

Lightning fatality risk map of the contiguous United States

William P. Roeder¹ · Benjamin H. Cummins² · Kenneth L. Cummins³ ·
Ronald L. Holle⁴ · Walker S. Ashley⁵

Received: 10 May 2015 / Accepted: 27 July 2015 / Published online: 1 August 2015
© Springer Science+Business Media Dordrecht 2015

Abstract A new method to calculate lightning fatality risk is presented in order to develop a way to identify the lightning risk in areas where lightning fatality data are not available. This new method uses GIS software to multiply lightning flash density and population density on a grid and display the results on a map. A comparison to the known lightning fatality data was done to verify the method. The method works well with a quadratic regression correlation coefficient as high as 0.864, although a hybrid quadratic/log-linear regression is preferred for various reasons despite having a slightly lower correlation coefficient (0.827). Given the good performance, the lightning fatality risk method may be useful for developing countries, where lightning fatality reports may not be reliable, to help guide where to allocate scarce resources for lightning safety initiatives.

Keywords Lightning · Lightning fatalities · Geographic Information System · GIS · Lightning safety · Lightning safety education

✉ Ronald L. Holle
rholle@earthlink.net

William P. Roeder
wroeder@cfl.rr.com

Benjamin H. Cummins
bhcummins@gmail.com

Kenneth L. Cummins
cummins@atmo.arizona.edu

Walker S. Ashley
washley@niu.edu

¹ Rockledge, FL, USA

² Resource Systems Group, Inc., Burlington, VT, USA

³ Department of Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

⁴ Holle Meteorology and Photography, Oro Valley, AZ, USA

⁵ Meteorology Program, Department of Geography, Northern Illinois University, Dekalb, IL, USA

1 Introduction

Lightning is the third leading source of storm deaths in the USA, killing an average of 51 people per year from 1984 to 2013 (NWS 2015). Lightning had been the second leading of storm deaths in the USA for many years, but tornadoes recently took second place (Roeder 2012). Lightning is also a significant source of storm deaths worldwide with an estimated average number of fatalities of up to 24,000 per year (Cardoso et al. 2011; Holle 2008; Holle and López 2003).

The geographical distribution of lightning fatalities in the USA is well known. The distribution by state has been extensively studied (Curran et al. 2000; Holle 2009, 2012; Roeder and Jensenius 2012) as well on a 60×60 km resolution map (Ashley and Gilson 2009). Such information is very useful in lightning safety education since tuning lightning safety to the local population is important (Roeder et al. 2011, 2012). Unfortunately, the geographical distribution of lightning fatalities in developing countries is not as well known, mostly due to fatalities in remote locations not being reported, identifying lightning as the cause after the fact, and lower-quality record keeping. Therefore, the purpose of this paper is the development of a new method to estimate the spatial lightning fatality distribution that may be useful in guiding lightning safety initiatives in developing countries. This method uses GIS software to multiply annual cloud-to-ground (CG) lightning flash density by population density and display the results on a map for easy visualization. This multiplication is assumed to be a first approximation for the risk of lightning fatalities. A lightning fatality risk map was created and verified for the contiguous United States (CONUS). This work was inspired by Geographical Information System (GIS) applications of lightning data by Gijben (2012). Other factors impacting the risk of lightning fatality are not included and are listed in Table 1. The availability of locations safe from lightning should also be considered, such as the quantification of urbanization by Gomes and Ab Kadir (2011). In addition, these results are annual averages, and the diurnal and seasonal distributions of lightning are not considered.

2 Lightning fatality risk map

The lightning fatality risk map for the CONUS was created by multiplying the annual CG lightning flash density (flashes/km² year) by the population density (people/km²). A representative annual lightning flash density for the CONUS is shown in Fig. 1, and the population density is shown in Fig. 2. The CG lightning flashes are from the National Lightning Detection NetworkTM (NLDN) which reports more than 90 % of all CG lightning flashes with a median location accuracy of better than 500 m (Cummins et al. 2006;

Table 1 Factors not considered in the lightning fatality risk map

Relative amount of time the local populace spend outdoors at risk from lightning
Amount of time spent in risky behaviors
Quick access to locations safe from lightning
Amount of time spent far from safety
Quantity and quality of lightning safety education
Changes in population density during the year such as from tourism or migration

