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.

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

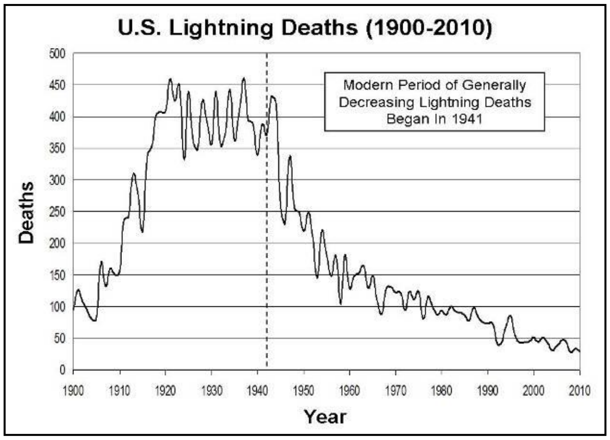

Figure-1. Annual lightning deaths in the U.S. from 1900-2010. The data from 1900-1991 are from Lopez and Holle (1998). The data from 1991-2010 are from National Weather Service (2011b).

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.

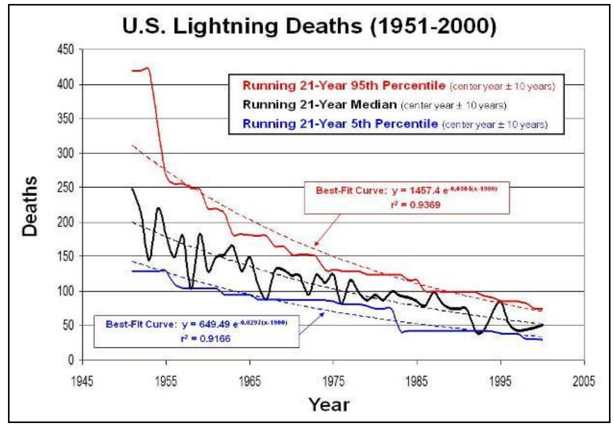

Figure-3. 5th percentile (blue) and 95th percentile (red) best-fit curves for the expected number of annual lightning deaths in the U.S. for 1950-2000.

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 95t h 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 ±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 he 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 ± 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.