Surrogate Data to Estimate Crop-Hail Loss

DAVID CHANGNON

Department of Geography, Northern Illinois University, DeKalb, Illinois

STANLEY A. CHANGNON

Changnon Climatologist, Mahomet, Illinois

(Manuscript received 5 August 1996, in final form 1 March 1997)

ABSTRACT

Crop-hail insurance loss data for 1948–94 are useful as measures of the historical variability of damaging hail in those 26 states where most crop damages occur. However, longer records are needed for various scientific and business applications, as well as information on potential losses in United States' areas without crop insurance. The long-term (1901 to present) data on hail-day incidences, as derived from National Weather Service historical station records, were investigated to determine if some form of a hail-day expression related well to the insurance losses. The areal extent of insured areas of Illinois, Texas, and Nebraska experiencing growing season frequencies of hail days matching or exceeding the once in 10-yr frequencies was found to have the best relationship with insured loss values. The computed correlation coefficients were +0.97 for Illinois, +0.73 for Texas, and +0.91 for Nebraska. These values appear to be a useful surrogate for 1) estimating pre-1948 loss values, 2) estimating loss values in areas with no insurance, and 3) further research involving other states with different crop and hail conditions.

1. Introduction

A problem faced by the crop insurance industry, and anyone seeking to assess the temporal frequency of severe weather, is the availability of long-term quality records of severe weather and the losses they create (Changnon et al. 1996). Knowledge of average and extreme conditions is paramount for designing adequate risk assessments for insurance rate making and for assessing climatic variability (Walter 1993). The crop-hail loss records objectively collected by the crop insurance industry are from 1948 to the present, but this is less than desired for many applications.

The goal of this study was to determine if there were some surrogate data that was closely related to the existing crop-hail loss data for the 1948–present period and that existed across the United States for years before and after 1948. If so, the surrogate data could be used 1) to estimate loss values before quality insurance records began in 1948, and 2) to estimate loss levels in areas where new crops were being introduced and no insurance records existed. Derived crop-hail loss values for years prior to 1948 could be combined with the 1948–94 values to assess whether recent values were more extreme than anything that had occurred in recorded history.

Past climatological studies of severe weather teach lessons about selecting data to be analyzed for such comparative purposes (Changnon 1982; Changnon et al. 1988). It is very difficult to obtain measures of the intensity of large storm events based on quantifiable physical measures such as their patterns of high wind speeds, the spatial distribution of hailstone sizes across large areas, or the areal extent of heavy rainfall amounts. This situation makes a physical comparison of storms distributed over many decades impossible. Furthermore, an important aspect of severe storms is that they are defined in two ways: 1) by their type (hurricane, hail, tornado, etc.) and physical characteristics (wind speed, size, duration, etc.), and 2) by the amount of damage they cause.

The 1948–94 crop-hail insurance data are valuable because they integrated hail activity with actual loss. However, incorporation of storm damages brings another uncertainty about past events: the ever-changing value of the dollar, changing crop strains, and varying amounts of insurance liability in an area. For many years, NOAA has accumulated data, on a state-by-state basis, of damages from severe weather in an annual publication entitled *Storm Data*. Unfortunately, data quality varies between states, and the amount of loss per event is expressed in only five broad classes, all factors that make it impossible to have precise numbers.

Corresponding author address: Dr. David Changnon, Meteorology Program, Department of Geography, Northern Illinois University, DeKalb, IL 60115.
E-mail: changnon@geog.niu.edu

