Reprinted from JOURNAL OF APPLIED METEOROLOGY, Vol. 15, No. 7, July 1976
American Meteorological Society
Printed in U. S. A.

Statistical Analyses of a Randomized Lightning Modification Experiment

ROBERT G. BAUGHMAN AND DONALD M. FUQUAY

PAUL W. MIELKE, JR.

Statistical Analyses of a Randomized Lightning Modification Experiment

ROBERT G. BAUGHMAN AND DONALD M. FUQUAY

USDA Forest Service, Northern Forest Fire Laboratory, Missoula, Mont. 59801

PAUL W. MIELKE, JR.

Department of Statistics, Colorado State University, Fort Collins 80523

(Manuscript received 5 June 1975, in revised form 14 May 1976)

ABSTRACT

Analyses and interpretations of data from a randomized lightning modification experiment yield strong evidence that AgI seeding of mountain thunderstorms reduces the frequency of lightning and alters several lightning characteristics. Statistical procedures include likelihood ratio tests based on the log-normal distribution and two nonparametric tests (Wilcoxon and sum-of-squared ranks tests). Specifically, the estimated reduction in the long continuing current intervals in cloud-to-ground flashes is about one-fourth (two-sided p-values of both nonparametric tests are less than 0.04). Further, seeding appears to reduce the frequency of cloud-to-ground flashes and the duration of lightning activity by about one-half (two-sided p-values of both nonparametric tests are all less than 0.03). These results could have important applications in forest fire management programs.

1. Introduction

Lightning is an important cause of forest fires in various regions of the world including North America. Annually, some 10 000 forest fires in the United States are caused by lightning. In years of intense lightning activity and dry weather, some 15 000 or more lightning fires may occur (Barrows, 1966). Severe lightning fire outbreaks are a major problem in wildland management. The development of weather modification techniques to attenuate lightning could provide a viable option for managing this critical forest fire problem (Fuquay, 1967).

Lightning modification experiments were carried out in western Montana by USDA Forest Service personnel during the 1960 decade (Project Skyfire). Various aspects of the experimentation, including results, have been previously reported (Fuquay, 1974). In general there was insufficient statistical evidence to support the hypothesis that cloud seeding has an effect on the frequency of lightning but there was strong statistical evidence supporting the hypothesis that seeding affects the characteristics of cloud-to-ground lightning.

The Forest Service and Colorado State University have undertaken further analyses of the lightning modification work. Reported here are specific findings from these analyses which indicate that seeding of mountain thunderstorms reduces the frequency of lightning and changes several characteristics of cloud-to-ground lightning.

2. Experimental design and description of variables

A single target area was used in the experimentation. The experimental unit, called an *operational day*, was defined as a unit of space and time that limited operations and observations to events that occurred within a specified target area between noon and midnight MST. The declaration of an operational day depended on a U. S. Weather Bureau forecast of a 30% chance of thunderstorm occurrence in the vicinity of the target area. A table of random numbers was used to assign a non-seed or seed treatment to an operational day. Once the treatment of an operational day was assigned on a random basis, the next operational day received the opposite treatment (i.e., the randomization scheme involved sequentially occurring pairs of operational days). The purpose of this randomization scheme was to prevent a long uninterrupted sequence of either non-seeded or seeded units.

The seed treatment was carried out using both ground-based and airborne silver iodide (AgI) generators. Clouds approaching or building within the test area were seeded from beneath the clouds near cloud-base level. The clouds were continuously seeded at an average rate of slightly greater than 2 kg of AgI per hour with periodic maxima in excess of 4 kg h^{-1} (Fuquay, 1974).

The initially planned statistical procedure was to include all lightning events occurring within the specified test area with the time constants of the operational day. The basis for accepting the lightning events for analysis was a field change greater than the expected median value corresponding to a distance of about 16 km. This criterion was used because previous tests with field meters in a nearby locale suggested this would be the maximum extent of reliable recording.

The first look at the data clearly indicated that the established field change criterion was excluding all but the closest discharges from the analyses. As a result, lightning from only a small portion of each storm was being accepted. It was obvious that only a very small, non-representative data sample would be available from the experiment. Further, the reliance on a forecast to specify an operational day resulted in a large number of operational days void of lightning being included in the analyses.

A further look at the data established that reliable recordings on lightning frequency were available to a range of 32–36 km and data on characteristics of each discharge to a range of 16–24 km. Thus, an alternative statistical procedure, based on an *individual storm* criterion, was established. The results reported here are based on criteria established after the completion of the field tests.

