A UNITED STATES
DEPARTMENT OF
COMMERCE
PUBLICATION

NOAA Technical Memorandum ERL WMPO-13

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Environmental Research Laboratories

On the Design and Evaluation of Cumulus Modification Experiments

PART I. Precipitation Correlations for Two Areas as Background for Cross-Over Experimentation in Florida

WILLIAM L. WOODLEY
JOYCE DONALDSON

PART II. On the Use of Predictors and Covariates

JOANNE SIMPSON
ANTHONY R. OLSEN
JANE C. EDEN

Weather
Modification
Program Office
BOULDER
COLORADO
March 1974

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Environmental Research Laboratories

NOAA Technical Memorandum ERL WMPO-13

ON THE DESIGN AND EVALUATION OF CUMULUS MODIFICATION EXPERIMENTS

PART I. PRECIPITATION CORRELATIONS FOR TWO AREAS AS BACKGROUND FOR CROSS-OVER EXPERIMENTATION IN FLORIDA

William L. Woodley
Joyce Donaldson

Experimental Meteorology Laboratory

PART II. ON THE USE OF PREDICTORS AND COVARIATES

Joanne Simpson
Anthony R. Olsen
Jane C. Eden

Experimental Meteorology Laboratory

Weather Modification Program Office
Boulder, Colorado
March 1974

TABLE OF CONTENTS

ABSTRACT

The main obstacle facing sound advances in weather modification is high natural variability coupled with the large expense and long effort required to obtain data. These problems become particularly severe when convective precipitation is the object of modification. In Florida, even on selected, relatively fair days meeting suitability requirements for seeding, cumulus rainfall varies naturally over two orders of magnitude. This spread maintains whether we are considering single isolated clouds or the six-hour rainfall in a target area of several thousand square miles. If there is only a small data sample, it is therefore possible for a few extreme values to dominate a seeded or control population and thence their comparison.

Over the past decade, the Experimental Meteorology Laboratory (EML) has conducted randomized dynamic seeding experiments on tropical cumuli, mainly in and near south Florida. The experiments have been designed and analyzed using numerical simulation, together with both Bayesian and classical statistics. The single cloud phase of the experimentation was brought to a definitive conclusion, with the results and procedures guiding development of a multiple cumulus seeding experiment in a 4900 mi^2 target. In the area experiment, randomization has been by days, with rainfall comparisons for both "floating" and total targets. Two preceding reports (Simpson, Eden, Olsen and Pezier, 1973; Simpson, Woodley, Cotton and Eden, 1973) summarize the effort up to this time.

A primary goal is to obtain an estimate of a seeding factor F, defined as the average multiplicative amount by which the seeding increases the rainfall in the seeded units relative to the controls. With several different analysis approaches, we found that for the "floating" targets the seeding factor may be large enough so that it can be evaluated with a reasonable amount of further experimentation of the same design, leading to about 20-30 pairs of cases in total. We specified how the number of cases required depends inversely on the seeding factor. Results indicated that in our total target area experiment, or in any case where a seeding factor of less than about 2 is anticipated, a prohibitively large sample of cases (perhaps 50-100 pairs) appears to be required. In addition, randomization by days has a further obvious weakness when evaluation is by radar, the calibration of which may vary from day to day and from one summer to the next.

To overcome these difficulties, new approaches to ex-

periment design, simulation or evaluation may be necessary. To explore possible improvements, EML has held series of meetings with prominent statisticians at intervals over the past few years. This report is the result of our first efforts to follow up the two most promising suggested improvements.

A solution to both the natural variation and the radar calibration problem would be provided if we could find two target areas within radar range which had nearly identical convective regimes that did not interact. Then we would select one area for seeding and the other for control on the same day using a randomized cross-over design. Thus, the basic question "what would nature have done without the seeding" is answered by the behavior of rainfall in the control area. The usefulness of this approach depends on the correlation of the natural rainfall in the two targets and its validity on the lack of mutual contamination. Part I of this report examines the prospects for this type of cumulus experimentation in south Florida.

An alternative approach is to devise a method for predicting the rainfall in the single target using one or more covariates. Clearly, the short-range forecasting of convective rainfall is a formidable task. Recently the Weather Service has advanced this problem, combining model predictions and statistical screening in a method called MOS or model output statistics (Klein, 1965; Klein, Lewis and Enger, 1959). In Part II of this report, we take the first steps in adapting this methodology to cumulus modification; we believe that the frontier so opened is an extremely promising one.

