2014 23rd International Lightning Detection Conference
18 - 19 March • Tucson, Arizona, USA
5th International Lightning Meteorology Conference
20 - 21 March • Tucson, Arizona, USA

An Update to Predicting the Number of U.S. Annual Lightning Fatalities

William P. Roeder
Private Meteorologist
Rockledge, FL, U.S.A

***Abstract*—An update to the statistical analysis of annual rate of lightning fatalities in the U.S. is presented. This provides an updated method to estimate the total number of expected lightning fatalities in the U.S. for any year in the near future and the expected number of fatalities on any day during any year. These estimates account for the declining rate of U.S. lightning fatalities over the past several decades. The uncertainty in these estimates is also provided for use as error bars and to allow hypothesis testing of the statistical significance of differences between any values of interest and the estimated number lightning fatalities. For example, one can estimate the probability that the observed lightning fatalities by any date in a year indicates that year is significantly above, below, or within expectations. The updated model predicts a median of 31.2 lightning fatalities in the U.S. for 2014 with a 95% confidence interval of 16.3 to 53.1 fatalities. The apparent sudden decrease in U.S. lightning fatalities beginning in 2008 is also discussed.**

Keywords—U.S. lightning fatality rate; lightning deaths

I. BACKGROUND

The number of lightning deaths in the U.S. has been declining for many decades (Roeder, 2013; Roeder, 2012; Holle, 2012; Ashley and Gilson, 2009; Holle et al., 2005; Lopez and Holle, 1998). As a result, the traditional running 30-year mean used by National Weather Service for sources of weather fatalities overestimates the current rate. For example, using the new updated method that takes into account the declining rate over the past decades, the expected mean U.S. lightning fatality rate for 2014 is 24.8 deaths (median is 31.2 deaths). The details of how these updated estimates were developed are shown later. For comparison, the running 30-year mean (1984-2012) of U.S. lightning fatalities is considerably higher, 53.3 deaths per year (National Weather Service, 2013). While the running 10-year mean (2004-2013) is better, 33.0 deaths per year (National Weather Service, 2013), it still overestimates the correct rate. In addition, such a short period average can be skewed by a single extreme event.

Corresponding Author: William P. Roeder,
e-mail: wroeder@cfl.rr.com

II. UPDATES TO 2012 STUDY

This paper updates the 2012 study of the inter-annual and intra-annual lightning fatalities in the U.S. (Roeder, 2013; Roeder, 2012). One of these papers (Roeder, 2012) is at Appendix-1 to provide background information and to detail the previous analysis for the convenience of the reader, and so the author may concentrate on the new information in this paper.

A. Updates to Inter-Annual Fatalities

The updates to the inter-annual study are relatively minor. These updates added the two latest years of observed U.S. lightning fatalities (2012-2013) and restricted the exponential best-fit curve to the last 30 years (1984-2013) to make the results more representative of current U.S. annual lightning fatality rates. The annual lightning fatalities in the U.S. are listed in Table-1 and shown graphically in Figure-1 along with a best-fit negative exponential curve.

Table-1.
Number of annual lightning deaths in the U.S. from 1900-2010. Data from 1984-1991 are from Lopez and Holle (1998) and data from 1992-2013 are from National Weather Service (2014).

Year	Deaths	Year	Deaths	Year	Deaths
1984	91	1994	69	2004	31
1985	85	1995	85	2005	38
1986	78	1996	53	2006	48
1987	99	1997	42	2007	45
1988	82	1998	44	2008	28
1989	75	1999	46	2009	34
1990	74	2000	51	2010	29
1991	73	2001	44	2011	26
1992	41	2002	51	2012	28
1993	43	2003	43	2013	23

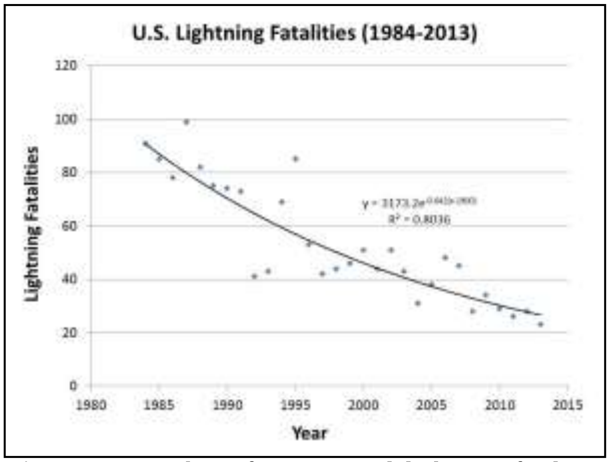

Figure-1. Number of U.S, annual lightning fatalities for 1984-2013 and best-fit negative exponential curve.

The best-fit exponential curve allows estimating the expected number of lightning fatalities for any year in the near future. The previous study provided a best-fit curve from 1941 onward, since that was when the onset of exponential decrease of lightning fatalities appeared to begin in the U.S. (Roeder, 2012 (Appendix-1)). The new best-fit negative exponential curve for the expected mean number of U.S. annual lightning fatalities covers 1984-2013 and is in (1).

$$y = 3173.2e^{-0.042(x-1900)} \quad (1)$$

$$r^2 = 0.80$$

where y = <u>mean</u> annual lightning fatalities, and x = year.

The median is preferred over the mean since the percentile distribution of this curve is asymmetric. The best-fit negative exponential curve for the median number of U.S. lightning fatalities (1994-2003) is in (2). The reason for the shorter period is discussed later.

$$y = 538.69e^{-0.025(x-1900)} \quad (2)$$

$$r^2 = 0.93$$

where y = <u>median</u> annual lightning fatalities, and x = year

The expected median number of lightning fatalities in the U.S. for 2014 is 31.2 deaths, slightly more than the expected mean (26.4 deaths). The subtraction of 1900 years in the exponents of both equations (1) and (2) is done merely for the convenience of keeping the first coefficients (in front of the exponential function) from being excessively large, e.g. 538.69 vs. 2×10^{23}.

Although linear regression fit the median almost as well ($r^2 = 0.92$), a negative exponential regression was chosen since the preceding years clearly had a negative exponential distribution (Roeder, 2012 (Appendix-1)) and so represents reality better. It is only in the later period when the lightning fatalities are in the tail of the curve and only appear to be quasi-linear. Also, the linear regression has the disadvantage of decreasing too much when extrapolated into the future. The best fit negative exponential curve will not have this disadvantage since the flattening tail of the distribution will be represented. For example, for 2025 the negative exponential curve predicts a median of 23.7 fatalities while the linear regression predicts 15.6 fatalities, which would be less than the expected 25th percentile and obviously unrepresentative of the intended median.