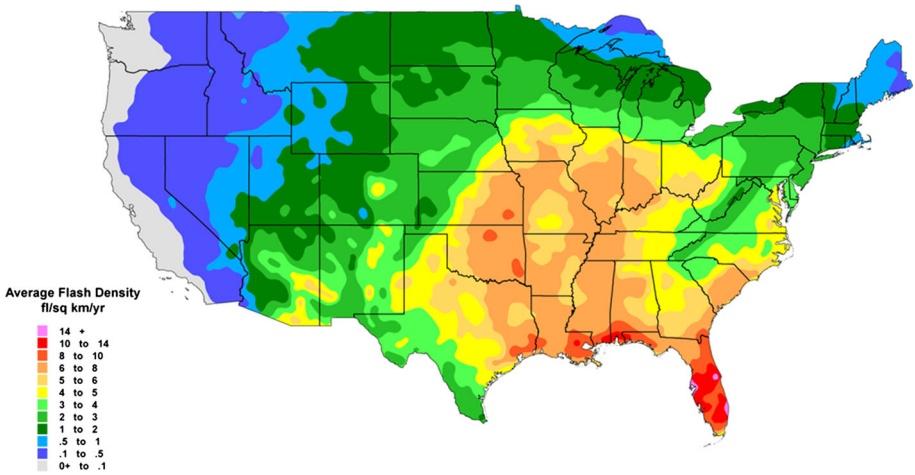


Fig. 1 CG lightning flash density (1997–2010) for the USA from the National Lightning Detection Network (Cummins et al. 2006; Cummins and Murphy 2009). The NLDN is owned and operated by Vaisala, Inc

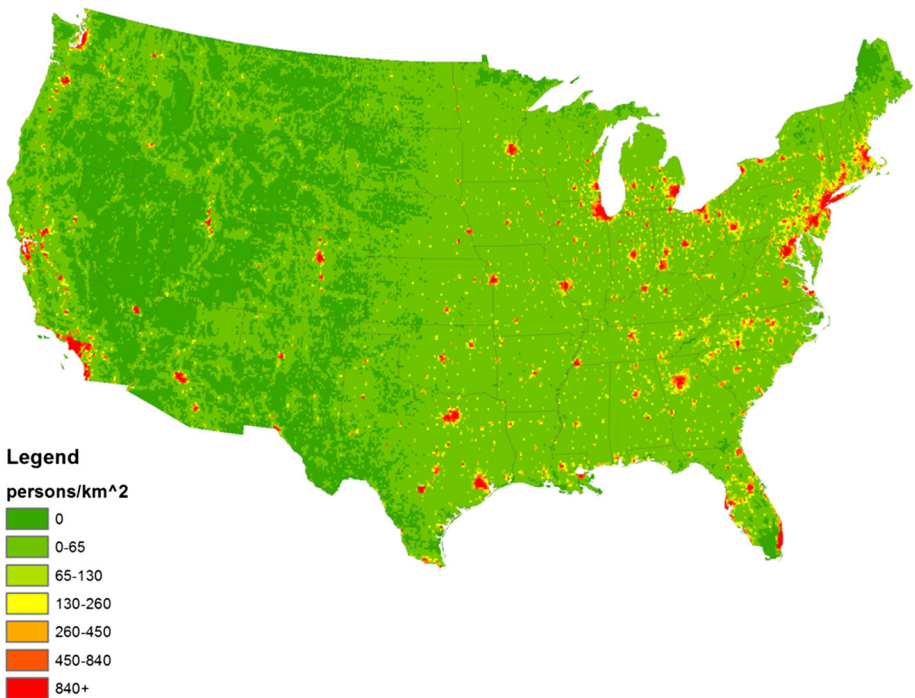


Fig. 2 2000 Population density for the continental USA from the 2013 US Census (www.census.gov)

Cummins and Murphy 2009) for the time period employed here. In this study, the lightning flashes from January 2003 through December 2012 were initially accumulated into 0.1 by 0.1 degree regions normalized by the number of observation years. For some analyses,

these data were accumulated into larger grids and spatially smoothed using a 3×3 Gaussian filter to filter out noise and more easily see overall patterns. The population density is from the 2000 National Historical Geographic Information System (<http://www.nhgis.org>), accumulated into the same 0.1×0.1 degree grids that were then employed for the flash density analysis. The resulting lightning fatality risk map is shown in Fig. 3, and some technical details are in Table 2. The results are normalized by grid area, which decreases with increasing latitude, to allow direct comparison of the gridded polygons. For comparison purposes, the observed lightning fatality locations (1959–2006) are included in Fig. 3 as dots.

The concept of the lightning fatality risk map seems reasonable: Lightning fatalities should be directly related to CG lightning flash density and population density. Others have made the same assumption but did so only for an entire country or state (Gomes and Ab Kadir 2011), not on a high-resolution grid. Another study identified the factors important for lightning fatalities in a region: number of flashes, number of people, and area (Gao et al. 2014). These factors can be combined into lightning flash density and population density, equivalent to the assumption in this study.

