Raingage Network Requirements from a Simulated Convective Complex Weather Modification Experiment

JAMES A. HEIMBACH, JR.[1]

University of North Dakota, Grand Forks, ND 58202

ARLIN B. SUPER

Water and Power Resources Service, U.S. Department of the Interior, Miles City, MT 59301

(Manuscript received 18 March 1980, in final form 14 July 1980)

ABSTRACT

A convective complex weather modification experiment was simulated using Monte Carlo techniques. The purpose was to estimate the optimum raingage density for evaluation of a possible future experiment. The data base consisted of radar volume scans made within 150 km of Miles City, Montana, during May–July 1977. A total of 103 convective complexes were identified and tracked from radar data.

Raingage networks of various densities were simulated under the lowest-tilt radar scans to estimate total rainfall accumulation from each complex. Randomly chosen rainfall amounts were increased by given percentages to simulate assumed seeding treatments. A Monte Carlo scheme yielded estimates of the number of experimental units required for various combinations of α- and β-probability levels, treatment effects (percentage of increases) and raingage densities. Applying these results to the numbers of operationally available convective complexes as a function of area gave estimates of the optimal spacing and seasons required to detect a treatment. The results suggest that an unacceptably long field experiment would be necessary to detect treatment effects of 50% or less without some stratification of precipitation data.

1. Introduction

The High Plains Cooperative Program (HIPLEX) is tasked with removing critical scientific uncertainties from summer convective weather modification aimed at precipitation enhancement on the High Plains of the United States. Toward that end, a randomized experiment, known as HIPLEX-1, is currently being conducted with cumulus congestus clouds. If this experiment is successful, it is anticipated that future work will concentrate on larger, more complicated cloud systems referred to as convective complexes. This paper presents a preliminary attempt to estimate the raingage network requirements for a possible future weather modification experiment dealing with convective complexes.

Schickedanz and Changnon (1970) investigated optimal techniques for detecting cloud seeding effects for hail suppression. Their primary concern was the specification of the best statistical design, e.g., random crossover. The issue of experimental duration was addressed by Schickedanz and Huff (1971) who considered experimental design as related to several stratifications. In both of these papers, data bases were generated from distribution fits to gage network storm samples or hail pad/observer estimates of hail streaks. This precluded the examination of sampling variance.

Silverman (1979)[2] examined the effects of sampling variance in a simulation of a convective complex experiment. Hypothetical raingage networks were repeatedly placed over simulated rainfall accumulation patterns. The number of storms required to reach specified α- and β-probability levels increased with gage spacing.

This paper considers the combined magnitudes of the sampling variance due to the gage network and the natural or "among-storms" variance. The work was based on rainfall patterns, estimated by radar, in the vicinity of Miles City, Montana.

2. Design and evaluation considerations

a. *Experimental design*

The intent was to simulate a convective complex (CC) weather modification experiment. Such an

---

[1] *Present affiliation:* Institute of Natural Resources, Montana State University, Bozeman, MT 59717.

[2] Silverman, B. A., 1979: On the sampling variance of precipitation gage networks and its effect on the evaluation of precipitation augmentation experiments. *Extended Abstracts; Seventh Conference on Inadvertent and Planned Weather Modification*, Banff, Alberta, Amer. Meteor. Soc., 116–117.

0021-8952/80/101176-08$06.00

experiment can be considered successfully concluded when the treatment effect δ is detected at a specified significance level of α, where α is the probability of a Type I error—the probability of rejecting a true null hypothesis. In other words, accepting a seeding effect when, in fact, there is none.

It was assumed that the next phase of HIPLEX would treat CC's and would employ a random experimental design with a seed:no-seed ratio of 1:1. It was further assumed, for planning purposes, that it would be operationally practical to maintain a network of 250 raingages, whatever the spacing, and that operations would be limited to within a 150 km radius of the Skywater 5.4 cm radar at Miles City. This simplified the problem of determining the optimal placement of 250 gages within this specified area for a random experimental design. Obviously, use of a different experimental design, different number of gages, different sized area, etc., might change the results of this experiment somewhat. The technique, however, could be applied to widely differing conditions.

The ultimate measure of success for a precipitation augmentation experiment is the increased volumetric production of rainfall measured at the surface. Choosing this as the response variable allowed time-dependent errors to be ignored.

