A Method for Estimating Crop Losses from Hail in Uninsured Periods and Regions

DAVID CHANGNON
Meteorology Program, Department of Geography, Northern Illinois University, DeKalb, Illinois

STANLEY A. CHANGNON AND SUZY S. CHANGNON
Changnon Climatologists, Mahomet, Illinois

(Manuscript received 1 July 1999, in final form 27 February 2000)

ABSTRACT

The insurance industry, insurance regulatory bodies, and scientists investigating climate change all desire long records of hail losses. Existing loss records for some states cover the 1948–present period; this span is helpful but is not long enough to define trends, possible fluctuations, and extremes adequately. The only other hail data with much longer records are the frequencies of hail days collected at National Weather Service stations since 1901, and a newly developed database for the major hail-loss states that contains hail-day data for 910 cooperative stations for 1901–94. This study tested two methods for estimating the historical loss values using hail-day data; one method produced modified hail-day values found to relate closely to loss values in the nation's 21 primary hail-loss states. The method involved modifying a station's hail-day values for each of the crop-season months using insurance-derived monthly hail intensity indices, resulting in an annual hail-intensity-weighted value. These weighted values of each year were combined using all stations in the crop regions of a state. The state-weighted annual indices were compared with the insurance loss values and yielded correlation coefficients of +0.60 or higher in 18 of 21 states; the resulting regression equations were used to estimate the loss values for the 1901–47 period. The temporal fluctuations and trends in the state hail intensity and loss values for 1901–94 showed major regional differences. States in the High Plains had increasing losses and greater variability with time, whereas states near the Great Lakes exhibited decreasing hail losses and variability with time. The approach can also be used to estimate loss values for areas for which historical loss values do not exist.

1. Introduction

The development of a new historical dataset of very long (1901–94) records of hail-day incidences at 1061 weather stations in the major hail regions of the United States offered an opportunity to examine the relationship of hail-day frequencies and hail losses to crops in major crop-loss states (Changnon and Changnon 1999). If a strong relationship could be defined using the insured state losses for 1948 (when hail insurance data collection began) to 1994, the historical crop–hail losses in the years prior to 1948 could be estimated. This estimate would allow the hail-insurance industry to examine the estimated state losses for the 1901–47 period and thus to obtain a better historical measure of trends, extremes, and potential loss fluctuations over time. A good relationship of hail days and crop loss would have a second benefit; that is, to estimate the past loss history and define the hail risk in areas that are undergoing new crop development and for which no insurance records exist. Both applications are useful in establishing future rates and planning to manage during extremes.

This study developed and tested two methods for estimating crop–hail loss values using modified hail-day values. One method was found to yield good relationships in the higher-loss states, and it was used to estimate the pre-1948 losses and to examine the long-term fluctuations in crop–hail losses for 1901–94.

2. Data and analysis

The research to define a relationship was based on two methods. Both methods required modifications of the raw hail-day values because the number of hail days in a year for any given state unto itself does not relate well to the magnitude of the state's crop–hail losses. Past research has shown that crop–hail losses occur on many days each year but that much of the year's total loss is concentrated in a few of these days (Changnon 1977). For example, in Illinois the annual average number of days with crop–hail losses is 47 and yet only 10