The final update to the inter-annual analysis added the 2.5th and 97.5th percentiles, in addition to the previous percentiles, so that the entire set is: 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97.5th percentiles. The two new percentiles allow two-tail hypothesis testing at the 95% confidence level. As in the 2012 study, these percentiles were estimated from calculating each percentile for the annual lightning fatalities for each year (1994-2003) ±10 years (Figure-2). The period ends at 2003 to allow for the +10 year part of the estimate since the available data ends in 2013. Next, a best-fit negative exponential curve was calculated for each percentile (Table-2). Normally, more than 20 data points would be desired for a best-fit curve and including earlier years was considered. However, in this case, the choice was made to use the smaller sample of the 1994-2003 period, rather than 1984-2003, to avoid some internal inconsistencies in the results due to some discontinuities in the data in 1993 and earlier. For example, using the 1984-2003 results led to the 2.5th percentile being larger than the 5th percentile in some future years. Using 1994-2003 also made the results more representative of the current lightning fatality rate. This may be important since there is evidence that the U.S. lightning fatality rate changed beginning in 2008. A graph of the median, and 2.5th and 97.5th percentiles (95% confidence interval) and the best-fit curves projected to 2013 is shown in Figure-3.

Due to the variability in the data and the closeness of some of the percentiles, some of the equations can produce internally inconsistent results, e.g. the predicted 2.5th percentile can be slightly larger than the 5th percentile for some years. In practical application, if one predicted percentile is inconsistent with another, consider using the more conservative of the two.

As an example, the prediction for 2014 is a median number of U.S. lightning fatalities of 31.2 with a 95% confidence interval of 16.3 to 53.1 fatalities. These estimates are calculated from the equations in Table-2. Note that for 2013, the last year for which data are available, the model would have predicted a median of 31.9 fatalities with a 95% confidence interval of 17.0 to 54.6 fatalities. This compares very well with the observed number of 23 lightning fatalities.

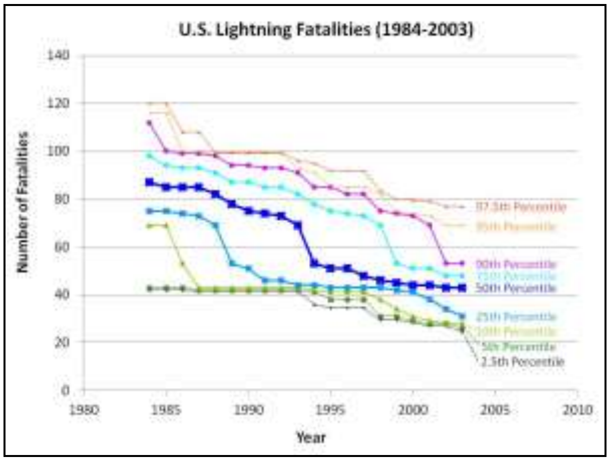

Figure-2 Running 21-year percentiles (center ± 10 years) of U.S. lightning fatalities for 1984-2003.

Table-2. Equations for various percentiles for the number of U.S. annual lightning fatalities (1994-2003). The 2.5th and 97.5th percentiles form a 95% confidence interval.

Percentile	Equation	r^2
2.5th	$y = 1953.40e^{-0.042(x-1900)}$	0.94
5th	$y = 5416.80e^{-0.052(x-1900)}$	0.93
10th	$y = 7507.10e^{-0.055(x-1900)}$	0.92
25th	$y = 1117.7e^{-0.034(x-1900)}$	0.74
50th	$y = 538.69e^{-0.025(x-1900)}$	0.93
mean	$y = 6616.5e^{-0.049(x-1900)}$	0.80
75th	$y = 1067.93e^{-0.029(x-1900)}$	0.90
90th	$y = 1950.16e^{-0.034(x-1900)}$	0.85
95th	$y = 1766.4.0e^{-0.032(x-1900)}$	0.95
97.5th	$y = 1153.00e^{-0.027(x-1900)}$	0.91

y = expected number of fatalities; x = year

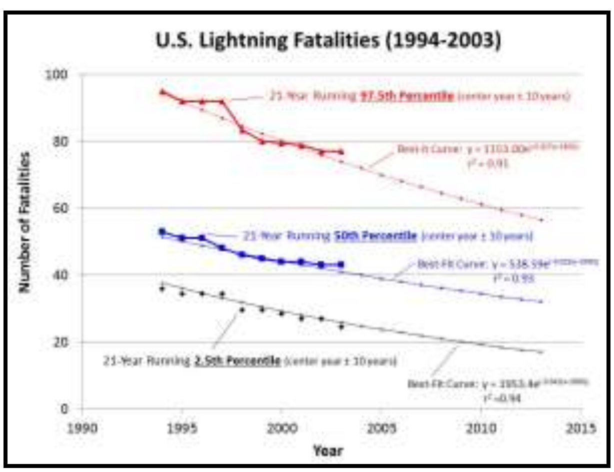

Figure-3. The 2.5th, 50th, and 97.5th percentiles of U.S annual lightning fatalities for 1994-2003 and best-fit negative exponential curves extrapolated to 2013.

There is an interesting pattern in the running 21-year percentiles (Figure-2). Many of the percentiles have a sudden drop in value that occurs in increasingly later years for increasing percentiles. For example, there is a sudden decrease in the 10th percentile between 1985-1897. There is a sudden decrease in the 25th percentile between 1988-1991. Likewise, the 50th percentile decreased suddenly 1993-1994; the 75th percentile in 1998-1999; and the 90th percentile in 2001-2002. One speculative explanation is that the lightning safety education efforts that intensified in the U.S. during that time (Cooper, 2012) began to improve the lightning safety behavior of an increasing amount of the population over time.

