Electric Power Systems Research 233 (2024) 110498

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Underestimating lightning risk due to multiple Ground Strike Point flashes

Sanele T. Gcaba *, Hugh G.P. Hunt

Johannesburg Lightning Research Laboratory, School of Electrical and Information Engineering, University of Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Keywords:
Ground flash density
Ground strike point
Lightning
Lightning location system
Map
Risk
SALDN

ABSTRACT

This paper presents an analysis of the lightning risk in South Africa using data detected and geolocated between July 2007–July 2016 from the Southern African Lightning Detection Network (SALDN), which is owned and operated by the South African Weather Service. The study specifically examines the Ground Flash Density (GFD) parameter, which is currently used to depict the lightning risk across the country to inform effective lightning protection. The research finds that the GFD parameter may underestimate the actual lightning risk to ground-based infrastructure, people, animals, and plants. A Ground Strike Point (GSP) algorithm is applied to the SALDN dataset to derive a more accurate Ground Strike Point density map of South Africa. The difference map results, obtained by comparing the GSP density with the GFD, reveals that the GFD parameter underestimates the lightning risk across the entire South African landscape, more significantly in areas between Mpumalanga and KwaZulu-Natal provinces. In these regions, the underestimation reaches a magnitude as high as 32 GSPs per square km per year. The study suggests that applying a constant scaling factor to the GFD parameter can serve as a reliable risk estimation tool when sufficient Lightning Detection Network (LDN) stroke data is unavailable or when there are constraints in data processing power and time due to the large dataset.

1. Introduction

Cloud-to-Ground (CG) lightning flashes may pose a devastating impact on ground-based infrastructure, people, animals, and plant life. In South Africa, the power utility (Eskom) attributed 24% of its transmission line faults to lightning flashes, and insurance claims due to lightning-related damages are of the order of R500 million per annum [1,2]. Lightning in South Africa is said to cause more deaths than any other natural event or phenomena, with a recorded mean rate of 1.5 and 8.8 per million deaths in urban and rural populations respectively [3,4]. In light of the impact of CG lightning flash, it has become crucial to minimize the losses due to lightning by providing lightning protection that is informed by accurate lightning risk assessment in a specific area. The IEC 62305–2 international standard provides a guideline for estimating the lightning risk within a particular area. In the standard, and at a high level, the risk (R_x) is quantified as shown in Eq. (1) below. The risk is a product of the number of dangerous events (N_x), implying lightning events which are currently calculated as CG flashes per square kilometer per year, the probability of damages to a structure (P_x) and the consequent loss (L_x) [5]. The flash density (N_G) map of South Africa, therefore, provides a geographic depiction of CG lightning risk across the country and has found use in various lightning protection design applications such as the Eskom utility transmission line designs through lightning performance analysis [1,6].

$$R_x = N_x \times P_x \times L_x \tag{1}$$

An example of the impact of CG flash on ground-based infrastructure is the lightning-induced overvoltages expected on power lines every year. Eq. (2) below shows the calculation of the expected flashovers (F) based on the Monte Carlo approach for lightning performance assessment for N lightning events [7]:

$$F = \frac{N_i + N_d}{N} A N_g \tag{2}$$

Where N_i and N_d are the number of direct and indirect lightning events, A is the striking area, and N_g is the yearly ground flash density. The map that is currently in use in South Africa for depicting the lightning risk provides an estimation that is based on the Southern African Lightning Detection Network (SALDN) stroke data which is grouped into flashes in accordance with the IEC 62858 [8]. The flash grouping criteria does not account for the possible multiple ground strike-points (GSPs) of a flash, and the criteria assume that all strokes of a flash can be adequately represented by the parameters of the first return stroke, thus assuming a single GSP per flash.

* Corresponding author.
E-mail address: sanelet.gcaba@gmail.com (S.T. Gcaba).

https://doi.org/10.1016/j.epsr.2024.110498
Received 4 August 2023; Received in revised form 3 May 2024; Accepted 15 May 2024
Available online 1 June 2024

Table 1
Literature findings on high-speed video observations of CG lightning flashes across different countries.

Study	Location	CG flashes	Multiplicity	N(GSP)	Average (GSP/Flash)
Poelman et al. [10], Schumann et. al [11] (2017–2018)	South Africa	490	3.8	626	1.29
Saba et al. [12] (2006)	Southeastern Brasil	233	3.8	235 (for 138 flashes)	1.7
Valine et al. [13] (1997)	Tucson, Arizona	386	2.8	558	1.45

2. Background

Negative CG flashes typically consist of three to five strokes with approximately half of the flashes striking the ground at more than one point [9]. The multiple ground terminations or strike points present an additional risk of a lightning flash striking ground-based infrastructure, people, animal, and plant life at more than a single point per flash. Positive CG flashes, on the other hand typically consist of single strokes lightning [9].

2.1. High-speed video observations

High-speed cameras have been instrumental in understanding the lightning phenomena, these have enabled the dissection and observation of lightning flashes into their finer details of individual strokes [10]; other applications of high-speed video observations include performance evaluation of lightning detection and location systems through ground truth observations as well as enabling the identification of ground strike-points of individual strokes. It is however practically impossible to have these high-speed cameras installed and focused at every location where the lightning incidents will take place considering the nature of lightning formation and propagation. On some of the most recent research, Poelman et al. compare ground truth observations for negative CG flashes from South Africa, Austria, Brazil, and the USA [10]. The results show a mean number of strokes per flash (multiplicity) of 3.67 and an average ground strike-point per flash of 1.56, this implies that ground strike points are typically 56% higher than the number of flashes. Table 1 depicts high-speed video studies that show that a single lightning flash typically has multiple GSPs and thus illustrating that the lightning flash density parameter and map may be underestimating the lightning risk.

The average ground strike-point per flash observed in South Africa, Brazil, and the USA as shown in Table 1 above is comparable with the range between 1.5 and 1.7 observed by Rakov and Uman [14].

2.2. Southern African lightning detection network

The South African Weather Service (SAWS) installed the South African Lightning Detection Network (SALDN) in 2005. The SALDN initially comprised of 19 LS7000 Vaisala inc. sensors. Improvements on the network saw the introduction of additional 6 sensors, taking the network to a total of 25 sensors by 2015. The lightning detection sensors across South Africa generally have a baseline of 150 km and there are 18 LS7000 sensors, 4 LS7001 sensors and 3 LS7002 sensors installed [15–17]. The performance of a lightning detection system is defined by its Detection Efficiency (DE) and Location Accuracy (LA), where DE refers to the number of lightning events (stroke or flash) detected versus the actual lightning events, while LA refers to the distance between the position of the located return stroke and the actual ground strike-point. Self-evaluation techniques used by such networks have generally shown that the SALDN maintains a median flash detection efficiency of 90% and an estimated median location accuracy of 500 m. The location accuracy is based on the assumption that the location errors follow a Gaussian distribution, and the most probable location is given by the 50% confidence ellipse with the estimated location of the stroke at the center of the ellipse. This implies that there are 50% chances that the actual location lies inside the ellipse with the semiminor, semimajor and ellipse angle given by the detection network [18]. The minimum system performance requirements for an LLS to derive lightning density as stipulated in the IEC 62858 [8] are as follows:

- The Flash DE should be at least 80%
- The median LA should be better than 500 m
- Classification accuracy of CG strokes against intra-cloud strokes should be at least 85%

However, there have also been a number of ground-truth performance evaluations of the SALDN, involving all towers. The first occurred in 2014, by *Hunt et al.* and compared SALDN flash reports with time-stamped photographs of upward lightning events to the Sentech tower, Johannesburg for the period of 2009 to 2013. A flash DE of 76% and median LA of 280 m was found [19]. As the camera resolution did not allow for individual strokes to be identified, stroke DE was not assessed. Fensham et al. [20] subsequently evaluated the SALDN performance using high-speed camera footage as the ground-truth data. The detected and geo-located SALDN stroke reports were compared with 43 high-speed camera captures of upward lightning event to the Hillbrow and Sentech towers and 206 high-speed video captures of downward lightning events in the Johannesburg region. The results showed that the network had flash DE of 55.8% and a stroke DE of 63.4% for upward lightning events (with a median location error of 59.2 m) but a flash DE of 93%, a stroke DE of 69% for downward lightning events (with a median location error of 124.4 m) [20,21]. Similarly, Poelman et al. showed the location of the SALDN was in the range of 110 m to 150 m in the Johannesburg region, utilizing the median error ellipse information [22]. These are the definitive ground-truth study for the SALDN, however, it is only at a specific location. In order to assess the performance of the SALDN over the entire country/network, the approach taken by Zhu et al. was implemented on the SALDN dataset. Tall towers around South Africa were identified and the algorithm was implemented. The results found a median location accuracy of 158 m. The location accuracy increased greatly in provinces in the center of the network, decreasing around the edges of the network as expected. Based on the above studies, a location accuracy of 150 m is considered as a reasonable estimate of the median location accuracy of the SALDN over the whole country/region [23].

Fig. 1 below shows a geospatial plot of a single flash from the 2017 SALDN stroke data correlated with ground-truth observations from high-speed cameras placed around the Johannesburg area. The figure shows a single flash with a total of six detected and geo-located lightning strokes, these were observed to have two ground strike points as illustrated with the red circles.

It can be seen from the figure that due to the inherent location errors, the SALDN detected stroke locations of a single flash do not line up to ground strike points as observed from the high-speed camera ground-truth data. Therefore, this implies that using the stroke data to estimate the lightning risk in a particular area may overestimate the number of ground strike points and consequently the lightning risk.

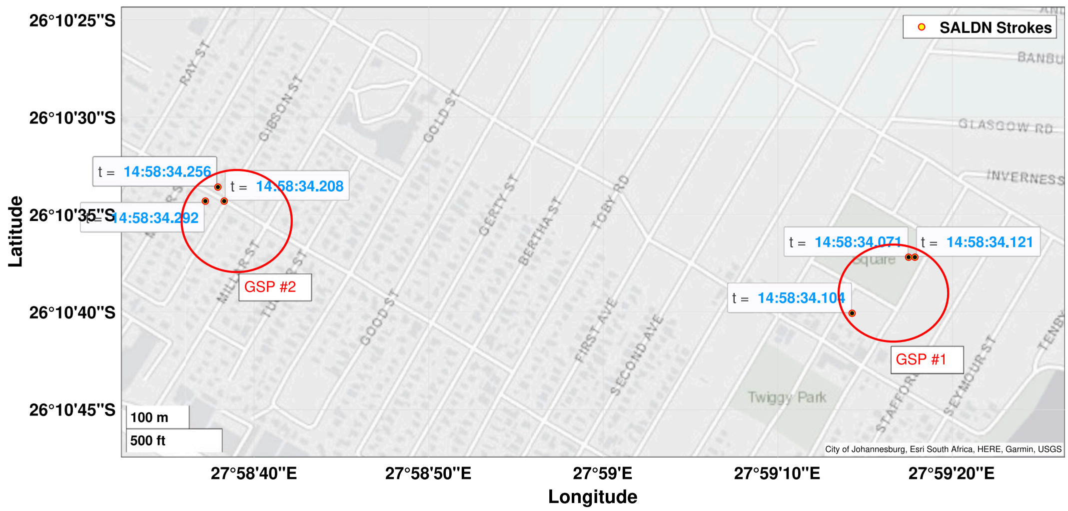


Fig. 1. Ground-Truth correlated single flash data with six strokes and two ground strike-points in Johannesburg.

2.3. Flash grouping algorithm and flash density maps

The IEC 62858 [8,24,25] prescribes a detailed standard framework for computing lightning location system-based lightning density, the standard guarantees the estimation of the lightning risk with an error that is less than 20%, and the following is the grouping criteria for a lightning stroke to form part of a flash:

- The stroke must occur within less or equal to 1 s after the first reported return stroke within the same flash
- The location of the stroke must be within less than or equal to 10 km from the first reported return stroke
- The time interval between successive strokes is less than or equal to 500 ms

The lightning flash is then characterized by the date, time, location, and peak current amplitude of the first reported stroke of the flash.

2.4. Ground strike-points from lightning location data

The observation that estimating the lightning risk using the flash density parameter could be underestimating the risk has seen some proposals to scale the flash density (N_G) by a constant factor to deduce the ground strike-point (N_{SG}), this is as per the Eq. (3) below [26].

$$N_{SG} = f \times N_G \quad (3)$$

Where f is proposed to be a scaling factor of 2. The proposed scaling factor is derived from the typical average number of GSPs per flash that has been observed and documented in the literature and presented in Section 2.1 above.

2.5. GSP algorithms

GSP algorithms are statistical clustering algorithms aimed at grouping individual strokes reported by a LLS into consistent clusters forming ground strike-points [27]. These clusters are typically arranged based on similarities that characterize a common ground strike-point such as the distance between the individual strokes of a flash and the relation between the error ellipses of the individual strokes of the flash. The first three algorithms found in literature cluster strokes using statistical k-means-based clustering methods. K-means clustering is the most widely used clustering and data analysis method [27]. The key difference between these k-means algorithms is the actual metric considered when clustering. The algorithms either make use of a euclidean distance between strokes, elliptical scaling, and/or the location errors defined by the error/confidence ellipses to create clusters defining common ground strike points.

Three GSP algorithms have been proposed in literature for grouping individual LLS detected and geolocated strokes of a flash into one or more GSPs:

1. The *Progressive K-Means* algorithm was initially introduced and used by *Pédeboy et al.* to identify multiple ground contacts of a flash using lightning stroke dataset from the French national lightning location system. The algorithm statistically clusters strokes of a flash using the distance metric between strokes. While typical k-means algorithms are not sufficient to cluster strokes into GSPs considering that the number of strike points (k) are initially not known, the *Progressive K-Means* algorithm iteratively increments k from 0 and clusters strokes of a flash to their nearest GSP [27].
2. Campos et al. [28] made further improvements to the progressive k-means by adding an additional layer that considers the lightning stroke error ellipse (semi-major and semi-minor axis) to further refine the accuracy when clustering the individual strokes of a flash into GSPs.
3. Matsui et al. [29] presented the most recent algorithm for clustering Japanese Lightning Detection Network (JLDN) strokes of a flash into their GSPs, the algorithm uses the distance between strokes of a flash and overlapping error ellipses of the respective strokes to define a GSP.