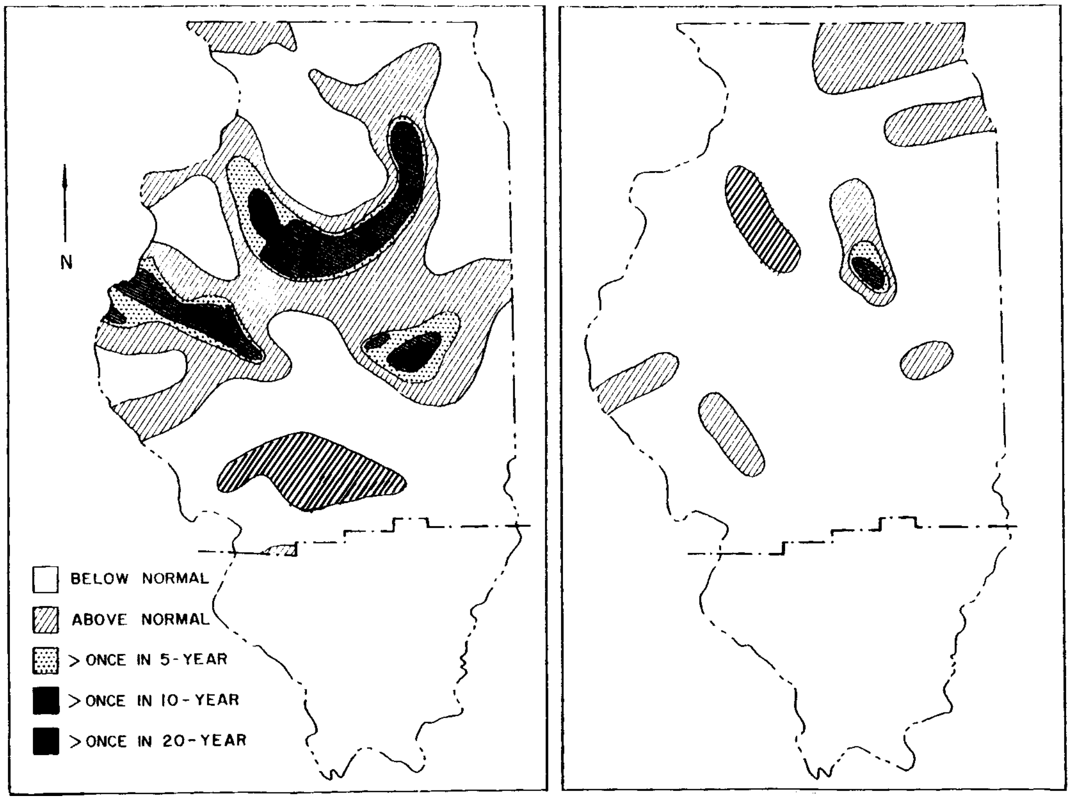

FIG. 1. Areas of Illinois with extensive insurance (located north of the dashed line) experiencing various hail-day values during the June–August period of 1953 (left) and 1970 (right).

Fortunately, the insurance industry began keeping detailed loss and liability records at the county level for the entire United States in 1948. The industry has developed an index that adjusts the losses to changing liability.

2. Data and analysis

The crop-hail insurance industry began a systematic collection of data on hail and wind losses to crops in 1948 (Roth 1949). The crop-hail insurance firms in the nation established and have since maintained an association that had the mission of collecting each firm's annual insurance data by location, the sales of premiums, amount of liability, and losses, and then summarizing these data for each county in each state (NCIS 1995). This database has been maintained by the National Crop Insurance Services (NCIS) and its predecessor agency, the Crop-Hail Insurance Actuarial Association (CHIAA). Together they have accumulated the daily and annual crop-hail loss data for each state since 1948 (CHIAA 1978). These data are based on at least 90% of all the crop-hail insurance coverage in the United States in each year.

TABLE 1. Hail values for selected years in Illinois. The areal extent of above average, the 5-yr, 10-yr, and 20-yr hail-day values are expressed as percentages of the insured area. Also shown are the annual loss cost values based on crop-hail damages.

Year	Above average	5 yr	10 yr	20 yr	Loss cost ($)
1953	46.7	13.4	7.7	1.1	3.02
1954	43.2	15.7	8.5	1.0	2.54
1955	15.3	0.8	0.5	0	0.44

The insurance data have also been adjusted by the NCIS for annual changes in liability (coverage), dollar values, and other factors by calculating the ''loss cost.'' The annual loss cost value, expressed in dollars per $100 of liability, for a given county or state (and a given crop or all crops grown in the state) is determined by dividing the annual losses by the annual liability and multiplying the resultant value by 100. The losses are defined as the dollar amount paid for loss under the policy conditions, while the liability is the dollar amount of insurance placed on the insured crop (Changnon et al. 1996). For example, a loss cost of $2.40 means that for the area and liability affected, the amount of $2.40 was paid for losses for each $100 of liability in force. Adjusting loss by amount of liability, or insured coverage, handles the effects due to changing dollar values, crop strains, and fluctuations in coverage, and makes the loss cost a valuable expression for assessing hail loss changes with time (Changnon and Changnon 1990).