B. Updates to Intra-Annual Fatalities

The updates to the intra-annual study were more substantial than those to the inter-annual study. The 2012 study provided the median dates for selected percentiles of U.S. lighting fatalities based on the dates of those fatalities from 2006-2011. The update adds the dates of the lightning fatalities from 2012 and 2013 to the analysis, a 33% increase in sample size. The data are from the NWS lightning safety database (NWS, 2014) and discussed in Jensenius (2014), Roeder and Jensenius (2013), Roeder and Jensenius (2012a), and Roeder and Jensenius (2012b). In addition, a logistic regression on the percentile dates was added for a more robust overall result. Finally, the 2.5th and 97.5th percentiles were added, to provide a 95% confidence interval, in addition to the previous percentiles.

The distribution of U.S. intra-annual lightning fatalities was determined as follows. Nine percentiles were calculated for each of 8 years 2006-2013: 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97.5th percentiles. The median date for each percentile was calculated from the 8 years (Table-2). A best-fit logistic curve was used to estimate the distribution of these percentiles during a typical year. Note that the middle of the U.S. lightning fatality year is 16 July.

A generalized logistic curve is shown in (3).

$$y = 100*(1/(1 + e^{-(\alpha + \beta*(x - \delta))})) \qquad (3)$$

where y = percentile (%),
x = day of the year,
e.g. 1 Jan = 1, 1 Feb = 32, 31 Dec = 365,
and α, β, δ are the distribution parameters.

A best-fit logistic curve cannot be determined analytically and so the best-fit curve was done manually. The three distribution parameters (α, β, δ) were iterated sequentially until the RMSE of the median percentiles minus the curve was minimized. Once the first round of optimized parameters was obtained, the sequence of iteration had to be repeated since a change in one parameter changes the optimum value of the other parameters. The sequence of iterative optimization was repeated until the parameters were accurate to four decimal places, i.e. changed by less than 0.00005. The best-fit logistic curve is in (4).

$$y = 100*(1/(1 + e^{-(0.0315 + 0.0420*(x - 195.4000))})) \quad (4)$$

where y = percentile (%),
and x = day of the year.

This best-fit logistic curve had a RMSE of 1.23% and a bias of 0.3%, indicating a good match to the intra-annual percentiles calculated from 2006-2013. This good match can also be seen visually in Figure-4.

Table-2. Percentiles for the intra-annual of lightning fatalities in the U.S. The percentiles are the median of the percentiles observed in each year for 2006-2013.

Percentile	Median Date
0	1 Jan
2.5th	21 Apr
5th	2 May
10th	22 May
25th	15 Jun
50th	16 Jul
mean (standard deviation)	14 Jul (45.8 days)
75th	8 Aug
90th	5 Sep
95th	16 Sep
97.5th	24 Sep
100th	31 Dec

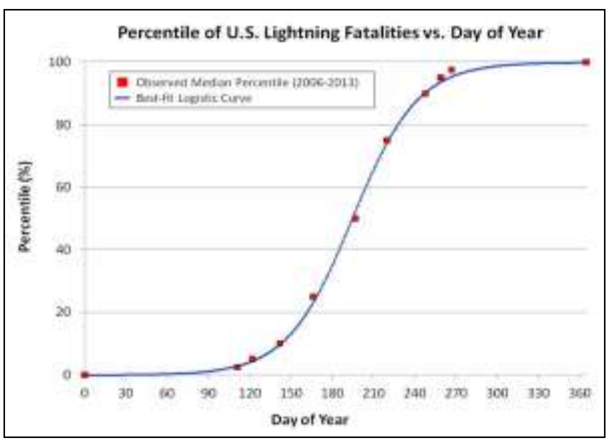

Figure-4. Best-fit logistic curve for the distribution of intra-annual lightning fatalities in the U.S.

The inter-annual and intra-annual analyses can be combined to predict the expected number of accumulated lightning fatalities and their error bars for any day of any year in the U.S. For example, the inter-annual analysis predicts an expected median of 31.2 lightning fatalities in the U.S. for 2014. Through 20 June, the intra-annual analysis indicates that 27.0% of the annual lightning fatalities should have accumulated. Therefore, the expected median number of U.S. lightning fatalities by the end of 20 June 2014 is 8.4 fatalities (27% of 31.2). As another example, the expected 2.5th and 97.5th percentiles of lightning fatalities for 2014 are 16.3 and 53.1 fatalities, respectively, and form a 95% confidence interval. Multiplying these numbers by the intra-annual percentile for 20 June (27.0%) gives a 95% confidence interval of 4.6 to 14.3 lightning fatalities for that date.

An EXCEL tool was built to automate the most useful parts of these analyses (Figure-5). The user needs to only enter the desired date in the yellow shaded cell and the rest of the data are calculated. This tool is useful for monitoring the development of lightning casualties in the U.S. as the year develops, which has some utility in lightning safety education. For example, around mid- 2013, it became apparent that 2013 could set a new record low for U.S. lightning fatalities, but that new record might not be statistically significant at the 95% level. This eventually verified with a new record low of 23 observed fatalities. The expected 5th percentile was 21.5 fatalities (using the previous version of the tool, since data for all of 2013 were not yet available), so the observed value was not statistically different from the expected value at the 95% significance level (1-tail hypothesis test), though it was close—if only two fewer lightning casualties had occurred, the new record would have been statistically significant. A separate EXCEL spreadsheet (not shown) was developed to calculate the probability of a new U.S. record low for the year occurring based on the observed number of lightning fatalities by the date of interest and the number of lightning fatalities observed in each year for 2006-2012.

III. POSSIBLE CHANGE IN U.S. LIGHTNING FATALITIES BEGINNING IN 2008

There is some evidence that the pattern of U.S. lightning fatalities may have changed beginning in 2008. The number of annual lightning fatalities in the U.S. for 2002-2013 is shown in Figure-6 (6 years of the changed fatality rate (2008-2013) and 6 years before the change (2002-2007)). A sudden drop in fatalities and year-to-year variability is apparent beginning in 2008. The drop in fatalities is more apparent if one considers 2004 to be a low outlier for 2002-2007. Linear regressions for the two periods are shown to help indicate the change and suggest the drop was 10.0 fatalities, using a transition year of 2007.5.