Due to the problems stated above, the lightning frequency data were actually analyzed in accordance with the following individual storm criterion. Individual storms (viz. periods of lightning activity occurring within an experimental unit) were identified as existing singly (and almost wholly) within the test area by radar and confirmed by field change measurements as producing at least one lightning discharge. Thus, individual storms were identified as the source of recorded lightning and as recipients of the prescribed treatment. Individual storms available for these analyses came from all data acquired during the period 1965 through 1967. Three morning storms which were non-seeded were unfortunately included in earlier analyses (Fuquay, 1974). These storms, which produced 3, 18 and 55 lightning discharges over storm durations of 10, 38 and 116 min, respectively, did not fit the criterion of an experimental unit (i.e., they could not have been seeded within this randomization scheme) and are not included here. Thus, the analyses of this paper involve 14 non-seeded and 12 seeded storms which occurred within experimental units (Table 1).

In this paper, any discharge or flash that does not reach earth is defined as intracloud (IC) and those between cloud and earth as cloud-to-ground flashes (CG). Those few discharges that could not be clearly classified as IC or CG are called indeterminates. The cloud-to-ground flashes containing continuing-current intervals are referred to as hybrid flashes and all others are called discrete flashes. Hybrid flashes with continuing current of 40 or more milliseconds are called long-continuing current (LCC) discharges (Fuquay *et al.*, 1972). The very luminous component of a discharge is called a return stroke. A complete review of lightning technology and terminology is given by Uman (1969).

TABLE 1. Lightning frequencies for non-seeded and seeded storms.

Storm identification	Date	Time interval (MST)	Type of lightning: Cloud-to-ground (CG)	Intra-cloud (IC)	Indeterminate	Total
			Non-seeded			
1	7/2/65	1711–1914	61	457	0	518
2	7/4/65	1234–1337	33	129	2	164
3	7/4/65	1350–1559	62	206	7	275
4	7/8/65	1718–1835	45	210	1	256
5	8/19/65	1619–1720	0*	24	1	25
6	8/19/65	1720–1915	30	99	4	133
7	7/12/66	1318–1632	82	437	2	521
8	8/4/66	1459–1619	10	57	0	67
9	9/7/66	1705–1821	20	119	13	152
10	9/12/66	1904–2243	358	357	40	755
11	7/3/67	1555–1742	63	226	7	296
			Seeded			
1	7/20/65	1610–1726	49	179	0	228
2	7/21/65	1357–1418	4	8	1	13
3	7/29/65	1450–1629	18	88	1	107
4	8/27/65	1840–2005	26	418	21	465
5	7/6/66	1716–1845	29	49	1	79
6	7/14/66	1353–1418	9	1	0	10
7	7/14/66	1459–1631	16	129	0	145
8	7/14/66	1714–1740	12	1	0	13
9	7/15/66	1523–1604	2	12	0	14
10	7/15/66	1644–1818	22	58	0	80
11	8/29/66	1357–1447	10	27	0	37
12	8/29/66	1640–1752	34	51	10	95

* The value 0.5 was substituted for 0 in application of log-normal statistics T_1 and T_2.

Variables were selected for testing included those considered to be important factors in producing lightning-caused forest fires. These variables are as follows:

- CG flash frequency (per storm)
- IC flash frequency (per storm)
- Total flash frequency (CG+IC+indeterminate flashes per storm)
- Duration of lightning activity (time from first to last flash per storm)
- Lightning intensity (ratio of total flash frequency to duration of lightning activity)
- Maximum number of CG, IC and total flashes for any continuous 5 and 15 min period within each storm
- Return stroke frequency (number of return strokes for discrete and hybrid flashes)
- Return stroke interval (time interval from first to last return stroke for discrete and hybrid lightning in milliseconds)
- Continuing current interval of LCC flashes (milliseconds).

A summary of lightning statistics is given in Table 2.

3. Statistical procedures

The statistical procedures used to analyze the data of this study include parametric estimation and test

TABLE 2. Lightning summary statistics.