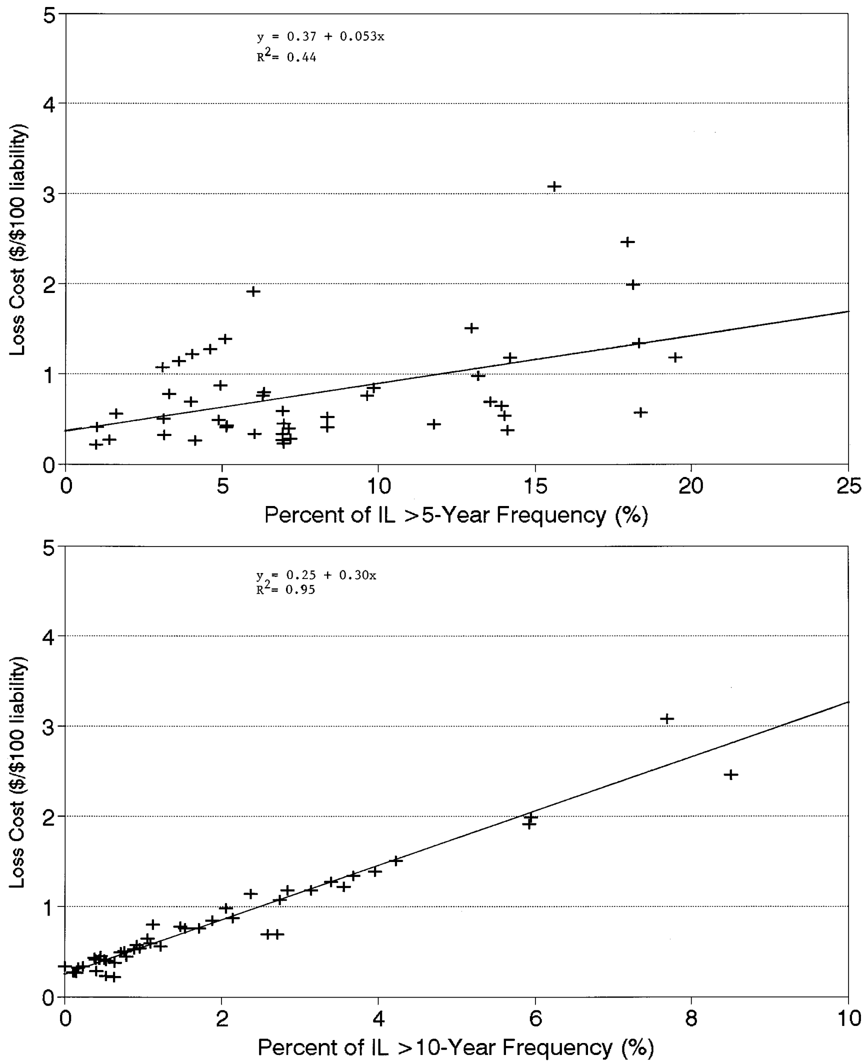

FIG. 2. The relationship between percent area of Illinois with extensive insurance experiencing more than (a) a 5-yr, (b) a 10-yr, and (c) a 20-yr hail-day frequency vs loss cost (dollars per $100 liability) for 1948–94.

The only form of historical hail data that could serve as a surrogate for crop-hail loss data are the hail-day records of the stations of the National Weather Service (NWS). The manned first-order stations have made routine measurements of hail occurrences since 1900. The volunteer observers of the cooperative substations in the nation were also asked to record days with hail, but many chose not to do so. A technique for assessing the quality of substation hail records was devised (Changnon 1967a). Application of this technique to substation records for 26 states where most of the nation's crop-hail losses occur showed that between 40 and 80 substations in each state had quality hail-day records that lasted 30 or more years (Stout and Changnon 1968). Thus, a reasonably dense array of stations is available for defining the incidence of hail across a state and for a long period of years dating back to 1901.

State loss cost values were compared with the frequency of hail days, as defined by observers at the NWS during the 1948–94 period. Various statistical measures of the growing season hail days were tested to determine if any measure related well to the loss cost values. The hypothesis underlying such a relationship was that the number of hail days across a state was related to the magnitude of the crop-hail loss in the growing season; that is, more hail days created more hail loss. Past studies of hail in Illinois supported this concept by showing that although 75% of the annual crop-hail loss occurred in 20% of the hail days, the number of extremely damaging days was closely linked to the number of days with hail (Changnon and Fosse 1981).

Data for three states with different hail regimes and different crops were selected for analysis: Illinois, Nebraska, and Texas. For each hail station and the growing

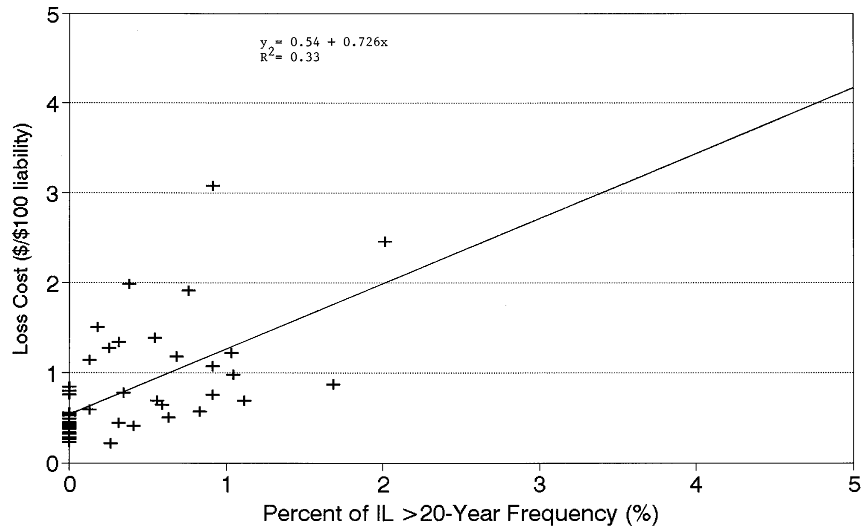

FIG. 2. (*Continued*)

season for the primary hail-damaged crops in a given state, the average number of hail days and the once in 5-, once in 10-, and once in 20-yr hail-day values were calculated. The once in 5-, 10-, and 20-yr values were arbitrarily chosen for this study; however, in hydrologic applications these values represent standard recurrence interval levels. Then for each year, the number of hail days at each station was expressed as meeting any or all of these four levels. The station (point) hail-day data were plotted on a state map for each year so that the areal extent of each condition, say the area with above average values, could be expressed using the Thiessen method for determining areas (Conrad and Pollak 1950). Maps for two years (Fig. 1) illustrate these areas for two quite different June–August periods, 1953 and 1970, in Illinois. There are notable differences in the areal extent of the four levels of hail frequency. The well-insured area is north of the line across the southern end of Illinois.

