The Design and Evaluation of Rainfall Modification Experiments

PAUL T. SCHICKEDANZ AND FLOYD A. HUFF

Illinois State Water Survey, Urbana

(Manuscript received 7 January 1971)

ABSTRACT

Storm rainfall data from dense raingage networks in Illinois were employed in a study to determine the length of time required to obtain significance for various increases in storm rainfall due to weather modification efforts. The primary purpose was to evaluate the effect of stratifying the storm data on the detection of seeding effects for a given design using highly accurate measurements of the rainfall parameters. It was also desired to evaluate the efficiency of various rainfall parameters and the efficiency of various statistical designs in detecting various increases. Results indicate that the length of experimentation necessary for detection of seeding effects varies according to weather type, precipitation type, rainfall parameter, and statistical design employed. Results also indicate that as the seeding-induced increase becomes large, the choice of stratification, rainfall parameter, and statistical design becomes less important. An evaluation procedure is recommended which incorporates desirable features from several of the designs, stratifications and rainfall parameters considered in this study. Although it is difficult to verify, a 20% increase in precipitation can be detected in a five-year experiment provided proper choices are made of weather types, statistical designs, data stratifications and rainfall parameters.

1. Introduction

One of the major problems encountered in precipitation modification projects has been the evaluation of the results. Evaluation of modest increases or decreases in surface precipitation during cloud seeding operations is exceedingly difficult because of the great natural variability of precipitation, in both space and time. The primary purpose of this paper is to evaluate the effect of stratifying the data on detection of seeding effects for a given design using highly accurate measurements of the rainfall parameters. A secondary purpose of this research was to evaluate various rain parameters as means of detection of seeding effects for a given design. A third purpose was to evaluate the efficiency of various statistical designs. It was also desired to establish theoretical frequency distributions that could be used in fitting storm precipitation data in the midwestern type of climate. The comparison of verification methods has been confined mostly to the use of standard statistical designs applied to surface precipitation measurements in the manner used by past investigators. Evaluation of the cloud model prediction techniques proposed by Simpson *et al.* (1967) and Weinstein and Davis (1968) is considered to be beyond the scope of the present study.

2. Basic data and definitions

a. Source of data

The east-central Illinois network (ECI) supplied the basic data used in this study. It consisted of 49 raingages arranged in a nearly uniform grid pattern in a 400-mi^2 rural area of relatively flat terrain in which elevations ranged from 650 to 910 ft MSL. The network was operated from 1955–66 with no significant changes in gage locations. This paper is primarily devoted to investigation of the use of areal storm precipitation data in the verification of weather modification experiments as determined by the most complete data sample available. The data for this evaluation were based on 1346 storms from the 12-year sample on the ECI network.

Also investigated was the use of area-depth measurement parameters consisting of the maximum point rainfall, mean rainfall gradient, and average rainfall depth over the sampling area. However, the area-depth analyses were based upon a sample of moderate to heavy rain storms consisting of 228 storms with durations ranging from 1–48 hr. This sample was chosen because it was already available from a previous study by Huff (1968). Thus, lighter storms are not included in the area-depth sample used in this study. However, it is believed that the method will be applicable through the intensity spectrum of storm rainfalls. The area-depth analysis is presented in this paper as a "special" or "pilot" study.

b. Definition of terms

Several terms used frequently in this paper are defined as follows. A *storm* is a precipitation period separated from preceding and succeeding precipitation on the sampling area by 6 hr or more. This definition has been found most suitable for separating storms resulting from different synoptic causes on the sampling networks.

The *warm season* includes the five months of May–September. The *cold season* encompasses the seven months of October–April.

Alpha (α) is the probability of asserting that there is a seeding effect present when in fact there is not. *Beta* (β) is the probability of asserting that there is no seeding effect present when in fact there is a seeding effect. The *power* $(1\text{-}\beta)$ of the test refers to the probability of detecting a seeding effect when the effect is present. Since the emphasis of this paper was placed on obtaining sample sizes for various data stratifications and rain parameters, sample sizes were computed only for $\alpha=0.05$ and $\beta=0.50$.

c. Data stratifications

Frequently, cloud seeding is not conducted on all types of storms, and seeding selectivity is likely to increase as knowledge advances in weather modification. Statistical verification of seeding results could be facilitated by data stratification if seeding effects vary substantially with the meteorological characteristics of storms as indicated by recent work of Braham and Flueck (1970). It has also been indicated by recent research that storms which naturally produce large quantities of precipitation may actually be decreased by seeding. Thus, two types of stratifications were used in these studies. The first was made according to synoptic and precipitation type for both the storm mean rainfall and the area-depth data. The second stratification was made according to the magnitude of the rainfall recorded in the network.

The seven basic synoptic weather types used in this study included four frontal types (cold, warm, static and occluded), prefrontal squall lines, air mass instability showers, and low center passages. The three major precipitation types used in this study were steady rain, rainshowers and thunderstorms. For details of the precipitation and synoptic classification methods, the reader is referred to Huff (1968). The sample size of each stratum for the storm mean rainfall and the area-depth data are shown in Table 1.

The areal mean precipitation was stratified into storms which had measurable amounts recorded at one or more raingages in the network ($\geqslant 0.01$ inch). Areal means for the qualifying storms were then stratified into those with network means $\geqslant 0.005$ inch (trace), >0.10 inch, and >1.00 inch. A special group consisted of the storm means derived from those gages recording amounts $\geqslant 0.01$ inch (portion of the area with measurable rainfall). In addition, the maximum point rainfall was determined for all qualifying storms. These storm means (four classes) were then stratified into warm and cold season storms. These stratifications were used with both the areal storm and area-depth data.

TABLE 1. Sample size (number) of storms in the various strata used in the areal storm rainfall and area-depth studies.

	Areal storm rainfall studies					
Stratification	Means $\geqslant 0.005$ inch	Means >0.10 inch	Means >1.00 inch	Gage amounts $\geqslant 0.01$ inch	Maximum point storm rainfall	Area-depth studies
Warm season storms						
Synoptic type	579	335	52	675	675	151
Low centers	47	15	5	52	52	9
Air mass	149	73	4	197	197	32
Cold fronts	156	94	15	176	176	39
Warm fronts	65	44	9	70	70	21
Occluded fronts	12	8	0	12	12	6
Stationary fronts	96	63	14	110	110	25
Squall lines	54	38	5	58	58	19
Precipitation type	578	335	52	673	673	151
Steady rain	74	34	7	84	84	10
Rainshowers	173	58	3	245	245	15
Thunderstorms	331	243	42	344	344	126
Cold season storms						
Synoptic type	598	312	42	671	671	77
Low centers	277	136	20	301	301	25
Air mass	17	4	2	25	25	5
Cold fronts	153	86	6	176	176	24
Warm fronts	35	26	2	40	40	9
Occluded fronts	45	21	1	46	46	5
Stationary fronts	61	31	7	72	72	6
Squall lines	10	8	4	11	11	3
Precipitation type	457	276	41	514	514	77
Steady rain	196	115	13	213	213	16
Rainshowers	140	61	7	179	179	22
Thunderstorms	121	100	21	122	122	39

d. Designs

Five statistical designs using the areal storm and area-depth data were used. These designs included: 1) *random-experimental*, which involves randomization of days over a single target area into seeded and nonseeded days with nonseeded days being the control; 2) *random-historical*, in which a random choice is made of days to be seeded over a single target with the historical record as control; 3) *continuous-historical*, in which all rain days within a given stratification are seeded with the historical record as a control; 4) *cross-over*, which requires seeding a target chosen at random with another area being the control (random interchange of target and control); and 5) *target-control*, in which all potential rain days are seeded in a fixed target area with a nearby area serving as a fixed control.