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

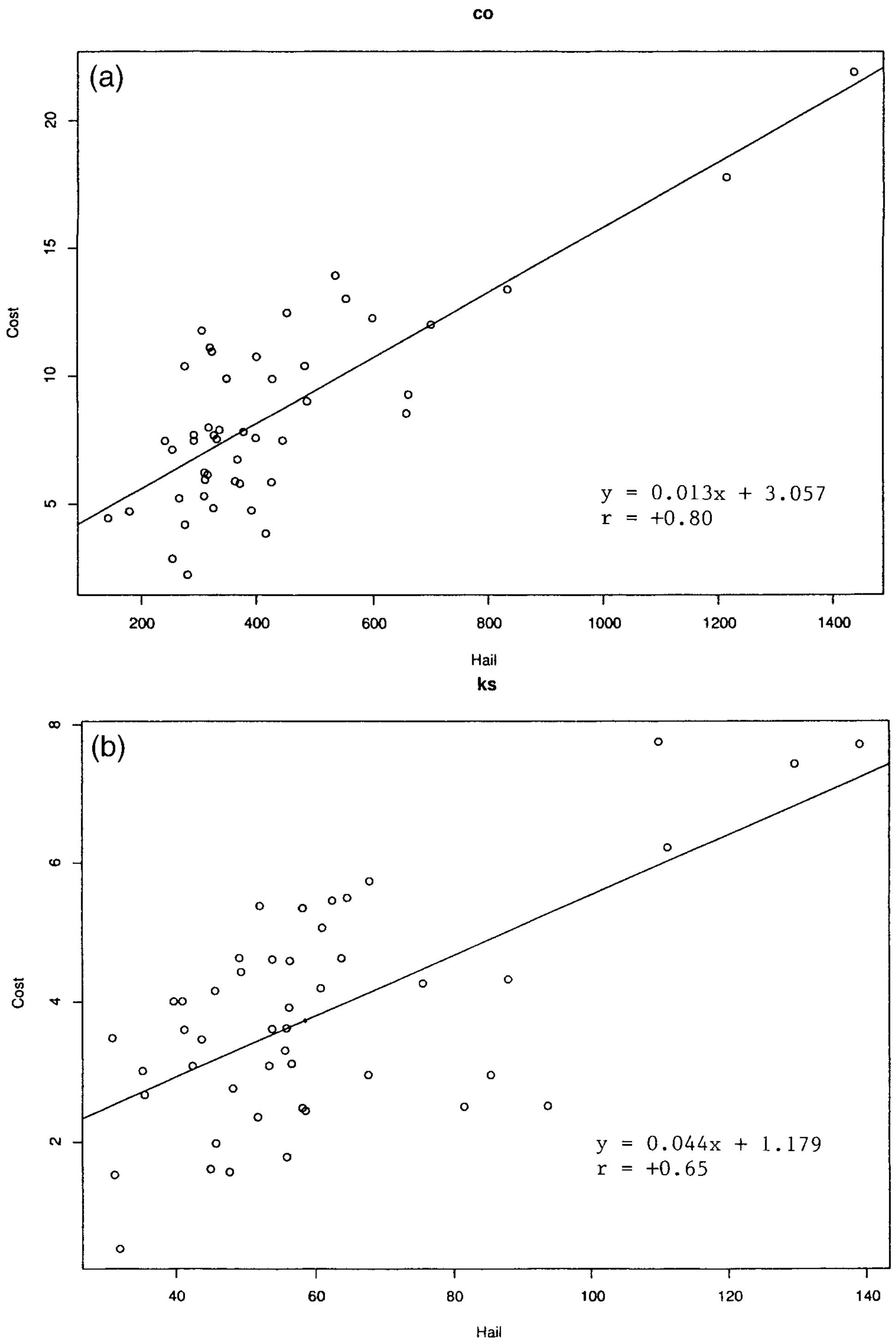

FIG. 1. The relationship of 1948–94 hail-intensity values to loss-cost values for (a) Colorado and (b) Kansas. Equations for predicted loss cost and correlation coefficients are provided.

days account for more than 80% of the total annual loss (Changnon 1967a). Hence, any method of relating crop–hail losses to the hail-day values had to modify the hail-day values in some manner to reflect the more extreme storm conditions during a growing season.

For temporal comparisons of losses in a state, the insurance industry uses the "loss cost." This metric is the annual amount of loss in dollars divided by the annual amount of liability in dollars, multiplied by $100. The loss cost adjusts for shifting dollar values as well