Expected U.S. Lightning Fataities

For Entire Year (POR: 1994-2003 (± 10 years))

Date	2.5th Percentile	5th Percentile	Median (50th Percentile)	95th Percentile	97.5th Percentile
14-Jul-14	16.3	14.4	31.2	46.0	53.1
			Mean		
			24.8		

For This Date (POR: 2006-2013)

Day Of Year	Percentile of Annual Lightning Fataliies (logistic regression)	Expected Number of U.S. Lightning Fatalities by this Date				
196	50.4	2.5th Percentile	5th Percentile	Median	95th Percentile	97.5th Percentile
		8.2	7.3	15.7	23.2	26.7
				Mean		
				12.5		

Figure-5. Example of EXCEL tool to calculate the expected number of U.S. lightning fatalities by a desired date and desired year and various error bars for those expected numbers.

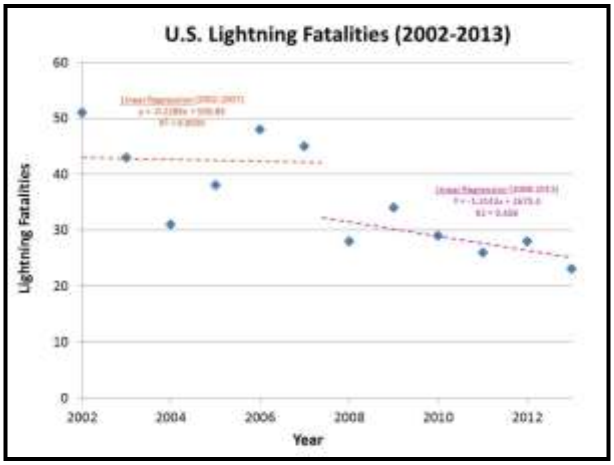

Figure-6. Annual lightning fatalities in the U.S. from 2002-2013. Note the sudden change to lower fatalities beginning in 2008. Linear regressions for 2002-2007 and 2008-2013 are shown to help indicate the change. The magnitude of the decrease was about 10.0 fatalities.

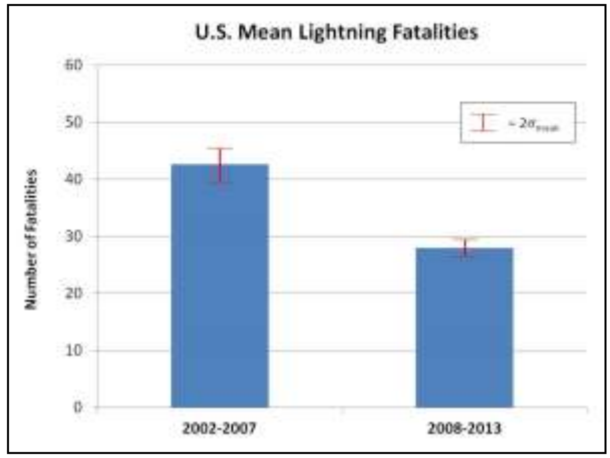

Figure-7. Mean, standard deviation, and standard deviation of the mean for 2002-2007 and 2008-2013. The error bars are <u>two</u> standard deviations <u>of the means</u> and show visually that means are very different statistically, as confirmed by the t-test ($p = 0.0067$).

The mean for 2002-2007 is 42.7 fatalities with a standard deviation <u>of the mean</u> of 2.95. For 2008-2013 the mean is 28.0 with a standard deviation <u>of the mean</u> of 1.48. These data are shown in Figure-7. A two-tailed Student's t-test with unequal variances for a null hypothesis that the means are equal is rejected with a p-value = 0.0067 (Wilks, 2006). This strongly supports the observation that the lightning fatality rate in the U.S. did indeed change beginning in 2008. A one-tailed hypothesis test assuming a lower 2008-2013 mean was not done since there is no a priori evidence to suggest a decrease vs. increase in lightning fatalities.

Another argument can be made indicating that the U.S. lightning fatality rate suddenly lowered beginning in 2008. Each of the most recent six years (2008-2013) was less than expected, as compared to the previously developed best-fit exponential curve (Roeder, 2012 (Appendix-1)). However, none of the individual annual differences from expectation was statistically significant, even though two of those years set new record lows. However, taken as a whole, those six consecutive years of fatalities below expectations are statistically significant at nearly the 98.4% level. If the distribution of fatalities was random about the expected values, the probability of six consecutive observations all below expectation is $(0.5)^6$, or 1.6%. If this trend continues, a new statistical model of the expected lightning fatalities should be developed, restricting the analysis to the 2008 and later years.

Some members of the lightning safety community have speculated on the cause of this apparent sudden change in U.S. lightning fatalities. The cause was first thought to have perhaps been due to the economic recession that began in 2007-2008—people may have been spending less time doing at-risk outdoor recreation to save money. However, if that were the cause, then the lightning fatalities should return toward expectations as the economy recovers. Since that return to expectations has not been observed (2013 was a new record low in U.S. lightning fatalities), the economic recession may not have been the primary cause of the change in lightning fatalities, or at least not solely the cause. Further speculation considered that the cause may have been a combination of the economic recession and NOAA's lightning safety education campaign that began in 2001 (Jensenius and Franklin, 2012; Jensenius and Franklin, 2008; Hodanish et al., 2008; Jensenius et al., 2008; Jensenius and Franklin, 2006a; Jensenius and Franklin, 2006b). The population may have been primed to change their lightning safety behavior by the education campaign, then once the recession precipitated a change in behavior, that change persisted even after the recession ended.

Another possible cause would be a shift in the amount of lightning in the U.S. However, it is assumed that a significant enough reduction in lightning activity has not occurred over the six year period. For example, there was an intense drought of the U.S. in 2012, especially the Mid-West, with an assumed drop in lightning activity, yet 2012 did not show a corresponding drop in lightning fatalities compared to the previous or subsequent year.

IV. FUTURE WORK

The analysis should be updated as more annual data accrue. This is especially true for the intra-annual distribution of lightning fatalities since only 8 years of data have been available so far (2006-2013).

The annual percentile distributions also could be improved. The choice of period record for the percentiles and ± interval over which to measure each annual percentile should be optimized to find the best balance between sample size and making the results more representative of the present.

The sudden drop in U.S. lightning fatalities in 2008 should be studied further. The annual rates should be monitored to see if they remain below expectation. If so, then a new model of expected lightning fatalities, representative of the new rate, should be developed. In addition, the annual lightning flash rate in the U.S. should be investigated to see if changes in flash rate help explain the changes in fatalities. Finally, social scientists should be consulted.