The areal extent of each category was measured and expressed as a percent of the state's insured area. For example, this showed that in 1953 (Fig. 1), 46.7% of the heavily insured area of Illinois had above-average numbers of hail days, 13.4% had values equal to or in excess of the once in 5-yr values, 7.7% had 10-yr values, and 1.1% had 20-yr or greater values. These four values were determined for each year during 1948–94 and were compared with the annual loss cost values.

3. Results

a. Relationship of loss costs and hail days in Illinois

The frequency of hail days for the June–August period was chosen for analysis because 96% of all reported crop-hail damage in Illinois occurs in these three months (Changnon 1967b). Since 98% of the insurance coverage is sold in the northern three-fourths of the state, the frequency of hail days was analyzed for this area of Illinois. This area comprises 44 200 square miles, located north of an east–west line through southern Illinois (Fig. 1). The various measures of the areal frequency of June–August hail days in this area were tested against the annual loss costs for the 1948–94 period. Table 1 presents the values for three years to illustrate the range of values found and the loss costs.

The series of 47 values of loss cost (1948–94) were compared with the 47 values for each of the four hail-day expressions (percentages) using Pearson's product moment correlation. The correlation coefficients were +0.63 for the above-average areas, +0.66 for the 5-yr frequency areas, +0.97 for the 10-yr areas, and +0.57 for the 20-yr areas (Figs. 2a–c). The poor correlation for the 20-yr areas was because many years had no 20-yr or greater values (a once in 20-yr event occurs less than three times on average in 47 yr). The strongest relationship was found to exist between annual loss cost and the amount of area of Illinois that experienced, in a given June–August period, hail days matching or exceeding the 10-yr frequency. Previous findings by Changnon (1977) and Changnon and Fosse (1981) indicated that the crop-hail damage occurred during a small number of hail days and that once a specific location exceeded its once in 10-yr number of hail days, the probability that an extremely damaging hail day occurred was greater than 70%. The pattern based on the 10-yr values at the 44 stations with long hail records is shown in Fig. 3. It reveals, for example, that in the Peoria area, the 10-yr frequency is slightly less than two hail days. In the higher incidence areas in southwestern and northwestern Illinois, the stations had values in excess of three hail days expected to occur once every 10 years.

A Pearson product moment correlation coefficient of +0.97 for the 10-yr values indicates a strong relation-

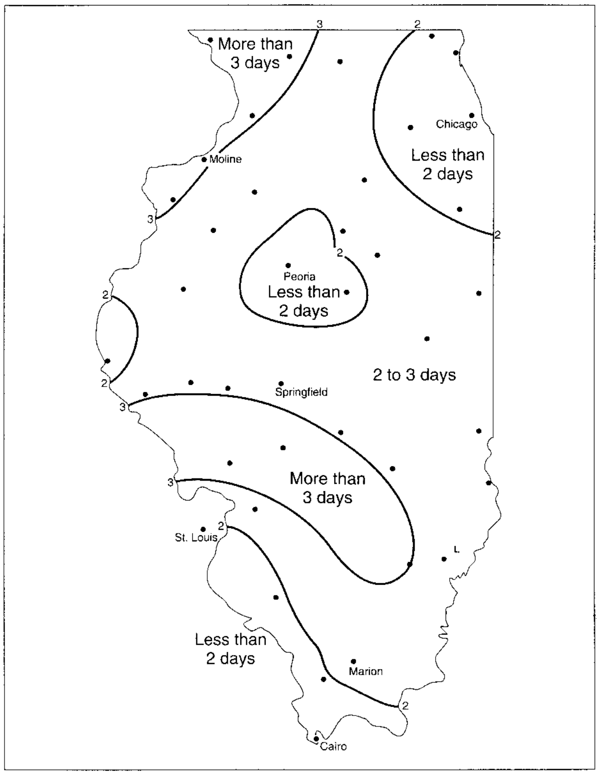

FIG. 3. The pattern based on the number of June–August hail days expected to occur at least once every 10 years at a point in Illinois.

ship, and the square of this coefficient is 0.95. This indicates that the areas with hail days equaling or exceeding the 10-yr values explain 95% of the variability found in the annual loss cost values in Illinois during the 1948–94 period (Fig. 4). The 47-yr period was divided into two periods, 1948–71 and 1972–94, to discern whether differences in this relationship existed for shorter time periods. Large differences in the correlation coefficients could be indicative of different severe weather regimes within the 47-yr period. The correlation coefficient for the early period was +0.96 and for the later period +0.99, indicating the relationship between the two variables was only slightly different when examined over shorter periods. A regression equation (below) based on this relationship was then used to calculate the annual loss costs for the 1901–47 period.

$$\text{annual loss cost} = 0.25 + 0.30\,(\%\text{ of insured area with } > 10\text{-yr hail days}). \quad (1)$$

