The Use of Network Lightning Data to Detect Thunderstorms near Surface Reporting Stations

RONALD M. REAP

Techniques Development Laboratory, Office of Systems Development, National Weather Service, NOAA, Silver Spring, Maryland

(Manuscript received 14 February 1992, in final form 5 June 1992)

ABSTRACT

Relationships between network lightning data and hourly thunderstorm observations were examined for the northeastern United States, Oklahoma, Florida, and the western United States to provide additional information on the possible effects of using lightning data to replace or supplement the hourly observations. Identification of thunderstorms for three of the four regions was found to agree closely with the hourly observations, provided the network reports were accumulated for a radius of 48 km or more about the station. The best agreement was found over Florida where high ground-flash densities resulted in a greater likelihood of both observer and network recording a given thunderstorm. In the immediate vicinity (8 km) of a station, use of lightning data from current national or regional networks would not provide observations comparable to the manual observations of thunderstorms due to the poor agreement between the two sets of observations at this radius. Selection of an 8-km radius would result in a decrease of nearly 75% in the number of thunderstorms detected by the network relative to that reported by the observer.

1. Introduction

The availability of cloud-to-ground lightning location data from automated networks across the continental United States has fostered the notion of using the lightning data to replace, or at least supplement, the hourly observations of thunderstorms taken at surface reporting stations. Recent meetings, such as the FAA Lightning Workshop (Federal Aviation Administration 1991), have focused interest on the problem of identifying appropriate geographical regions around surface reporting stations or airports where network lightning data could be used to replace or supplement the scheduled hourly observations. The purpose of this article is to provide additional information to aid in the formulation of appropriate algorithms for using lightning data to identify thunderstorm activity in the vicinity of surface stations and air terminals with a degree of accuracy acceptable to operational users, primarily commercial and general aviation interests. Such algorithms could, for example, transform the raw lightning data into an automated message or observation alerting users to the presence, approximate location, intensity, and movement of thunderstorms relative to a given station. The research described in this article is, in effect, an expanded version of a previous study by Reap and Orville (1990) on the relationships between network lightning locations and surface hourly observations of thunderstorms for the northeastern United States. In the present study, these relationships are examined for three different lightning detection networks and four separate geographical regions.

The renewed interest in accurate reporting of lightning near air terminals, as a hazard in itself and as an indicator of other convective hazards, is prompted by the fact that many of the newer aircraft contain composite materials and fly-by-wire systems making them more vulnerable to lightning strikes. Ground-based operations, such as aircraft refueling and baggage handling, are also extremely sensitive to lightning, and would greatly benefit from improved reporting of nearby thunderstorm activity.

2. Data collection

In the present study, lightning location data and surface hourly observations were examined for four geographical regions, namely, the northeastern United States, Oklahoma, Florida, and the western United States. Lightning data for the Northeast and Florida were obtained from the automated network that was operated by the State University of New York at Albany (SUNYA); data for Oklahoma were obtained from the National Severe Storms Laboratory (NSSL) experimental network; and data for the western United States were obtained from the Bureau of Land Management (BLM) network used for wildfire management. Each of the networks is a gated, wideband magnetic direction-finder system that determines the location of a

---

Corresponding author address: Ronald M. Reap, Techniques Development Laboratory, 1325 East-West Highway, Silver Spring, MD 20910.

TABLE 1. Summary of hourly observations and ground-flash counts by geographical region. Flash counts represent number of flashes occurring within an 80-km radius of surface station.

	Northeastern United States	Florida	Oklahoma	Western United States	Total
Flashes	1 735 826	2 597 648	837 098	905 244	6 075 816
Hourly observations	983 995	624 303	96 719	686 563	2 391 580
Number of stations	132	34	12	122	300
Sample years	1985–86	1987–90	1985–86	1983–84	—

ground strike by triangulating directions to strikes measured by two or more stations (Krider et al. 1976, 1980). Detailed information on the direction finders, network configuration and geographical coverage, detection efficiency, regional climatological characteristics of lightning, etc., can be obtained for the northeastern United States from Reap and Orville (1990), for the western United States from Reap (1986), and for Oklahoma from Reap and MacGorman (1989). Similar information for Florida, based on current research efforts, will be briefly described in this article.