In this paper the term "sampling variance" is used to denote the variance attributed to the influence of gage network variations on estimates of total storm rainfall production. The familiar term "within-groups" (-storms) variance refers to gage-to-gage differences in rainfall amounts (plus measurement error). In a uniform network the storms total is the number of gages N times the average rainfall measured by the gages. Therefore, the sampling variance is N times the variance of means. Since the variance of means is estimated by the within-storms variance/N, the sampling variance is, in this case, equivalent to the within-storms variance.

b. Statistical techniques

The Monte Carlo technique has been used extensively in the design of weather modification experiments. Schickedanz and Decker (1969) used it to examine the effects of among-storms variance of Illinois daily rainfall on experimental duration. Schickedanz and Changnon (1970) applied the Monte Carlo technique to hail, and Olsen and Woodley (1975) to rainfall in the Florida Area Cumulus Experiment. These papers also addressed the effect of among-storms variance on experimental duration. Silverman (1979)[2] applied the Monte Carlo technique in his study of rainfall sampling variance. In his analysis, the among-storms variance from actual CC rainswaths in the Miles City area was compared to the simulated sampling variance and found to be far larger.

In this paper among-storms and sampling variances are left combined. This is a realistic approach because in an actual experiment the relative magnitudes can only be estimated at best. The degree of influence of the sampling variance is accessed by varying the simulated gage network spacing. The variance due to measurement errors by the raingages is ignored since it is assumed that it is an insignificant contribution.

In the techniques described in more detail below, randomly seeded experimental units (CC's) are added to develop an increasingly large sample until a specified α-significance is achieved for a particular treatment effect. This process is repeated to obtain a distribution of the number of CC's needed to reach the α-probability level, thereby allowing the power of the test, $1-\beta$, to be specified. β is the probability of a Type II error—the probability of accepting a false null hypothesis, i.e., rejecting a seeding effect when one actually exists.

The use of a nonparameterized statistical test in this process eliminates the restraint of requiring a distribution to be fit to the data base and cumbersome transforms of the data.

3. The simulated data base

a. Definition of convective complexes

A CC is considered to be a convective system larger than a field of cumulus congestus but smaller than a squall line (usually one or a few cumulonimbus in close proximity). The CC data base used in this study was obtained by the Skywater 5.4 cm radar located at Miles City, during May–July 1977. Volume scans were obtained each 5 min, from 1–12° of tilt, whenever echoes $\geq$20 dBz existed within 150 km between 1030 and 2400 MDT daily. Pairs of PPI displays were computer-generated for each volume scan with each print character ~4 km on a side. One PPI display portrayed the maximum echo tops and the other the maximum reflectivity factors (dBz) from all tilts for each horizontal location. These displays were used to manually identify and track each CC using 10 dBz as the echo boundary.

To qualify as a CC, a radar echo had to exceed 30 dBz, and reach an altitude of 9 km MSL or higher sometime during its lifetime, but not necessarily simultaneously. The separation between CC's had no absolute lower limit, but was subject to the interpretation of the analyst. Sometimes weak echoes (10–15 dBz) from adjoining CC's, probably representing cirrus anvils, would even touch for a short period with no evidence of merger of the CC's. Usually, however, individual CC's were separated by many kilometers for most or all of their lifetimes.

Many CC's were verified as such by the review of logs, aircraft and ground-based photography, and satellite imagery. In fact, it was this review, correlated with the PPI displays, that suggested the 30 dBz and 9 km MSL thresholds used in the definition.

Many CC's were excluded because they did not develop entirely within the 150 km radius area of radar coverage, moved within 25 km of the radar where ground clutter is a problem so no data are recorded, developed after dark, or merged with another CC in their first 40 min of existence with radar reflectivity ≥20 dBz. A total of 103 CC's formed the final sample.

b. Rainswath samples

Rainfall accumulation patterns for each of the 103 CC's was estimated from digitized radar data measured at the lowest (1°) tilt. The Z-R relationship used was developed near Miles City in western North Dakota:[3]

$$Z = 155R^{1.88}. \tag{1}$$

c. Raingage network simulations

A computer image of any simulated uniform raingage network can be constructed by specifying three parameters:

$$\left.\begin{aligned} \Delta x &= \text{north-south gage spacing} \\ \Delta y &= \text{west-east gage spacing} \\ \Delta z &= \text{eastward shift of successive rows} \end{aligned}\right\} . \tag{2}$$

Fig. 1 illustrates the use of these parameters. For a hexagonal array

$$\left.\begin{aligned} \Delta y &= 2\Delta z \\ \Delta x &= \sqrt{3}\Delta y/2. \end{aligned}\right\} . \tag{3}$$

The area represented by one gage is $\Delta x \Delta y$ which is the same as would be found by applying the Thiessen (1911) polygon.