Variable	Non-Seeded storms Maximum	Minimum	Median	Seeded storms Maximum	Minimum	Median
Maximum lightning occurrence in a 5-min interval						
CG flashes	32	0*	10	13	1	4.5
IC flashes	52	8	24	64	1	8.5
Total flashes	57	8	31	76	4	12.0
Maximum lightning occurrence in a 15-min interval						
CG flashes	77	0*	18	19	1	8.0
IC flashes	102	11	54	175	1	21.0
Total flashes	116	11	73	194	7	23.0
Number of return strokes per discrete flash	13	1	4	9	1	3.5
Number of return strokes per hybrid flash	17	1	6	9	1	4.0
Time interval (ms), first to last return stroke, CG-discrete	910	30	200	540	25	128.0
Time interval (ms), first to last return stroke, CG-hybrid	1430	15	380	720	15	250.0
Continuing current interval (ms) CG-hybrid (LCC)	520	40	180	250	40	120.0

* The value 0.5 was substituted for 0 in application of log-normal statistics T_1 and T_2.

procedures based on the log-normal distribution and two nonparametric tests.

The log-normal distribution was selected as the basis for the parametric estimate of change due to treatment because of its apparent ability to describe a variety of meteorological and other types of data (Mielke and Johnson, 1974) and its ease of application. The density function of the log-normal distribution is given by

$$f(x)=\begin{cases}[\alpha x(2\pi)^{\frac{1}{2}}]^{-1}\exp\{-[\ln(x/\beta)]^2/(2\alpha^2)\}, & 0<x\\ 0, & x\leqslant 0\end{cases}$$

where the shape parameter (α) and scale parameter (β) are positive constants. The moments about zero of a random variable, say X, obeying this log-normal distribution are given by

$$E\{X^\nu\}=\beta^\nu\exp(\nu^2\alpha^2/2).$$

Let $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$ denote random samples of observations associated with non-seeded and seeded storms, respectively, which occurred during experimental units. The pairs of log-normal distribution parameters corresponding to the x's and y's are denoted by (α_N,β_N) and (α_S,β_S), respectively. The likelihood ratio test statistic for testing the null parameter space (null hypothesis)

$$\omega_1=\{\alpha_N=\alpha_S=\alpha,\ \beta_N=\beta_S=\beta\}$$

against the alternative parameter space (alternative hypothesis)

$$\Omega_1=\{\alpha_N=\alpha_S=\alpha,\ \beta_N,\ \beta_S\}$$

is given by

$$T_1=(m+n)[\ln(\hat{\alpha}^2_{\omega_1})-\ln(\hat{\alpha}^2_{\Omega_1})],$$

where

$$\hat{\alpha}^2_{\omega_1}=\frac{1}{m+n}\left[\sum_{i=1}^m(\ln x_i)^2+\sum_{i=1}^n(\ln y_i)^2-\frac{1}{m+n}\left(\sum_{i=1}^m\ln x_i+\sum_{i=1}^n\ln y_i\right)^2\right],$$

$$\hat{\alpha}^2_{\Omega_1}=\frac{1}{m+n}\left[\sum_{i=1}^m(\ln x_i)^2+\sum_{i=1}^n(\ln y_i)^2-\frac{1}{m}(\sum_{i=1}^m\ln x_i)^2-\frac{1}{n}(\sum_{i=1}^n\ln y_i)^2\right].$$

If the actual parameter space is ω_1, then the approximate distribution of T_1 is chi-square with 1 degree of freedom (cf. Wilks, 1962, pp. 402–422). The maximum likelihood estimate of the ratio β_S/β_N is given by

$$\hat{\beta}_S/\hat{\beta}_N=\exp\left[\frac{1}{n}(\sum_{i=1}^n\ln y_i)-\frac{1}{m}(\sum_{i=1}^m\ln x_i)\right].$$

In particular, $100[(\beta_S/\beta_N)-1]$ may be interpreted as the percentage increase due to seeding (a negative value denotes a decrease). In addition, the likelihood ratio test statistic for testing the null parameter space (null hypothesis)

$$\omega_2=\{\alpha_N=\alpha_S=\alpha,\ \beta_N,\ \beta_S\}$$

against the alternative parameter space (alternative hypothesis)

$$\Omega_2=\{\alpha_N,\alpha_S,\beta_N,\beta_S\}$$

is given by

$$T_2=(m+n)\ln(\hat{\alpha}^2_{\omega_2})-m\ln(\hat{\alpha}^2_N)-n\ln(\hat{\alpha}^2_S),$$

where

$$\hat{\alpha}^2_{\omega_2}=\hat{\alpha}^2_{\Omega_1},$$

$$\hat{\alpha}^2_N=\frac{1}{m}\left[\sum_{i=1}^{m}(\ln x_i)^2-\frac{1}{m}\left(\sum_{i=1}^{m}\ln x_i\right)^2\right],$$

$$\hat{\alpha}^2_S=\frac{1}{n}\left[\sum_{i=1}^{n}(\ln y_i)^2-\frac{1}{n}\left(\sum_{i=1}^{n}\ln y_i\right)^2\right].$$

If the parameter space in question is ω_2, then the approximate distribution of T_2 is also chi-square with 1 degree of freedom. In summary, statistic T_1 is used to examine scale parameter differences between two populations which, by assumption, have a common shape parameter. Statistic T_2 is used to investigate the assumption that the same two populations have a common shape parameter.