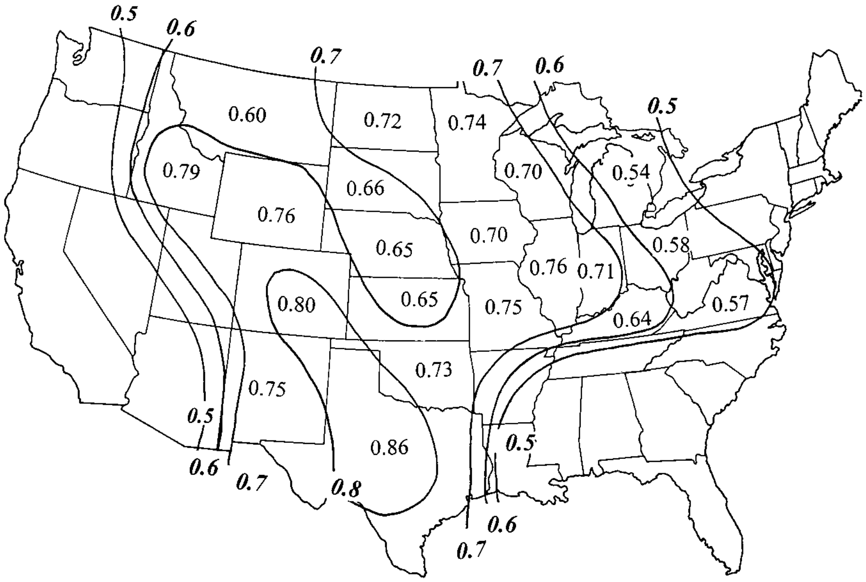

FIG. 2. The weighted hail-day–loss-cost correlation coefficients for the 21 states. All r values are positive.

as shifting coverage and is the industry's basis for risk assessment and rate making. State loss-cost values for 1948–94 were employed in the relationship analysis, which was based on the 27 states that experience most U.S. crop–hail losses (Changnon and Stout 1967).

The hail-day data available consisted of the monthly number of days with hail during the 1901–94 period, as recorded at 151 first-order stations and 910 cooperative substations (Changnon and Changnon 1999). These stations had quality hail-day values for periods ranging from 20 up to 94 yr and were concentrated in the 27 states in which most of the crop–hail losses occur. Because the crop–hail loss value to be used was an annual value for an individual state, the hail-day expressions developed for comparison were also based on some form of yearly value. The hail-day data had been assessed using a well-established technique (Changnon 1967b). The assessment of 16 000 cooperative substations in 27 states revealed that 910 (5%) had quality data and only a few cooperative substations had more than 50 yr of quality hail data during the 1901–94 period (Changnon and Changnon 1999). In each of the 27 states with such data, there were 30–40 stations with 20–50 years of good data during the 94-yr period. In any given year, there were typically 25–30 stations with good data. Prior studies of spatial aspects of hail reveal that sampling adequacy is a function of the frequency of hail, with more sampling points needed in high–hail frequency states such as Wyoming (Changnon 1971). Moreover, to obtain a good measure of the areal extent of damaging hail requires a dense sampling network of one point per square kilometer (Changnon 1968). Hence, the 30–40 quality-data stations in a given state did not measure all the hail, and many more stations would have improved the measure and the relationship between insurance loss values.

TABLE 1. The relationship of the annual weighted hail-day values to the loss-cost values for the 21 states with the highest hail-day frequencies during the 1948–94 period.

State	Correlation coefficient	Std error	Equation for predicted loss cost
Colorado	+0.80	2.266	$3.057 + 0.013x$
Idaho	+0.79	0.682	$0.493 + 0.055x$
Illinois	+0.76	0.390	$0.041 + 0.161x$
Indiana	+0.72	0.305	$0.310 + 0.065x$
Iowa	+0.69	0.827	$-0.140 + 0.130x$
Kansas	+0.65	1.205	$1.179 + 0.044x$
Kentucky	+0.64	1.146	$2.560 + 0.044x$
Michigan	+0.54	0.638	$0.430 + 0.079x$
Minnesota	+0.74	1.075	$0.631 + 0.057x$
Missouri	+0.74	0.448	$0.808 + 0.033x$
Montana	+0.61	1.856	$2.223 + 0.027x$
Nebraska	+0.65	1.211	$0.498 + 0.033x$
New Mexico	+0.75	2.184	$1.862 + 0.020x$
North Dakota	+0.72	1.328	$1.134 + 0.020x$
Ohio	+0.58	0.501	$-0.028 + 0.199x$
Oklahoma	+0.73	1.932	$1.410 + 0.025x$
South Dakota	+0.66	1.333	$2.196 + 0.024x$
Texas	+0.86	1.117	$2.564 + 0.014x$
Virginia	+0.57	1.292	$2.161 + 0.120x$
Wisconsin	+0.70	0.796	$-0.066 + 0.109x$
Wyoming	+0.76	2.789	$-1.053 + 0.017x$

a. Method one

The months when hail losses to crops occur in each state were identified and were used to define the period for a comparative study. The hail-day values for these periods at each station were used to calculate a period-of-record average, and the values were also ranked to