There were eight years with large loss costs (greater than \$1.50) during the early 47-yr period (1901–47), while only five years since 1947 have experienced large loss costs (Fig. 4). The average loss cost for the 1901–47 period was \$0.79, and that for 1948–94 was nearly identical, \$0.81, indicating that the mean loss cost had not changed when comparing the two long periods.

b. *Relationship of loss costs and hail days in Texas*

The study of hail days and insurance-derived loss costs in Texas was identical to that done for Illinois. After assessing the crop-hail loss records for Texas, the March–November frequency of hail days was chosen for analysis. The cotton crop, which is Texas' primary target for hail damage, can experience hail damage from early March until late November. The crop-hail insur-

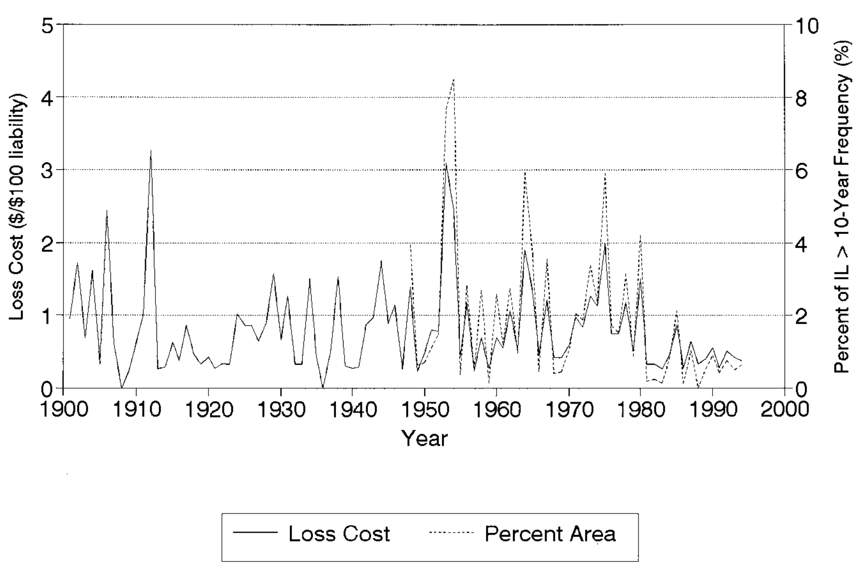

FIG. 4. Illinois annual crop-hail loss costs: 1901–47 calculated and 1948–94 observed, and 1948–94 percent area of Illinois with extensive insurance experiencing more than a 10-yr hail-day frequency.

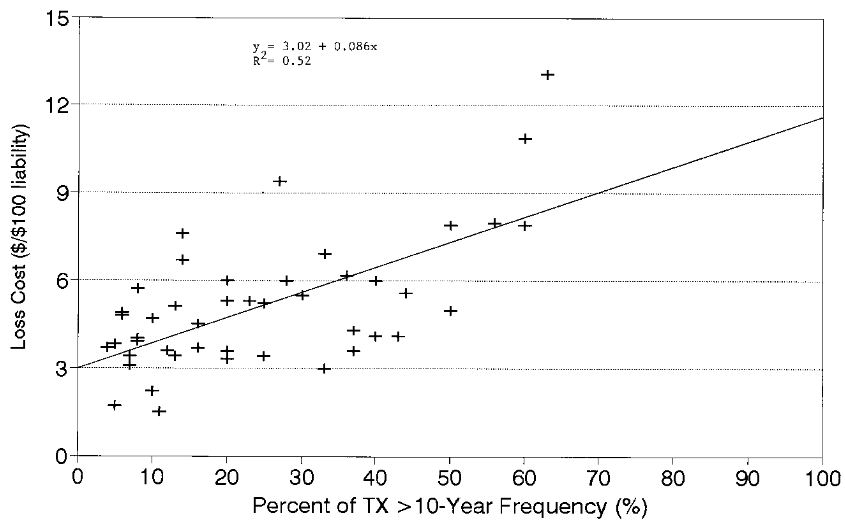

FIG. 5. The relationship between percent area of Texas with extensive insurance experiencing more than a 10-yr hail-day frequency vs loss cost (dollars per $100 liability) for 1948–94.

ance coverage is concentrated in 51 counties of Texas with 41 in and near the Panhandle, 4 in the San Antonio area, and 6 in the extreme south near Brownsville. Data from the 28 stations with quality hail data within these regions were mapped for each year, and the areal extent of the hail-day frequencies were expressed as a percent of the heavily insured area.

The analysis of the relationship of the Texas loss cost values for 1948–94 and the hail-day values involved the area with above-average values and the areas with 5-, 10-, and 20-yr frequencies, as in Illinois. The resulting correlations showed the above-average value coefficient was +0.48, the 5-yr was +0.61, the 10-yr was +0.73, and the 20-yr was +0.56. As in Illinois, the 10-yr values had the best relationship with loss costs. However, as is shown in Fig. 5, the Texas data appear to be much more scattered than that for Illinois (Fig. 2c). The 10-yr hail-day values in Texas (Fig. 6) are much greater than those in Illinois (Fig. 3). For example, at Lubbock the 10-yr frequency value is seven hail days, and in the high hail incidence areas of north Texas most stations have values in excess of five hail days.