Poelman et al. [22,30] assessed the algorithms and found that they all perform at success rates up to 90% against ground truth observations in Austria, Brazil, France, Spain, South Africa and the United States. In other words, the algorithms correctly categorize strokes as ending in either previous or new locations 90% of the time.

3. Methodology

3.1. LDN dataset clustered into flashes

The study utilized a 9-year dataset encompassing SALDN lightning strokes detected and geolocated by the SAWS between July 2007 and July 2016. It is worth noting that during this period, the SALDN underwent sensor upgrades highlighted in Section 2.2 to improve system performance. The improved SALDN technology impacts the DE and LA, enhancing the dataset reliability and analysis; however, the old dataset used was the sole pre-processed data accessible to the authors from the SAWS during analysis. It is expected that the SALDN dataset consists

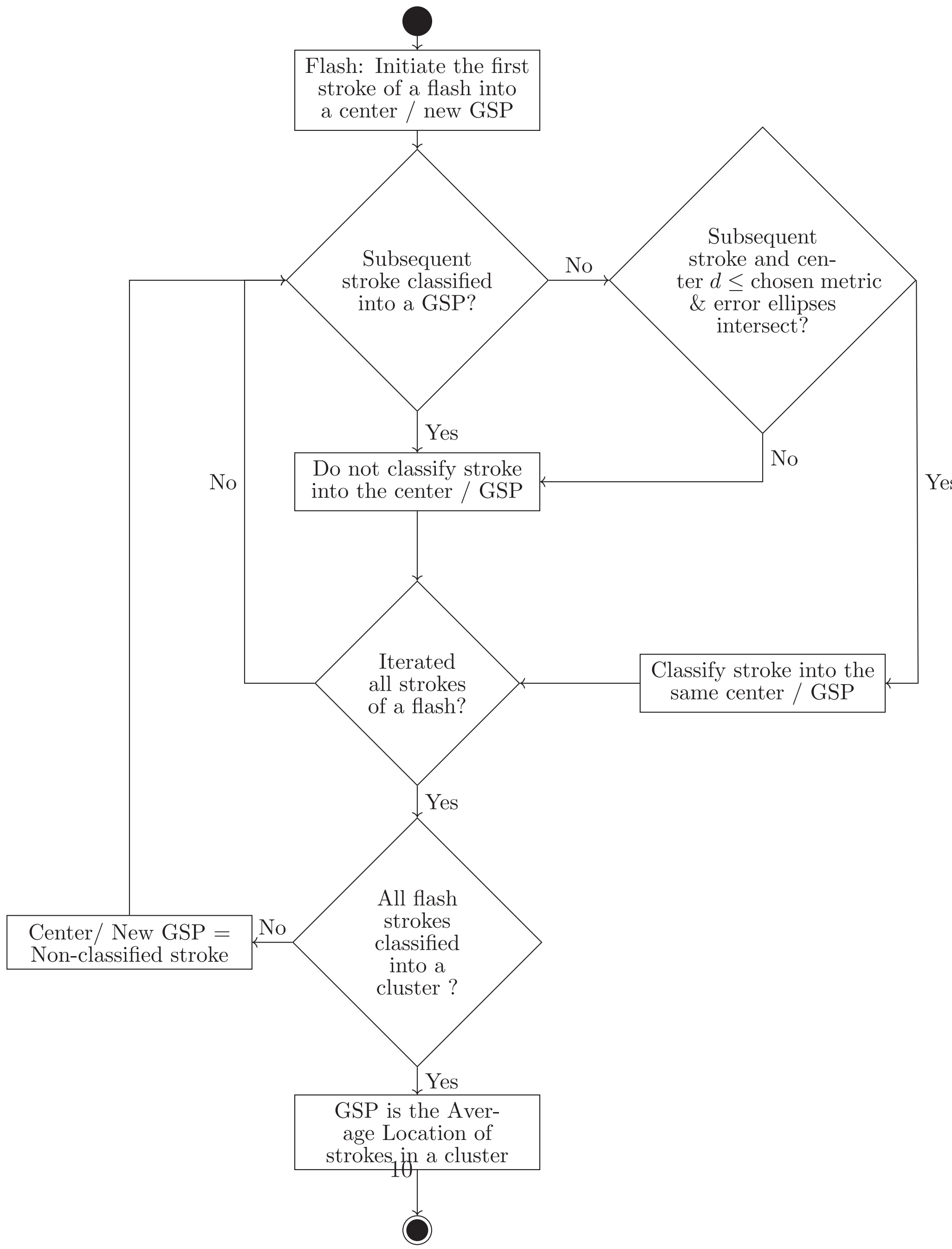


Fig. 2. Lightning strokes to ground strike-point clustering algorithm.
Source: Adapted from Matsui et al. [29].

of an insignificant number of records of upward lightning events. This consideration is based on the fact that only 1% of CG events are upward flashes initiated from tall structures located on the ground [31]. The CG flashes are selected from the dataset by removing all the intracloud flashes, further processing of the data for the proposed application requires only CG flashes. The LDN stroke data is then clustered into flashes using the criteria presented in Section 2.3 above.

3.2. Cluster strokes into GSPs

The algorithm proposed by Matsui et al. [29], was adapted and used for the current investigation. When compared with the other algorithms labeled 1 to 3 described in Section 2.5, it was chosen due to its simplicity while maintaining the ability to consider both the distance metric and error ellipse data for deriving clusters of better-refined ground strike points from lightning location system data. The adapted version of the algorithm is presented in Fig. 2.

The algorithm in Fig. 2 operates on LDN stroke data that is clustered into flashes, each stroke of a flash is iterated through and clustered into a ground strike-point where the distance is within a defined limit and the error ellipses of the strokes intersect. The distance (d) between subsequent strokes is calculated using the haversine formula as in Eq. (4) below.

$$haversin(\frac{d}{R}) = haversin(\varphi_2 - \varphi_1) + cos(\varphi_1)cos(\varphi_2)haversin(\lambda_2 - \lambda_2) \quad (4)$$

And

$$haversin(\theta) = sin^2(\frac{\theta}{2}) \quad (5)$$

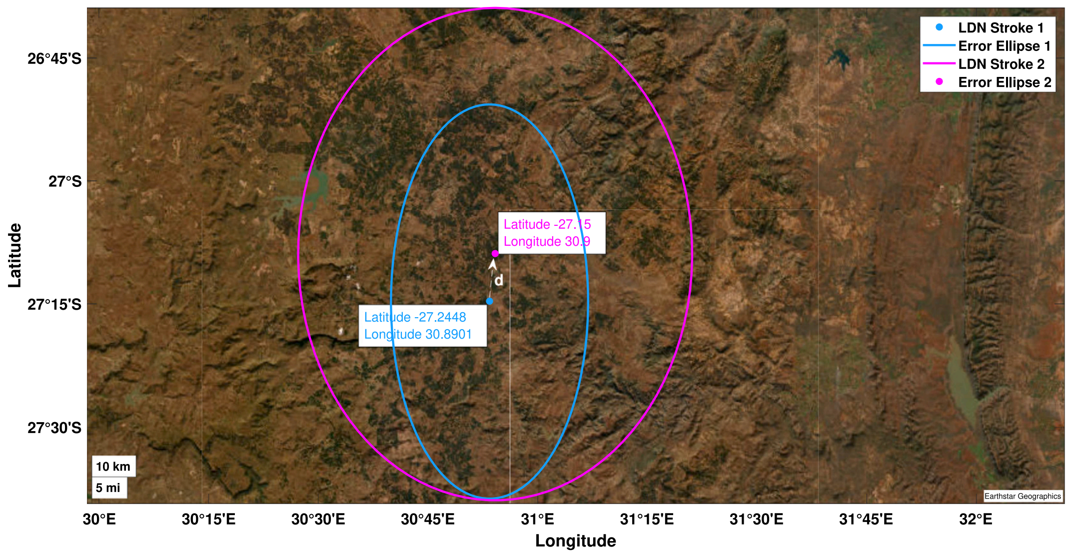


Fig. 3. Illustration of how strokes within a flash and their respective error/confidence ellipse data are compared to define a single ground strike-point.