3 Verification

Although reasonable, the concept of the lightning fatality risk map is new and verification is required. Fortunately, a dataset of observed lightning fatalities is available (Ashley and Gilson 2009). Some details of this observed lightning fatality map are in Table 3 and

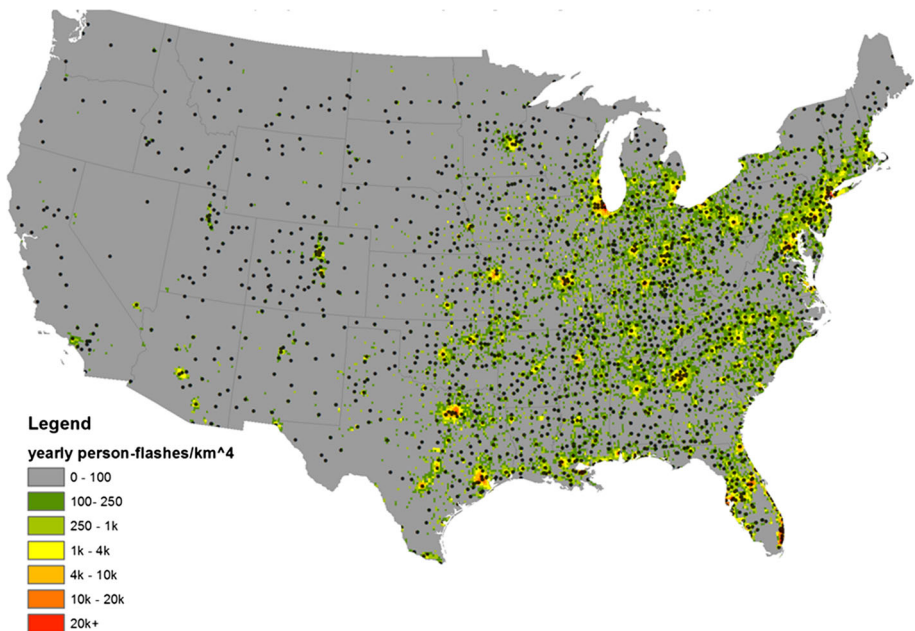


Fig. 3 Lightning fatality risk for the CONUS. Lightning fatality risk is the product of CG mean lightning flash density (2003–2012) and population density (2000). Details of the map are in Table 2. The *black dots* are the individual lightning fatalities (1959–2006), which are included for visualizing spatial correlation

Table 2 Technical details of the new CONUS calculated lightning fatality risk map (Fig. 3)

GIS software	ArcMap v10.0
Grid spacing	0.1° × 0.1° lat/lon (~ 10 × 8 km)
Population data	US census (2000) ^a
Lightning flash data	NLDN (2003–2012) (Cummins et al. 2006; Cummins and Murphy 2009)
Map projection	Albers equal-area conic projection
Smoothing	None
Lightning fatalities with no locations (1959–2006)	291 (<i>not plotted</i>) (6.3 % of total)
Lightning fatalities with inexact locations	714 (15.5 % of total)

^a Socioeconomic Data and Applications Center, Columbia University (<http://sedac.ciesin.org>)

Table 3 Technical details of the previous CONUS observed lightning fatality map (Fig. 4; Ashley and Gilson 2009) used as part of the verification of the calculated lightning fatality risk map

GIS software	ArcGIS 9.3
Grid spacing	60 × 60 km
Period of record	1959–2006
Number of fatalities	4408 (290 not plotted due to no location provided)
Map projection	Albers equal-area conic projection
Smoothing	3 × 3 low-pass Gaussian filter
Lightning fatalities with inexact locations	714 (15.5 % of total)
Lightning fatalities with no locations	290 (<i>not plotted</i>) (6.6 % of total)

shown in Fig. 4. The period of the lightning fatality data is older and longer than the lightning flash density. This was done to provide a sufficient sample size of lightning fatalities, and the gridded lightning fatality database was the only one available. However, while differences would result, especially from population shifts during this period, the general pattern should be representative. It should be noted that 6.6 % of the lightning fatalities had no location and were not plotted. In addition, 15.5 % had uncertain locations, e.g., being recorded at the county seat even though the fatality may have occurred anywhere in the county, and introduce a small amount of variability into the map that affects verification of the lightning risk map.

The verification included both subjective and objective components. The subjective verification was a visual comparison of the lightning fatality risk maps with the known lightning flash density and population density across the USA. The subjective verification showed the lightning fatality risk maps behaved as expected, including discussion of six state and regional maps for parts of the USA (see Roeder et al. 2014 for details).