The two nonparametric tests include the 2-sample tie-adjusted Wilcoxon test (associated normalized test statistic denoted by T_3) and the 2-sample tie-adjusted sum-of-squared ranks test (associated normalized test statistic denoted by T_4). These tests have previously been discussed with respect to evaluating weather modification experiments (Grant and Mielke, 1967; Mielke, 1967, 1972; Duran and Mielke, 1968; Mielke *et al.*, 1970). Specifically, positive (negative) values of T_3 and T_4 are associated with increases (decreases) in the observations of a particular variable during the seeded storms.

TABLE 3. Statistical description of the measured variables.

Variable*	$\hat{\beta}_S/\hat{\beta}_N$	Test statistics T_1**	T_2**	T_3**	T_4**
CG flash frequency (11, 12)	0.45	2.21 (0.138)	3.72 (0.054)	−2.18 (0.029)	−2.46 (0.014)
IC flash frequency (11, 12)	0.18	6.80 (0.009)	5.96 (0.015)	−2.49 (0.013)	−2.36 (0.018)
Total flash frequency (11, 12)	0.27	6.93 (0.008)	0.73 (0.392)	−2.52 (0.012)	−2.49 (0.013)
Duration of lightning activity (time from first to last flash per storm) (11, 12)	0.54	7.58 (0.006)	1.08 (0.299)	−2.25 (0.025)	−2.27 (0.023)
Lightning intensity (ratio of total flash frequency to duration of lightning activity) (11, 12)	0.51	4.56 (0.033)	0.25 (0.618)	−1.97 (0.049)	−1.81 (0.071)
Max 5 min CG flash frequency (11, 12)	0.53	2.99 (0.084)	1.41 (0.234)	−2.41 (0.016)	−2.56 (0.010)
Max 15 min CG flash frequency (11, 12)	0.50	2.40 (0.121)	1.48 (0.224)	−2.22 (0.026)	−2.47 (0.014)
Max 5 min IC flash frequency (11, 12)	0.38	4.97 (0.026)	8.73 (0.003)	−1.88 (0.060)	−1.35 (0.177)
Max 15 min IC flash frequency (11, 12)	0.32	4.78 (0.029)	6.61 (0.010)	−1.91 (0.056)	−1.54 (0.124)
Max 5 min total flash frequency (11, 12)	0.44	4.83 (0.028)	3.70 (0.055)	−1.97 (0.049)	−1.56 (0.120)
Max 15 min total flash frequency (11, 12)	0.43	4.45 (0.035)	0.96 (0.328)	−2.00 (0.045)	−1.78 (0.074)
Return stroke frequency, per CG-discrete (118, 26)	1.02	0.01 (0.915)	0.38 (0.540)	0.09 (0.931)	−0.06 (0.951)
Return stroke frequency, per CG-hybrid (141, 25)	0.72	6.12 (0.013)	2.07 (0.150)	−2.85 (0.004)	−2.99 (0.003)
Return stroke interval, CG-discrete (interval from first to last return stroke) (96, 22)	0.74	3.07 (0.080)	0.28 (0.596)	−1.62 (0.105)	−1.33 (0.182)
Return stroke interval, CG-hybrid (134, 24)	0.66	2.82 (0.093)	0.24 (0.624)	−1.96 (0.050)	−1.94 (0.053)
Continuing current interval, CG-hybrid (LCC) (for continuing current intervals >40 ms) (134, 19)	0.75	3.69 (0.055)	0.071 (0.789)	−1.76 (0.039)	−1.78 (0.038)

* Values in parentheses are the number of non-seeded events, followed by the number of seeded events.
** p-values are shown in parentheses.