We are presenting this material at the current early stage for the following two reasons: first, other major cumulus modification experiments are presently in their design stage and we believe that our work should be of considerable benefit in obtaining optimal designs; second, we are developing and setting out a procedure for analysis of the 1973 Florida area experiment, using the 1970-1971-1972 data to construct the framework.

ON THE DESIGN AND EVALUATION OF CUMULUS MODIFICATION EXPERIMENTS

PART I

Precipitation Correlations for Two Areas as Background for Cross-over Experimentation in Florida

William L. Woodley and Joyce Donaldson

1. MOTIVATION

The Experimental Meteorology Laboratory (EML) has conducted a series of multiple cloud seeding experiments over a target area in south Florida (Simpson, Woodley, Olsen and Eden, 1973). These experiments have been based on the random experimental design which involves randomization of days over a single target area into seeded and nonseeded days with nonseeded as the control. An alternate design under consideration was the cross-over target control which requires random interchange of target and control areas among seeding days. The cross-over approach with a target and control area within the day was appealing in that this design minimizes the "noise" of natural rain variability inherent in the random experimental design (where the seed and nonseed cases are on different days). Further, cross-over procedures require less time to verify a particular seeding effect than does the random experimental design (Schickedanz and Huff, 1971). The disadvantages of the cross-over design are the contamination problem between target and control and the possibility of other effects (e.g. reduced insolation from cirrus canopy;

altered low level wind field, etc.) on the control area because of seeding in the nearby target.

The rejection by EML of the cross-over design in favor of the random experimental approach was dictated by practical considerations. It was impossible to select two land areas within range of the formerly used research radar that were free of blind cones produced by obstructions to the radar beam (hatched areas, fig. 1). However, it is now possible to make estimates of areal rainfall using the WSR-57 radar of the National Hurricane Center (Herndon, Woodley, Miller, Samet and Senn, 1973), and this radar has no obstructions to the energy radiated by its antenna. Because of this development, EML is re-examining cross-over procedures. A first step is the determination of how well the rainfall is correlated in the two areas selected for cross-over. Results are presented in this paper.

2. METHOD

Two land areas within range of the WSR-57 radar that might be used for cross-over experimentation were defined (fig. 2). A zone only seven miles in width buffers the two areas. The larger area (A) is the EML research area that covers approximately 4900 mi^2 and the smaller area (B) covers 2600 mi^2. The percentage coverage of echoes or rain was computed for each area for the nine hours of 1600 to 2400 GMT

