when the LLS classifiy CG in IC flashes. The minimum CA requiered by the standard is 85%.

The minimum requirements for detection efficiency, location and classification accuracies are respectively 80%, 500 meters and 85%. These quality indicators are in the Section 4 of the IEC 62858 document that also gives an overview of the different techniques used in LLS performance validation.

III. THE LLS LIGHTNING DATA STRUCTURE

The "stroke data" can be considered as a raw data since it is the less structured information delivered by the LLS directly derived from the measurements of the electromagnetic signals generated by the return strokes. A stroke observed by an LLS is then fully consistent with the natural phenomenon of the return stroke. Based on the stroke data a flash grouping algorithm may be used to aggregate individual return stroke in a mother entity so called the flash [8]. Thus, this algorithm creates a more synthetic dataset that is called the "flash data" aiming at describing the lightning flash phenomenon that can be observed by any human observer. The human eye cannot distinguish the individual return strokes occurring in a channel because of the very small inter stroke delays as it integrates the light produced by all strokes in one common discharge. In general, LLS operators use a similar algorithm aiming at grouping strokes according to their separation distance and delay. All the strokes located at a maximum distance of 10 km from each other within 500 ms delay are part of the same flash.

The time difference between the last and the first stroke in the same flash cannot exceed 1 s as it is specified in the Section 4 of the IEC 62858. It is important to note that by convention, LLS operators have decided that a flash is defined by the characteristics (date, time, location, peak current amplitude...) of its first stroke. As a result, the flash dataset is a subset of the stroke dataset only made of the first return strokes after the flash grouping algorithm processing.

Fig. 1. A photography showing a forked flash where one can see
two different paths for return strokes hitting the ground in different places.

Of course, this works nicely for lightning flashes exhibiting only one ionized channel to the ground, but a problem arises when considering the so-called “branched” or “forked lightning flashes”. In most of the cases the dart leader that prepares the occurrence of a subsequent stroke uses the existing conducting path created by the stepped leader/first stroke sequence. Sometimes the dart leader only partly follows the preceding return stroke channel and creates its own path to the ground producing a multiple GSP flash (see figure 1) [9]. In average, between a third to half the multiple strokes downward -CG flashes exhibit more than one contact to the ground depending on the terrain and the season. According to some reports based on high speed video observations found in literature, the mean separation distance between the different GSP in such flashes is 1.8 km [10][11][12][13][14]. This result is consistent with the statistics computed based on the results of the GSP algorithm that will be presented hereafter [15].

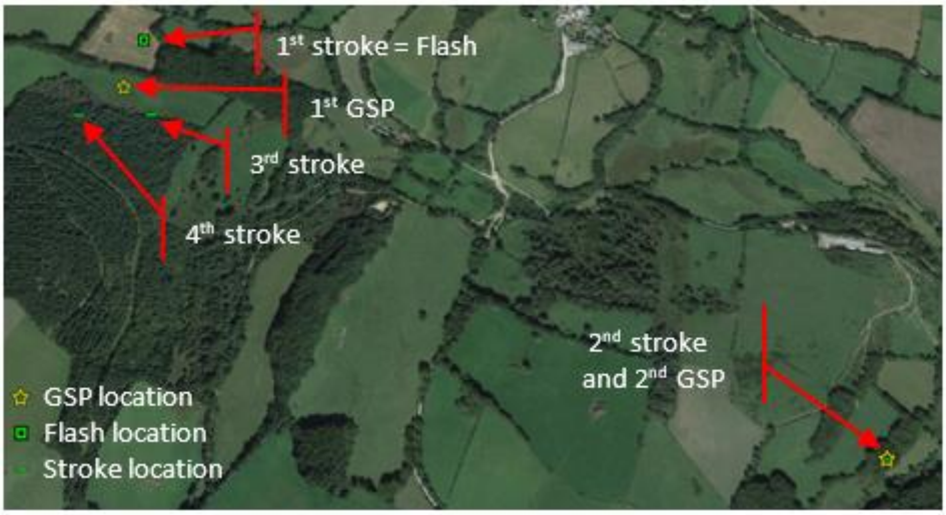

Fig. 2. An example derived from lightning data collected by Météorage that corresponds to a flash exhibiting 2 GSP and 4 strokes. One can clearly see the two different clusters of strokes in the upper-left and lower-right corners of the image. The 1st stroke intiated the 1st GSP and by definition gives the location of the flash. The 2nd stroke occurred in a 2nd GSP at 1.2 km from the previous one. Then the 3rd and 4th strokes used the ionized channel created by the 1st stroke in the 1st GSP. The scattering of the stroke locations in the 1st GSP is due to the location errors committed by the LLS because of sensor measurement errors. As the distance between the strokes is of th same order as the LLS location accacy, the strokes are considered having used the same channel leading to the GSP. Therefore the position of the GSP can be computed as the average of the stroke locations.