4. Results and discussion

The log-normal likelihood procedures and nonparametric tests discussed in Section 3 are used to analyze each of the variables described in Section 2. The results of these analyses are given in Table 3. Each p-value (probability of having a more extreme test statistic than the observed value under the null hypothesis that seeding has no effect) in Table 3 is two-sided (i.e., the seeding effect could be either an increase or a decrease). With two exceptions (return stroke frequency and interval for each discrete CG discharge), the observed seeding effect on each variable is an estimated decrease in which the two-sided p-values are at most 0.06 for at least two of the three cogent test statistics (T_1,T_3,T_4). This is strong evidence that seeding reduced both the frequency of lightning occurrence and the duration of lightning activity, and also altered characteristics of individual flashes.

In applications of test statistic T_1, the common shape parameter assumption appeared questionable for some discharge frequency data (five realized p-values for test statistic T_2 are less than 0.1, all involving discharge frequency data). Note that the return stroke frequency and interval for each discrete CG discharge and the return stroke frequency, interval and continuous current interval for each hybrid CG discharge are nested observations within the non-seeded and seeded storms. Thus, the non-seeded and seeded event frequencies corresponding to these observations differ from the non-seeded and seeded storm frequencies.

These results have strong implications for the application of cloud seeding to lightning fire prevention or reduction. Recent evidence indicates that forest fires in western Montana forests are ignited primarily by LCC discharges (Fuquay *et al.*, 1972). Further it is suggested that the probability of ignition increases proportionally with the continuing current duration (Fuquay, 1974). All three test statistics (T_1,T_3,T_4) indicate that seeding reduces the length of continuing current intervals for hybrid LCC flashes by about one-fourth (the two-sided p-values of T_3 and T_4 are both less than 0.04). In addition, seeding appears to reduce the frequency of CG flashes by more than one-half. We conclude that a properly conducted seeding program would very likely provide a means for selectively reducing forest fire ignitions in western Montana forests.

There is a close association between rainfall and lightning occurrence. A positive correlation between lightning frequency and precipitation amount has been reported (Battan, 1965; Kinzer, 1972). The significant reduction in lightning reported here (i.e., an observed reduction in the duration of lightning activity by nearly one-half with two-sided p-values of T_1, T_3, T_4 all less than 0.03) implies that seeding could decrease precipitation from the seeded cumulonimbus clouds. This possibility should receive consideration in situations where the amount of rainfall from seeded clouds is of importance.

REFERENCES

Barrows, J. S., 1966: Weather modification and the prevention of lightning-caused forest fires. *Human Dimensions of Weather Modification*, W. R. Sewell, Ed., Res. Paper 105, Dept. Geogr., The University of Chicago, 169–182.

Battan, L. J., 1965: Some factors governing precipitation and lightning from convective clouds. *J. Atmos. Sci.*, **22**, 79–84.

Duran, B. S., and P. W. Mielke, 1968: Robustness of sum of squared ranks test. *J. Amer. Statist. Assoc.*, **63**, 338–344.

Fuquay, D. M., 1967: Weather modification and forest fires. *Ground Level Climatology*, Washington, D. C., AAAS, 309–325.

——, 1974: Lightning damage and lightning modification caused by cloud seeding. *Weather and Climate Modification*, W. N. Hess, Ed., Wiley, 604–612.

——, A. R. Taylor, R. G. Hawe and C. W. Schmid, 1972: Lightning discharges that caused forest fires. *J. Geophys. Res.*, **77**, 2156–2158.

Grant, L. O., and P. W. Mielke, 1967: A randomized cloud seeding experiment at Climax, Colorado, 1960–65. *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 5, University of California Press, 115–131.

Kinzer, G. D., 1972: Cloud-to-ground lightning versus radar reflectivity in Oklahoma thunderstorms. NOAA Tech. Memo. ERL NSSL-59, 24 pp [Available from NTIS, COM-73-10050].

Mielke, P. W., 1967: Note on some squared rank tests with existing ties. *Technometrics*, **9**, 312–314.

——, 1972: Asymptotic behavior of two-sample tests based on powers of ranks for detecting scale and location alternatives. *J. Amer. Statist. Assoc.*, **67**, 850–854.

——, L. O. Grant and C. F. Chappell, 1970: Elevation and spatial variation effects of wintertime orographic cloud seeding. *J. Appl. Meteor.*, **9**, 476–488; corrigendum, **10**, 842.

——, and E. S. Johnson, 1974: Some generalized beta distributions of the second kind having desirable application features in hydrology and meteorology. *Water Resources Res.*, **10**, 223–226.

Uman, M. A., 1969: *Lightning*. McGraw-Hill, 264 pp.

Wilks, S. S., 1962: *Mathematical Statistics*. Wiley, 644 pp.