Rainswaths were superimposed on simulated networks. Each gage's accumulation was estimated soley from the radar range bin (0.5 km by 1.0°) located directly above it. Each simulated network was placed under each CC once; i.e., there was no random shifting about to derive multiple samples because this would not simulate an actual experiment. Ten hexagonal raingage configurations were used. Except for the 41.6 km² per gage, the area represented by each gage was double the area of that for the next smaller simulated network. The simulated networks had a sufficient areal extent to include all CC rainswaths. "Ground-truth" rainfall accumulations in this simulation were assumed to be the contributions summed from all range bins.

It is recognized that the use of radar-estimated rainfall amounts as "true" values at each raingage

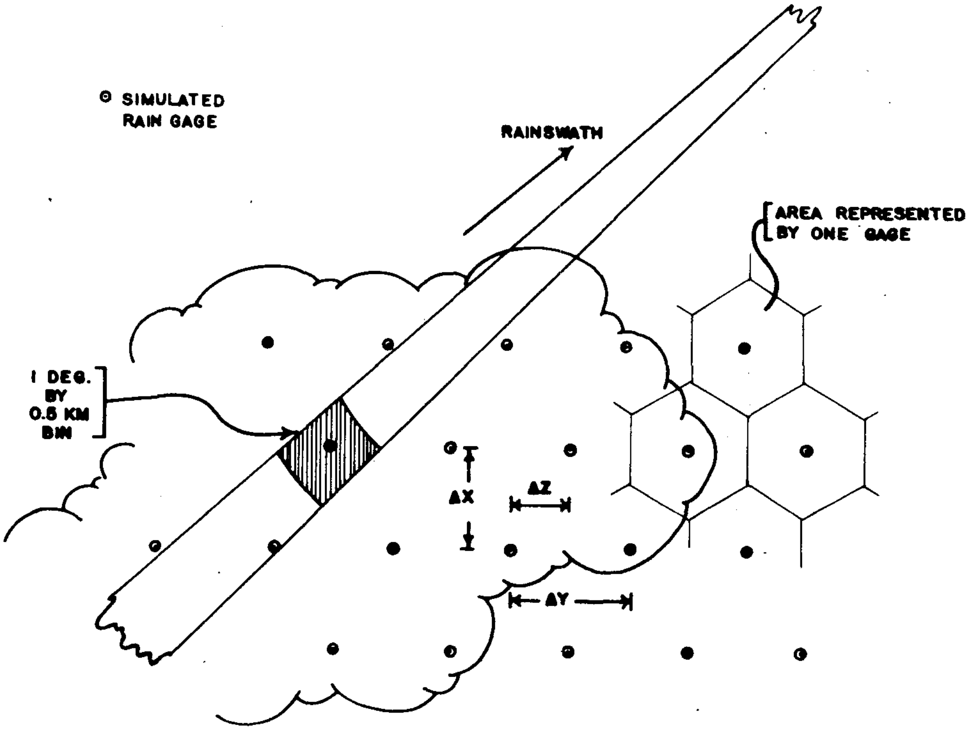

FIG. 1. Configuration of convective complex sampling.

---

[3] Smith, P. L., Jr., D. E. Cain and A. S. Dennis, 1975: Derivation of an R-Z relationship by computer optimization and its use in measured daily rainfall. *Preprints 16th Radar Meteorology Conf.*, Houston, Amer. Meteor. Soc., 461–466.

location is an oversimplification, especially as only a single climatological Z-R relationship was used. This may increase or decrease the simulated sampling and among-storms variances. While this effect is believed to be limited, there is no known way of evaluating it with the available data set.

Fig. 2 depicts the details of a particular CC sampled by the various networks used with the gage spacing and total rainfall accumulations noted. The CC was missed entirely by the 2771 km^2 per gage network and nothing more than a trace was detected for 1386 km^2 per gage. Although definition is lost, the total accumulation estimate remains reasonably stable through about 346 km^2 per gage. This example was classed as a small storm.

d. Characteristics of simulated CC samples

To estimate the impact of storm size, the 103 CC sample was divided into two halves based on the "ground truth" median of 2.47×10^5 m^3 precipitation accumulation. Consideration by size (total rain volume) should suggest how much the experimental duration might be reduced if an accurate forecasting or nowcasting ability was developed.