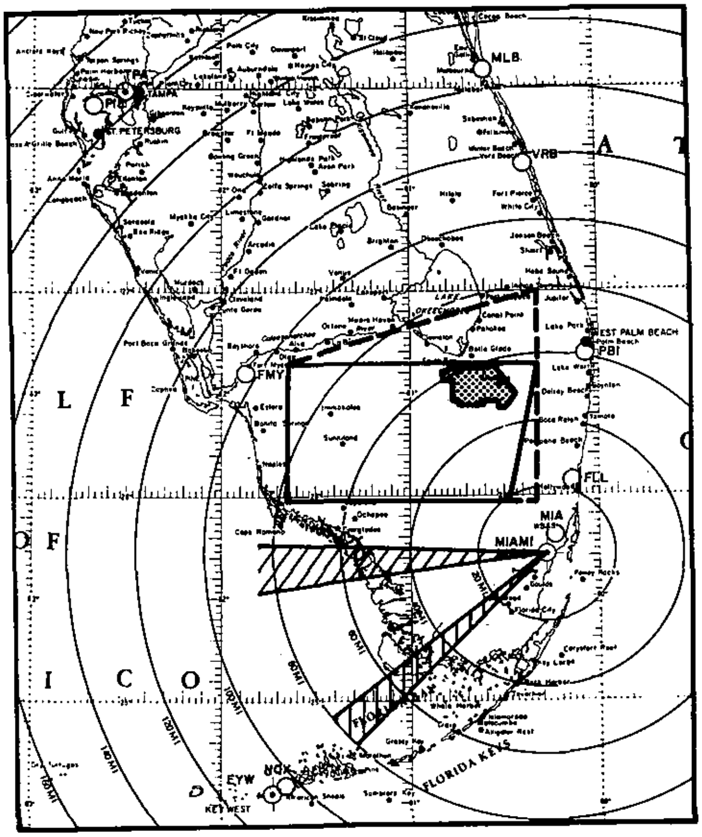


Figure 1. Map of south Florida showing the EML seeding target in 1970 (smaller area with solid line) and in 1971 through 1973 (larger area with solid and dashed lines). The stippled region is the EML meteorological network. The hatched sectors represent the blind regions due to obstructions to the S-band radar of the Radar Meteorology Laboratory of the University of Miami.

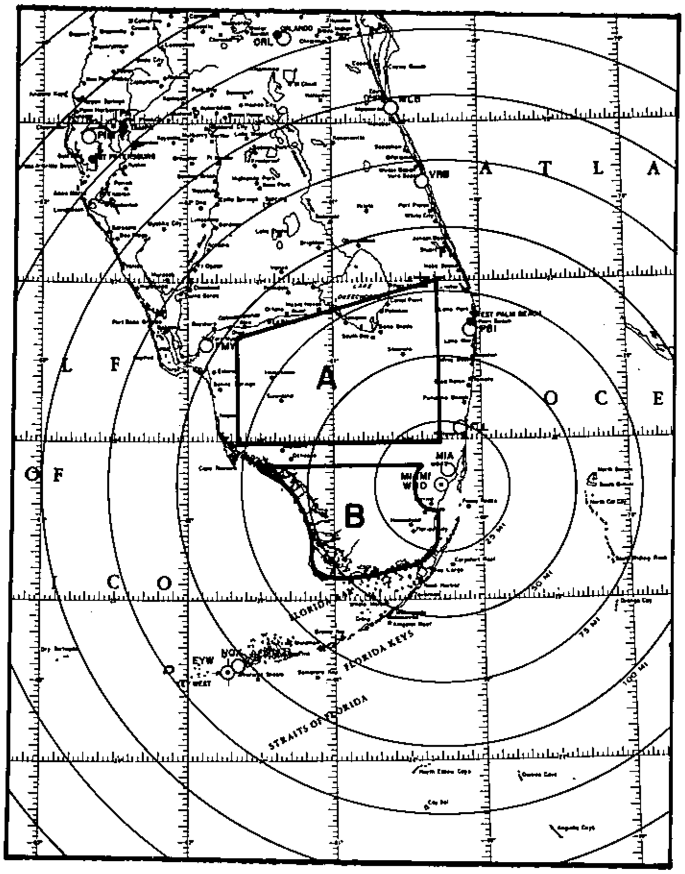


Figure 2. Two areas within range of the WSR-57 radar of the National Hurricane Center that might be used for cross-over experimentation in Florida. Area A is the EML seeding target shown in figure 1.

and the hourly calculations for the two areas were correlated after meteorological stratification. Those hours in which neither area had an echo were excluded. A total of 110 days were used in the study. These days were selected from April through August in 1968, 1970, 1971 and 1972, depending on whether a run of the EML cumulus model was available for the day.

The stratification of the days was made on the basis of the meteorological suitability factor (MSF), $S\text{-}N_e$, that is used by EML to determine the suitability of a day for seeding experimentation (Simpson and Woodley, 1971). The S is the maximum seedability (difference in kilometers between seeded cloud top height and unseeded cloud top height) predicted by the EML cumulus model using the 1200 GMT Miami radiosonde and a hierarchy of horizontal cloud sizes, and N_e is the number of hours between 1300 and 1600 GMT with S-band radar echoes in the target. The maximum value of N_e is 3.0. When the MSF is zero or negative, the day is usually dry ($S = 0$ and $N_e = 0$) or wet (S is small; $N_e = 3.0$ and $S\text{-}N_e \leq 0$), respectively. EML scientists believe that seeding conditions optimize with an increase in the MSF; a value of 1.00 to 1.50 is the minimum value of $S\text{-}N_e$ a day can have and still be acceptable for experimentation.