e. Limitations

The Illinois study was restricted to areas of 50–550 mi^2 for which data from dense raingage networks were available to define accurately the time and space properties of precipitation. Although the study was restricted to areas of county size or smaller, seeding capability on areas of this size must eventually be determined for efficient agricultural application. Furthermore, water supply augmentation from seeding frequently involves treatment of relatively small basins. For example, in Stall's (1964) study of low flows of Illinois streams for use in impounding reservoir design, the drainage areas for the 164 basins used had a median area of 170 mi^2 and only 25% of the basins had areas >600 mi^2. Light to moderate droughts, in which the potential benefits from cloud seeding could be very substantial, sometimes encompassed only a few hundred square miles. It is believed that the results presented here for the areas in Illinois are applicable to any area east of the Rockies with the exception of coastal and mountainous areas.

3. Areal storm precipitation

Theoretical frequency distributions were fitted to the area storm precipitation data, and the distributions were used to obtain sample size requirements for the various rain parameters, stratifications and designs. In these statistical design studies, both the non-classical (sequential) and the classical (non-sequential) analyses were employed. The components of the particular test being used were computed for the nonseeded distributions, and then with assumed changes in the distributional parameters a sample size was computed through algebraic relations. These values were computed for the log-normal 1- and 2-sample tests.

a. Theoretical frequency distributions

The log-normal distribution was fitted to the areal storm precipitation data for the rain parameters and data stratifications shown in Table 1. A non-truncated log-normal distribution was used for the samples of storm mean rainfall $\geqslant 0.005$ inch, storm mean rainfall in which gage amounts >0.01 inch, and for maximum point rainfall. However, for the stratified samples involving mean storm rainfall >0.10 and >1.00 inch, a truncated distribution was obtained by making the transformation $(X-0.10)$ and $(X-1.00)$ on the truncated samples of areal mean rainfall >0.10 and >1.00 inch, respectively. The log-normal mean and variance were then estimated from the transformed samples in each case. The Kolmogorov-Smirnov "goodness of fit" test was applied when sample sizes were <40 and the chi-square test when sample sizes were $\geqslant 40$ (Schickedanz and Changnon, 1970). The resulting probability values are shown in Table 2. When all warm season storms were considered as one sample, a poor fit was obtained for all five types of data. All probabilities were <0.01 except for areal mean rainfalls >1.00 inch for which the probability was 0.04. When all cold season storms were grouped together in one sample, a good fit was obtained for means >0.10 and >1.00 inch, but poor fits were obtained for the other cases.

At this point, further analysis was restricted to precipitation and synoptic types which are more likely to be seeded during a rain modification experiment. A most likely candidate in the Midwest is the summer air mass storm which tends to be more isolated and slower moving than other synoptic types. This was the type of storm seeded in the Whitetop experiment (Braham, 1966). Among precipitation types, the thunderstorm was selected since it is the major source of water during the warm season. Since frontal storms are the major producers of warm season rainfall among the various synoptic types, cold fronts were selected for analysis also. During the cold season, low centers and steady rain were selected since they are major producers of winter precipitation in the Midwest. The average number of storms for each of these types, based upon the network sample of storms, is shown in Table 3.

b. Tests of hypothesis

The storm was used as the experimental unit in the various designs. The number of storms required to obtain significance in a rain modification experiment was computed for the various designs and tests in these statistical design studies. Both the non-classical (sequential) and the classical (non-sequential) analyses were employed. For purposes of sample size computations, it was assumed that all the cases in Table 3 were log-normal distributed although there were 6 out of 20 with probabilities <0.05. Under the assumption that the data were log-normal distributed, various increases of $\delta=0.05$, 0.10, 0.20, 0.40, 0.60 and 0.80 were assumed and applied to the nontransformed precipitation data. The corresponding scale change was

TABLE 2. Probability* values for the "goodness of fit" test for the log-normal distribution of areal storm rainfall.

Stratification	Means $\geqslant 0.005$ inch	Means >0.10 inch	Means >1.00 inch	Gage amounts >0.01 inch	Maximum point storm rainfall
Warm season storms	<0.01	<0.01	0.04	<0.01	<0.01
Synoptic type					
Low centers	0.36	>0.20	>0.20	0.18	0.12
Air mass	0.41	0.64	>0.20	0.25	<0.01
Cold fronts	0.15	0.43	>0.20	0.26	0.02
Warm fronts	0.04	0.23	>0.20	0.16	0.19
Occluded fronts	>0.20	>0.20	**	>0.20	0.18
Stationary fronts	0.42	0.10	0.19	0.23	<0.01
Squall lines	0.03	>0.20	>0.20	<0.01	<0.01
Precipitation type					
Steady rain	0.73	>0.20	>0.20	0.35	0.46
Rainshowers	0.15	0.71	>0.20	0.51	0.13
Thunderstorms	<0.01	0.02	0.08	0.02	0.01
Cold season storms	0.04	0.76	0.59	0.01	<0.01
Synoptic type					
Low centers	>0.20	0.73	>0.20	0.09	0.62
Air mass	0.15	>0.20	**	>0.20	>0.20
Cold fronts	0.04	0.09	>0.20	0.03	0.14
Warm fronts	0.20	>0.20	**	0.21	0.02
Occluded fronts	0.22	>0.20	**	0.54	0.85
Stationary fronts	0.75	>0.20	>0.20	0.93	0.30
Squall lines	>0.20	>0.20	>0.20	>0.20	>0.20
Precipitation type					
Steady rain	0.08	0.44	>0.20	0.10	0.20
Rainshowers	<0.01	<0.01	>0.20	<0.01	0.05
Thunderstorms	0.54	0.10	>0.20	0.43	0.40

* Probability that the observed differences between the data sample and the given distribution could have occurred by random chance.