Because forked flashes are not rare, some LLS operators like Météorage have developed a GSP identification algorithm that intends to detect the different downward channels inside a flash [14][15][16]. Nowadays, it exists two main algorithms, the first one based on the individual stroke electrical parameters and the like the rise time, the decay time and the peak current amplitude, and the second algorithm that uses the separation distance between strokes [15][16]. These algorithms are presented in the annex B of the IEC 62858. As an example, figure 2 shows the results of the GSP clustering algorithm developed by Météorage using the separation distance between strokes.

Based on the GSP algorithm, it is possible to obtain from an LLS three dataset with a different level of abstraction. The lowest level corresponding to the raw data observed by the LLS is the stroke data. Then, strokes are classified in clusters representing the ionized channel or downward branches in a flash leading to a GSP. The GSP location is computed as the average of the position of the strokes that used the ionized channel. Finally, the flash is the top level in the hierarchical lightning data structure.

It is interesting to note that the suggested lightning data structure presented here below can be completed with an upper level corresponding to the "lightning cell dataset". This new object results from the grouping of all the flashes in consistent clusters whose contours and characteristics describe the thundercloud as observed from an LLS.

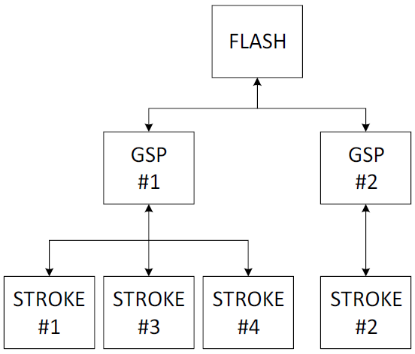

Fig. 3. This example shows the corresponding lightning data structure derived from the real data presented in figure 2. One can see the flash, GSP and stroke datasets are organized in horizontal layers each one corresponfoing to a new level of abstraction.

IV. DISCUSSION

The IEC 62858 standard is a great improvement in terms of lightning data quality management for both LLS operators and engineers in charge of lightning data applications design. The former has now a common framework to evaluate the performances of their systems and the latter benefit from a more reliable information (e.g. limitations) about the data to be used in his application. However, some comments can be done on the status of the document.

The test and validation of a manufactured product are, in general, performed by a third-party laboratory authorized to certify the product in accordance to a given standard requirements. It must be noted that today LLS operators cannot rely on such a laboratory to perform the evaluation of the performance of their system. In practice, the setup of LLS validation techniques is rather difficult to achieve because it requires specific skills and material. This is also true for the operators themselves and only few of them can run an operational monitoring of their system performances. Most of the time, the sensor technology is validated based on test networks deployed in small areas by the manufacturer and the LLS operator use these results to prove their performances. Of course, this is not valid for the IEC 62858 that states that the operational LLS must be validated and not only the technology.

The IEC 62858 standard recommend using the "flash data" for N_g calculation. However, as a flash location is determined by the first stroke, the N_g will underestimate the risk in case of forked lightning flashes. On the contrary, using the "stroke data" will overestimate the risk as because of the inevitable location errors all the strokes, including those in the same channel, have a different computed position. As a result, a multi strokes flash with a single channel will virtually increase the number of strikes to the ground. To improve the N_g estimate, the annex A of the IEC 62858 standard stipulates that the GSP data is to be used when available because it better reflects the real lightning threat as it considers all the contacts of a flash to the ground. Today, only few LLS operator can deliver the GSP data in agreement with the IEC 62858 requirements and therefore compute the N_{sg} that is the lightning density based on GSP data. Because this dataset seems to be promising and several methods are available, the maintenance team MT16 has started a work on the comparison and validation of the GSP identification algorithms [17]. When this work is done, the LLS operators can choose the best algorithm and compute the GSP data.