The surface hourly observations of thunderstorms used in the present study were derived from hourly reports taken at National Weather Service (NWS), FAA, and military stations within the contiguous United States. Mean station spacing for the region east of the Rocky Mountains is about 115 km. All stations follow the NWS procedure where a thunderstorm is considered to begin either 1) when thunder is heard, or 2) when overhead lightning or hail is observed and the local noise level is such as might prevent hearing thunder. A thunderstorm is considered to have ended 15 min after the last occurrence of any of the above criteria. A "special observation" can be issued any time during the hour when thunder is first heard at the station. The data used in this study, however, contain only hourly observations that are generally taken at about 10 min before the hour. Therefore, the surface observations correspond to the period 10–25 min prior to the hour, accounting for the required 15 min elapsed time after the last peal of thunder is heard.

In the analysis procedure, the lightning data were extracted from a network lightning archive for the 25-min interval extending from 0 to 25 min prior to the hour. This interval extends slightly beyond the period of the hourly observations and to some extent takes into account the natural reluctance of observers to drop thunder from their observations when there is distant lightning beyond the audible range of thunder (Maier et al. 1984).

The data sample for the hourly observations and lightning location reports for the four geographical areas of interest are summarized in Table 1. The ground-flash counts represent the number of flashes occurring within an 80-km radius of the surface stations, the maximum radius used in our analysis. In all, 6 075 816 ground flashes and 2 391 580 hourly observations were evaluated for 300 reporting stations in the four regions of interest. The results and conclusions presented in section 4 are, therefore, based on a very large data sample.

3. Data analysis

The procedure for comparing the lightning location data and surface hourly observations requires the computation of statistical measures from 2 × 2 contingency tables constructed for each station. To construct the tables, lightning reports for a specific hour were scanned with respect to each surface station reporting during that particular hour in order to accumulate flash-count totals for a specified radius R from the station, where R ranges from 8 to 80 km. In effect, flash-count totals were accumulated for circular regions of varying radius centered at each station. Finally, contingency tables were constructed for each radius. A sample 2 × 2 contingency table is shown in Table 2. Composite contingency tables were then constructed by summing the x, y, z, and w values in Table 1 for *all stations in a particular region.* From the composite tables, we were able to plot and examine relationships based on the regional data samples.

4. Results and conclusions

The results and conclusions presented in this section apply to observational data obtained from the lightning networks under study. It is possible they might not apply to other types of lightning locating systems or network configurations. Results of our comparison of the lightning data and hourly surface observations are shown in Fig. 1, where the relative frequency $F(O)$ for the four geographical regions expresses the fraction of observer-reported thunderstorms that also had two or more coincident lightning flashes recorded by the network within a given radius of the observing station. Two or more flashes were used because isolated, single

TABLE 2. Contingency table used in statistical analysis.

Hourly observations	Network		
	Lightning	No lightning	Total
Thunderstorm	x	y	$x + y$
No thunderstorm	z	w	$z + w$
Total	$x + z$	$y + w$	$x + y + z + w$

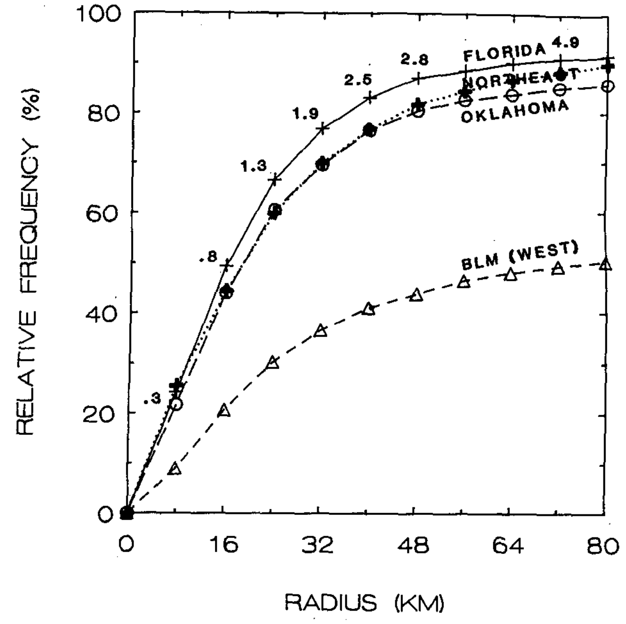

FIG. 1. Relative frequency $F(O)$ as a function of distance from station. Here, $F(O)$ is the fraction (%) of observer-reported thunderstorms that were also recorded by the network. The ratio of network-reported thunderstorms to observer-reported thunderstorms is given by the numbers plotted next to the data points.