The mean rainfall accumulations $\bar{y}$ were rather consistant through the denser gage spacings for the smaller half, the larger half, and all 103 CC's combined (see Fig. 3). The reductions from "ground truth" to 10.8 km^2 per gage resulted from missing a few extreme range bin values. The variances s^2

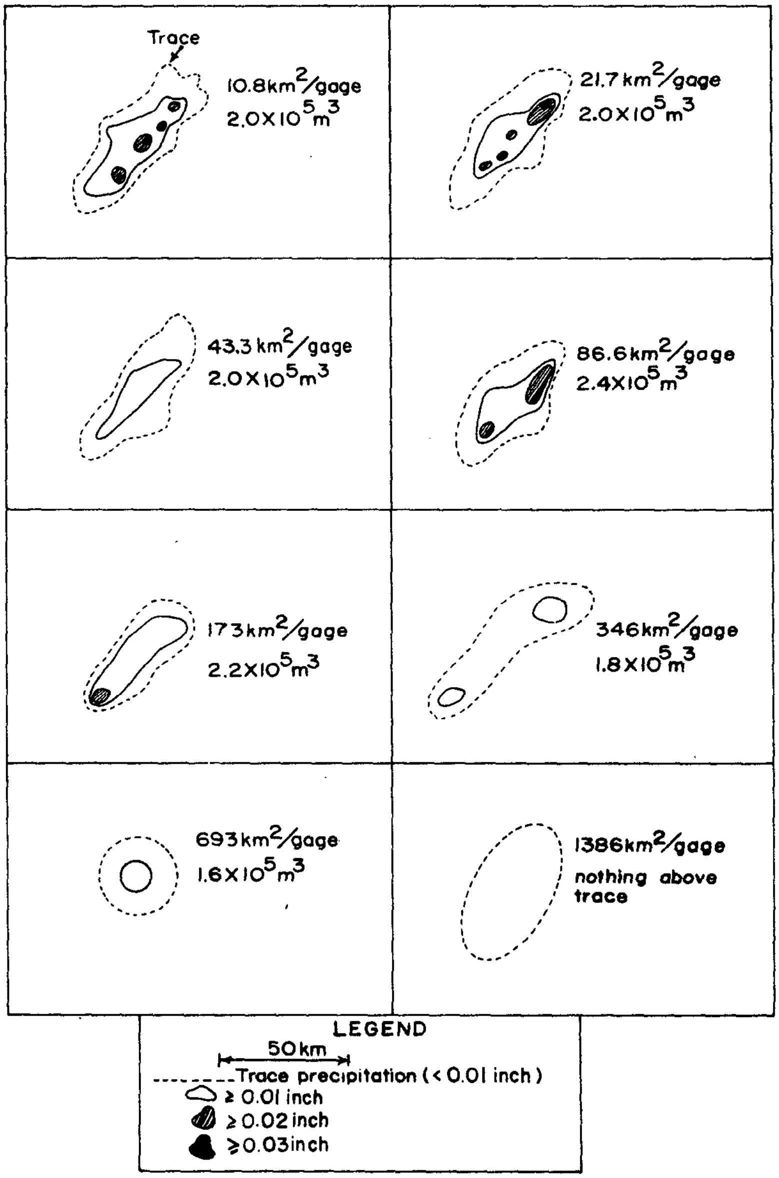

FIG. 2. Simulated sampling of 1 CC for various raingage densities.

which represent among-storms plus sampling variances, also decreased for the same reason. The decrease of $\bar{y}$ at large gage spacings is due to increasingly more CC's being missed entirely. From about 100–800 km^2 per gage, $\bar{y}$ increases for the larger CC's (also reflected in combined 103 samples) because the rainfall per gage changes little while the area per gage increases. The variances generally show an increase for wide gage spacings. The coefficient of variation, $s^2/\bar{y}^2$, which is an index of required sample size, shows a sharp increase for gage spacings greater than about 300 km^2. For the smaller half of the CC's this occurs before the 100 km^2 point.

4. CC's required to detect a treatment effect

The Wilcoxon (Mann-Whitney) rank-sum test (Noether, 1967) was applied to test for significance of treatment effect. Sample CC's were generated by randomly choosing a storm and then randomly seeding or not seeding it (1:1 ratio). Each CC was then "returned" to the data base so it could be chosen, at random, again. The rank-sum test was first applied after 15 samples were generated in this fashion, and then in increments of five samples until the required one-tailed α-level was reached. If more than 350 samples were needed, the process was halted to conserve computer time.

The convergence of four simulated experiments to an α-level of 0.05 (=1.645 standard-normal deviates of Wilcoxon rank-sum test statistic) is shown in Fig. 4. Rainfall accumulations summed from all range bins were used in this figure. An assumed treatment effect of $\delta = 25\%$ was added to the total rainfall volume of each randomly "seeded" CC. This figure exemplifies the hazards of detecting a seeding treatment with a large noise level. In one case only 25 storms were required, whereas one required over 350. If two of these experiments were terminated early, the conclusion would be a negative treatment effect!

This process of achieving a specified α-level was repeated 100 times for each assigned value of δ to give a frequency distribution of the number of CC's required. This was used to estimate $\beta(= 1 - \text{power})$. The number of storms required to reach a β-level was the value separating $100(1 - \beta)\%$ of the totals from the remainder of the 100 replications.