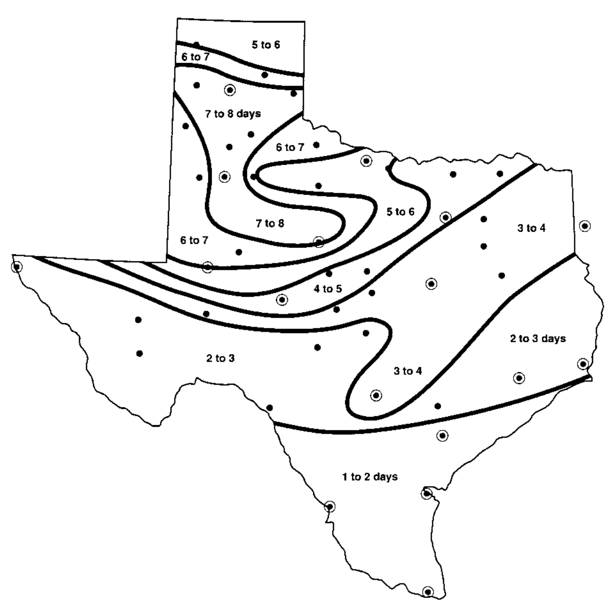

FIG. 6. The pattern based on the number of March through November hail days expected to occur at least once every 10 years at a point in Texas.

The annual loss cost and once in 10-yr hail-day values for Texas are plotted in Fig. 7. Examination of their fluctuations from 1948 to 1994 shows a reasonably good relationship. For example, the loss costs were exceptionally high in 1949, 1979, 1982, 1987, 1989, and 1992, and the percent of the heavily insured area of Texas experiencing extreme numbers of hail days was also quite high, 50% or more of the insured area. However, in four years (1950, 1962, 1968, and 1986) the losses were high and the area of hail days relatively low. The correlation coefficient of +0.73 indicates that the hail-day values at the 10-yr level (expressed as a percent of the insured area) explain 52% of the variability found in the Texas loss costs. Splitting the 1948–94 period into two smaller periods (1948–71 and 1972–94) the correlation coefficients determined from the relationship of the two variables were +0.69 for the early period and +0.76 for the later period. Because the correlation coefficients are weaker than those found for Illinois, it indicates that other unknown variables influence the loss cost values in Texas. Despite the weaker correlation coefficients, this is a sufficiently good relationship to justify developing a regression equation for predicting loss costs using the areal extent with 10-yr frequencies of hail days.

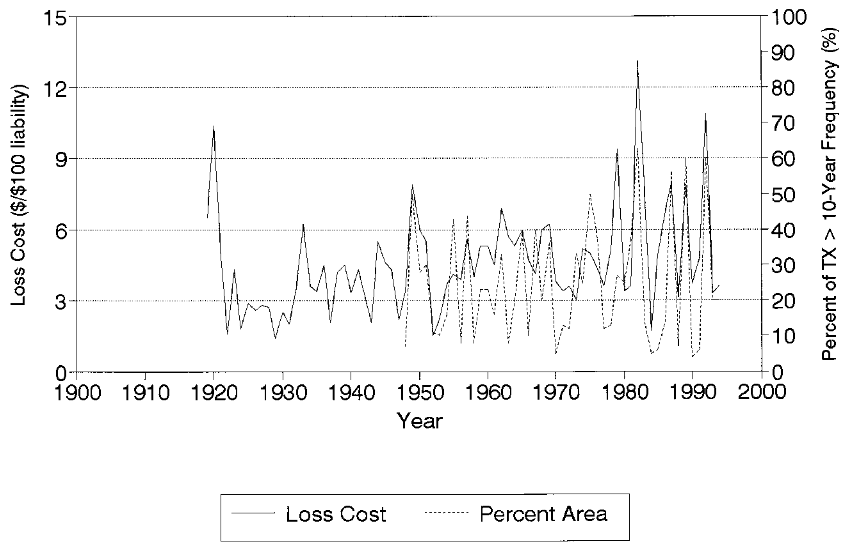

FIG. 7. Texas annual crop-hail loss costs: 1919–47 calculated and 1948–94 observed, and 1948–94 percent area of Texas with extensive insurance experiencing more than a 10-yr hail-day frequency.

annual loss cost

$$= 3.02 + 0.086$$

$$(\% \text{ of insured area with } >10\text{-yr hail days}). \qquad (2)$$

The average loss cost value for the pre-1947 period was $3.72, while the average value in the period since 1947 was $5.11 (Fig. 7). This increase could be related to changes in the types of severe weather that are related to the production of hail; however, further analysis is required to confirm this notion.