flashes can be unreliable indicators of thunderstorm occurrence (Reap and Orville 1990). In any case, the results obtained from duplicate computer runs using a threshold of one or more flashes did not significantly differ from those based on two or more flashes. The relative frequency in Fig. 1 is computed from $F(O) = x/(x + y)$, where x and y are variables from Table 2. Except for the BLM network, Fig. 1 shows a good correspondence between thunderstorms observed at the station and ground strikes recorded by the network, provided a radius of 48 km or more is taken about the station. In effect, the larger radii take into account the errors and biases inherent in both sets of observations, as discussed in detail by Reap and Orville (1990). Similarly, Fig. 2 gives the relative frequency $F(L)$ for the four geographical regions, where $F(L)$ is the fraction of cases where the lightning network recorded two or more ground flashes and the observer also reported a thunderstorm. The relative frequency in Fig. 2 is computed from $F(L) = x/(x + z)$, where x and z are variables from Table 2. Except for the BLM network, the curves in Fig. 2 are very similar and indicate a significant decrease of $F(L)$ with range. This result is not surprising since the observer encounters greater difficulty in identifying more distant thunderstorms. Additional details on the variations of $F(O)$ and $F(L)$ with respect to varying flash thresholds, time-of-day, flash density, and peak signal strength for the Northeast can be obtained from Reap and Orville (1990).

At various workshops and meetings concerned with the possibility of replacing the surface thunderstorm observations with lightning data, it has been suggested that lightning data for relatively small radii (8 km) be used to indicate a thunderstorm in the immediate vicinity of the station or air terminal. Except for the BLM network, Fig. 2 would seem to support this idea since it indicates fairly good agreement at the 8-km radius where 72%–76% of the network-detected thunderstorms were also recorded by the observer. As shown in Fig. 1, however, the lightning networks indicate two or more ground strikes within an 8-km radius in only 22%–24% of the cases where an observer at the station reported a thunderstorm. The poor agreement in Fig. 1 at the 8-km radius is due, in part, to the mean lightning location error for the networks, which is estimated to be about 5 km for the SUNYA and NSSL networks and considerably larger for the BLM network. In addition, a significant number of the observer-reported thunderstorms lie beyond the 8-km radius for which the lightning reports were accumulated since this radius is well below the audible range of thunder (20 km) and the visual range of lightning. This is especially true at night when the observer is better able to see more distant lightning against the dark background sky (Reap and Orville 1990). Sole reliance on lightning data obtained from current networks to indicate a thunderstorm in the immediate vicinity (8 km) of the station or terminal would, therefore, result in a decrease of nearly 75% in the number of thunderstorms relative to that reported by the human observer. From the standpoint of aviation safety, this trade-off could possibly cause operational problems. In contrast, at 48 km