Figs. 5 and 6 illustrate the number of CC's required to satisfy various α, β and δ constraints. In several examples, more than the maximum allowable 350 storms were needed. Fig. 5 presents examples for all 103 cases for $\alpha = 0.05$ and $\delta = 100\%$. It can be seen that increasing the area per gage also increases the number of CC's required over most of the range considered. However, the increase in required CC's is rather gradual until about 300 km^2 per gage is exceeded.

For Fig. 6 the 103 CC's were stratified into two halves according to total rainfall volume. As might be anticipated, gage spacing is more important with the smaller storms, becoming critical for values

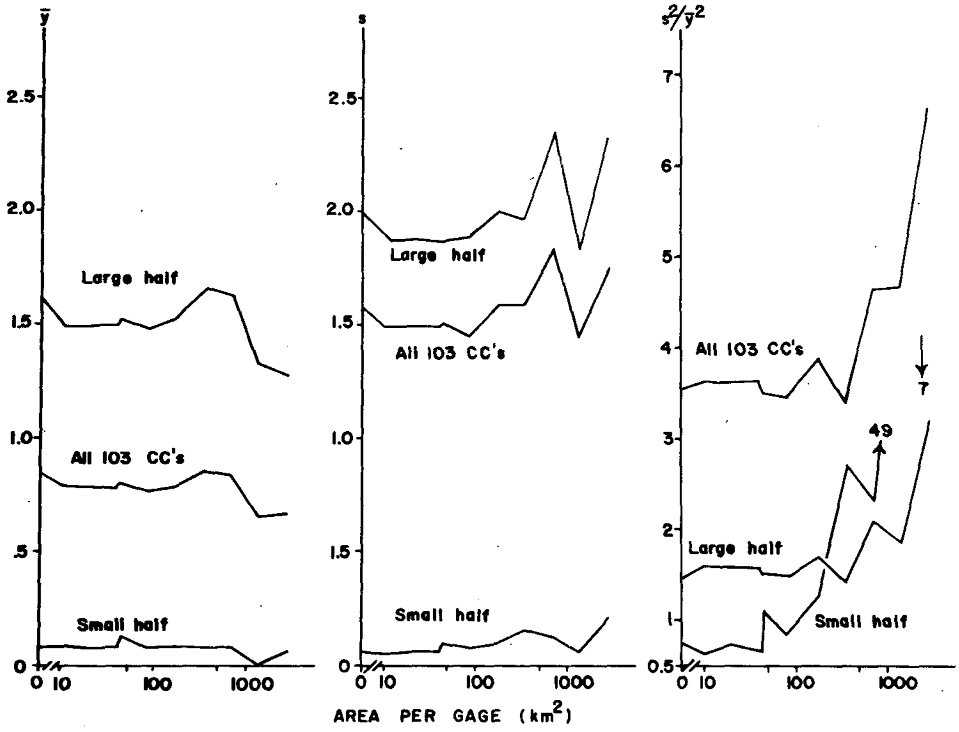

FIG. 3. Means $\bar{y}$, standard deviations $s(10^6\ \text{m}^3)$, and coefficients of variation $s^2/\bar{y}^2$, for rainswath accumulation versus raingage spacing.