** Sample size too small for computation.

made on the transformed scale by the addition of the logarithm of $(1+\delta)$. It was assumed that seeding would not affect the variances of the distributions. Relationships given by Schickedanz *et al.* (1969) were used to determine the size of sample necessary to obtain significance for the various increases for the log-normal 1- and 2-sample tests. Sample sizes were computed for the continuous-historical, the random-historical, the random-experimental, the target-control, and the crossover designs. For use in the computation of sample size required in the target-control and crossover design, the correlation coefficient r between network storm means in two adjacent areas was assumed to be 0.75. This value was based upon a previous study of correlation between adjacent areas (Huff and Schickedanz, 1970). In a sequential test, the sample size was computed for the random-historical and continuous-historical designs through use of the 1-sample test.

c. Numerical results

This section provides a comparison of the various designs, tests, and types of data measurements used in the areal-storm study. This comparison is made to select the most appropriate design and type of data measurement. However, the final choice cannot be based on pure sampling numbers alone. Also to be considered are the relation of the particular data sample

TABLE 3. Average number of storms per year expected for selected data types, storm types, and season of storm mean rainfall (based on 12-year sample).

Storm rainfall	Warm season Air mass	Warm season Thunderstorms	Warm season Cold fronts	Cold season Steady rain	Cold season Low centers
Areal mean $\geqslant 0.005$ inch	12.4	27.6	13.0	19.6	26.6
Areal mean >0.10 inch	6.1	20.3	7.8	10.5	12.8
Areal mean >1.00 inch	0.3	3.5	1.3	1.1	1.7
Areal mean for gage $\geqslant 0.01$ inch	16.4	28.7	14.7	20.1	29.0
Maximum point rainfall	16.4	28.7	14.7	20.1	29.0

and specific area to the assumptions involved in each test and design, the cost of operation, and the measurement capabilities required for a particular design in relation to the field program.

Fig. 1 shows the number of storms required (detection ability) to obtain significance for a 20% increase in precipitation under assumed seeding effects for various designs and data types for air mass storms. The graph suggests that the stratification of the areal means produced strata in which the storms were more homogeneous than the entire sample. This is reflected in the fact that fewer storms are required to detect a 20% increase with the samples in which means were greater than 0.01 and 1.00 inch than with the sample which included all means >0.005 inch.

Approximately the same number of storms are required when all storms with areal means $\geqslant 0.005$ inch are used as when areal means based on gage amounts >0.01 inch are used. As the design becomes less efficient (requires more storms) the difference between types varies considerably. For example, the difference in sampling requirements between maximum rainfall and all storms with means $\geqslant 0.005$ inch for the crossover design is only 18 storms (82 vs 100), whereas for the experimental-random design it is 144 storms (662 vs 806). A most important point brought out in Fig. 1 is that the choice of design in detecting a 20% increase in air mass storms is a very important factor. Although the random-experimental design requires the longest verification time, it should be noted that in the statistical sense, it is the most valid, since no historical data are required and practically no assumptions are violated.

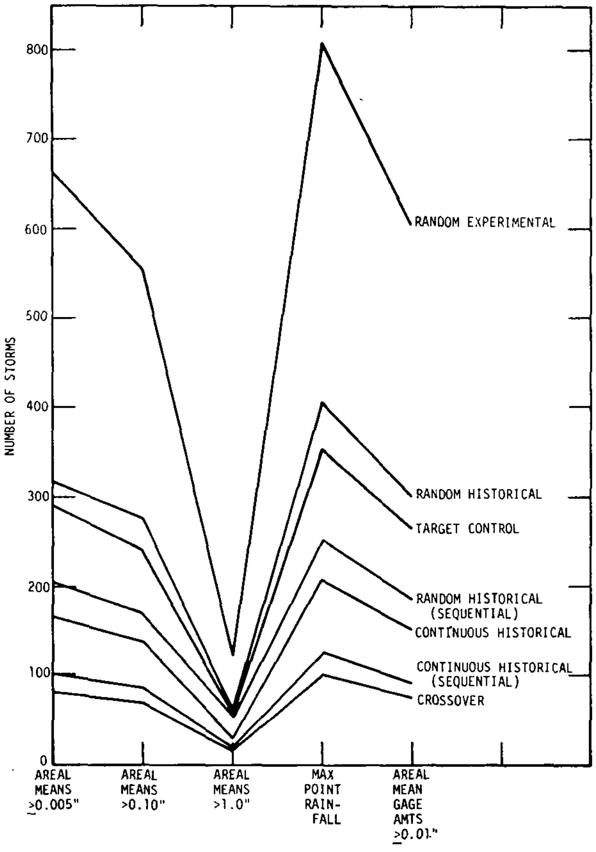

FIG. 1. Number of storms required to obtain significance with 20% increase in warm season air mass storms ($\alpha=0.05$, $\beta=0.20$).

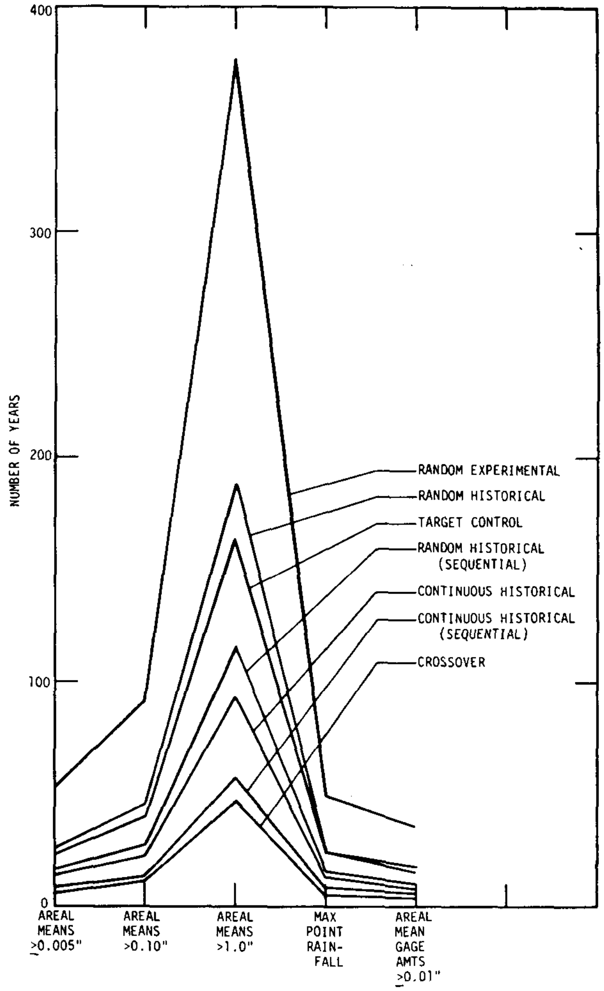

FIG. 2. Number of years required to obtain significance with 20% increase in warm season air mass storms ($\alpha=0.05$, $\beta=0.50$).