V. SUMMARY

An update to the statistical analysis of annual rate of lightning fatalities in the U.S. was presented, including an update to the method estimating the total number of expected lightning fatalities in the U.S. for any year in the near future and the expected number of fatalities on any day during any year. The uncertainty in these estimates was also updated for use as error bars and to allow hypothesis testing. The model predicts a median of 31.2 lightning fatalities in the U.S. for 2014 with a 95% confidence interval of 16.3 to 53.1 fatalities. The apparent sudden decrease of U.S. lightning fatalities beginning in 2008 was discussed.

VI. REFERENCES

Ashley, W. S., and C. W. Gilson, 2009: A reassessment of U.S. lightning mortality, *Bulletin of the American Meteorological Society*, **90**, Oct 09, 1501-1518

Cooper, M. A., 2012: A brief history of lightning safety efforts in the United States, *4th International Lightning Meteorology Conference*, 4-5 Apr 12, 8 pp.

Hodanish, S. J., K. Torgerson, J. S. Jensenius, M. A. Cooper MD, M. Utley, and W. P. Roeder, 2008: Leon the Lightning Safety Lion says: "When Thunder Roars, Go Indoors!"--NOAA's Efforts Regarding Children's Lightning Safety, *3rd Conference on Meteorological Applications of Lightning Data*, Paper 2.15, 4 pp.

Holle, R. L, 2012: Recent studies of lightning safety and demographics, *4th International Lightning Meteorology Conference*, 4-5 Apr 12, 15 pp.

Holle, R. L., R. E. Lopez, and B.C. Navarro, 2005: Deaths, injuries, and damages from lightning in the United States in the 1890s in comparison with the 1990s, *Journal of Applied Meteorology*, **44**, 1563-1573

Jensenius, J., 2014: A detailed analysis of recent lightning deaths in the United States, *5th International Lightning Meteorology Conference*, 20-21 Mar 14, 8 pp.

Jensenius, J., and D. Franklin, 2012: A review of NOAA's lightning safety awareness campaign and its impact across the United States, *4th International Lightning Meteorology Conference*, 4-5 Apr 12, 8 pp.

Jensenius, J., and D. B. Franklin, 2008: Lightning kills – play it safe: NOAA's continuing efforts to educate the public on the dangers of lightning, *2nd International Lightning Meteorology Conference*, 24-25 Apr 08, 7 pp.

Jensenius, J., and D. B. Franklin, and S. Hodanish, 2008: Lightning kills – play it safe: NOAA's efforts to educate the public on the dangers of lightning, *3rd Conference on Meteorological Applications of Lightning Data*, Paper 5.2, 5 pp.

Jensenius, J., and D. B. Franklin, 2006a: Lightning kills – play it safe: NOAA's efforts to educate the public on the dangers of lightning, *19th International Lightning Detection Conference*, 24-25 Apr 2006, 7 pp.

Jensenius, J., and D. B. Franklin, 2006b: Lightning kills – play it safe: NOAA's efforts to educate the public on the dangers of lightning, *1st International Lightning Meteorology Conference*, 26-27 Apr 2006, 7 pp.

Lopez, R. E., and R. L. Holle, 1998: Changes in the number of lightning deaths in the United States during the twentieth century, *Journal of Climate, Vol. 11, No. 8*, Aug 98, 2070-2077

National Weather Service, 2013: NOAA, National Weather Service Office of Climate, Water, and Weather Services, 1325 East West Highway, Silver Spring, MD 20910, *www.nws.noaa.gov/om/hazstats.shtml*, 2013

National Weather Service, 2014: NOAA, National Weather Service Lightning Safety, 1325 East West Highway, Silver Spring, MD 20910, *www.lightningsafety.noaa.gov/more.htm*, 2014

Roeder, W. P, 2013: Estimating the expected number of U.S. lightning fatalities during a year, throughout a year, and comparison to other storm phenomena, *6th Conference on Meteorological Applications of Lightning Data*, Paper P1.4, 6-10 Jan 2013, 1 pp.

Roeder, W. P., 2012: A statistical model for the inter-annual and intra-annual fatalities from lightning in the U.S. and comparison to other storm phenomena, *4th International Lightning Meteorology Conference*, 4-5 Apr 12, 6 pp.

Roeder, W. P., and J. Jensenius, 2013: NOAA's new lightning fatality database and Applications in lightning safety education, *6th Conference on Meteorological Applications of Lightning Data*, Paper P1.3, 6-10 Jan 2013, 1 pp.

Roeder, W. P., and J. Jensenius, 2012a: A new high-quality lightning fatality database for lightning safety education, *International Lightning Meteorology Conference*, 4-5 Apr 12, 9 pp.

Roeder, W. P., and J. Jensenius, 2012b: Lightning safety applications of the new NOAA lightning fatality database, *37th Annual Meeting of the National Weather Association*, P2.30, 10 pp.

Wilks, D. S., 2006: Statistical methods in the atmospheric sciences, 2nd edition, *Academic Press, Inc.*, Vol. 91 International Geophysics Series, pp. 627

APPENDIX-1

(Roeder, 2012)

2012
22nd International Lightning Detection Conference
2 - 3 April • Broomfield, Colorado, USA
4th International Lightning Meteorology Conference
4 - 5 April • Broomfield, Colorado, USA

A Statistic Model For The Inter-Annual And Intra-Annual Fatalities From Lightning In The U.S. And Comparison To Other Storm Phenomena

William P. Roeder
Private Meteorologist
Rockledge, FL

1. Introduction

The number of lightning deaths in the U.S. has been declining for many decades. As a result, the traditional 30-year mean used by National Weather Service for most climatological averages overestimates the current rate. The 30-year mean for lightning fatalities in the U.S. (1980-2009) is 57 deaths per year (National Weather Service, 2011a). However, the expected U.S. lightning fatality rate for 2012 is 31.7 deaths, as shown later. The 10-year climatological mean is more representative with 41 deaths per year (2000-2009) (NWS, 2011a). However, such short period averages can be misleading, especially if skewed by a single extreme event. Others have tried to prove the impact of improved lightning safety public education since the early 1990s by showing that successive periods have smaller slopes (Lengyel, 2004). However, since the lightning fatalities are following a curve and started declining well before the 1990s, proving the impact of public education this way is difficult.

This paper provides a best-fit negative exponential curve for the U.S. annual lightning deaths since the early 1940s. This will allow estimating the expected number of lightning deaths for any year in the near future. The annual 95th percentile confidence interval for any year is also provided. This will allow estimating if the actual number of lightning deaths is statistically significant. Finally, estimates for the median, 25th percentile, and 75th percentile of the expected annual lightning deaths is also provided.