Subjective analysis of the Florida map shown here (see Fig. 5) demonstrates how one might use lightning fatality risk maps and how small-scale features can be revealed using a map with fine grid resolution. Florida has the highest lightning fatality rate in the USA, the highest lightning flash density, and some of the sharpest gradients of population density. The highest flash rates in the USA are across central Florida (Fig. 1). Population centers in the Florida lightning risk map are Tampa/St. Petersburg and Orlando, both with more than 10 M people flashes/year km⁴. Similarly the high population density of the Miami area is

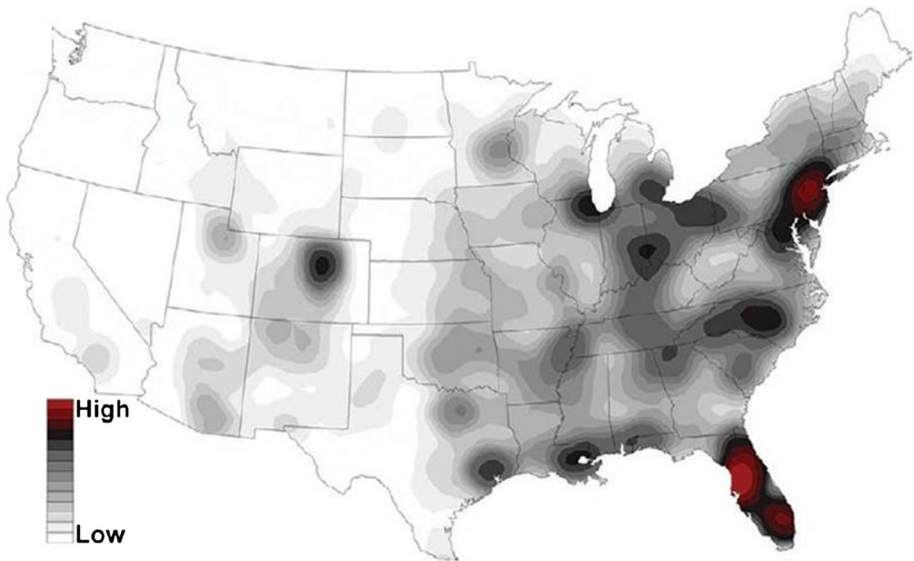


Fig. 4 Number of observed lightning fatalities in the USA (1959–2006) smoothed on a 60×60 km grid (Ashley and Gilson 2009). This is the ground truth for verification of the new lightning fatality risk maps presented here. Details of map in Table 4

evident, even though that area has less lightning activity than central Florida. The city of Jacksonville is also apparent, although it lies in an area of relatively lower lightning activity. Port Charlotte in southwest Florida can be seen as a region with more than 400 K annual people flashes/year km^4 . All of these areas show one or more spatially proximate fatalities. The areas of low lightning fatality risk in Florida are also apparent. There is a rapid decrease in population density southeast of Orlando due to rural areas and swamps. Although the lightning flash rate remains high, the drop in lightning fatality risk due to the much lower population density is shown in Fig. 5. In addition, the extremely rapid decrease in population west of Miami/Ft. Lauderdale is also indicated.

The objective verification is a grid-by-grid comparison of the calculated lightning fatality risks with the observed annual lightning fatality density across the CONUS. As discussed previously, fatalities are from Ashley and Gilson (2009). Although the periods of record for the datasets differ, we assume the general conclusions are still valid (2003–2012 for the lightning fatality risk, 1959–2006 for the lightning fatalities, 2010 for the population). Since the map of lightning fatalities in Fig. 4 had extensive smoothing, this objective verification was done using Ashley and Gilson's original data set of observed lightning fatalities. The lightning fatality risk (population density \times CG annual lightning flash density) and annual lightning fatality were placed on the same grid spacing and the same smoothing applied. A Gaussian smoothing function was used with the scale factor based on various numbers of grid spaces. Eight combinations were analyzed: two grid spacings with four smoothing functions each (Table 4). Various regression analyses were done to determine the behavior between the calculated lightning fatality risk and the observed lightning fatalities. The correlation coefficients for the best-fit linear regression and best-fit quadratic regression are shown for each of the eight cases. A linear regression of observed annual lightning fatality on calculated lightning fatality risk was then