Fig. 1 is a presentation of the detection ability of a particular design in relation to data type. However, the number of storms per year or season varies considerably among data types; therefore, one type of storm may have a high detection ability, and yet require a long time (real time) to obtain significance for a given increase in rainfall. Fig. 2 is a presentation of the number of years required to obtain significance based on the average number of storms per year from Table 3. In real time, the air mass sample of areal means >1.00 inch requires the longest amount of time to obtain significance for a 20% increase because of the infrequent occurrence of storms of this magnitude. In this real-

TABLE 4. Correlation between mean and maximum precipitation on a 49-gage, 400-mi^2 network for several synoptic storm types in the warm season.

Storm type	Correlation coefficient
Low centers	0.98
Air mass	0.87
Cold fronts	0.93
Warm fronts	0.93
Occluded fronts	0.92
Static fronts	0.86
Squall lines	0.89

time situation, the mean ($\geqslant$0.005 inch) and maximum rainfall are nearly the same for the crossover design, but the maximum rainfall becomes a slightly better verifier than the mean as the design becomes less efficient.

At this point it should be emphasized that the use of maximum rainfall as a verifying parameter is possible through use of a dense network (such as the East Central Illinois used here), which provides a reliable estimate of this parameter. In normal climatic networks, measurement of the areal maximum would be too inaccurate for use. An excellent correlation between mean and maximum rainfall was obtained on our network of 49 gages in 400 mi^2 as shown in Table 4.

Fig. 3 shows the number of storms required for different synoptic and precipitation types. These graphs indicate that differences in sampling requirements between synoptic and precipitation types become more pronounced as the design efficiency decreases. The detection requirements vary the most between data types for summer air mass storms and the least with thunderstorms. Overall, steady rain in the cold season has the lowest sampling requirements.

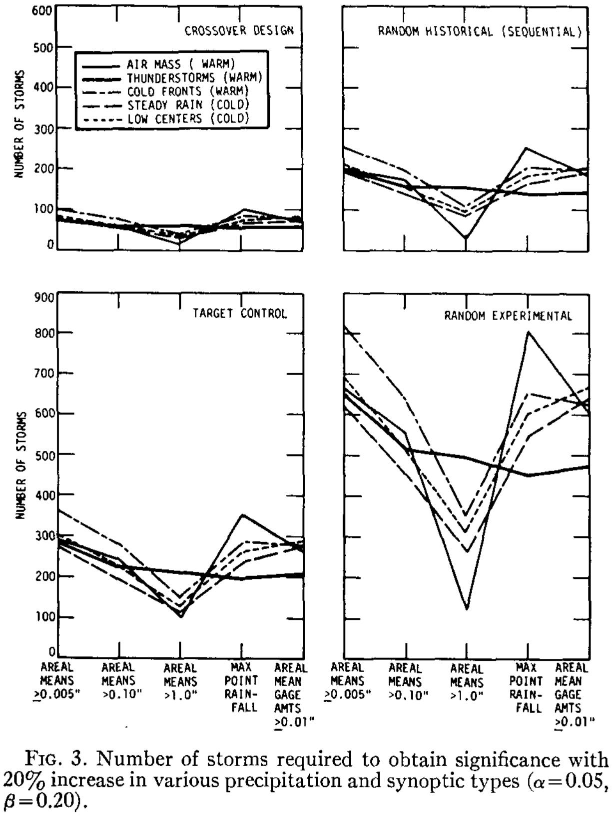

FIG. 3. Number of storms required to obtain significance with 20% increase in various precipitation and synoptic types ($\alpha=0.05$, $\beta=0.20$).

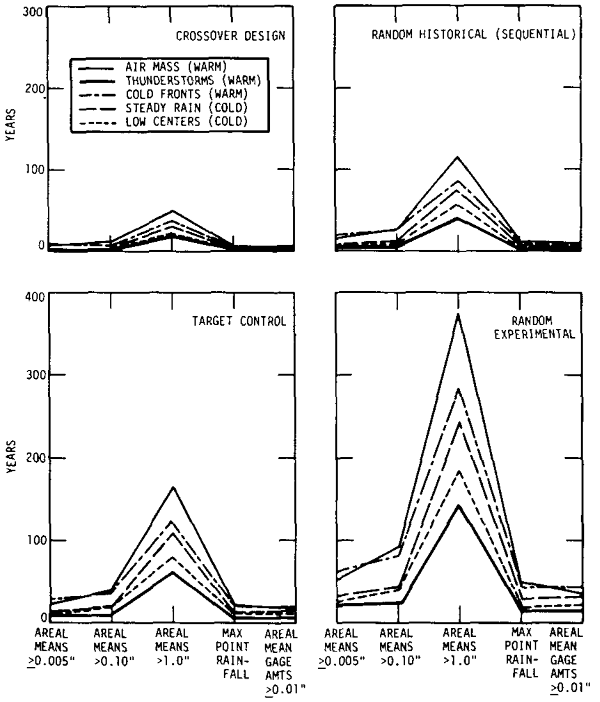

FIG. 4. Number of years required to obtain significance with 20% increase in various precipitation and synoptic types ($\alpha=0.05$, $\beta=0.50$).

Fig. 4 shows the real-time requirement for the various synoptic and precipitation types. Whereas the detection capability was variable between data types in Fig. 3, a clear-cut trend emerges in the real-time graphs of Fig. 4. For all designs, the air mass storm in summer requires the most time, followed by cold fronts (warm season), steady rain (cold season), low centers (cold season). Thunderstorms (warm season) require the least time because of their much greater frequency (see Table 3).

Fig. 5 shows the number of years required to obtain significance for various increases in areal mean rainfall for different synoptic and precipitation types according to statistical design. It is immediately obvious why there has been controversy in many weather modification experiments of five years or less. For the random-experimental design, more than 80% increase in air mass storms is required to obtain significance in a five-year experiment.