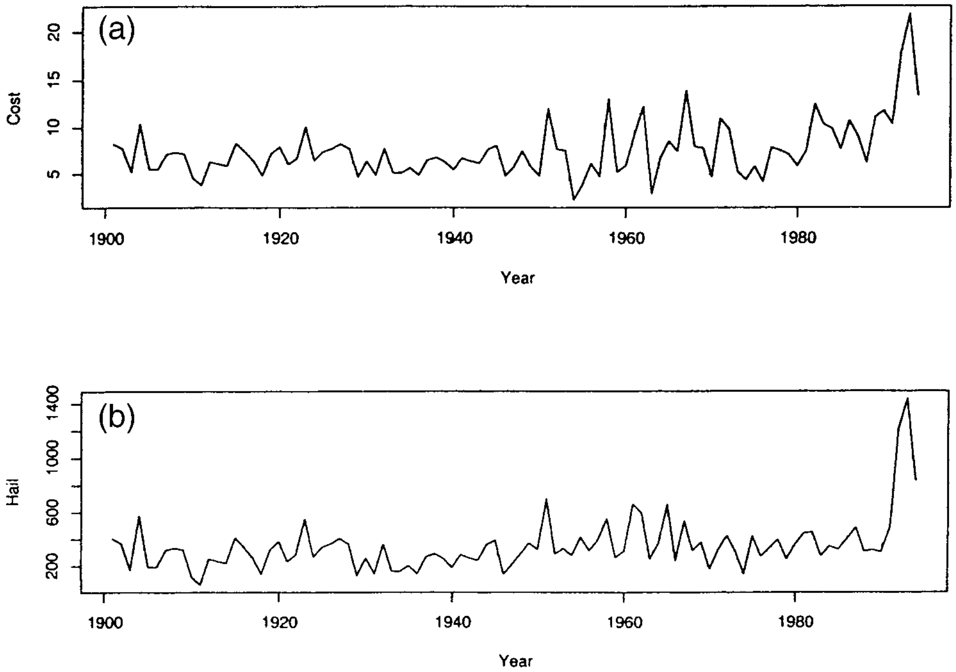

FIG. 3. The 1901–94 curves of (a) loss-cost values and (b) weighted–hail index values for Colorado.

determine the values expected to occur once in 5 yr and those expected to occur once in 10 yr. For example, in Illinois, the crop–hail loss season is May–October (Changnon and Stout 1967), and the hail-day values for an Illinois station for this period in each year became the basis for calculating the station's average, 5-yr value, and 10-yr value. This ''extreme-frequency'' approach was based on an early probing test, using data for three states, that had suggested that the areal extent of higher hail frequencies in a crop area related well to the crop-loss costs (Changnon and Changnon 1997).

The next analytical step for a given state involved constructing hail maps for each year during the 1901–94 period. Ninety-four annual maps showing the weather stations that had reached or exceeded their annual average value, 94 annual maps showing the stations that had met or exceeded their 5-yr return values, and 94 annual maps showing the stations achieving 10-yr values were developed. Isopleth analysis was then used to delineate the areas on each map that met each condition. For example, in 1918, Missouri had 22 670 km^2 of average or higher values, 9870 km^2 with 5-yr values, and 1012 km^2 with 10-yr values. Those areal values for noncropped areas in each state were not used in calculating the state total area at risk for the three criteria. These annual area totals were then expressed as a percent of the total cropped value in the state meeting each of these three criteria. These percentage values for 1948–94 were used in the statistical comparative analysis with the crop–hail loss costs for 1948–94. Three comparisons were made with the annual loss-cost values, including the percentages with average, 5-yr, and 10-yr values (Changnon and Changnon 1997).

b. *Method two*

The second method of defining a relationship between hail-day frequencies and loss costs was based on modifications of hail-day values (in a state) by using crop–hail loss values that measured the loss differences during the growing season (Changnon and Stout 1967). The extent of crop damage varies greatly during the growing season because of shifts in a crop's susceptibility to hail damage and to seasonal changes in storm intensity (Changnon 1977). For example, the average loss cost per hail day in North Dakota in July is $1.07 as compared with only $0.25 in June. This difference means that a day with hail in July on average causes 4.3 times the loss caused by the typical hail day in June. These monthly loss costs for each state integrate the seasonal changes in storm intensity and in crop susceptibility and they were used as indices to multiply the number of hail days in each month during the growing season. For states with two primary crops, the monthly loss costs were combined to form a single average monthly value (Changnon and Changnon 1999). Thus, this approach weighted the hail-day values.