Although the different entities available in the lightning dataset (e.g. stroke, GSP or flash) describes different physical processes occurring in a lightning flash their characteristics look very similar as they are all characterized by a date, time, geographical position and peak current estimate. Then, a special attention must be paid to make sure the data is relevant in respect of the targeted application or service. If not, the results obtained could be misleading and affect the decision-making process. For instance, using the stroke data to compute the N_g will conduct to an overestimation of the real lightning risk resulting in a more expensive lightning protection system. On the contrary, using flash data for lightning fault correlation might lead to misclassify a fault because only the first and not all the strikes to the ground is considered. As a conclusion, end users and LLS operators must cooperate on the specification of the user application, both actors bringing up their respective knowledge .

BIBLIOGRAPHY

[1] "Cloud-to-Ground Lightning Parameters Derived from Lightning Location Systems - The Effects of System Performance", CIGRE Report 376, 2010.

[2] Rakov V. A.,"Lightning phenomenology and parameters important for lightning protection", IX SIPDA 2007, Iguaçu – Brasil

[3] C. Bouquegneau, "The Need for an International Standard on Lightning Location Systems", in International Lightning Detection Conference and International Lightning Meteorology Conference (ILDC/ILMC), 2014.

[4] "International Standard IEC 62858 Edition 1, Lightning density based on lightning location systems (LLS) – General principles", International Electrotechnical Commission, 2015, ISBN 978-2-8322-2820-3

[5] A. Nag et al., "Evaluation of U.S. National Lightning Detection Network performance characteristics using rocket-triggered lightning data acquired in 2004–2009," J. Geophys. Res. Atmos., vol. 116, no. D2, p. D02123, 2011.

[6] W. Schulz, G. Diendorfer, S. Pedeboy, and D. R. Poelman, "The European lightning location system EUCLID - Part 1: Performance analysis and validation," Nat. Hazards Earth Syst. Sci., vol. 16, no. 2, pp. 595–605, 2016.

[7] W. Schulz, S. Pedeboy, C. Vergeiner, E. Defer, and W. Rison, "Validation of the EUCLID LLS during HyMeX SOP1," in International Lightning Detection Conference and International Lightning Meteorology Conference (ILDC/ILMC), 2014.

[8] Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and 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, doi:10.1029/98JD00153.

[9] L. Z. S. Campos, "On the mechanisms that lead to multiple ground contacts in lightning," PhD Thesis, INPE, Brazil, 2016.

[10] Fleenor S.A., C. J. Biagi, K.L. Cummins, E.P. Krider,"Characteristics of cloud-to-ground lightning in warm-season thunderstorms in the great plains", ILDC2008.

[11] Saba M. F., M. G. Ballarotti and O. Pinto Jr, 2006: Negative cloud-to-ground lightning properties from high speed video observations, JGR, vol 111, D03101.

[12] Stall C, K.L. Cummins, E.P. Krider, J. Cramer (2009), Detecting Multiple Ground Contacts in Cloud-to-Ground Lightning Flashes. Journal of Atmospheric & Oceanic Technology, vol. 26 (11), pp. 2392-2402, November 2009.

[13] Thottappillil R., V. A. Rakov, M. Uman, W. Beasley, M. Master and D. Shelukhin, 1992: Lightning subsequent stroke electric field peak greater than the first stroke peak and multiple ground terminations, JGR, 97(D7), 7503-7509.

[14] K. L. Cummins, "Analysis of Multiple Ground Contacts in Cloud-to-ground Flashes Using LLS Data: The Impact of Complex Terrain," in International Lightning Detection Conference and International Lightning Meteorology Conference (ILDC/ILMC), 2012

[15] S. Pedeboy, "Identification of the multiple ground contacts flashes with lightning location systems," in 22nd International Lightning Detection Conference and 4th International Lightning Meteorology Conference (ILDC/ILMC), 2012.

[16] L. Z. S. Campos, K. L. Cummins, and O. J. Pinto, "An algorithm for identifying ground strike points from return stroke data provided by Lightning Location Systems," in Asia-Pacific Conference on Lightning (APL), 2015, pp. 475–478

[17] S. Pédeboy and W. Schulz, "Validation of a ground strike point identification algorithm based on ground truth data," in International Lightning Detection Conference and International Lightning Meteorology Conference (ILDC/ILMC), 2014.

[18] W. Schulz, S. Pedeboy, and M. M. F. Saba, "LLS Detection Efficiency of Ground Strike Points," in Lightning Protection (ICLP), 2014 International Conference on, 2014, no. 2, pp. 381–384.