This paper also provides the distribution of lightning fatalities during the year.

Finally, a brief comparison of lightning fatality rates in the U.S. to fatalities from other storm deaths is also provided. Lightning appears to have fallen to the third leading cause of storm deaths in the U.S. from second place. Tornadoes appear to have become the second leading cause of storm deaths.

2. Inter-Annual Pattern of Lightning Fatalities

The distribution of year-to-year lightning deaths in the U.S. is discussed. A negative exponential best-fit curve is provided that overcomes the shortfalls in running averages and allows estimating the expected number of lightning fatalities in any year in the near future. The errors bars also allow hypothesis testing if the lightning fatality total for that year is significantly different from that expectation.

2.1 Background (1900-2010)

The number of annual lightning deaths in the U.S. from 1900-2010 is shown in Figure-1 and listed in Table-1. The data from 1900-1991 are from Lopez and Holle (1998), the data from 1992-2010 are from National Weather Service (2011b). Visual inspection shows three main sections: 1) generally increasing lightning deaths from 1900 to about 1920, 2) a plateau from about 1921 to 1940, and 3) generally decreasing lightning deaths from 1941-2010. The increasing period from 1900 to about 1920 is likely an artifact of poor reporting. The statistical analysis will deal only with the decreasing death rates from 1941 to 2010.

Visual inspection of Figure-1 indicates that the lightning death rates began a general decline sometime between 1938 to 1944. With no evidence for which of these dates to refer as the start of the decline, the statistical analysis will use the middle of this period, 1941, as the start of the decline in U.S. lightning deaths.

Visual inspection of Figure-1 indicates that the lightning death rates began a general decline sometime between 1938 to 1944. With no evidence for which of these dates to refer as the start of the decline, the statistical analysis will use the middle of this period, 1941, as the start of the decline in U.S. lightning deaths.

Figure-1. Annual lightning deaths in the U.S. from 1900-2010. The data from 1900-1991 are from Lopez and Holle (1998). The data from 1991-2010 are from National Weather Service (2011b).

2.2 Negative Exponential Curve Fitting (1941-2010)

Visual inspection of Figure-1 indicates that the lightning death rates began a general decline sometime between 1938 to 1944. With no evidence for which of these dates to refer as the start of the decline, the statistical analysis will use the middle of this period, 1941, as the start of the decline in U.S. lightning deaths.

The number of lightning deaths per year in the U.S. since 1941 is shown in Figure-2. The decrease appears to be following a negative exponential curve. Nine best-fit curves were tried including linear, polynomial curves of order two through six, exponential, power curves, and logarithmic curves. The negative exponential curve and logarithmic curves had the best correlation coefficient, being mathematically equivalent. The negative exponential curve is used since it is easier to visualize. The best-fit negative exponential curve is given by the following equation:

$$y = 1182.00e^{-0.0323(x-1900)}$$

$$r^2 = 0.9166$$

where y = expected number of lightning deaths in the U.S.,
and x = desired year.

Table-1.
Number of annual lightning deaths in the U.S. from 1900-2010. Data from 1900-1991 are from Lopez and Holle (1998) and data from 1992-2010 are from National Weather Service (2011b).

Year	Deaths	Year	Deaths	Year	Deaths
1900	95	1937	460	1974	112
1901	127	1938	396	1975	124
1902	108	1939	390	1976	81
1903	92	1940	340	1977	116
1904	81	1941	388	1978	98
1905	80	1942	372	1979	87
1906	169	1943	432	1980	94
1907	133	1944	419	1981	87
1908	161	1945	268	1982	100
1909	150	1946	231	1983	93
1910	156	1947	338	1984	91
1911	235	1948	256	1985	85
1912	243	1949	249	1986	78
1913	309	1950	219	1987	99
1914	279	1951	248	1988	82
1915	219	1952	212	1989	75
1916	337	1953	145	1990	74
1917	355	1954	220	1991	73
1918	402	1955	181	1992	41
1919	408	1956	149	1993	43
1920	408	1957	180	1994	69
1921	459	1958	104	1995	85
1922	425	1959	183	1996	53
1923	449	1960	129	1997	42
1924	333	1961	149	1998	44
1925	440	1962	153	1999	46
1926	364	1963	165	2000	51
1927	348	1964	129	2001	44
1928	425	1965	149	2002	51
1929	391	1966	110	2003	43
1930	358	1967	88	2004	31
1931	439	1968	129	2005	38
1932	355	1969	131	2006	47
1933	372	1970	122	2007	46
1934	442	1971	122	2008	28
1935	362	1972	94	2009	34
1936	409	1973	124	2010	29

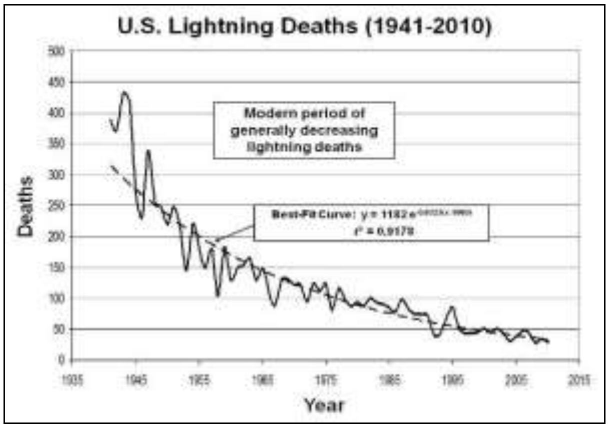

Figure-2. Annual lightning deaths in the U.S. from 1941-2010. Data are same as in Figure-1.

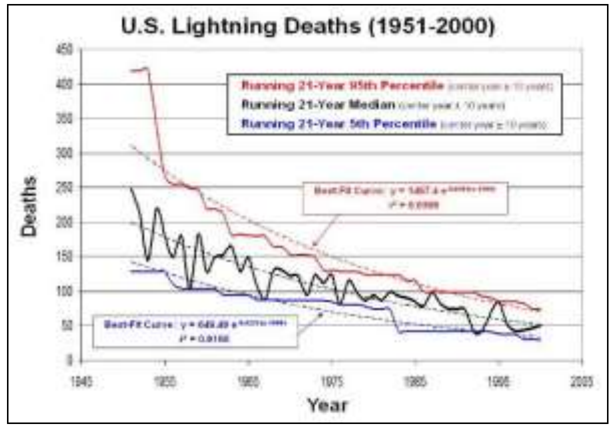

Figure-3. 5th percentile (blue) and 95th percentile (red) best-fit curves for the expected number of annual lightning deaths in the U.S. for 1950-2000.