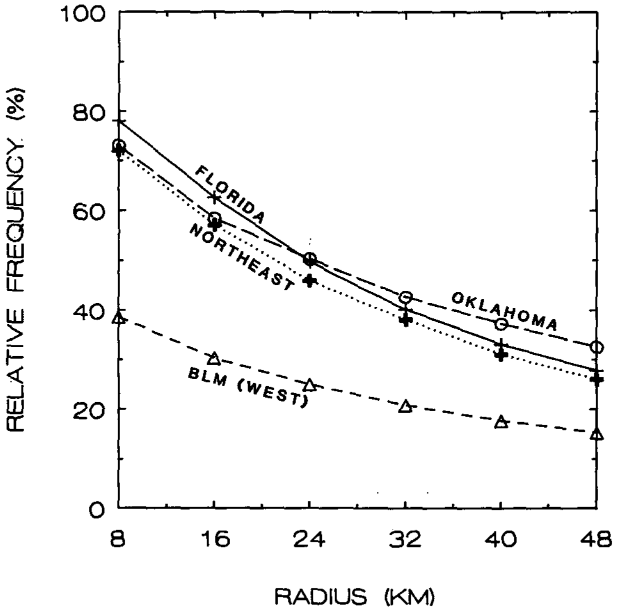

FIG. 2. Relative frequency $F(L)$ as a function of distance from station. Here, $F(L)$ is the fraction (%) of network-reported thunderstorms that were also recorded by the observer.

(30 mi) the agreement between the observer and network data is between 80%–88% for most networks, values that are more likely to be accepted for operational use by the aviation community.

The remaining 12%–20% of the cases in Fig. 1 where lightning is not detected by the networks at the 48-km radius could result, in part, from the occurrence of storms with only intracloud lightning or weak cloud-to-ground strikes not detected by the network. These cases could also be related to the fact that "end-of-thunder" times in the station records tend to lag the actual lightning (Maier et al. 1984). Observers are usually very busy during thunderstorms and often forget to close out the thunderstorm observation due to the press of other duties or find it difficult to remember the time of the last thunder in anticipation of it being the last. Observers might also be hesitant to regard a given thunder as being the last when distant lightning is observed beyond the audible range of thunder.

The variations between networks in Fig. 1 are also of significant interest and represent one of the principal findings in this article. In general, the results in Fig. 1 confirm and expand upon the basic results presented by Reap and Orville (1990) as represented by the curve in Fig. 1 for the northeastern United States. In a sense, the curves for the BLM and Florida regions can be viewed as representing both "worst-case" and "best-case" scenarios, respectively, with the curves for the Northeast and Oklahoma lying slightly below that for Florida. The 1983–84 data from the BLM network were obtained early in the network's operation before extensive site error corrections were applied to the data and all current direction finders were in place. The BLM network is primarily operated in support of wildfire detection and suppression activities in the western United States and has proven extremely valuable in this regard (Krider et al. 1980). Direction-finder spacing is, however, quite large compared to the other networks, with a resulting loss in detection efficiency within the BLM network. Detection efficiency is defined as the percentage of the total number of cloud-to-ground strikes that the network records, and is a function of range. Due to the nature of the BLM operation, resources were not available to effect the rigorous quality control evident in the SUNYA and NSSL networks, which were designed primarily for research and development purposes. The BLM network has been upgraded since the 1983–84 period for the data used in the present study. Results from the current network could, therefore, possibly change from that shown in Figs. 1 and 2. In addition, most of the lightning activity within the BLM network tends to occur over elevated terrain where few surface stations are located (Reap 1986). The low thunderstorm frequency over the valleys, where most of the surface stations are located, coupled with relatively large location errors and reduced detection efficiency resulting from the wide spacing of direction finders, acts to degrade the relationships in Figs. 1 and 2 for the BLM network relative to the other networks.

In contrast, the curve for Florida in Fig. 1 represents a best-case scenario in the sense that it is probably close to the upper limit of the accuracy expected from the networks under study. Direction-finder spacing within the SUNYA network in Florida is fairly small, with six direction finders located within the state. In addition, ground-flash densities in parts of Florida are the highest in the country (Orville 1991). As shown by Reap and Orville (1990), the frequency that a thunderstorm is reported by an observer could increase by as much as 35%–40% for storms producing ten or more flashes within an 8-km radius of the station. In effect, regions with high flash accumulations tend to exhibit a higher correlation between surface stations and the lightning network in terms of locating thunderstorms near the station. This phenomenon is clearly illustrated in Fig. 3a, where the individual station relative frequencies $F(O)$ are plotted for circular regions with a radius of 8 km about each station. Comparison with Fig. 3b, the cumulative ground-flash density (flashes per square kilometer) for the 1987–90 warm seasons, indicates a similar pattern in both fields, especially near Tampa, Florida. Analysis of daily thunderstorm relative frequency, not shown here, reveals that the number of thunderstorm days for Tampa and vicinity is similar to that found in many areas in central and south Florida. The ground-flash maximum near Tampa in Fig. 3b, therefore, represents greater flash activity per thunderstorm day in this region, which, as noted by Reap and Orville (1990), acts to increase the likelihood of both the observer and network recording a given thunderstorm. Despite the near-optimum conditions found in Florida, the agreement between observer-reported thunderstorms and lightning detected by the network is still quite low at the 8-km radius (Fig. 1). The agreement improves to about 50% at the 16-km radius, but does not reach values of 80% or more until 40 km.