3. RESULTS

The mean percentage coverage of echoes with time for the two areas is presented in figure 3 as a function of S-N_e. The smaller region has the greater percentage of its area covered by echoes early in the afternoon, but the situation is reversed later. Surprisingly, the mean coverage of echoes changes little as a function of S-N_e, averaging 10 to 15 percent.

The correlation of the percentage coverage of echoes with time in the two areas as a function of S-N_e is presented in figure 4. With the exception of a brief period in early afternoon, the correlation of the precipitation coverages for the two areas decreases with an increase in S-N_e. The correlations are greatest on disturbed days (generally corresponding to $S - N_e < 0.00$) and least on "fair" days (corresponding to $S - N_e \geq 1.50$). Ironically, the days best suited for seeding experimentation are those days on which the two-area correlations are minimal. Still a modest positive correlation between the two areas is better than none at all. With zero correlation, cross-over is reduced to a target-control experiment within the day. This is still preferable to the random experimental design because two experiments are obtained for each day of operations using a within-day target-control instead of the one day obtained with the random experimental design.

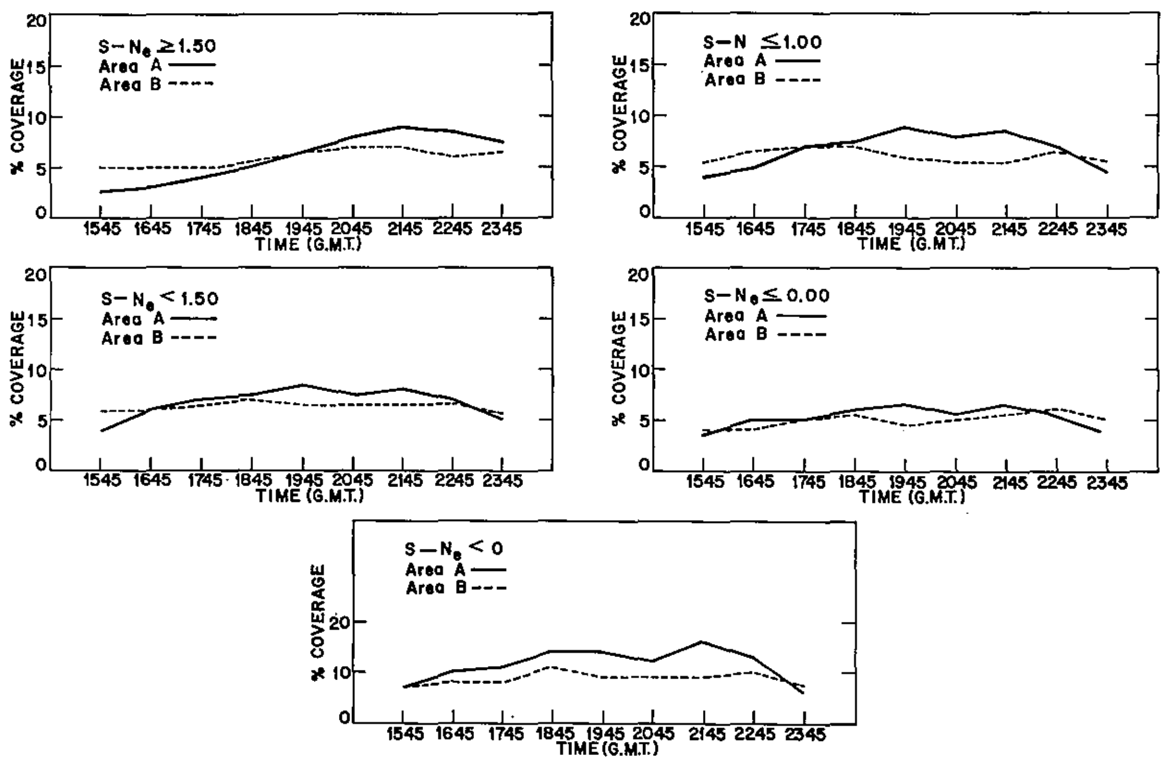


Figure 3. The mean percentage coverage of echoes with time for areas A and B as a function of $S\text{-}N_e$.