c. Relationship of loss costs and hail days in Nebraska

A comparable analysis was done for Nebraska. Ninety-one percent of the state has extensive crop-hail insurance coverage, and the primary period for crop-hail losses (wheat and corn) is May through September. Quality hail data were found for 43 stations in the well-insured areas, and the areal extent of above-average and 5-, 10-, and 20-yr hail-day frequencies for the May–September period was plotted and mapped for each year during 1948–94. The map of the 10-yr hail-day frequencies (Fig. 8) shows values generally increasing from an average of four along the Missouri River to more than nine in the panhandle. These values were compared with the annual loss costs. The correlation analysis produced coefficients of +0.52 for above-average areas, +0.71 for 5-yr areas, +0.91 for 10-yr areas, and +0.59 for the 20-yr areas. Figure 9 presents a graph with the 47 10-yr values and the loss costs, revealing a strong relationship. Again, when the 1948–94 period is split, the correlation coefficient of the later period was determined to be slightly higher when compared to the

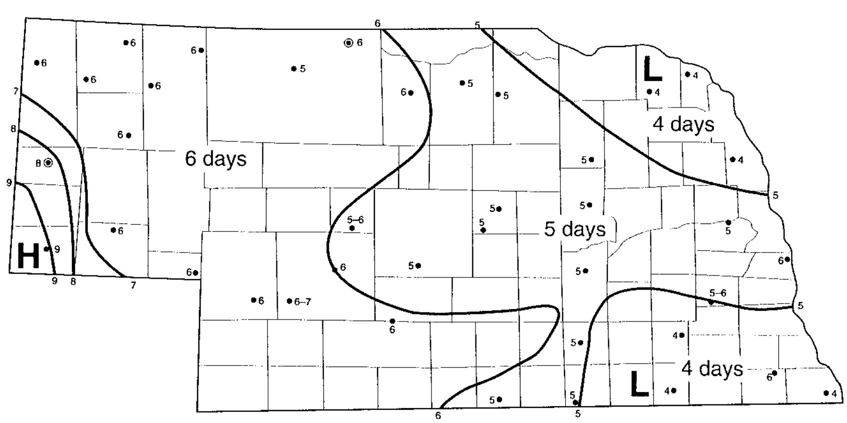

FIG. 8. The pattern based on the number of May through September hail days expected to occur at least once every 10 years at a point in Nebraska.

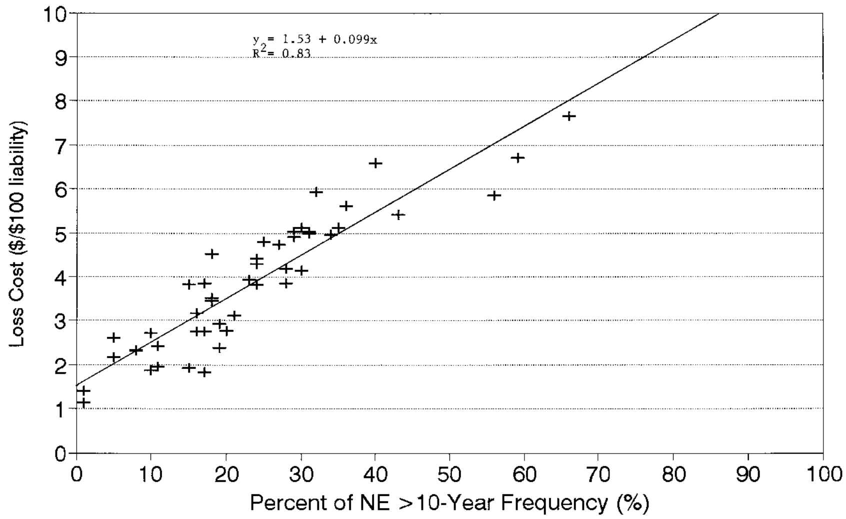

FIG. 9. The relationship between percent area of Nebraska with extensive insurance experiencing more than a 10-yr hail-day frequency vs loss cost (dollars per $100 liability) for 1948–94.

early period (+0.95 vs +0.86). The 10-yr values explained 83% of the variation found in the loss costs over the 47-yr period. The equation expressing their relationship is

$$\text{annual loss cost} = 1.53 + 0.099 (\text{\% of insured area with} >10\text{-yr hail days}). \quad (3)$$

This equation was used with the historical hail-day values for 1901–47 to estimate the annual loss cost values for these years (Fig. 10). This resulted in an average loss cost of $3.22 for 1901–47, compared with an average of $3.85 for 1948–94. Nebraska experienced low hail days during the 1920s and 1930s. However, two estimated early annual values exceeded the highest during 1948–94. The loss cost value calculated for 1932 was $7.85 and that for 1942 was $8.52, as compared to a post-1948 high of $7.73 in 1992.