For the case of steady rain and low centers during the cold season, 60% increases would be detectable in a five-year period using the random-experimental design. An increase of 80% would not be detected in cold

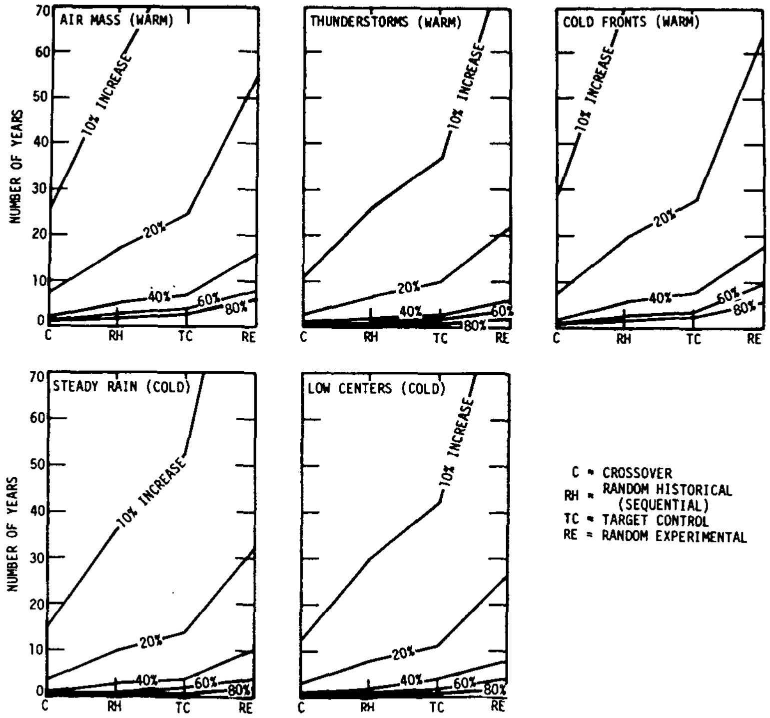

FIG. 5. Number of years required to obtain significance for various increases in areal mean rainfall (means $\geqslant 0.005$ inch) according to synoptic type, precipitation type, and experimental design ($\alpha = 0.05$, $\beta = 0.50$).

front storms and air mass storms using the random-experimental design. For steady rains and low centers, a 20% increase would be detected in five years using the crossover design. A 40% increase would be detected in summer air mass and cold front storms and a 20% increase would be detected within five years for warm season thunderstorms with this design. From Fig. 5 it is evident that the larger the seeding effect the less important the choice of design becomes.

The crossover design requires the smallest number of years to obtain significance. The random-historical design using a sequential test requires less than the target-control design, but for those increases which are detectable in a five-year experiment the difference between the two is a maximum of one year. An important conclusion from this graph is that if one desires to detect increases of 40% or less in the mean rainfall within a five-year experiment, some design other than random-experimental must be used.

TABLE 5. Number of years required to obtain significance combining all storms in warm and cold seasons ($\alpha = 0.05$, $\beta = 0.50$).

Sampling design	Years required for given increase (%) Warm season 20	40	60	Cold season 20	40	60
Random-historical (non-sequential)	8	2	1	6	2	1
Crossover	2	1	1	2	1	1
Random-experimental	16	5	2	13	4	2

As indicated earlier, the crossover requirements are based upon a correlation coefficient of 0.75 between target and control precipitation. This assumes a distance of ~30 mi between target and control. As the distance increases, the correlation decreases and sampling length would increase. For example, with a correlation coefficient of 0.50, the sampling time is approximately double that indicated in the preceding discussions.

Under some circumstances, it would be desirable to seed all types of storms during a period of deficient precipitation in Illinois. As indicated in Table 2, the log-normal distribution does not fit the data with the desired accuracy when all storms are combined into one sample. However, for purposes of comparison with the stratified data, the application of sample size equations for the log-normal test will give an approximation of the required sample sizes that might be expected if the nonparametric U test were used (Schickedanz and Huff, 1970). Table 5 provides estimates of the sampling requirements for three designs during the warm and cold seasons for assumed seeding-induced increases of 20, 40 and 60%. It is assumed that all storms will be seeded and mean rainfall is the verifying parameter.

At this time it should be pointed out that the length of sampling period discussed in this section is based upon seeding of all storms within a given stratification. Obviously, this would not normally occur in most seeding experiments. The estimates in the graphs and tables

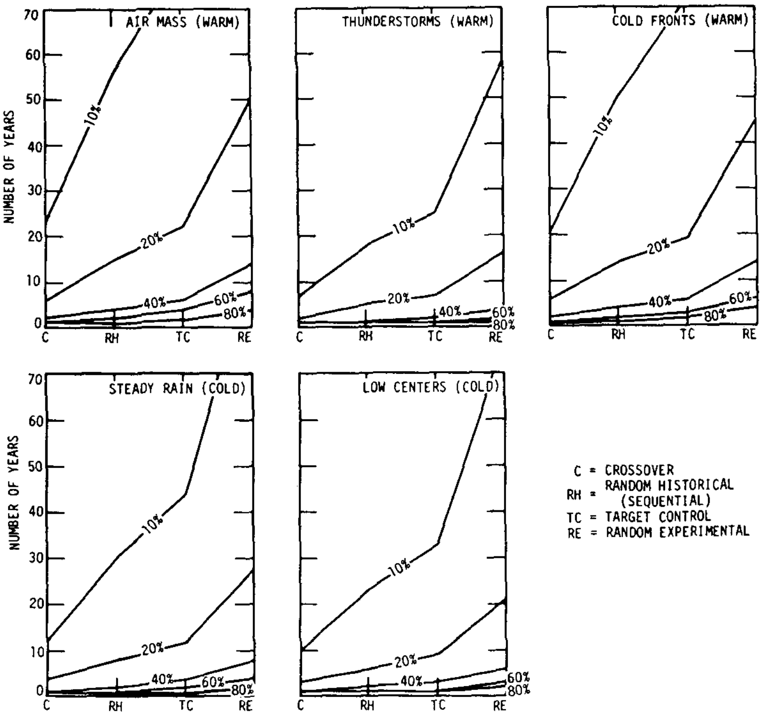

FIG. 6. Number of years required to obtain significance for various increases in maximum point rainfall according to synoptic type, precipitation type, and experimental design ($\alpha=0.05, \beta=0.50$).

can be readily modified to provide sampling lengths under any estimated seeding efficiency. For example, if an investigator estimates only 50% of the naturally occurring storms will be included in his seeding experiment, then the sampling length is doubled. If only 25% are to be included, the sampling time is four times the number obtained from the graphs.

Fig. 6 depicts the same information as Fig. 5 except maximum rainfall is used as the precipitation measurement rather than mean rainfall. The number of years required are generally similar to those for mean rainfall although the maximum rainfall is a slightly better predictor. This difference is largely due to the fact that the average number of storms for maximum precipitation is slightly greater than for mean rainfall (Table 3) in which only storms with areal means $\geqslant 0.005$ inch were used.

In general, the areal storm precipitation data lead to the following conclusions:

1) For the five types of precipitation and synoptic types emphasized in this study, the log-normal distribution provided a better fit than the gamma distribution.

2) The crossover design requires the least amount of time to obtain significance for a given increase, followed by the continuous-historical (sequential), continuous-historical (non-sequential), random-historical (sequential), fixed-target-control, and random-historical (non-sequential). The most time is required by the random-experimental design.