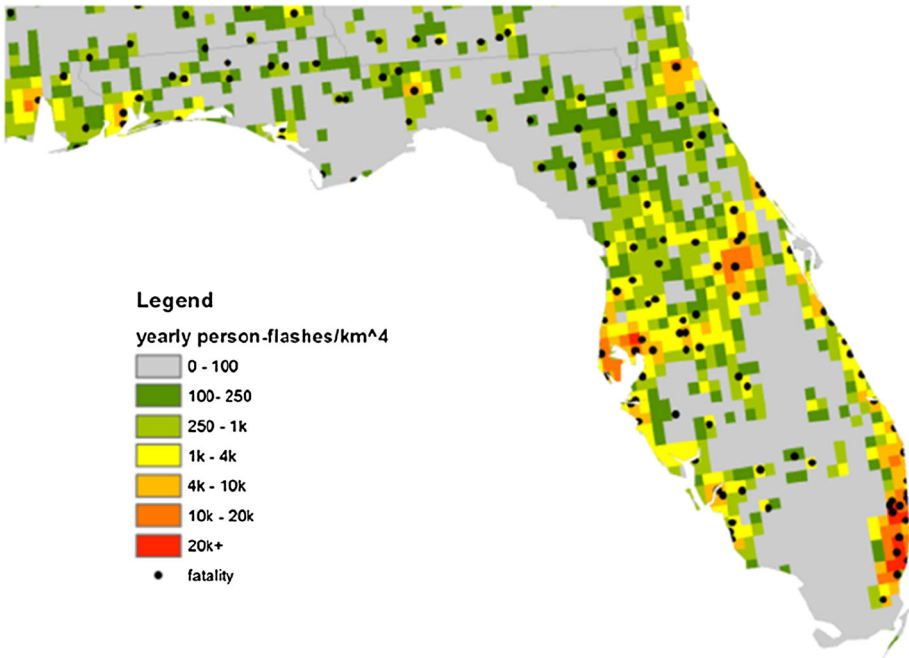


Fig. 5 Lightning fatality risk map for Florida. The *black dots* are the observed lightning fatalities

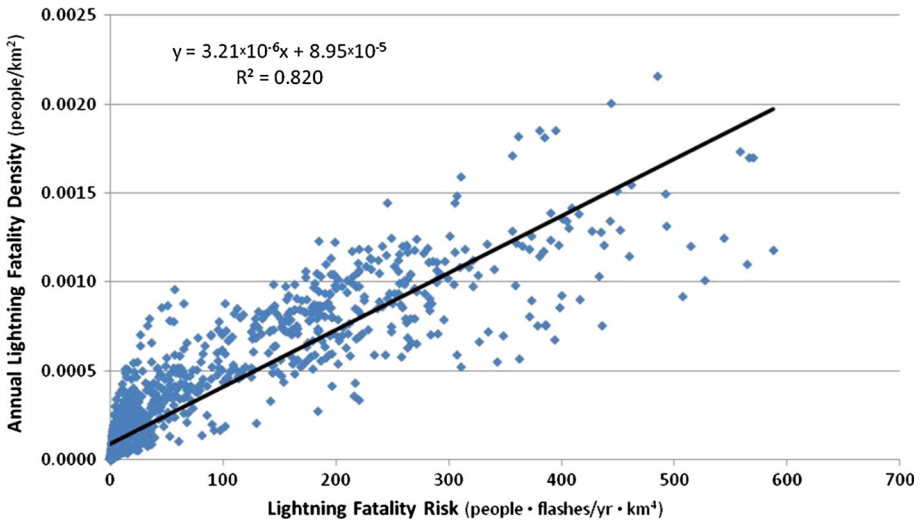


Fig. 6 Lightning fatality density versus lightning fatality risk. The linear regression is shown for the best linear regression (1.0° grid resolution, 1.5 grid smoothing)

performed on each of the eight grids (Table 4). The best linear regression was the 1.0° latitude/longitude grid with the Gaussian smoothing of 1.5 grid spaces. The equation for this regression is at (1) and is shown graphically in Fig. 6.

Table 4 Correlation coefficients (r^2) from the linear regressions and quadratic regressions of observed lightning fatality versus calculated lightning fatality risk for eight combinations of grid spacing and smoothing functions

Grid spacing (°lat/lon)	0.5°				1.0°			
	None	0.5	1.0	1.5	None	0.5	1.0	1.5
Grid smoothing (no. of grid spacings)								
Linear regression	0.515	0.617	0.738	0.767	0.644	0.716	0.807	0.820
Quadratic regression	0.564	0.667	0.794	0.815	0.710	0.775	0.8423	0.864

Highest r^2 in both cases were for the larger grid and largest smoothing (1° and 1.5° grid spacing, respectively) and shown in bold above

$$\begin{aligned}
 y &= 3.21 \times 10^{-6}x - 8.95 \times 10^{-5} \\
 r^2 &= 0.820
 \end{aligned}
 \tag{1}$$

where x = lightning fatality risk (people flashes/year km⁴), y = annual lightning fatality density (fatalities/km²).

Linear regression through the origin was considered but not used. While one might assume that zero lightning flash density or zero population density would lead to zero lightning fatalities, that assumption does not consider people traveling to outdoor areas for recreation, which would not be assessed in the population density since people are counted where they live. Since the assumption of intersection at the origin cannot be made a priori, regression through the origin is not justified.