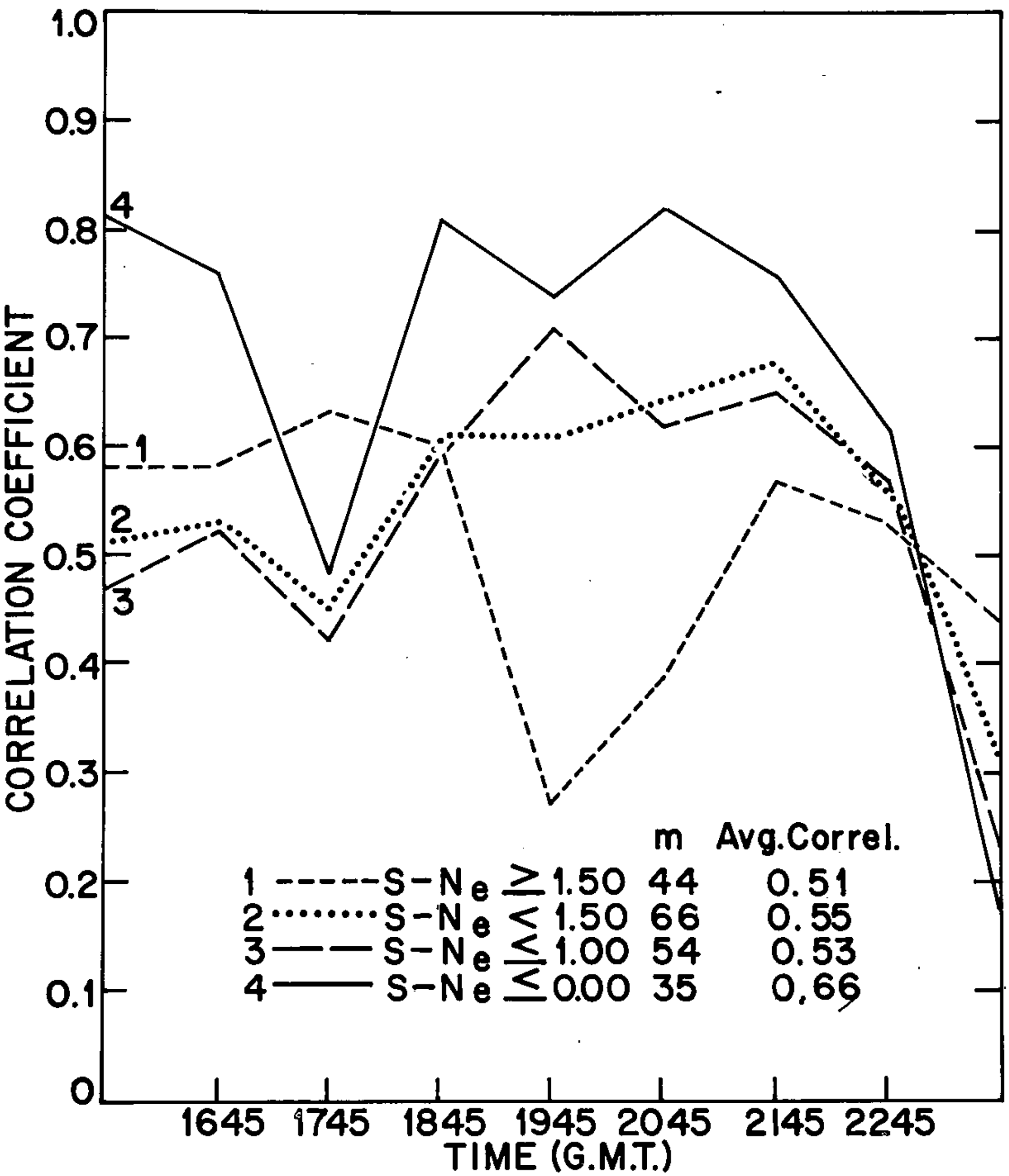


Figure 4. The correlation with time of the percentage coverage of echoes in areas A and B as a function of $S-N_e$. The number of days (m) in each category and the average correlation for the nine hour period are also shown.

In this study we also examined the correlations of the precipitation coverages in the EML target (area A) on successive days. This was done so that the intraday precipitation correlations for areas A and B (discussed above) could be compared with the interday correlations for area A. The pairings for the correlations were formed only for those successive days when $S - N_e \geq 1.50$, the EML suitability criterion for an experimental day. For example, suppose days A through G were under consideration and that $S - N_e \geq 1.50$ on days BCD and FG. The pairs for correlation for this data set are BC, CD, and FG.