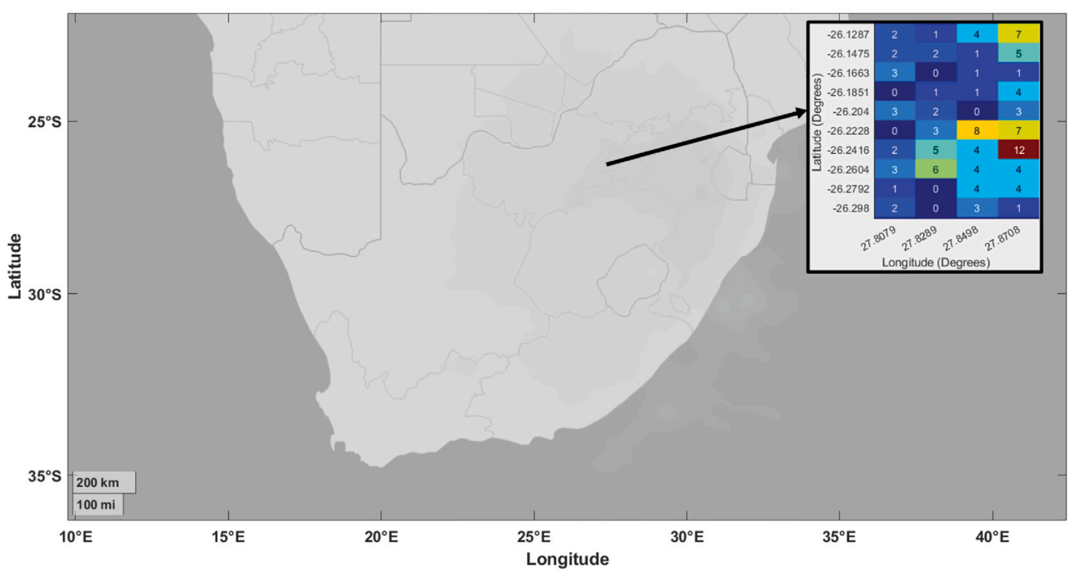


Fig. 4. Illustration of how the flash and GSP density map is derived from counting the number of flashes and GSPs per square kilometer area across South Africa.

Where: d = distance between two geographic coordinates in km
R = Radius of the earth in km
$\varphi_{1,2}$ = Latitude coordinates
$\lambda_{1,2}$ = Longitude coordinates

Choosing the distance metric to be used as the threshold must be based on the location accuracy of a network. This is because two strokes in the same location may be geolocated at distance from each other due to location accuracy errors — if the distance threshold metric in the applied GSP algorithm is less than this median location accuracy, strokes that should be grouped into 1 GSP will be separated as 2 different GSPs incorrectly.

For the purposes of this study, a distance threshold metric of 150 m is selected based on the median location accuracy of the SALDN over the whole of South Africa as discussed in Section 2.2.

Fig. 3 is an illustration of two LDN strokes and their intersecting error ellipses, this is the criteria used as a comparison method between flash strokes confidence/error ellipses to define a single GSP. The respective error/confidence ellipses are compared and checked to see if they intersect with each other. If the calculated distance (d) is within the defined threshold and the respective error ellipses are found to intersect with each other, they are then classified to form a single GSP defined as the average of the individual stroke locations.

3.3. Density mapping

In order to map the geographical distribution of the lightning ground flash density and GSP density, the study area bounded between the *22 °S to 35 °S* latitude and *17 °E to 33 °E* longitude illustrated in Fig. 4 was used. The area is effectively modeled into 2 km × 2 km cells and the Figure illustrates how the density map is derived by counting the number flashes or GSPs on a square kilometer area by following the below described steps.

- Step 1: A count of the number of flashes and GSPs respectively per 2 km × 2 km grid cell must be obtained for a period of some years (in this case, 9 years). This involves identifying all flashes and GSPs inside the 2 km × 2 km grids over the period and counting them.
- Step 2: The number of flashes and GSPs in the grid cells from Step 1 are divided by the area of the grid cells respectively to get

the square kilometer count. For the 2 km × 2 km area, the results are divided by 4.

- Step 3: Finally the results from Step 2 are divided by the number of years in the period under observation. In this case, 9 years.

A 2 km × 2 km grid resolution is chosen based on the official flash density map of South Africa as developed by Evert et al. [15]. The size of South Africa makes this the best grid cell to develop a visually relevant map. Similarly, utilizing this grid size allows comparison with the previously developed maps. The output of these steps provide the count of lightning flashes and GSPs per square kilometer per year in the gridded map of South Africa. These results graphically represent the distribution of lightning risk based on density.

3.4. Difference density arrays

To assess the extent to which the lightning density risk is underestimated when the GFD parameter is used compared to the GSP density parameter, the difference between the two density arrays needs to be evaluated across South Africa. This is done as follows:

$$Diff_{density} = |N_{sg} - N_g| \quad (6)$$

- The difference density ($Diff_{density}$) depicted in Eq. (6) above is calculated by subtracting the GFD array (N_g) from the GSP density array (N_{sg}), these arrays are bins of Flash and GSP counts that are located at the same geographic grids tessellated as 2 km × 2 km grid cells.
- The second difference density is computed between the N_g parameter that is scaled by a factor 1.56 as discussed in Section 2.1 and the N_{sg} parameter from GSP clustered stroke data. The method for computing the difference is illustrated in Fig. 5 below.