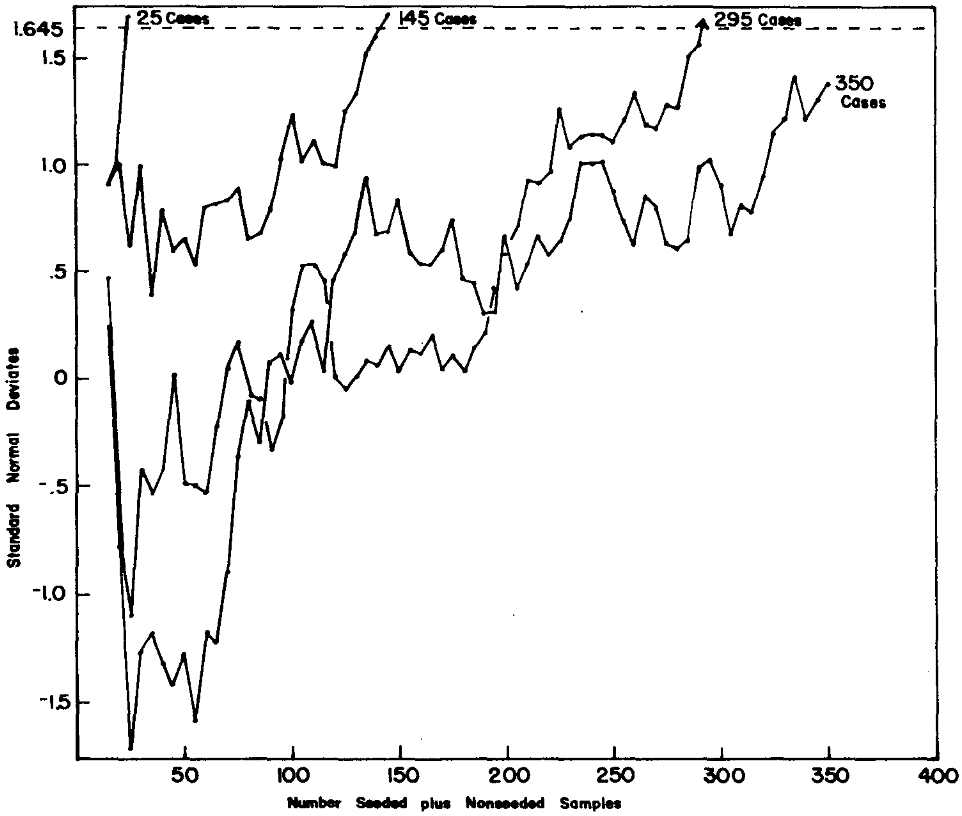

FIG. 4. Convergence of four simulated experiments to a one-tailed α-level of 0.05. (These examples used all 103 CC's derived from the complete radar footprint. The treatment effect was 25%).

> 100 km^2 per gage. For the larger storms, gage spacing has little effect until several hundred kilometers per gage is exceeded. For both large and small storms, a δ value of only 25% yields very large numbers of CC's required.

Comparison of the $\beta = 0.1$ case of Fig. 5 with case A of Fig. 6 leads to implications concerning the value of forecasting for a CC weather modification experiment. If a capability could be developed to accurately forecast, before seeding, whether a particular CC would become large or small, the number of cases required for a seeding experiment could be substantially reduced. For 100 km^2 per gage, $\alpha = 0.5$, $\beta = 0.1$ and $\delta = 100\%$, about 240 CC's would be required according to Fig. 5 where all cases were considered. Roughly half as many cases would be needed for the same conditions if the smaller CC's of Fig. 6 were the experimental unit. The situation is even better for the larger storms; however, it might be argued that they are unlikely to be the experimental unit in the near future due to complexities of achieving a large change in rainfall by seeding.

Speculating, then, that the smaller CC's would likely be the preferred experimental unit, Fig. 6 shows it would be desirable to have the area per gage less than about 100 km^2. Otherwise, the number of CC's required increases rapidly.

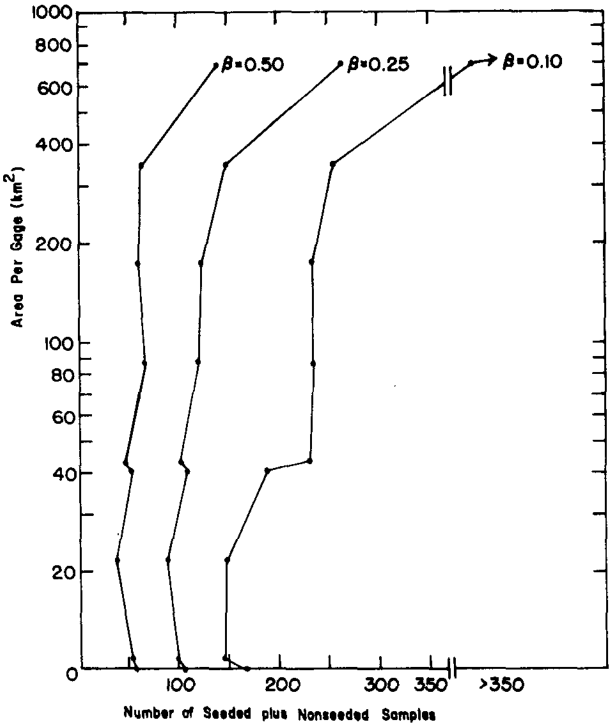

FIG. 5. Numbers of CC's required to reach example β-probability levels for $\alpha = 0.05$ and $\delta = 100\%$ using entire 103 CC data base.