2.3 Error Bars (1941-2010)

A running 21-year 5th and 95th percentile was calculated for each year from 1951 to 2000 centered on each year ± 10 years. The analysis is limited to 1950 through 2000, to allow for the running ± 10 running years to 1941 and 2010, respectively. A best-fit negative exponential curve was chosen for both the 5th and 95th percentile curves (Figure-3). Each of these best-fit curves is given by the following equations:

$$y = 649.49e^{-0.0297(x-1900)}$$

$$r^2 = 0.9178$$

where y = 5th percentile of lightning deaths, and x = desired year,

and

$$y = 1457.40e^{-0.0304(x-1900)}$$

$$r^2 = 0.9369$$

where y = 95th percentile of lightning deaths, and x = desired year.

The 5th through 95th percentiles represent a 90% confidence interval for the number of expected annual lightning deaths. The confidence interval decreases in later years as the expected number of lightning deaths also decreases. For the actual observed number of lightning deaths for any year, this confidence interval can be used to perform a one-tail hypothesis test at the 95% significance level. One-tail hypothesis tests are the most likely to be performed to answer if the observed lightning deaths for any particular year are statistically significantly more than expected or less than expected, as opposed to a two-tailed test to answer if the observed number of lightning deaths is different than expected, i.e. either more or less than expected. The running ±10-year medians, 25th percentiles, and 75th percentiles were also calculated for 1950-2000. Their best-fit negative exponential curves were also calculated. The results of these analyses are summarized in Table-2.

Table-2.
Statistical analyses of the number of lightning deaths in the U.S.

Metric	Best-Fit Curve*	Correlation Coefficient (r^2)	Period Of Record
Number of Lightning Deaths	$y = 1182.00e^{-0.0323(x-1900)}$	0.9178	1941-2010
95th Percentile	$y = 1457.40e^{-0.0304(x-1900)}$	0.9369	1950-2000
75th Percentile	$y = 986.35e^{-0.0277(x-1900)}$	0.9646	1950-2000
Median	$y = 878.88e^{-0.0285(x-1900)}$	0.9580	1950-2000
25th Percentile	$y = 767.46e^{-0.0290(x-1900)}$	0.9568	1950-2000
5th Percentile	$y = 649.49e^{-0.0297(x-1900)}$	0.9166	1950-2000

y = the metric being calculated for any desired year, x = the desired year.
For example, the number of lightning deaths expected in the U.S. for 2012 = $1182.00e^{-0.0323(2012-1900)}$ = 31.7 deaths.

The record low of 26 lightning deaths in the U.S. occurred in 2011. The expected distribution of lightning fatalities for 2011 is shown in Table-3. Since the actual number of lightning deaths was larger than the 5th percentile, the record low was not statistically significant at the 95% significance level (one-tail test), although it was close. The distribution of expected lightning fatalities in 2012 is in Table-4.

Table-3.
Expected distribution of 2011 U.S. lightning fatalities.

Metric	Expected Value
Mean	32.8
95th Percentile	49.9
75th Percentile	45.6
Median	37.2
25th Percentile	30.7
5th Percentile	24.0

Table-4.
Expected distribution of 2012 U.S. lightning fatalities.

Metric	Expected Value
Mean	31.7
95th Percentile	48.4
75th Percentile	44.3
Median	36.1
25th Percentile	29.8
5th Percentile	23.3

2.4 Causes of Decline in Lightning Deaths

There has been much speculation as to reasons for the decreasing number of annual lightning deaths in the U.S. Some of the proposed reasons include increased indoor plumbing and rural electrification that have made buildings safer from lightning, and industrialization and urbanization that have led most people to spend less time working outdoors. This trend may be somewhat compensated by increased outdoor recreation. The increased and improved lightning safety public education may also have played a role, especially since the early 1990s and even more so starting with the national Lightning Safety Awareness Week beginning in 2001. However, it is difficult to ascribe the amount of change in lightning deaths to the various factors at various times over the years, though it is odd that all these various causes at various times yield a fairly consistent rate of decline over so many decades. Fortunately for this analysis, the cause(s) of the declining annual number of lightning deaths is not important, only the consistent decline that can be analyzed statistically.

3. Intra-Annual Pattern of Lightning Fatalities

The distribution of lightning deaths during the year in the U.S. is presented (Table-5). This analysis was based on the new NOAA database that contains 210 fatalities and covers the period from 2006-2011 (Roeder and Jensenius, 2012). The inter-quartile range (25th and 75th percentiles) is frequently provided in statistical analysis. The author prefers to also provide the inter-decile range (10th and 90th percentiles) to describe the distribution more fully. The 2.5th, 5th, 95th, and 97.5th percentiles were also provided to describe the lightning fatality distribution throughout the year more fully.

Table-5.
Distribution of intra-annual U.S. lightning fatalities (2006-2011).

Metric	Date
Mean (standard Deviation)	10 Jul (46.5 days)
2.5th Percentile	21 Apr
5th Percentile	2 May
10th Percentile	11 Jun
25th Percentile	15 Jun
50th Percentile (median)	14 Jul
75th Percentile	3 Aug
90th Percentile	5 Sep
95th Percentile	16 Sep
97.5th Percentile	27 Sep

The distribution is very non-Gaussian so the percentiles are a better indicator of the distribution than the mean and standard deviation. In particular, there is an extremely rapid increase in the first half of June. This rapid increase is presumably due to the large increase of lightning flash rate in the U.S. during Jun (Holle and Murphy, 2010) and the increase of outdoor recreation due to the warmer temperatures and summer vacations.

The lower percentiles are very sensitive to some outliers in Jan-Mar of some years. Excluding just the two earliest lightning fatalities in the 6-year database from the analysis, one in Jan and one in Feb, makes the 2.5th and 5th percentiles about a week later. Excluding the five earliest fatalities in Jan-Mar makes those percentiles about 2 weeks later.