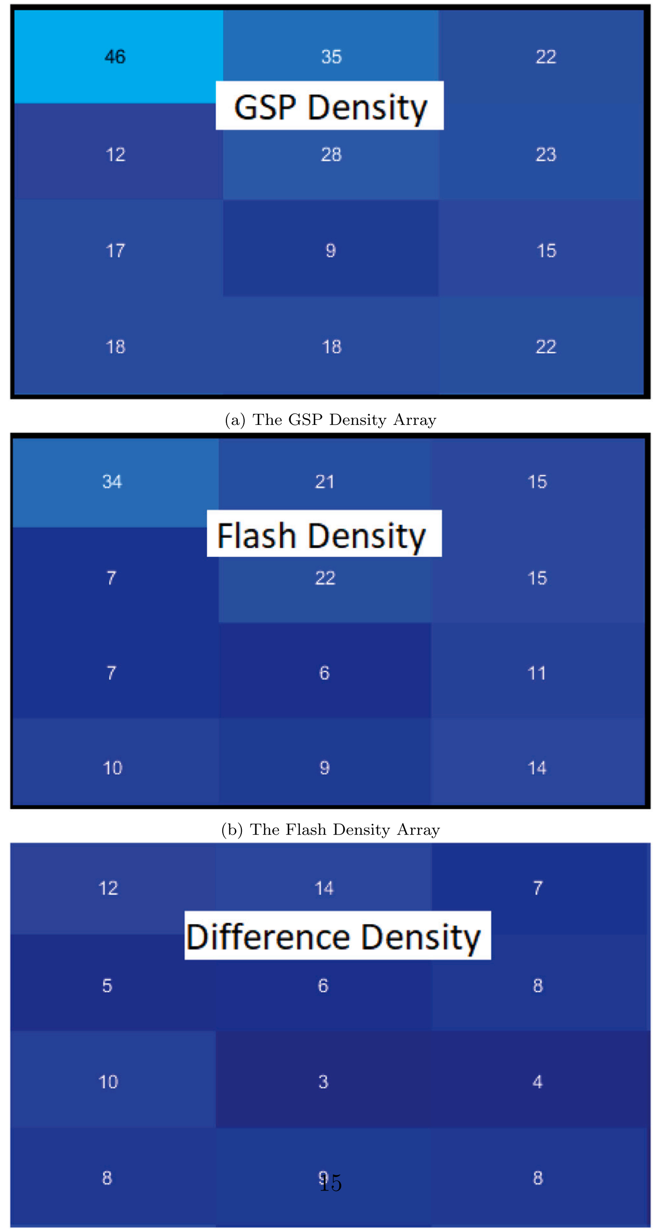
4. Results

The subsequent sections detail the outcomes of implementing the described methodology on the SALDN dataset. This involves setting the distance threshold parameter to 150 m, reflecting the SALDN's estimated location accuracy. It then applies a scaling factor of 1.56, derived from findings in [10].

4.1. GSP and GFD difference density map with a 150 m distance threshold parameter

Fig. 6 below geographically depicts the difference between the GSP density per square km and the GFD for the 9-year period between July 2007 and July 2016. The difference density map in Fig. 6(c) shows a relatively high difference in lightning density in areas around the central-eastern parts of South Africa, particularly around the borders of KwaZulu-Natal, Free State, and the Mpumalanga province. The regions around the borders of Mpumalanga and KwaZulu-Natal, which show high GFD and GSP densities, also exhibit an increase in difference density per square kilometer per year and the difference density color bar and lightning distribution pattern demonstrates that the difference between the GSP density plot and GFD increases in a proportional pattern with the GSP and GFD pattern respectively. The plot in Fig. 6(c) shows that no difference or very low difference corresponds to the white color on the color bar, an increase in density is illustrated by the color yellow and light red to dark red which illustrates an increase in density difference proportionally in that order. Table 2 below shows a statistical summary of the difference dataset.

The difference in Table 2 shows a minimum density of 0 per square kilometer per year, indicating that there are regions in which the flash to GSP ratio is 1 to 1 - likely due to many singles stroke flashes. The standard deviation of 2.94 indicates that the differences have a relatively high variability around the mean, suggesting that the differences between GSP and GFD vary widely.



(a) The GSP Density Array

(b) The Flash Density Array

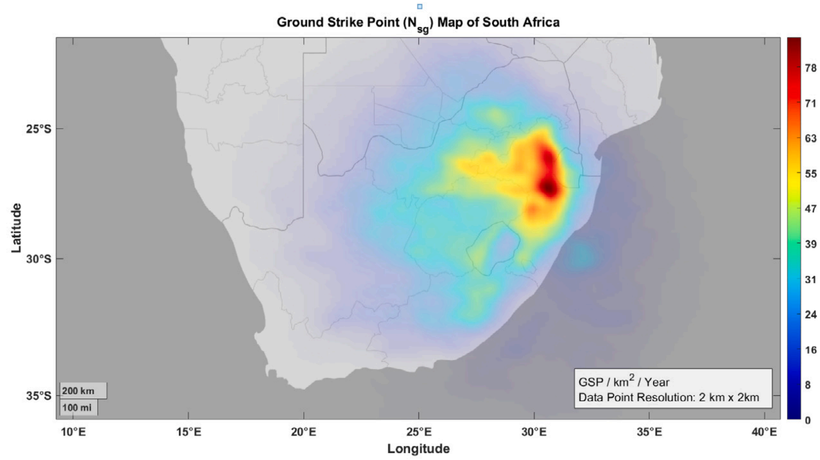
(c) The Difference Density Array

Fig. 5. The comparison between the GSP density and Flash density Arrays (GSP density - GFD).

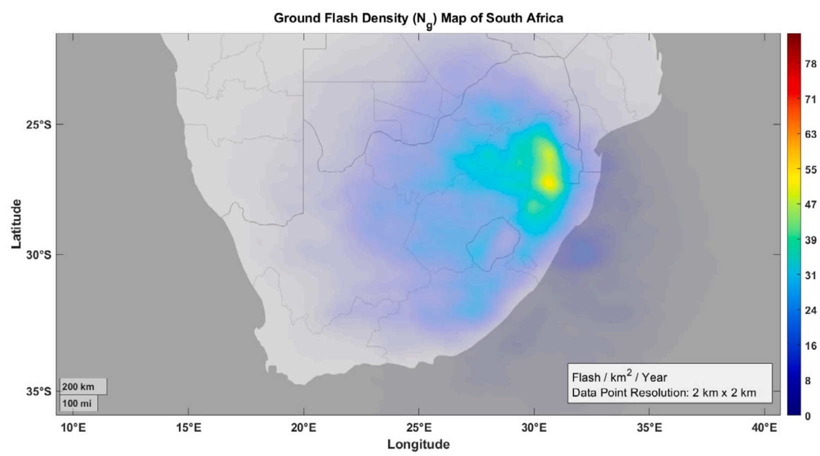
Table 2
Difference ($|N_{sg} - Ng|$) statistical data from the 9-Year SALDN dataset.

Difference ($|N_{sg} - N_g|$)/km^2/year	Value
Min	0
Arithmetic Mean	1.99
Median	0.67
Max	32
Standard deviation	2.94

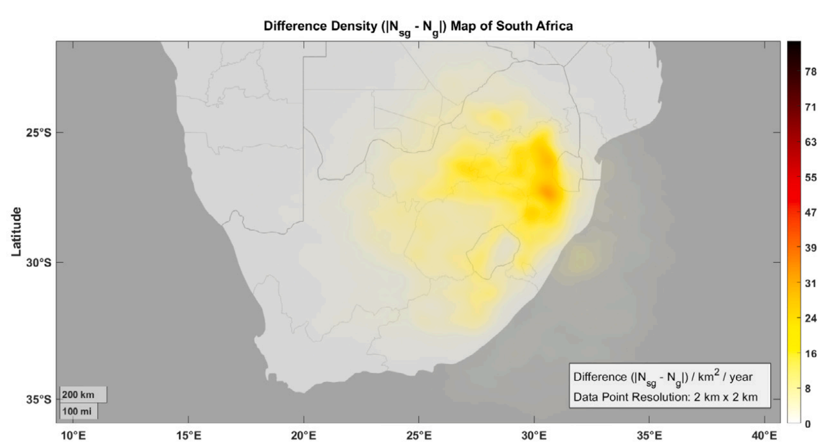
Overall, the statistics suggest that there are significant differences between GSP density and GFD maps, with a typical positive difference of around 1.99 but with some extreme values that are much larger. The results show a maximum difference of 32 ground contact points per square kilometer per year at the geographic location 27.2100°S, 30.4600°E in Mpumalanga. This means that on average, there are approximately 32 more GSPs per year than flashes in this region - which