The linear regression with the best correlation coefficient (grid spacing of 0.5°, 1.5 grid Gaussian smoothing; Fig. 6) appears to have a systematic bias. At lower lightning fatality risk, the fatalities appear to trend toward being above the linear regression. At higher risk, the fatalities appear to occur more frequently below the regression and with larger differences from the regression than the points above the regression. This suggests that a nonlinear regression may give better results, perhaps a best-fit log-linear or quadratic polynomial.

A log-linear regression was tried since it would have the desired properties to match the plotted data: approaches zero fatalities toward zero risk, monotonically increasing fatalities and higher risk, asymptotically approaches an upper limit. Unfortunately, the log-linear regression was dominated by the large number of lower risk values and so yielded a r^2 of only 0.650, significantly worse than the linear regression ($r^2 = 0.820$). Of course, risk values of zero had to be excluded to allow the log-linear regression, reducing the number of data pairs to 1204.

A quadratic regression yielded an r^2 of 0.864, slightly better than the linear regression (r^2 of 0.820) for this grid spacing and smoothing. More importantly, the quadratic regression did not have the systematic bias of the linear regression. This quadratic regression equation is at (2) and shown graphically at Fig. 7.

$$\begin{aligned}
 y &= 5 \times 10^{-9}x^2 + 5 \times 10^{-6}x + 6 \times 10^{-5} \\
 r^2 &= 0.864
 \end{aligned}
 \tag{2}$$

where x and y are as in (1).

However, care must be taken in using the quadratic regression to higher values of lightning fatality risk. At risk values higher than about 475 people flashes/year km⁴, the predicted lightning fatalities will decrease at higher risk, which is contrary to expectation.

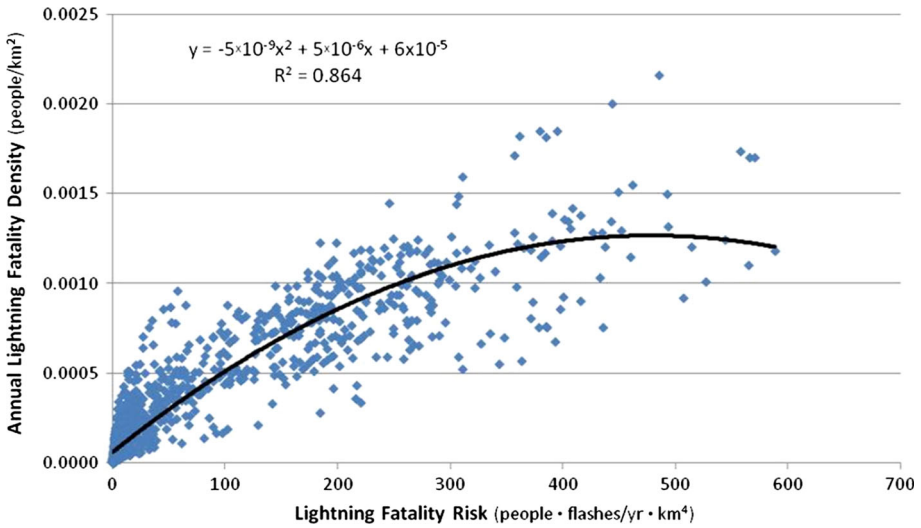


Fig. 7 Lightning fatality density versus lightning fatality risk. The quadratic regression is for the 1.0° grid resolution, 1.5 grid smoothing

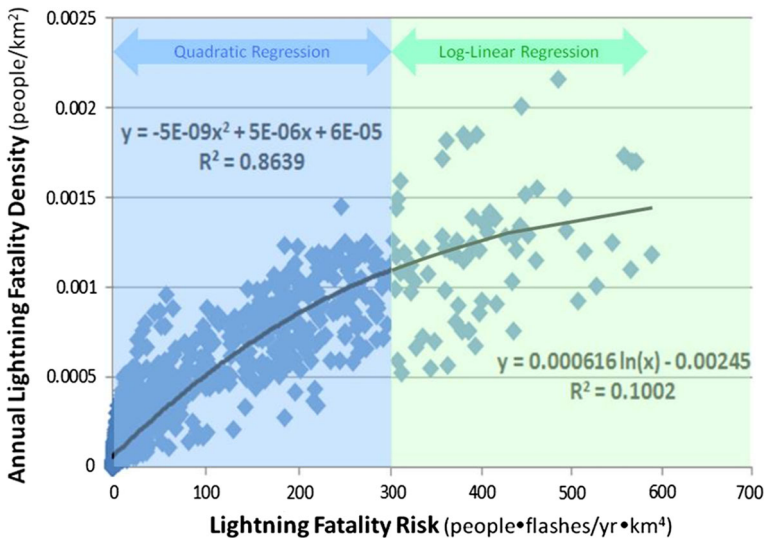


Fig. 8 Lightning fatalities versus lightning fatality risk. The hybrid quadratic/log-linear regression is shown (1.0° grid resolution, 1.5 grid smoothing). The correlation coefficient, weighted by number of data point, is 0.827

It should be noted that care must also be taken with the linear regression since it tends to underestimate the lightning fatalities at lower risk and significantly overestimate the fatalities at higher risks.