The curves in Fig. 1 for the Northeast and Oklahoma are probably typical of the relationship between the hourly surface observations and lightning data that exist over most of the country. The curves are almost identical up to about 48 km, beyond which the northeastern curve indicates slightly better agreement. This departure is probably due to the fact that some of the surface stations in Oklahoma used in our analysis lie outside of the network formed by the four NSSL direction finders, thereby slightly degrading the results at the longer ranges.

The ratio of network reports to observed thunderstorms with increasing distance is also of interest and is plotted along the curve for Florida in Fig. 1. Values for the Northeast and Oklahoma are not plotted, but are very similar. At radii of 48–80 km, over three-to-five times as many thunderstorms are detected by the lightning network. This increase is not particularly surprising since the region sampled at 50 km, for ex-

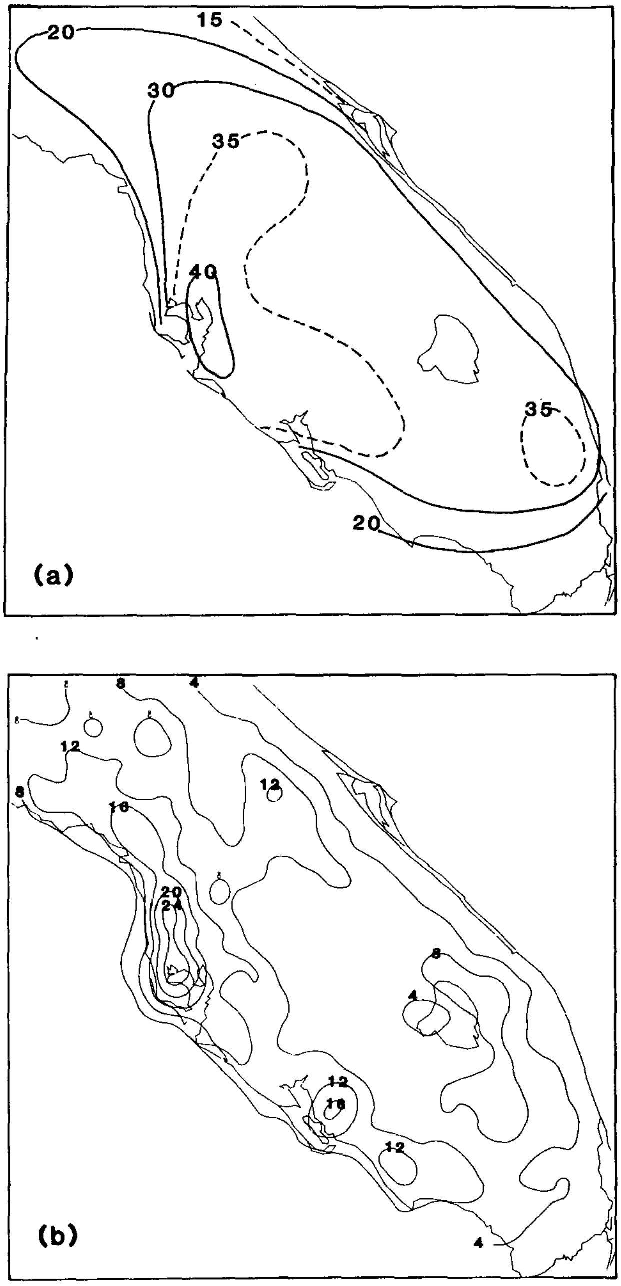

FIG. 3. (a) Relative frequency $F(O)$ over central and south Florida for circular regions with 8-km radius centered at stations, where $F(O)$ is the fraction (%) of observer-reported thunderstorms that were also recorded by the network; (b) ground-flash density (flashes per square kilometer) for all flashes in the 1987–90 warm-season sample with no correction for system detection efficiency.

ample, is about six times larger than that at 20 km, the upper limit for thunder audibility. Only at a radius of about 20 km, where the ratio equals unity, are the two observing systems compatible in the sense that they report, on average, the same number of thunderstorms.