At each weather station, the weighted monthly values in a year were summed to derive a value labeled as the annual ''intensity index.'' For example, in June 1948, Urbana, Illinois, had two hail days, and this total was multiplied by $0.50 (the June loss cost for Illinois). In

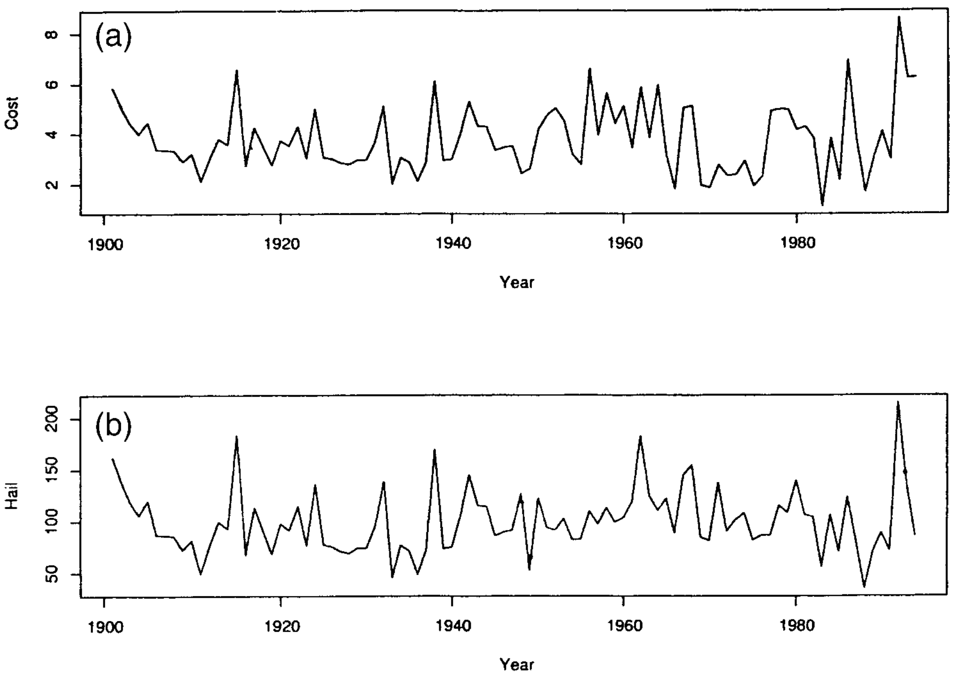

FIG. 4. Same as Fig. 3 but for Nebraska.

July, the station had one hail day that was modified by the loss cost of \$0.95. Their sum was \$1.95 (2 × 0.5 plus 1 × 0.95), and, with no other hail days in 1948, the yearly intensity index for 1948 became 1.95 at Urbana. This process was repeated at all stations, and, in a given state, these intensity indices for a given year were summed to obtain a statewide annual index total. This total was divided by the number of hail-reporting stations to develop an adjusted yearly intensity value. This adjustment was necessary because the number of stations with quality hail-day data did not remain the same in every year. The resulting state annual intensity