A hybrid quadratic/log-linear regression was carried out in order to overcome the shortfall of the quadratic regression noted above. This hybrid uses a quadratic regression at

lower lightning fatality risk and a log-linear regression at higher risk values. The risk threshold for changing the regressions was manually selected to avoid discontinuities in value and slope between the two regressions. A lightning fatality risk threshold of 300 people flashes/year km⁴ was selected based on visual examination of the graphs. The log-linear regression was adjusted slightly to completely avoid discontinuities. The combined r^2 for both regressions, weighted by the number of data pairs in each regression, was 0.827. This weighted combined correlation coefficient was done to allow traditional interpretation, i.e., 1.0 is perfect correlation. The equation of the hybrid regression is provided below and shown graphically in Fig. 8.

$$\begin{aligned} \text{If } x \leq 300 \text{ people flashes/year km}^4 : \\ y = -5 \times 10^{-9}x^2 + 5 \times 10^{-6}x + 6 \times 10^{-5} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{If } x > 300 \text{ people flashes/year km}^4 : \\ y = 6.16 \times 10^{-4} \ln(x) - 2.45 \times 10^{-3} \\ \text{combined } r^2 = 0.827 \end{aligned}$$

where x and y are as in (1).

The hybrid quadratic/log-linear regression has a slightly higher correlation coefficient than the best linear regression, but a slightly lower correlation coefficient than the quadratic regression. It overcomes the systematic bias of the linear regression and avoids the undesired behavior of the quadratic regression of decreasing fatalities at higher risk. Finally, it asymptotically approaches an upper limit as is consistent with visual inspection of the data. Therefore, this hybrid quadratic/log-linear regression is recommended over the other regressions, even though it does not have the highest correlation coefficient.

4 Future work

The lightning fatality risk map presented here is a preliminary attempt to establish and verify the new method. Since it verifies well in the USA, the most important future work is to extend the method to other countries. This would first be done preferably where the pattern of lightning fatalities is already well known. It would also be useful if those countries had a wide range of CG flash density and a wide range of population density. These additional verifications would be beneficial in justifying the application of this method in developing countries where lightning fatalities may not be reported well. This is especially important since the main motivation for this work was to help guide lightning safety efforts in developing countries. Once some further verification of the method is made, it will be straightforward to create a lightning fatality risk map for developing countries. A global lightning fatality risk map could also be developed that can be used to inform lightning safety efforts anywhere in the world.

The lightning fatality risk map is for annual lightning risk. It may be useful to apply the same method to monthly or seasonal maps. Likewise, studying diurnal patterns of lightning fatality risk may be useful. The hybrid quadratic/log-linear regression had some subjectivity in its development. A more objective method might be pursued that could be more easily applied in contexts where verification data are not readily available.

The CG lightning flash density was used in constructing the lightning fatality risk map; however, the density of ground contact points would be more appropriate. This is not the

same as the stroke density, since in flashes with multiple strokes, the subsequent strokes often strike the same point and represent little additional risk. However, the subsequent strokes also often contact the ground elsewhere (Valine and Krider 2002), often a few km away, and so represent substantial additional risk. Unfortunately, the number of ground strike points is not reported by most lightning detection systems. However, it can be inferred from stroke detection systems, as demonstrated by Cummins (2012). This was not feasible for a study on this spatial geographical scale, but may be important for a more geographically focused study.

Another important factor in lightning fatalities is the behavior of the local population. People that spend more time outside, especially during lightning activity, or cannot or will not seek safety when lightning threatens have a greater likelihood of becoming a lightning fatality. If the data were available or inferable, variations in behavior could be included as another multiplicative factor in the construction of lightning fatality maps, perhaps as a percent of time spent at risk. However, in areas without lightning-safe locations such as some parts of the developing world, the variations in behavior would not be important.

Finally, the lightning fatality risk map presented here assumes that the population density is constantly at the reported grid point. However, there are areas with significant population change throughout the year, such as those related to tourism. In addition, local populations may move out of the immediate area during lightning season. In some developing countries, migration may also be an important factor. Future research should engage social scientists to develop sources and databases that may assist with such research.

5 Summary

A new method to estimate the risk of lightning fatality was developed. This method uses a GIS to multiply annual CG lightning flash density and population density to estimate the risk of lightning fatality and visualize the results on a map. This method was applied to the contiguous US and verified against the observed lightning fatalities. The method verifies well with the preferred hybrid quadratic/log-linear regression model having a correlation coefficient as high as $r^2 = 0.827$ for the 1.0° lat/lon grid with 1.5 grid point Gaussian smoothing. The hybrid quadratic/log-linear regression is preferred since it overcomes many of the problems of the other regressions even though one of those other regressions had a slightly higher correlation coefficient. Further refinements to the lightning fatality risk maps are possible and discussed.