(a) The GSP Density Map of South Africa using a distance metric of 150 m and intersecting error ellipses

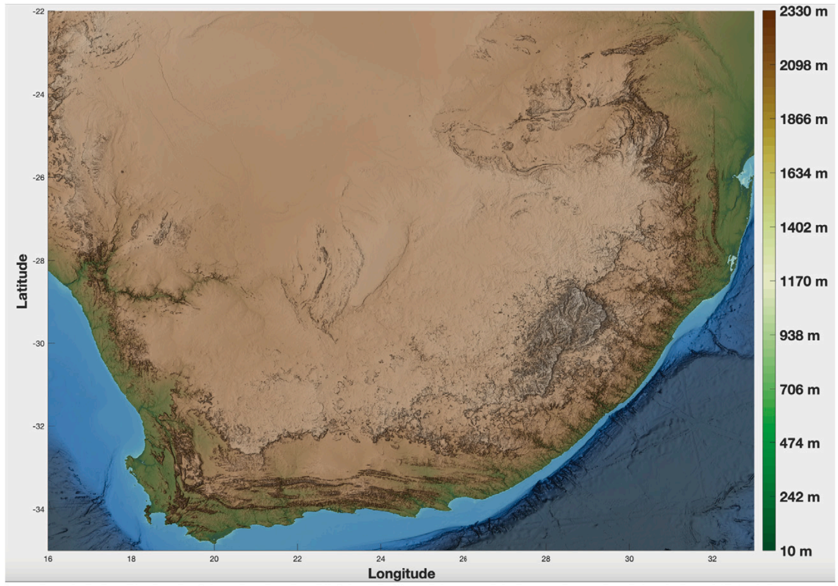


(b) The Flash Density Map of South Africa from the SALDN stroke data spatially and temporally clustered into Flashes

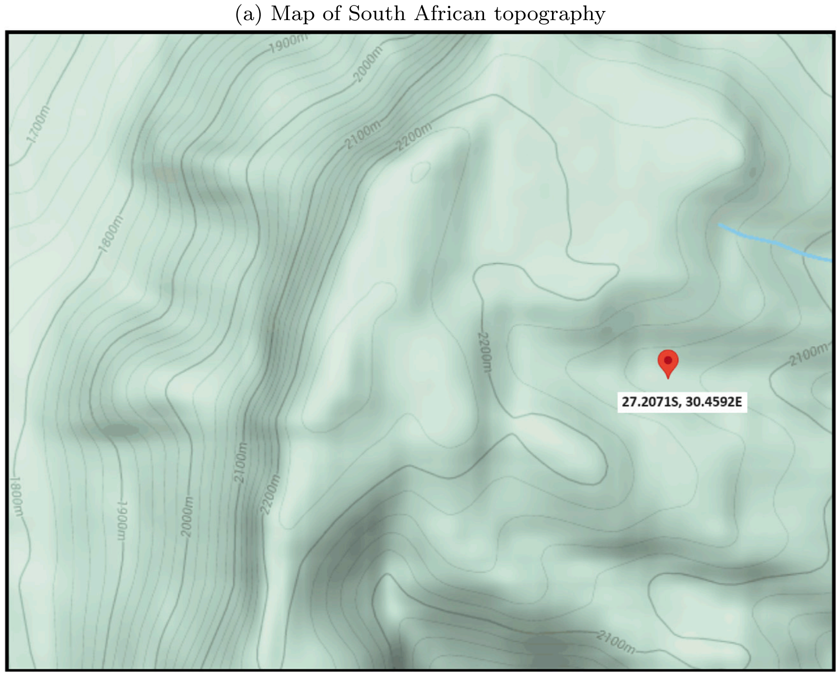


(c) Difference Density Map

Fig. 6. The comparison between the GSP density and Flash density using a difference map ($|N_{sg} - N_g|$)



(a) Map of South African topography



(b) Altitude around the areas with the highest difference density for the 9-year period.

Fig. 7. A topographical illustration of the South African Landscape (a) the topography across South Africa, ranging from 0 to 2300 m above sea level. (b) Altitude (above 2000 m) around the areas with the highest difference density for the 9-year period.

is in the order of 50 flashes per year, and increases to 80 if we include the GSP contacts, approximately 1.5 times greater. A scaling factor of 2 would greatly overestimate the GSP density in this region. This location is 38 km south-west of Piet Retief, which has historically shown a high lightning density. Fig. 7(a) shows the map of South African topography, and it can be seen that the regions showing high lightning Flash and GSP density are at higher altitudes relative to areas with low Flash and GSP density. In fact, it appears the highest difference between flash and GSP density occurs just off the peak of the main mountain ranges. In particular the altitude around the location with the highest difference between the lightning Flash and GSP density ranges between 1800 m to 2200 m as shown in Fig. 7(b).

It is evident from the difference density plot that the GFD parameter, and consequently, the South African GFD density map as a lightning risk depiction tool, underestimates the amount of lightning risk when compared with the GSP density map. The extent of lightning risk underestimation is illustrated by the difference density map in Fig. 6(c), where the value of underestimation consistently follows the GSP pattern.

Fig. 8 shows a flash density plot of excluding all flashes with a multiplicity greater than 1. The density follows a similar pattern to what is seen in both the Flash density and the GSP density plots.

4.2. GSP and scaled GFD difference density map

Fig. 9 below shows the GSP density map of South Africa, which is derived from the GFD map scaled by a constant scaling factor of 1.56 as discussed in the methodology. This scaling factor is based on the

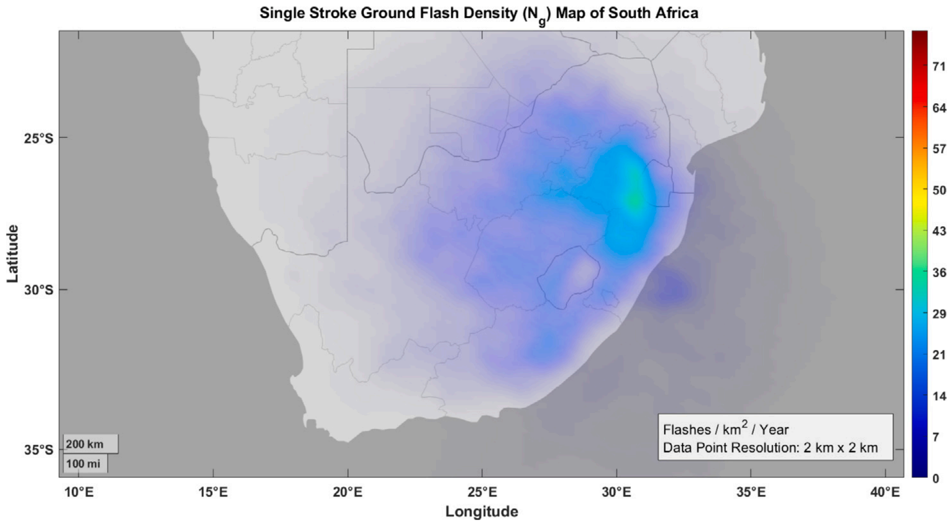


Fig. 8. Lightning flash density over South Africa for single stroke flashes for the 9 year period.