5. Summary

The relationships between observed thunderstorms and lightning data were examined for four geographical regions within the contiguous United States. In a data sample containing 6 075 816 ground flashes and 2 391 580 hourly observations, identification of thunderstorms by network lightning data for three of the four regions was found to agree closely with the hourly observations taken at surface stations, provided the network reports were accumulated for a radius of 48 km or more about the station. The continuing improvement in the relationships beyond the audible range of thunder (20 km) was most likely related to the errors, biases, and uncertainty inherent in both sets of observations. The best agreement between the hourly observations and lightning data was found over central and south Florida, where a relatively dense network of direction finders coupled with high ground-flash densities in many areas resulted in a greater likelihood of both observer and network recording a given thunderstorm.

Use of lightning data from current operational networks to indicate thunderstorms in the immediate vicinity (8 km) of a station or terminal would not provide observations comparable to the manual observations of thunderstorms due to the poor agreement between the two sets of observations at this radius. Selection of an 8-km radius would result in a decrease of nearly 75% in the number of thunderstorms detected by the network relative to that reported by the observer under the current system. At 48 km, the agreement between the network and observer increased to 80%–88%, values that are much more likely to be accepted by the aviation community. Future networks with much smaller direction-finder spacings, such as the experimental network at the Kennedy Space Flight Center in Florida (Lopez and Holle 1986), could possibly provide the necessary resolution at the 8-km radius. It is also possible that improved single-site lightning sensors at air terminals or surface stations could result in improvements at the shorter ranges (8 km) that would allow sole use of lightning data in place of the surface observations.

Acknowledgments. Special thanks are extended by the author to K. Demerjian and R. Pyle of the Thunderstorm Analysis Center, State University of New York at Albany (SUNYA), for providing the lightning data for Florida, to D. MacGorman of the National Severe Storms Laboratory (NSSL) for providing the lightning data for Oklahoma, and to the BLM for providing the lightning data for the western United States used in this study.

REFERENCES

Federal Aviation Administration, 1991: *Proc. FAA Lightning Workshop,* Washington, D.C., FAA, 29 pp. [Available from J. F.

Canniff, DOT/VNTSC, DTS-53, Kendall Square, Cambridge, MA 02142.]

Krider, E. P., R. C. Noggle, and M. A. Uman, 1976: A gated wideband magnetic direction finder for lightning return strokes. *J. Appl. Meteor.*, **15,** 301–306.

——, ——, A. E. Pifer, and D. L. Vance, 1980: Lightning direction finding systems for forest fire detection. *Bull. Amer. Meteor. Soc.*, **61,** 980–986.

Lopez, R. E., and R. L. Holle, 1986: Diurnal and spatial variability of lightning activity in northeastern Colorado and central Florida during the summer. *Mon. Wea. Rev.*, **114,** 1288–1312.

Maier, L. M., E. P. Krider, and M. W. Maier, 1984: Average diurnal variation of summer lightning over the Florida Peninsula. *Mon. Wea. Rev.*, **112,** 1134–1140.

Orville, R. E., 1991: Lightning ground flash density in the contiguous United States—1989. *Mon. Wea. Rev.*, **119,** 573–577.

Reap, R. M., 1986: Evaluation of cloud-to-ground lightning data from the western United States for the 1983–84 summer seasons. *J. Climate Appl. Meteor.*, **25,** 785–799.

——, and D. R. MacGorman, 1989: Cloud-to-ground lightning: Climatological characteristics and relationships to model fields, radar observations, and severe local storms. *Mon. Wea. Rev.*, **117,** 518–535.

——, and R. E. Orville, 1990: The relationships between network lightning locations and surface hourly observations of thunderstorms. *Mon. Wea. Rev.*, **118,** 94–108.