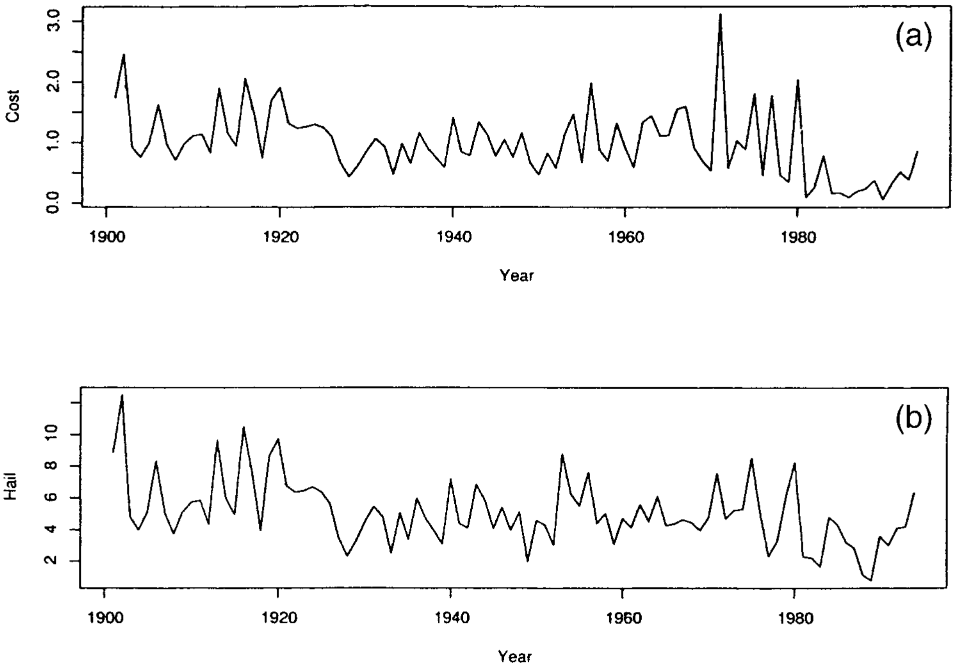

FIG. 5. Same as Fig. 3 but for Ohio.

TABLE 2. The temporal distribution of the five highest loss-cost values for the 21 states during 1901–95.

State	1901–10	1911–20	1921–30	1931–40	1941–50	1951–60	1961–70	1971–80	1981–90	1991–95
CO				1	1	2			1	
ID						1	2	1	1	
IL			1		1	1	1	1		
IN		1				2		2		
IA			1	1		1	1	1		
KS		1			2	1				1
KY						2	1	1		1
MI				1		2	2			
MN			1			2	1	1		
MO						1	2		2	
MT	1	1					1	1		1
NE		1		1		1			1	1
NM	1				2			1	1	
ND					1	2		1		1
OH	1	2				1		1		
OK	1					2	1			1
SD						1	1	2		1
TX					1			1	2	1
VA			2		1	1	1			
WI	1		1		1		2			
WY			1		1	1		1		1
Sum	5	6	7	4	11	24	16	15	8	9

indices for the 1948–94 period were compared with the annual crop–hail loss-cost values.

3. Evaluation of the methods for defining relationships

The comparison of 1948–94 annual loss costs with the areal extent of average, 5-yr, and 10-yr values (method one) provided a wide variety of outcomes for the 27 states tested. The 10-yr values provided the highest correlations in all but eight states (Oregon, Washington, Idaho, South Carolina, North Carolina, Virginia, Georgia, and Michigan). In these eight states, the best correlations were with the 5-yr values, but they were relatively weak, all being less than +0.2. Correlation coefficients *r* greater than +0.6 were found for the 10-yr values in Kansas, Nebraska, Texas, Illinois, Missouri, and North Dakota. The best correlations in the other 13 states were found to be relatively weak, less than +0.4. The average-area values had the poorest relationships in all states. The results did not provide a technique that could adequately estimate the pre-1948 loss-cost records in 21 of the 27 states.

TABLE 3. Computed slopes of state hail-intensity values for 1901–94 period. Level of significance is in percent.

State	Slope (change in values per year)	Level of significance (%)
Colorado	+2.769	0.1
Idaho	+0.034	—
Illinois	−0.005	—
Indiana	−0.033	10
Iowa	+0.017	—
Kansas	+0.251	0.1
Kentucky	+0.076	—
Michigan	−0.023	—
Minnesota	+0.016	—
Missouri	+0.028	—
Montana	−0.402	10
Nebraska	+0.098	—
New Mexico	+0.848	5
North Dakota	+0.358	10
Ohio	−0.031	0.1
Oklahoma	−0.128	—
South Dakota	+0.258	10
Texas	+1.966	0.1
Virginia	−0.015	—
Wisconsin	−0.037	—
Wyoming	+0.834	—

The comparative analysis based on the state intensity indices (method two) provided better relationships in all states. Figure 1 presents the scatterplots of the annual intensity values for 1948–94 and the related crop–hail loss costs for Colorado and Kansas. These plots show that most years had low intensity and loss-cost values, with a few years of high values. The correlation coefficients were +0.80 for Colorado and +0.65 for Kansas. Extreme-value analytical approaches might have done a better job of fitting the extremes, but the insurance interests were comfortable with analyses and results based on standard correlation analysis of the two datasets.