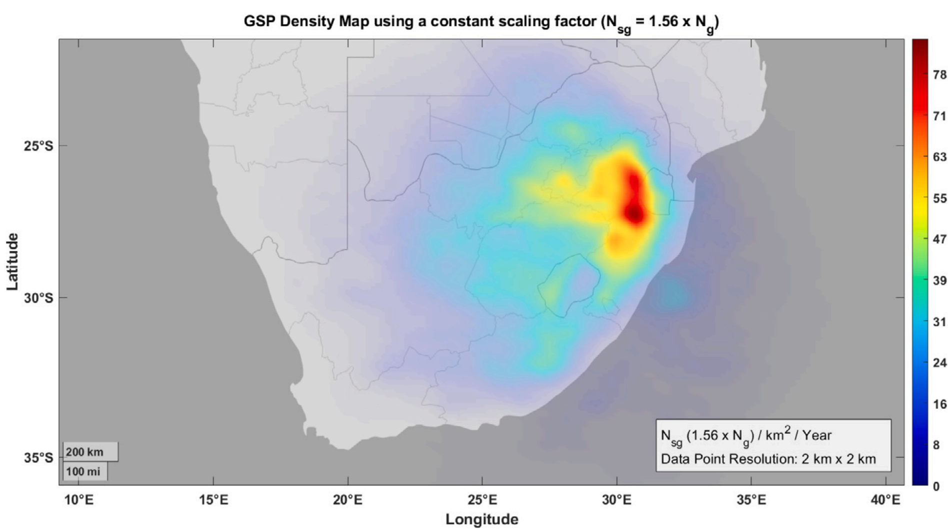


Fig. 9. GSP Density Map using a constant scaling factor of 1.56 based on ground-truth observations.

Table 3
Difference ($|N_{sg} - 1.56 \times N_g|$) statistical data from the 9-year SALDN dataset.

Difference ($|N_{sg} - 1.56 \times N_g|$)/km^2/year	Value
Min	0
Arithmetic Mean	0.41
Median	0.18
Max	12
Standard deviation	0.62

average ground truth observations across Austria, Brazil, South Africa, and the USA, as presented in [10].

Fig. 10 below shows the difference between the GSP density derived from a constant scaled GFD and the GSP density derived from the SALDN data over the same 9-year period.

The difference shown in Fig. 10 for a constant scaling factor is low across the South African landscape when compared with that presented in Fig. 6(c).

The statistics in Table 3 characterize the difference dataset for a constant scaling factor. As seen in the table, the arithmetic mean suggests that, on average, the two values being compared have a difference of about 0.41. This is small and suggests that using a scaling factor does not significantly overestimate (or underestimate) the GSP density over the whole region, even though it may do so in specific locations or areas. This is also indicated by the median value of 0.18.

As expected, the maximum value of 12 is much smaller than the previous dataset where the maximum was 32. The standard deviation of 0.62 is also much smaller than the previous dataset where it was 2.94. The low difference shown in Fig. 10 indicates that a constant scaling factor based on LDN ground truth observations can provide a reasonable estimate of the lightning risk where:

- Sufficient LDN Flash data to derive the GSP density from GSP clustering algorithms is not available.
- There is no sufficient time or processing power to derive GSP density from first principles using the GSP clustering algorithms.

However, to achieve a difference that is approximately zero between the LDN-based GSP density and the GFD scaled density, and to improve the lightning risk estimation, a square kilometer area-specific scaling factor matrix can be used instead of a constant. This is necessary to account for the area-specific and multiple variables that can affect the GSP density, such as elevation.

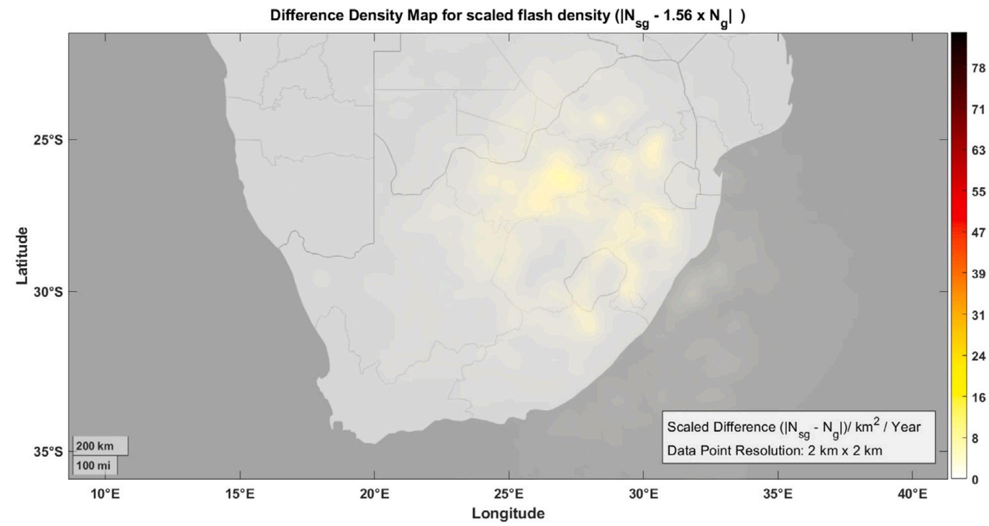


Fig. 10. Difference Density map between the scaled flash density ($N_{SG} = 1.56 \times N_G$) and the LDN-derived GSP density.

5. Conclusion

In this study, a GSP density map of South Africa obtained through a GSP clustering algorithm is presented. The authors then aimed to compare the GFD parameter with the GSP density parameter to evaluate the extent to which the former underestimates the lightning risk. For this purpose, the authors used a 9-year dataset of the SALDN lightning strokes detected and geolocated by the SAWS between July 2007 and July 2016, which was presented on a 2 km *x* 2 km gridded map of South Africa. The results showed that the GFD parameter consistently underestimates the lightning risk across the South African landscape, in proportion to the GFD and GSP patterns. The highest value of underestimation was 32 per square kilometer per year and was recorded in Mpumalanga. The study also showed that a constant scaling factor based on LDN ground-truth observations can provide sufficient lightning risk estimation where LDN stroke data is not available to derive GSPs using the GSP clustering algorithms and where there is no sufficient data processing time and power for the large dataset. Additionally, the author suggests that improving the risk estimation using a constant scaling factor may require a square kilometer area-specific scaling factor matrix that takes into account multiple factors, such as elevation, that have been shown to affect the GSP density.