3) In terms of years, the difference between continuous-historical (sequential), continuous-historical (non-sequential), random-historical (sequential), and fixed-target-control is, in general, two years or less when a seeding increase of 40% or more is present in the data. However, the difference between the crossover and random-experimental design may be large.

4) The larger the seeding increase becomes, the less important is the choice of design and type of data measurement.

5) Mean and maximum rainfall require approximately the same amount of time to obtain significance.

6) The stratification of mean areal precipitation into samples of greater than 0.10 and 1.00 inch increases the amount of time to obtain significance.

7) To detect increases $\geqslant 40\%$ in areal mean storm rainfall, some design other than random-experimental must be used.

4. Area-depth data

The area-depth study described in this section is limited in application because the sample was restricted as a result of time and personnel restrictions. Thus, this is to be considered as a pilot or "feasibility" study into the desirability of using area-depth data in the verification of weather modification experiments.

a. General comments

The use of area-depth relations in the design and evaluation of precipitation modification experiments is a logical extension of the use of storm mean precipitation. Area-depth curves, frequently used in hydrological analyses, provide a simple mathematical expression of the spatial distribution of precipitation in a sampling area. The curve provides a measure of the mean and maximum precipitation, rainfall gradient, and skewness of the distribution. Thus, area-depth analyses provide indirect information on physical changes produced in precipitation systems (Huff, 1968). The primary purpose of this investigation was to determine whether the area-depth parameters (maximum rainfall and rainfall gradient) and the associated rainfall-depth (mean rainfall) can be used together to provide evidence in addition to that available simply from the one rainfall variable (mean rainfall).

If so, their use can facilitate verification and improve reliability of statistical evaluations of seeding results beyond that achieved with analyses of a single parameter, point or areal mean rainfall, which has been used most frequently in past experiments. The specific definition of the area-depth curve, the method of determination, and the rationale of using the square-root form of the area-depth equation is described elsewhere (Huff and Schickedanz, 1970).

Theoretical frequency distributions were fitted to the area-depth data, and the distributions were used to obtain sample size requirements for the five statistical designs (listed in the design section). The sample sizes were computed by the same methods used for the areal storm precipitation data.

TABLE 6. "Goodness of fit" probability* values for the log-normal distribution of area-depth data.

Stratification	Maximum point rainfall $\bar{y}_{\max}$ (inches)	Integrated rainfall gradient S (inches mi^{-1})	Average rainfall depth $\bar{y}$ (inches)
Warm season	<0.01	0.49	<0.01
Synoptic type			
Low centers	0.20	>0.20	>0.20
Air mass	>0.20	>0.20	0.08
Cold fronts	0.03	0.16	0.05
Warm fronts	>0.20	>0.20	>0.20
Occluded fronts	>0.20	>0.20	>0.20
Stationary fronts	>0.20	>0.20	0.10
Squall lines	>0.20	>0.20	>0.20
Precipitation type			
Steady rain	>0.20	>0.20	>0.20
Rainshowers	>0.20	0.14	>0.20
Thunderstorms	<0.01	0.88	0.09
Cold season	0.06	0.05	0.17
Synoptic type			
Low centers	0.08	>0.20	>0.20
Cold fronts	>0.20	>0.20	>0.20
Warm fronts	0.13	>0.20	>0.20
Precipitation type			
Steady rain	>0.20	>0.20	>0.20
Rainshowers	>0.20	>0.20	>0.20
Thunderstorms	>0.20	>0.20	>0.20

* Probability that the observed differences between the data sample and the given distribution could have occurred by random chance.

b. Theoretical frequency distributions

The log-normal distribution was fitted to the intercepts ($y_{\max}$) and the slope parameters (S) of the 228 storms. The log-normal distribution was also fitted to the areal mean rainfall ($\bar{y}$). For purpose of distribution fitting and subsequent statistical analyses, the absolute values of the S's were used. The "goodness of fit" test was applied and the resulting probability values are shown in Table 6.

In general, the data are well described by the log-normal distribution. For the individual types, cold fronts (warm) and thunderstorms (warm) are the only cases in which the "goodness of fit" probability is less than 0.05. Furthermore, one would expect 1 out of 20 distributions to have a "goodness of fit" probability <0.05. For the maximum rainfall data, there are 2 out of 16 distributions which have probabilities <0.05. The factors were viewed as justification for using the log-normal parameters in the subsequent calculations of sample size.

When all warm season storms were grouped together, the slope data were still described satisfactorily by the log-normal distribution, whereas the maximum and mean data were not. However, when cold season storms were grouped together, all three types of data were described adequately by the log-normal distribution. Thus, it would appear that the physical makeup of the different types may be distinctly dissimilar in summer storms but quite similar in winter storms.

Various groupings of synoptic and precipitation types for the warm season were then made in order to increase the number of storms in the various stratifications. Also, since seeding of more than one type of storm is likely to occur in either scientific experiments or commercial operations, it is of interest to determine which types are homogeneous enough to be grouped into one sample. For $y_{\max}$, the synoptic types of warm, cold and stationary fronts can be grouped together to form another sample. For mean rainfall, the low centers can also be grouped with the warm, cold and stationary fronts to form one sample. For precipitation types, steady rains and rainshowers can be grouped together for both $y_{\max}$ and $\bar{y}$. The slope parameters were similar enough to permit a single grouping of all the warm season parameters. Probabilities of the above samples being log-normal were all >0.05. In the area-depth study, the storm was again used as the experimental unit, and the size of sample necessary for the detection of the specified increase in the area-depth parameters due to seeding effects was computed for the five experimental designs.

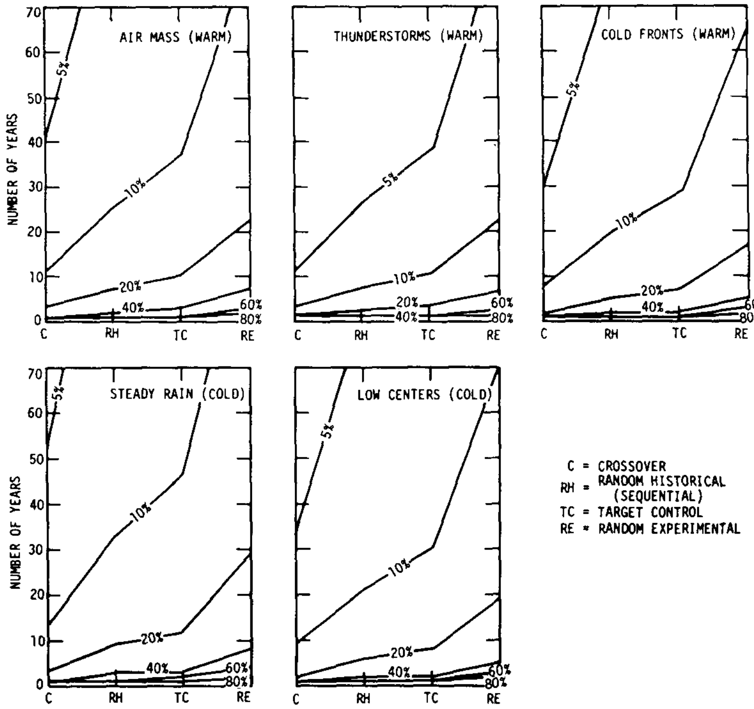

FIG. 7. Number of years required to obtain significance for various increases in maximum rainfall according to synoptic type, precipitation type, and experimental design for the area-depth data ($\alpha = 0.05$, $\beta = 0.50$).