The national pattern based on the states with correlation coefficients of 0.5 or higher (Fig. 2) reveals that the coefficients in the Rocky Mountains, High Plains, and Midwest all exceed +0.6. Many states had coefficients greater than +0.7, meaning that the weighted hail-day values converted to intensity indices explained more than half of the variability in the loss costs for 1948–94. Correlation coefficients of 0.6 are considered to be insignificant for many applications, but the in-

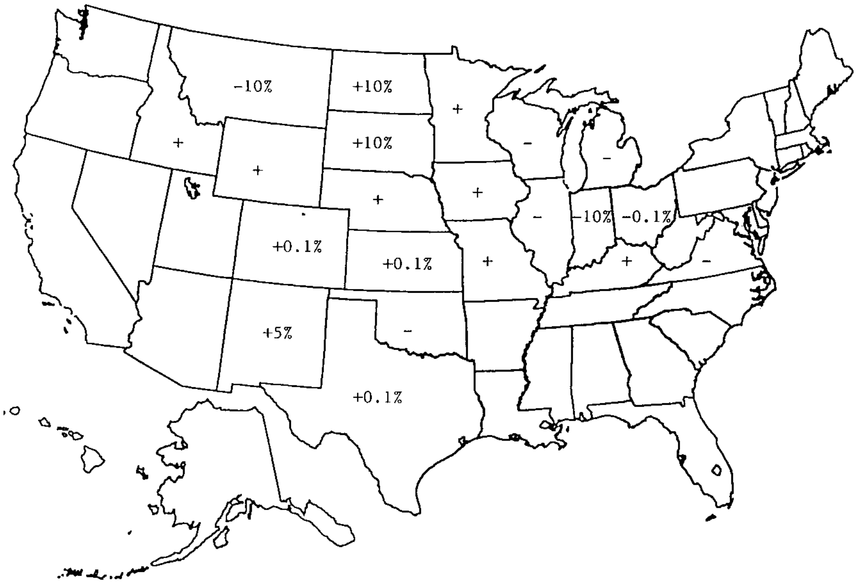

FIG. 6. The 1901–94 slopes (+ or −) associated with the hail-intensity values for the 21 states. Numerical value represents level of significance. Actual slope values are presented in Table 3.

surance industry felt that explanation of 35% of the variability was useful for their applications. Low correlations, +0.2 to +0.4, were found in Washington, Oregon, and four southeastern states (Georgia, North Carolina, South Carolina, and Tennessee). These six states have low hail-day frequencies and the lowest loss costs of the 27 states studied. Other unmeasured hail factors in these states affect the amount of loss. Table 1 presents the correlation coefficients, the standard errors of estimate, and the regression equations developed for the 21 states with the highest hail-day frequencies. Clearly, this approach for estimating the historical loss costs did not explain all the loss-cost variability in 1948–94, but was superior to method one and could be used in many states to estimate historical loss costs.

TABLE 4. Change in mean state hail-intensity values from the earlier (1901–47) to the later (1948–94) period. Level of significance is in percent.

State	Change in mean value	Level of significance (%)
Colorado	+143.21	1
Idaho	+3.60	—
Illinois	+0.36	—
Indiana	−1.02	—
Iowa	+0.90	—
Kansas	+14.25	1
Kentucky	+4.91	—
Michigan	−0.43	—
Minnesota	+1.46	—
Missouri	+2.55	—
Montana	−15.89	—
Nebraska	+9.28	—
New Mexico	+48.81	5
North Dakota	+24.06	5
Ohio	−1.21	1
Oklahoma	+4.29	—
South Dakota	+13.61	10
Texas	+82.16	1
Virginia	+0.49	—
Wisconsin	−0.07	—
Wyoming	+23.54	—

4. Applications

The historical (1901–47) loss-cost values were predicted using the equations shown in Table 1. The estimated values for 1901–47 (1919–47 in Texas and Wyoming) were then combined with the 1948–94 values to create a series of loss-cost values for 1901–94 for these 21 states (with *r* values $>+0.5$). Figures 3–5 present the 1901–94 curves of the weighted hail index (lower graph) and the loss-cost values for three states. Tests of variability showed that Colorado (Fig. 3) and other states in the southern plains (Kansas, New Mexico, Oklahoma, and Texas) had greater late-period (1948–94) variability. States in the northern plains, as illustrated by Nebraska (Fig. 4), had no temporal difference in their variability. Hail indices and loss costs near the Great Lakes, as illustrated by Ohio (Fig. 5), showed greater early-period variability (1901–47) than in the last 47 years. Experience in dealing with insurance users of such results indicates that any effort to generate spatial and temporal estimates of loss using hail-day data had to combine it with state-scale insurance-derived loss data.