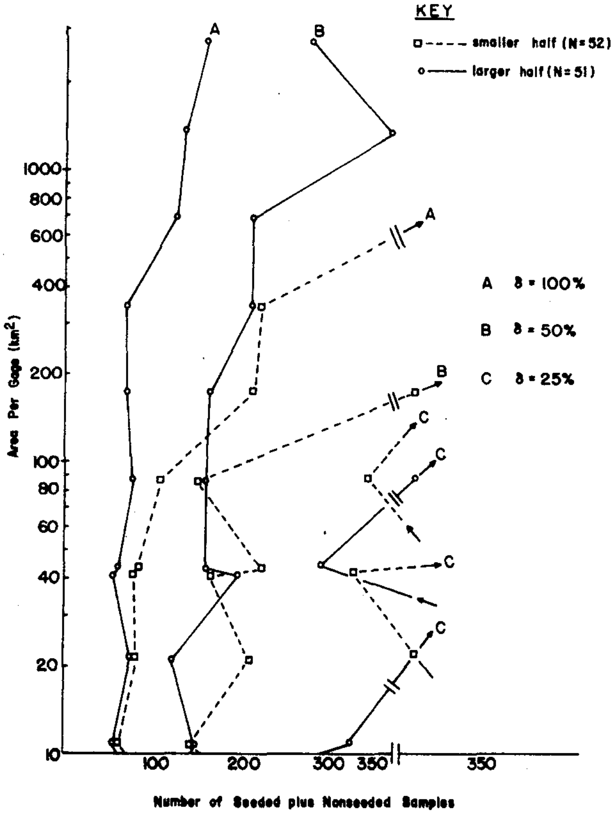

FIG. 6. Numbers of CC's required to reach $\beta = 0.10$ for $\alpha = 0.05$ using data base stratified by size and three treatment effects.

5. Seasons required to detect a treatment effect

Each of the 1977 CC's were evaluated to estimate which would likely have been operationally available had a seeding experiment been conducted. The PPI presentations for each volume scan were examined in chronological sequence for each day, and the experimentally suitable CC's were manually selected. To be a suitable experimental candidate, the CC could neither merge nor leave the 150 km radius of radar coverage within an hour of its first echo. Further, candidates had to be separated by more than 2 h as it was assumed that only one CC would be treated at a time, and that selection and treatment would require 2 h. The resulting numbers of CC's for the most active fractions of radar scope coverage are shown in Table 1.

TABLE 1. Operationally available convective complexes at Miles City, MT, 1977.

Radar scope coverage	⅛	¼	½	Full
Area of coverage (km²)	8875	17 750	35 500	71 000
Area per gage for 250 gage network (km²)	36	71	142	284
Convective complexes available:				
All 103 CC's	19	25	33	45
Small half	13	18	24	34
Large half	13	18	24	34

Interpolation from plots such as Figs. 5 and 6 yielded the numbers of CC's required for the four-gage densities listed in Table 1. Division by the number of CC's operationally available during the 1977 season (assumed typical) resulted in an estimate of seasons required to complete a seeding experiment. For example, using the unstratified data set (Fig. 5) and the probability levels $\alpha = 0.05$, $\beta = 0.10$ with $\delta = 100\%$, 10 seasons are required with 250 gages spaced at 36 km² per gage, 10 also for 71, 8 for 142 and 6 for 284 km² per gage. For these constraints, the optimal density for a 250 gage network would be 284 km² or more per gage; i.e., the gages should be spaced over the entire area of radar coverage. A summary of experimental seasons required for various conditions is given in Table 2.

In those cases requiring over 350 samples for all spacings, only a minimum number of seasons could be specified. This was more than seven years for the entire 103 CC sample and more than ten for the two halves.

6. An alternate approach

If the central limit theorem holds and the among-storms plus sampling variances are assumed equal for both seeded and non-seeded CC rainfall accumu-

TABLE 2. Numbers of experimental seasons required to detect δ-treatment effect and optimal areas over which to space 250 gages. (Number seasons/area expressed as fractions of 150 km radius radar coverage with *E* being the entire area. *I* implies optimal area indeterminant.)

		δ = 25%			δ = 50%			δ = 100%		
		All 103	Large half	Small half	All 103	Large half	Small half	All 103	Large half	Small half
$\alpha = 0.05$	$\beta = 0.10$	≥8/*I*	≥11/*I*	≥11/*I*	≥8/*I*	6/*E*	10/¼	6/*E*	3/*E*	6/¼
	$\beta = 0.25$	≥8/*I*	8/*E*	13/¼	≥8/*I*	3/*E*	6/¼	4/*E*	2/*E*	4/⅛
	$\beta = 0.50$	≥8/*I*	3/*E*	4/*E*	≥8/*I*	2/*E*	3/¼	1/*E*	1/*E*	2/¼
$\alpha = 0.10$	$\beta = 0.10$	≥8/*I*	10/*E*	14/½	11/½	4/*E*	8/¼	5/½	2/*E*	4/¼
	$\beta = 0.25$	11/½	5/½	7/¼	5/*E*	3/*E*	4/¼	3/½	2/*E*	3/*I*
	$\beta = 0.50$	4/½	2/*E*	3/*E*	2/*E*	1/*E*	2/¼	1/*E*	1/*E*	1/*E*

Table 3. As in Table 2 except required number of storms estimated using Eq. (4).