4. Summary

The investigations of the Illinois, Texas, and Nebraska hail data reveal that the amount of area experiencing an excessive number of hail days during the year, and occurring within the areas of insurance coverage in each

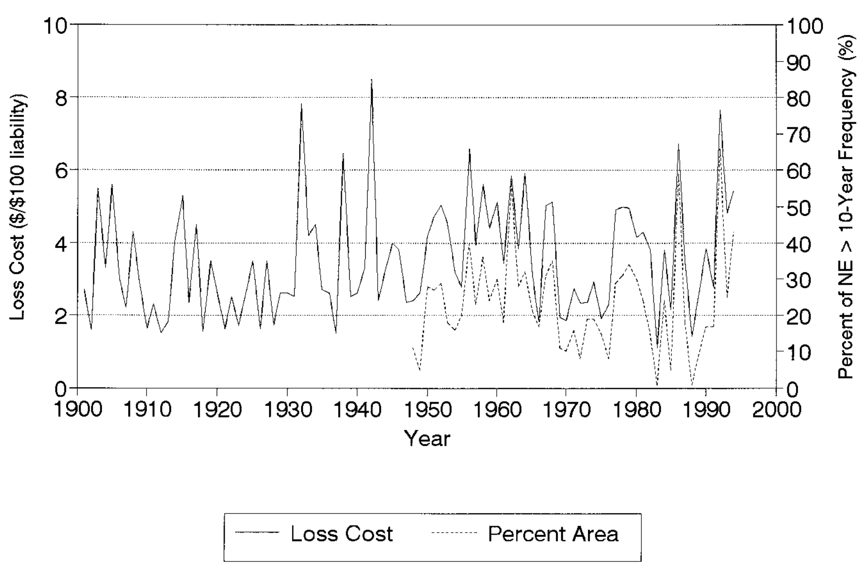

FIG. 10. Nebraska annual crop-hail loss costs, 1901–47 calculated and 1948–94 observed, and 1948–94 percent area of Nebraska with extensive insurance experiencing more than a 10-yr hail-day frequency.

state, relates well to the annual crop-hail loss cost value for the states. The year-to-year fluctuations in the NWS hail-day frequencies, data collected independently from the crop-hail loss data, support the validity of the fluctuations in loss cost values. The tests for three states revealed that the area with once in 10-yr hail-day values related best to loss cost, with the areal extent of 5-yr values rated second best. The hail-day areas explained 95% of the variations in loss cost values in Illinois, 52% of that in Texas, and 83% of that in Nebraska. One reason for the lower relationship in Texas were the years when large hail losses occurred but with relatively few storms.

The relationships appear sufficiently high to allow the estimation of loss costs for years before insurance records began (1948) and which have hail-day records. This means that the estimation of annual values back to 1901 is possible. The relationships also appear sufficiently strong to also allow the estimation of loss costs in uninsured areas where new crops are being grown for the first time. Such areas will have hail-day data, and hail-day loss cost relationships for the crops in question could be defined based on the same crop-loss relationships as defined elsewhere. The results are sufficiently encouraging to consider further investigations of hail data from other states.

Acknowledgments. This study was supported under Subaward UCAR S95-53410 of the University Corporation for Atmospheric Research and NSF Grant ATM-9508038. The authors wish to thank Ray Fosse, Richard Roth, and the anonymous reviewers for their thoughts and comments on the project. Special thanks go to Ms. Suzy Changnon, who edited the draft report.

REFERENCES

Changnon, S. A., 1967a: Method for evaluating substation records of hail and thunder. *Mon. Wea. Rev.,* **95,** 209–212.

——, 1967b: Areal-temporal variations of hail intensity in Illinois. *J. Appl. Meteor.,* **6,** 536–541.

——, 1977: The scales of hail. *J. Appl. Meteor.,* **16,** 626–648.

——, 1982: User beware: The upward trend in tornado frequencies. *Weatherwise,* **35,** 64–69.

——, and E. R. Fosse, 1981: Impacts and use of climatological information in the hail insurance industry. *Proc. Climate and Risk Conference,* Cambridge, MA, Mitre Corporation, 1–28.

——, and J. Changnon, 1990: Use of climatological data in weather insurance. *J. Climate,* **3,** 568–576.

——, D. Changnon, and R. Pyle, 1988: Thunder events and cloud-to-ground lightning frequencies. *J. Geophys. Res.,* **93,** 9495–9502.

——, D. Changnon, E. R. Fosse, D. Hoganson, R. J. Roth, and J. Totsch, 1996: Impacts and Responses of the Weather Insurance Industry to Recent Weather Extremes Rep. Changnon Climatologist CRR-41, Mahomet, IL, 166 pp. [Available from Changnon Climatologist, Mahomet, IL 61853.]

CIS, 1995: National Crop Insurance Services 1994 Rep. Overland Park, KS, 25 pp. [Available from National Crop Insurance Service, Overland Park, KS 66210.]

Conrad, V., and L. W. Pollak, 1950: *Methods in Climatology.* Harvard University Press, 459 pp.

HIAA, 1978: Crop-Hail Insurance Statistics Rep. Chicago, IL, 72 pp. [Available from National Crop Insurance Service, Overland Park, KS 66210.]

Roth, R. J., Sr., 1949: Crop-hail insurance in the United States. *Bull. Amer. Meteor. Soc.,* **30,** 56–58.

Stout, G. E., and S. A. Changnon, 1968: *Climatography of Hail in the Central United States.* Crop-Hail Insurance Actuarial Association Research Rep. 38, Chicago, IL, 46 pp. [Available from National Crop Insurance Service, Overland Park, KS 66210.]

Walter, A. S., 1993: U.S. crop insurance: Trends, status and profitability. *Seminar on Crop Insurance,* Toronto, ON, Canada, BEP International, 11 pp.