c. Numerical results

The number of years required to obtain significance for various increases in maximum point rainfall derived from area-depth data is shown in Fig. 7. Warm season thunderstorms require the least amount of time to obtain significance mostly because of the large number of storms expected in this category (Table 1). Increases of 60% or more would be detected for all designs and all weather types within a five-year period. Increases of 40% would be detected within five years for all synoptic types with all designs with the exception of the random-experimental design in air mass (warm) and steady rain (cold). For thunderstorms, 20% increases would be detected in five years with every design except the random-experimental which would take six years to detect a 20% increase. For air mass storms, steady rain and low centers, the 20% change would be detected only by the crossover design, but in the cold front storms the 20% increase would be detected in both the crossover and the random-historical designs.

Similar graphs were constructed with the area-depth parameters for the slope parameter and for the mean rainfall. For air mass storms, the slope parameter requires the least amount of time to obtain significance followed by maximum rainfall and mean rainfall. For thunderstorms, cold front and steady rain, the least amount of time is required with maximum rainfall, followed by the mean and then the slope. For low centers, maximum rainfall requires the least amount of time followed by the slope parameter and the mean. The fact that the relation between the detection power of the data types varies according to the weather type that is seeded makes the use of additional parameters even more important.

The maximum rainfall parameter was more efficient than the mean rainfall for two reasons. First, it produced smaller sample sizes for the same percentage increase than the other parameters. Second, if the natural relationship between the parameters remains the same during seeding, then the percentage changes in the maximum rainfall witll be larger than in the other parameters (Huff and Schickedanz, 1970). If the relations between the parameters were not the same in the seeded sample, this change could be used 1) as a test of the effectiveness of the seeding, and 2) as an indication of the type of physical changes being produced. Thus, if in a seeding experiment one chose to use mean rainfall for the chief source of verification, then the maximum rainfall would provide extra information at a smaller significance level, and the slope parameter would provide extra information at a larger significance level because of the varying degrees of efficiency between the mean rainfall and the two parameters.

Conclusions from results of the area-depth study included the following:

1) Practically all of the area-depth rainfall stratifications used in this study can be described by the log-normal distribution.

2) There is a high degree of correlation between the mean and maximum rainfall ($\geqslant 0.92$ in the various stratifications) in the storm sample used in this study.

3) In general, maximum rainfall requires less time to obtain significance than mean rainfall for the area-depth sample used in this study.

4) The relative rank order of the two area-depth parameters and mean rainfall in respect to the power of detection varies according to synoptic type and other factors. Thus, it is very informative to use all three parameters in a verification program to provide additional information.

5) Many of the synoptic and precipitation types can be combined to reduce the amount of time required to obtain significance.

6) The crossover design requires the least amount of time to obtain significance for a given increase, followed by the continuous-historical (sequential), continuous-historical (non-sequential), random-historical (sequential), fixed-target-control, and random-historical (non-sequential). The most time is required by the random-experimental design.

7) The larger the seeding increase, the less important the choice of design becomes.

5. Summary and conclusions

a. Discussion of optimum data, test, design and data type

A summary of the more pertinent results from the area-storm data is shown in Table 7. However, the selection of the optimum design cannot be based upon sample size requirements alone. Table 8 is a comparison of selected relative advantages and disadvantages of the various designs. For example, Table 7 implies that the crossover design, under ideal conditions, can verify with small sample sizes. However, the crossover design requires a buffer area because of the danger of contamination, and in the Israeli experiment (Gabriel, 1970) there is strong evidence that seeding in one experimental area may have contaminated the seeding in the other experimental area as well as the buffer zone.

All designs involving historical data are handicapped by the likelihood that the density of raingages will be different in the experimental period than in the historical period.

Thus, the final decision must be based upon sampling requirements as well as feasibility of being able to meet realistically all assumptions. The decision in regard to

TABLE 7. Comparison of number of years to obtain significance for various designs, tests and type of data measurement ($\alpha=0.05, \beta=0.50$).

Design and data type	Air mass [warm] (% increases)			Thunderstorms [warm] (% increases)			Cold fronts [warm] (% increases)			Steady rain [cold] (% increases)			Low centers [cold] (% increases)		
	20	40	60	20	40	60	20	40	60	20	40	60	20	40	60
Crossover															
Areal means ⩾0.005 inch	7	2	1	3	1	1	8	2	1	4	1	1	3	1	1
Areal means >0.10 inch	12	3	2	3	1	1	10	3	2	6	2	1	5	2	1
Areal means >1.0 inch	48	15	6	18	5	3	36	11	6	31	9	5	23	7	4
Maximum point rainfall	6	2	1	2	1	1	6	2	1	4	1	1	3	1	1
Areal mean gage values ⩾0.01 inch	5	1	1	2	1	1	6	2	1	4	1	1	3	1	1
Random-historical (sequential)															
Areal means ⩾0.005 inch	17	5	3	7	2	1	20	6	3	10	3	1	8	2	1
Areal means >0.10 inch	28	8	4	8	2	1	25	7	4	13	4	2	13	4	2
Areal means >1.0 inch	115	33	18	44	13	7	88	26	13	75	22	11	58	17	9
Maximum point rainfall	15	4	2	5	1	1	14	4	2	8	2	1	6	2	1
Areal mean gage values ⩾0.01 inch	11	3	2	5	2	1	13	4	2	10	3	1	7	2	1
Target-control															
Areal means ⩾0.005 inch	24	7	4	10	3	2	28	8	4	14	4	2	11	4	2
Areal means >0.10 inch	40	11	6	11	4	2	36	11	5	19	5	3	18	5	3
Areal means >1.0 inch	164	48	24	62	18	10	124	37	19	107	31	16	81	25	12
Maximum point rainfall	22	6	4	7	2	1	19	6	3	12	4	2	9	3	1
Areal mean gage values ⩾0.01 inch	16	4	3	7	2	1	19	5	3	14	4	4	10	3	2
Random-experimental															
Areal means ⩾0.005 inch	54	16	8	22	6	4	64	18	10	32	10	4	26	8	4
Areal means >0.10 inch	92	26	14	26	8	4	82	24	12	44	12	6	40	12	6
Areal means >1.0 inch	375	110	54	142	42	22	284	84	44	244	70	36	186	56	28
Maximum point rainfall	50	14	8	16	4	2	44	14	6	28	8	4	21	6	3
Areal mean gage values ⩾0.01 inch	36	10	6	16	4	2	44	12	6	32	10	8	22	6	4

TABLE 8. A general comparison of selected advantages and disadvantages of various designs.