The inter-annual distribution of intra-annual fatalities can be combined with the inter-annual analysis (section-2) to estimate the expected number of lightning fatalities in the U.S. throughout any year. This can be used to evaluate if a particular year is ahead of schedule, on schedule, or behind schedule for the expected number of lightning fatalities. An example for 2012 is shown in Table-6.

Table-6.
Distribution of expected U.S. lightning fatalities for 2012. The total expected is 31.7 deaths.

Metric	Expected Number of Fatalities
21 Apr (2.5th percentile)	0.8
2 May (5th percentile)	1.6
11 Jun (10th percentile)	3.2
15 Jun (25th percentile)	7.9
14 Jul (50th percentile)	15.9
3 Aug (75th percentile)	23.8
5 Sep (90th percentile)	28.5
16 Sep (95th percentile)	30.1
27 Sep (97.5th percentile)	30.9
31 Dec (100th percentile)	31.7

4. Comparison to U.S. Fatalities from Other Storm Phenomena

Comparison of the past thirty years of U.S. fatalities from other storm phenomena indicates that floods remain the leading source and hurricanes remains in fourth place. The fatality data are from *www.nws.noaa.gov/om/hazstats.shtml*.

However, it appears that tornadoes have now become the second leading source of storm deaths, replacing lightning, which has now become third place. The 30-year tornado fatality is 56 deaths per year for 1981-2010. Although the tornado fatality rate shows large year-to-year variability, it also appears to be stationary, showing no upward or downward trend. This average annual tornado fatality rate is above the 95th percentile expected lightning fatality rate for 2011. Since the decreasing lightning fatality rate was suspected a priori to be lower than the tornado rate, a one-tail hypothesis test is appropriate. Therefore, the data suggest that the U.S. tornado fatality rate has become statistically significantly larger than the current lightning fatality rate at the 95th percent level. The first year that the 95th percentile for lightning deaths was less than the long-term average for tornadoes was 2006. This suggests that 2006 was the first year that we could be confident that lightning had become the third leading source of storm deaths in the U.S., with tornadoes replacing it as the second leading source.

Lightning may also be dropping below hurricanes in the rank order of lightning deaths in the U.S. The 30-year (1981-2010) average number of hurricanes is 47 deaths per year. This is more than the expected number of lightning deaths (32.8) in 2011, but it is still less than the 95th percentile estimate for lightning deaths in 2011 (49.9), so we cannot be statistically confident that hurricanes has surpassed lightning in the rank order of storm deaths. The stationarity of the hurricane deaths has not been considered.

Lightning may even be dropping below straight-line winds in the rank order of storm deaths in the U.S. The 10-year (2001-2010) average number of wind deaths is 41 deaths per year. This is more than the estimated number of deaths from lightning for 2011, but less than the 95th percentile estimate, so we are not statistically confident that winds have surpassed lightning in the rankings. The 10-year mean for winds is used since the 30-year mean is not available. The stationarity of the winds deaths has not been considered.

5. Future Work

The current analysis is for the absolute number of lightning deaths. Repeating the analysis with the per capita lightning deaths may yield a more reliable estimate of the expected number of lightning deaths. This may also be more instructive for examining the behavioral changes for the declining rate of lightning deaths. Likewise, study of the intra-annual lightning fatality by region may be instructive. For example, the distribution may be more compressed in the Rocky Mountain States, starting later and ending sooner than in the Southeast U.S. Also, the lightning fatalities in the Pacific Northwest may occur later in the year as compared to the national average. However, the sample size for lightning fatalities in the Pacific Northwest may not justify this analysis.

The 95% confidence intervals for the annual number of lightning deaths were estimated by curve fitting the running $\pm$10-year 5th and 95th percentiles. This was done to approximate percentile regression. A statistical analysis software package would allow true percentile regression.

Statistical analysis may be able detect subtle changes in the rate of decline in lightning deaths, such as with the start of increased and improved lightning safety in the early 1990s and especially after the start of national Lightning Safety

Awareness Week in 2001. However, visual inspection of the data suggests that the data may be too noisy to detect such changes.

The analysis of fatalities from other storm phenomena should be expanded to perform a formal hypothesis test if the tornado fatality rate is larger than the lightning fatality rate to include a p-value. In addition, the stationarity of the tornado fatality rate should also be hypothesis tested. Likewise, the flood and hurricane fatality rate should be hypothesis tested for being greater than and less than the lightning fatality rate, respectively, along with p-values. Also, the stationarity of hurricanes and winds should be verified.

Finally, the lightning fatality analysis should be updated as more years of data occur. For example, 2011 should be added.

6. Summary

The lightning deaths in the U.S. from 1941-2010 were analyzed to provide the expected number of lightning deaths for any year. This was done with a best-fit negative exponential curve on the number of lightning deaths from 1941-2010. The running $\pm$ 10 year 5th and 95th percentiles were calculated for 1950-2000 and best-fit negative exponential curves calculated to provide the 90% confidence interval for the expected number of U.S. lightning deaths for any year. The median, 25th percentile, and 75th percentiles best-fit curves were also calculated. Finally, the percentile distribution of lighting fatalities within a year was calculated from the new 2006-2011 NOAA database.

7. References

Holle, R. L., and M. L. Murphy, 2010: Monthly distributions of U.S. NLDN cloud-to-ground lightning, *3rd International Lightning Meteorology Conference*, 21-22 Apr 2010, 13 pp.

Lengyel, M. M., 2004: Lightning casualties and their proximity to surrounding cloud-to-ground lightning, *M.S. Thesis*, University of Oklahoma, Norman, 68 pp.

Lopez, R. E., and R. L. Holle, 1998: Changes in the number of lightning deaths in the United States during the twentieth century, *Journal of Climate, Vol. 11, No. 8*, Aug 98, 2070-2077

National Weather Service, 2011a: NOAA, National Weather Service Office of Climate, Water, and Weather Services, 1325 East West Highway, Silver Spring, MD 20910, *www.nws.noaa.gov/om/hazstats.shtml*, 2011

National Weather Service, 2011b: NOAA, National Weather Service Lightning Safety, 1325 East West Highway, Silver Spring, MD 20910, *www.lightningsafety.noaa.gov/more.htm*, 2011

Roeder, W. P., and J. Jensenius, 2012: A new high-quality lightning fatality database for lightning safety education, *International Lightning Meteorology Conference*, 4-5 Apr 12, 9 pp.