An important concern of the crop–hail-insurance industry is the long-term fluctuations and trends, and, to this end, the temporal aspects of the loss-cost values for 1901–94 were assessed. Table 2 shows the temporal distribution of the five highest loss-cost values during the 1901–95 period. Such extremes and their time distributions are of great concern for hail-risk applications in the insurance industry. The total occurrences per decade show a majority came in the 1951–60 period (24 of the 105 peak values), and the fewest were in the Dust Bowl era of 1931–40. Examination of the state distributions over the 1901–95 period reveals two general kinds: one with most peaks in the last 45 yr (1951–95), and the other with a more even distribution. The states with a predominance of peaks in recent decades were found in two areas: the southern tier (Kentucky, Missouri, Oklahoma, and Texas) and the north (Minnesota, North Dakota, and South Dakota). Only five states (Kansas, New Mexico, Ohio, Virginia, and Wisconsin) had more peaks before 1951 than afterward. These results have important climate and business implications regarding the long-term variability of such extremes during the twentieth century, peaking in midcentury (1941–80), and with geographic regions exhibiting different distributions.

The temporal trends of the hail-intensity values of each state for the 1901–94 were also assessed. The Student's t test was used to determine whether the change in slope was significant (Table 3). Colorado and Texas experienced the greatest temporal increases in hail intensity. The evaluation of slope terms spatially identified some homogeneous regions of decrease and increase (Fig. 6). Most states in the High Plains experienced a significant increase in hail intensity over the 94 yr, whereas states near the Great Lakes experienced decreasing hail intensity. These results generally agree with those found in a long-term analysis of hail days at first-order weather stations (Changnon and Changnon 2000).

The two-tailed t test was used to determine whether the early mean in state hail-intensity values (1901–47) was different from the latter mean (1948–94). Changes in mean values and levels of significance are shown in Table 4. As expected, most states that exhibited a significant change in slope also had a significant change in mean hail-intensity value.

5. Conclusions

Two methods for estimating historical crop–hail loss costs using 1901–94 records of hail days at 1061 stations in 27 states were developed and tested. The method based on modifying the monthly hail days using monthly crop–hail loss-cost values to derive statewide hail-intensity indices was found to provide the best relationships. The method developed can be successfully used for estimating historical (1901–47) loss costs in the 21 states where the greatest hail damages to crops occur. It also offers a means for using historical hail-day values in uninsured areas to estimate the potential loss costs.

The relationship of annual weighted hail-day values to loss-cost values was good ($r > +0.60$) for states in the Rockies, High Plains, and Midwest. For these 21 states, loss-cost values could be estimated for the 1901–47 period. Weaker relationships were found in states with lower hail incidence.

The 1901–94 trends in weighted hail-day values indicated that the states in the High Plains have experienced significant increases in hail intensity, and the variability has also become greater. States near the Great Lakes have experienced decreases or little change in hail intensity.

Acknowledgments. This research and paper were funded by National Oceanic and Atmospheric Administration Grant NA76GP0439. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any subagencies. The authors thank Carl Lonnquist for his help with the hail-day database development and the anonymous reviewers for their helpful comments.

REFERENCES

Changnon, D., and S. A. Changnon, 1997: Surrogate data to estimate crop-hail loss. *J. Appl. Meteor.,* **36,** 1202–1210.
Changnon, S. A., 1967a: Areal-temporal variations in hail intensity in Illinois. *J. Appl. Meteor.,* **6,** 536–541.
——, 1967b: Method of evaluating substation records of hail and thunder. *Mon. Wea. Rev.,* **95,** 209–212.
——, 1968: Effect of sampling density on the areal extent of damaging hail. *J. Appl. Meteor.,* **7,** 518–521.
——, 1971: Means of estimating areal hail-day frequency. *J. Wea. Modif.,* **3,** 154–159.
——, 1977: The scales of hail. *J. Appl. Meteor.,* **16,** 626–648.
——, and G. Stout, 1967: Crop-hail intensities in the United States. *J. Appl. Meteor.,* **6,** 542–548.
——, and D. Changnon, 1999: Developing a hail database for the United States. CRR-47, Changnon Climatologist, 51 pp. [Available from Changnon Climatologist, 801 Buckthorn Circle, Mahomet, IL 61853.]
——, and ——, 2000: Long-term fluctuations in hail incidences in the United States. *J. Climate,* **13,** 658–664.