		δ = 25%			δ = 50%			δ = 100%		
		All 103	Large half	Small half	All 103	Large half	Small half	All 103	Large half	Small half
$\alpha = 0.05$	$\beta = 0.10$	43/*E*	24/*E*	27/½	11/*E*	6/*E*	7/½	3/*E*	2/*E*	2/½
	$\beta = 0.25$	27/*E*	14/*E*	17/½	6/*E*	4/*E*	5/½	2/*E*	1/*E*	2/½
	$\beta = 0.50$	14/*E*	8/*E*	9/½	4/*E*	2/*E*	3/½	1/*E*	1/*E*	1/½
$\alpha = 0.10$	$\beta = 0.10$	33/*E*	19/*E*	21/½	9/*E*	5/*E*	5/½	3/*E*	2/*E*	2/½
	$\beta = 0.25$	20/*E*	11/*E*	12/½	5/*E*	3/*E*	3/½	2/*E*	1/*E*	1/½
	$\beta = 0.50$	9/*E*	5/*E*	5/½	3/*E*	2/*E*	2/½	1/*E*	1/*E*	1/½

lations, then a simple approach outlined by Davis (1956) is applicable:

$$N = \frac{(z_{1-\alpha} - z_\beta)^2\sigma^2}{\mu^2\delta^2 p(1-p)} . \qquad (4)$$

Here N is the number of seeded plus nonseeded cases required to detect a δ-treatment effect at α- and β-probability levels. The standard normal deviates for the specified probability levels are represented by z. In this simulation the randomization factor p is 0.5 and the population variance σ^2 is assumed equal to the sample variance s^2.

Table 3 summarizes the optimal gage spacing determined through the application of (4). Approximately twice the number of experimental seasons were needed to detect a seeding effect of δ = 25% as compared to Table 2, approximately the same were required for δ = 50% and half the number of seasons for δ = 100%. Extrapolation of Table 2's indefinite values based on this comparison gives prohibitively large numbers of required seasons.

Table 3 is redundant in specifying optimal coverage for the data sets. The use of (4) does not invoke the random and multiple selection of CC's. Rather, the entire sample is always involved through the $\bar{y}$ and s statistics.

7. Discussion

This paper describes the simulated sampling of a convective complex seeding experiment by raingages. Several conclusions can be drawn. The among-storms or natural variance was much larger than the sampling variance. This generally resulted in little variation in the number of CC's required to reject the null hypotheses below ~300 km^2 per gage, at which point the sampling variance began to have a significant influence.

The usefulness of developing procedures to predict total CC rainfall accumulation was exemplified by stratification of the larger and smaller storms. This significantly reduced the number of seasons required, particularly for the larger half of the CC's.

For the modest treatment effect anticipated, it has been shown that an unacceptably long experiment would be needed to achieve acceptable α- and β-probability levels. One alternative would be to ease the α and β restrictions; however, the results would then be suspect. For this reason, it has been recommended that the Miles City HIPLEX raingage network deployment be postponed until covariates are found which reduce the among-storms variance.

Acknowledgments. The authors would like to express their appreciation to Dr. Paul Mielkie, Jr., for the fruitful discussions on several points in this experiment and to Ms. Ann Parker for her help in the analysis. This research was supported by the Office of Atmospheric Resources Management, Water and Power Resources Service, U.S. Department of the Interior, under Contract 14-06-D-7577 with the Montana Department of Natural Resources and Conservation.

REFERENCES

Davis, O. L., Ed., 1956: *The Design and Analysis of Industrial Experiments.* Hafner, 636 pp. (see pp. 24–31).

Noether, G. E., 1967: *Elements of Nonparametric Statistics.* Wiley, 104 pp. (see pp. 31–39).

Olsen, A. R., and W. L. Woodley, 1975: On the effect of natural rainfall variability and measurement errors in the detection of seeding effect. *J. Appl. Meteor.*, **14**, 929–938.

Schickedanz, P. T., and W. L. Decker, 1969: A Monte Carlo technique for designing cloud seeding experiments. *J. Appl. Meteor.*, **8**, 220–228.

——, and S. A. Changnon, 1970: The design and evaluation of hail suppression experiments. *Mon. Wea. Rev.*, **98**, 241–251.

——, and F. L. Huff, 1971: The design and evaluation of rainfall modification experiments. *J. Appl. Meteor.*, **10**, 502–514.

Thiesson, A. J., 1911: Precipitation averages for large areas. *Mon. Wea. Rev.*, **30**, 1082–1084.