CRediT authorship contribution statement

Sanele T. Gcaba: Writing – original draft, Software, Investigation, Formal analysis. **Hugh G.P. Hunt:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hugh GP Hunt reports financial support was provided by National Research Foundation.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work is based on the research supported in part by the National Research Foundation of South Africa and their support of the Johannesburg Lightning Research Laboratory through the Thuthuka programme (Unique Application No.: TTK23030380641) and by DEHNAFRICA and their support of the Johannesburg Lightning Research Laboratory. The authors would like to thank the South African Weather Service (SAWS) for their support and for providing the SALDN data used in this paper, specifically Morné Gijben and Michelle Hartslief.

References

[1] L. Peter, F. Mokhonoana, Lightning detection improvement FALLS brought to eskom's transmission line design and fault analysis, in: Proceedings of the 21st International Lightning Detection Conference and 3rd Lightning Meteorology Conference, 2010, pp. 19–22.
[2] R. Evert, G. Schulze, Impact of a new lightning detection and location system in south africa, in: 2005 IEEE Power Engineering Society Inaugural Conference and Exposition in Africa, 2005, pp. 356–363.
[3] B. Bhavika, The influence of terrain elevation on lightning density in South Africa (Ph.D. thesis), University of Johannesburg, 2007.
[4] I.R. Jandrell, G. Saayman, E. Trengrove, R. Blumenthal, Lightning medicine in South Africa, Afr. J. Health Prof. Educ. 102 (7) (2012) 625–626.
[5] IEC 62305-2:2010, IEC 62305-2:2010, Protection Against Lightning - Part 2: Risk Assessment, Standard, International Electrotechnical Commission, Geneva, CH, 2010.
[6] S. Gcaba, H. Hunt, Ground strike point density map of South Africa, in: 2022 36th International Conference on Lightning Protection, ICLP, 2022, pp. 1–5.
[7] M. Nicora, D. Mestriner, M. Brignone, R. Procopio, E. Fiori, A. Piantini, F. Rachidi, Estimation of the lightning performance of overhead lines accounting for different types of strokes and multiple strike points, IEEE Trans. Electromagn. Compat. 63 (6) (2021) 2015–2023.
[8] IEC62858:2015, IEC 62858:2015, Lightning Density Based on Lightning Location Systems (LLS) – General Principles, Standard, International Electrotechnical Commission, Geneva, CH, 2015.
[9] V. Rakov, Lightning parameters for engineering applications (keynote speech), in: 2010 Asia-Pacific International Symposium on Electromagnetic Compatibility, IEEE, 2010, pp. 1120–1123.
[10] D.R. Poelman, W. Schulz, S. Pedeboy, D. Hill, M. Saba, H. Hunt, L. Schwalt, C. Vergeiner, C.T. Mata, C. Schumann, et al., Global ground strike point characteristics in negative downward lightning flashes–Part 1: Observations, Nat. Hazards Earth Syst. Sci. 21 (6) (2021) 1909–1919.
[11] C. Schumann, H.G. Hunt, J. Tasman, H. Fensham, K.J. Nixon, T.A. Warner, M.M. Saba, High-speed video observation of lightning flashes over Johannesburg, South Africa 2017–2018, in: 2018 34th International Conference on Lightning Protection, ICLP, IEEE, 2018, pp. 1–7.
[12] M. Saba, M. Ballarotti, O. Pinto Jr., Negative cloud-to-ground lightning properties from high-speed video observations, J. Geophys. Res.: Atmos. 111 (D3) (2006).
[13] W.C. Valine, E.P. Krider, Statistics and characteristics of cloud-to-ground lightning with multiple ground contacts, J. Geophys. Res.: Atmos. 107 (D20) (2002) AAC–8.

[14] V.A. Rakov, M.A. Uman, Lightning: Physics and Effects, Cambridge University Press, 2003, pp. 1–12.
[15] C. Evert, M. Gijben, Official South African lightning ground flash density map 2006 to 2017, in: Inaugural Earthing Africa Symposium and Exhibition, Thaba Eco Hotel, Johannesburg, 2017, pp. 5–9.
[16] K.L. Cummins, M.J. Murphy, J.V. Tuel, Lightning detection methods and meteorological applications, in: IV International Symposium on Military Meteorology, 2000, pp. 26–28.
[17] M. Gijben, The South African lightning detection network, 2021, https://www-cdn.eumetsat.int/files/2021-03/The%20South%20African%20Lightning%20Detection%20Network.pdf. (Accessed: 12 May 2022).
[18] M. Gijben, The lightning climatology of South Africa, S. Afr. J. Sci. 108 (3) (2012) 1–10.
[19] H. Hunt, Y. Liu, K. Nixon, Evaluation of the South African lightning detection network using photographed tall tower lightning events from 2009–2013, in: 2014 International Conference on Lightning Protection, ICLP, IEEE, 2014, pp. 1746–1751.
[20] H. Fensham, C. Schumann, H. Hunt, J. Tasman, K. Nixon, T. Warner, M. Gijben, Performance evaluation of the SALDN using highspeed camera footage of ground truth lightning events over Johannesburg, South Africa, in: 2018 34th International Conference on Lightning Protection, ICLP, IEEE, 2018, pp. 1–5.
[21] H. Fensham, H.G. Hunt, C. Schumann, T.A. Warner, M. Gijben, The Johannesburg lightning research laboratory, part 3: Evaluation of the South African lightning detection network, Electr. Power Syst. Res. 216 (2023) 108968, URL https://www.sciencedirect.com/science/article/pii/S0378779622010173.
[22] D.R. Poelman, W. Schulz, S. Pedeboy, L.Z. Campos, M. Matsui, D. Hill, M. Saba, H. Hunt, Global ground strike point characteristics in negative downward lightning flashes–Part 2: Algorithm validation, Nat. Hazards Earth Syst. Sci. 21 (6) (2021) 1921–1933.
[23] N. Tanase, T. Manape, H.G. Hunt, Country-wide evaluation of the South African lightning detection network location accuracy, in: 2023 31st Southern African Universities Power Engineering Conference, SAUPEC, 2023, pp. 1–6.
[24] S. Pédeboy, Introduction to the IEC 62858: Lightning density based on Lightning Locating Systems, in: Proceedings of the ILPS, 2018.
[25] G. Diendorfer, Some comments on the achievable accuracy of local ground flash density values, in: International Lightning Detection Conference, sn, 2008.
[26] C. Bouquegneau, The need for an international standard on lightning location systems, in: 23rd International Lightning Detection Conference, 2014, pp. 18–19.
[27] S. Pédeboy, Identification of the multiple ground contacts flashes with lightning location systems, in: International Lightning Detection Conference, sn, 2012.
[28] L. Campos, On the mechanisms that lead to multiple ground contacts in lightning (Ph.D. thesis, Doctorate thesis), Graduate Course in Space Geophysics, Instituto . . . , 2016.
[29] M. Matsui, K. Michishita, S. Yokoyama, Characteristics of negative flashes with multiple ground strike points located by the Japanese lightning detection network, IEEE Trans. Electromagn. Compat. 61 (3) (2019) 751–758.
[30] D.R. Poelman, H. Kohlmann, W. Schulz, S. Pedeboy, L. Schwalt, Ground strike point properties derived from observations of the European lightning location system EUCLID, in: 2023 12th Asia-Pacific International Conference on Lightning, APL, 2023, pp. 1–5.
[31] N. Watanabe, Characterization and Modeling of Upward Lightning at the Gaisberg Tower (Ph.D. thesis), 2021.