The main motivation for developing the lightning fatality risk method is to assist lightning safety efforts in developing countries. Since the method works well for the USA, it may be useful in some developing countries where the geographical distribution of actual lightning fatalities may not be well documented. Given that the distribution of CG lightning can be reasonably well determined from the various global lightning detection networks, or other sources if available, and if the distribution of population density is also known, then a GIS can be used to create lightning fatality risk maps for those countries. These maps could then be used to guide lightning safety efforts in those countries to be more cost-efficient and more effective by spending funds on areas where it is most needed and by tailoring the efforts to the people living in that area.

Acknowledgments The authors thank Vaisala, Inc. for use of the NLDN data and for Fig. 1. We also thank two anonymous reviewers who contributed to the improvement of this manuscript.

References

- Ashley WS, Gilson CW (2009) A reassessment of US lightning mortality. *Bull Am Meteorol Soc* 90:1501–1518
- Cardoso I, Pinto O Jr, Pinto IRCA, Holle R (2011) A new approach to estimate the annual number of global lightning fatalities. 14th international conference on atmospheric electricity, Rio de Janeiro, Brazil
- Cummins KL (2012) On the relationship between terrain variations and LLS-derived lightning parameters. 31st international conference on lightning protection, Vienna, Austria
- Cummins KL, Murphy M (2009) An overview of lightning locating systems: history, techniques, and data uses, with an in-depth look at the US NLDN. *IEEE Trans Electromagn Compat* 51:499–518
- Cummins KL, Cramer JA, Biagi CJ, Krider EP, Jerauld J, Uman MA, Rakov VA (2006) The US National Lightning Detection Network: post upgrade status. 2nd conference on meteorological applications lightning data, Atlanta, GA, American Meteorological Society
- Curran EB, Holle RL, López RE (2000) Lightning casualties and damages in the United States from 1959 to 1994. *J Clim* 13:3449–3453
- Gao Y, Zhang Y, Zhang W, Zhang Y (2014) Characteristics of lightning casualties and vulnerability evaluation regionalization in China. 5th international lightning meteorology conference, Tucson, AZ, Vaisala
- Gijben M (2012) Lightning climatology of South Africa with a special focus on lightning risk maps. 4th international lightning meteorology conference, Broomfield, CO, Vaisala
- Gomes C, Ab Kadir MZA (2011) A theoretical approach to estimate the annual lightning hazards on human beings. *Atmos Res* 101:719–725
- Holle RL (2008) Annual rates of lightning fatalities by country. 2nd international lightning meteorology conference, Tucson, AZ, Vaisala
- Holle RL (2009) Lightning fatalities and fatality rates by US state. International conference on lightning static electricity, Pittsfield, MA, Paper GME-3
- Holle RL (2012) Recent studies of lightning safety and demographics. 4th international lightning meteorology conference, Broomfield, CO, Vaisala
- Holle RL, López RE (2003) A comparison of current lightning death rates in the US with other locations and times. International conference on lightning static electricity, Paris, France, paper 103-34 KMS
- NWS (2015) Natural Hazards Statistics, National Weather Service Office of Climate, Water, and Weather Services. <http://www.nws.noaa.gov/om/hazstats.shtml>. Update 10 June 2015
- Roeder WP (2012) A statistical model for the inter-annual and intra-annual fatalities from lightning in the US and comparison to other storm phenomena. 4th international lightning meteorology conference, Broomfield, CO, Vaisala
- Roeder WP, Jensenius J (2012) A new high-quality lightning fatality database for lightning safety education. 4th international lightning meteorology conference, Broomfield, CO, Vaisala
- Roeder WP, Holle RL, Cooper MA, Hodanish S (2011) Communicating lightning safety effectively. 5th conference on meteorological applications lightning data, Seattle, WA, American Meteorological Society
- Roeder WP, Holle RL, Cooper MA, Hodanish S (2012) Lessons learned in communicating lightning safety effectively. 4th international lightning meteorology conference, Broomfield, CO, Vaisala
- Roeder WP, Cummins BH, Ashley WS, Holle RL, Cummins KL (2014) Quantifying lightning fatality risk. 32nd international conference on lightning protection, Shanghai, China
- US Census (2013) www.census.gov
- Valine WC, Krider EP (2002) Statistics and characteristics of cloud-to-ground lightning with multiple ground contacts. *J Geophys Res* 107(D20):4441. doi:10.1029/2001JD001360