Design	Advantages	Disadvantages
Crossover	1. Uses only experimental data 2. Randomization in design 3. Requires small sample sizes	1. Difficult to position target and control in order to avoid contamination
Continuous-historical (sequential)	1. Requires relatively small sample sizes	1. Historical data sample may not be adequate 2. Sequential test is sensitive to trends 3. Lack of randomization may produce bias
Continuous-historical (non-sequential)	1. Requires relatively small sample sizes	1. Historical data sample may not be representative 2. Lack of randomization may introduce bias
Random-historical (sequential)	1. Requires relatively moderate sample sizes 2. Randomization present	1. Historical data sample may not be representative 2. Sequential test is sensitive to trends
Target-control	1. Requires sample sizes smaller than that of random experimental	1. Requires a high degree of correlation between target and control
Random-historical (non-sequential)	1. Randomization present in design	1. Historical data sample may not be representative of experimental period
Random-experimental	1. Only experimental data are used in analysis 2. Randomization is present in design	1. Often requires large sample sizes

the violation of assumptions must be, to a large degree, subjective in nature. The only way to answer the question of the effect of violation of assumptions other than subjectively would be to conduct repeated Monte Carlo trials in which various degrees of assumption violation are built into the method and then observe the total effect on the samples thus generated.

b. Recommended design and evaluation procedures

The crossover design will provide verification of seeding effects on surface precipitation quicker than the other statistical designs discussed in this study. Therefore, if assumptions are met, it is the first choice for the experimental design. However, it must be recognized that difficulties may arise in application of the crossover method when the characteristics of the precipitation distribution and prevalent types of storm systems in various areas of the country are observed. For example, the majority of Illinois precipitation results from extensive systems of relatively long durations associated with fronts and squall lines. Mesoscale circulations associated with such systems and interaction between neighboring elements can make the contamination problem difficult to overcome. From a practical standpoint, however, these storm systems should be subjected to seeding if substantial benefits are to be obtained from a seeding program.

If the contamination problem becomes too acute, however, it will be necessary to use a design employing randomization over a single target area, referred to in this study as the random-experimental design. However, this design results in the longest verification time of those investigated. To overcome the difficulties, the following alternative procedure is recommended.

First, the experiment should be developed as if one were going to use the experimental-random design, thus utilizing the lack of assumptions and the presence of randomization in the model. Most likely, the observations over the network for the period of the experiment will be more dense than those available from the historical record. However, every area will have some stations which have operated over a relatively long period of time. The nonseeded days can then be compared to the historical record as a test, in addition to the usual test comparison between the random samples of seeded and nonseeded days during the experimental period. Also, the nonseeded days can then be compared with the historical data in a test for trend or the representativeness of the historical record in relation to the experimental period. It is also possible that the historical nonseeded comparison could be used to remove trend from the data. Furthermore, even though the experiment would not be conducted in a sequential method, sequential tests cluld be applied and monitored along with the other tests. Thus, it would be possible by the use of additional tests and considerations to reduce the sampling size considerably.

The area-depth parameters could also be computed and used as another source of data to increase further the confidence in the results and to aid in reducing sample sizes. Thus, although the design would not be operated in the usual manner described for a sequential test, a stepwise sequential procedure could be used in which the results are being evaluated at various periods of time with the purpose of terminating the experiment when the accumulative weight of all tests indicate that the effect has been detected.

In addition, work is currently being performed on the possibility of reducing sample size requirements in the random-experimental design by utilizing predictor variables derived from the meteorological sounding. Research is also under way to determine the applicability of results from areas of less than 600 mi^2 to areas of larger size.

Recommendations concerning location of an Illinois experiment, precipitation measurement requirements, type of raingages, allied weather data, seeding method, time of year, and time of day have been presented by Huff (1970).

Acknowledgments. Appreciation is expressed to Mrs. Marion B. Busch who performed many of the computations used in the various analyses and to S. A. Changnon, Jr., for his review of this paper.

The research was supported by funds from the Section on Atmospheric Sciences of the National Science Foundation under NSF Grant GA-1360. Additional support came from the state of Illinois.

REFERENCES

Braham, Roscoe, Jr., 1966: Project Whitetop: Design of the experiment. Final Rept. Parts 1 and 2, University of Chicago, 156 pp.

——, and J. A. Flueck, 1970: Some results of the Whitetop experiment. *Preprints of Papers, Second National Conference on Weather Modification*, Santa Barbara, Calif., Amer. Meteor. Soc., 26–35.

Gabriel, K. R., 1970: *The Israeli Experiment 1961–67: Final Statistical Tables and Evaluation*. Jerusalem, Hebrew University, 47 pp.

Huff, F. A., 1968: Area-depth curves—A useful tool in weather modification experiments. *J. Appl. Meteor.*, **7**, 940–943.

——, 1970: Rainfall evaluation studies. Final Rept., Part 1, NSF Grant GA-1360, Illinois State Water Survey, 53 pp.

——, and P. T. Schickedanz, 1970: Rainfall evaluation studies, Part 2. NSF Grant GA-1360, Illinois State Water Survey, 224 pp.

Schickedanz, P. T., and Floyd A. Huff, 1970: An evaluation of downwind seeding effects from the Whitetop experiment. *Preprints of Papers, Second National Conference on Weather Modification*, Santa Barbara, Calif., Amer. Meteor. Soc., 180–185.

——, and Stanley A. Changnon, Jr., 1970: The design and evaluation of hail suppression experiments. *Mon. Wea. Rev.*, **98**, 242–251.

——, S. A. Changnon, Jr., and C. G. Lonnquist, 1969: A statistical methodology for the planning and evaluation of hail suppression experiments in Illinois. Final Rept., Part 2, NSF Grant GA-482, Illinois State Water Survey, 140 pp.

Simpson, J., G. W. Brier and R. H. Simpson, 1967: Stormfury cumulus seeding experiment, 1965: Statistical analysis and main results. *J. Atmos. Sci.*, **24**, 508–521.

Stall, John B., 1964: Low flows of Illinois streams for impounding reservoir design. Bull., 51, Illinois State Water Survey, Urbana, 395 pp.

Weinstein, A. I., and L. G. Davis, 1968: A parameterized numerical model of cumulus convection. Rept. No. 11, NSF GA-777, Dept. of Meteorology, Pennsylvania State University, 43 pp.