After pairing of the days, correlations were calculated for each hour from 1545 to 2345 GMT and for the average precipitation coverage in this nine hour period. As before, hours when both components of a pair were zero were not used. Forty days were used to form the pairs. Results of the correlations by hour are shown in figure 5; the number of pairs used to calculate the correlation coefficient are indicated in parentheses above each plotted point. The correlations are small and inconsistent, suggesting that one cannot have any confidence that the precipitation coverage at any time on one day will bear any resemblance to that on the next, even if both days satisfy the criterion $S - N_e \geq 1.50$.

The situation does not change for the average coverage for the nine daylight hours. The correlation for this period

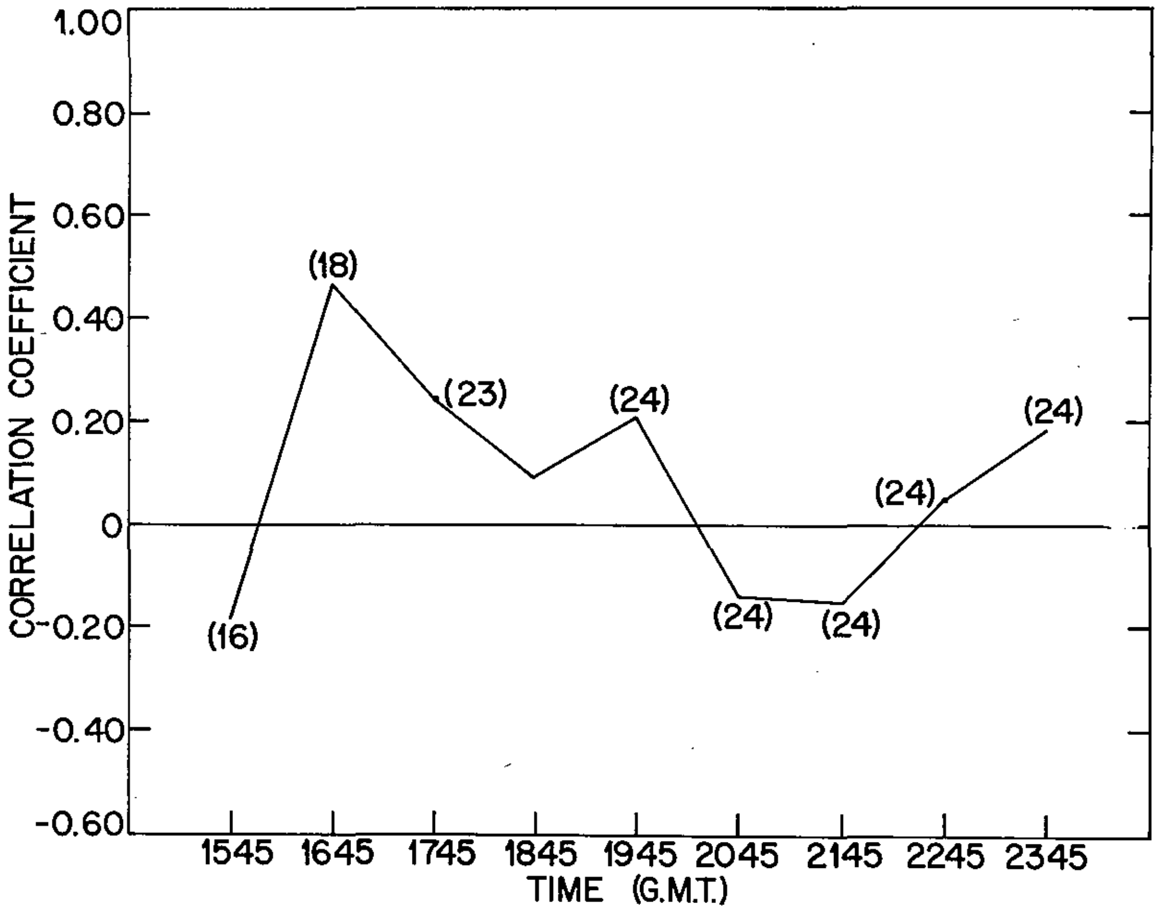


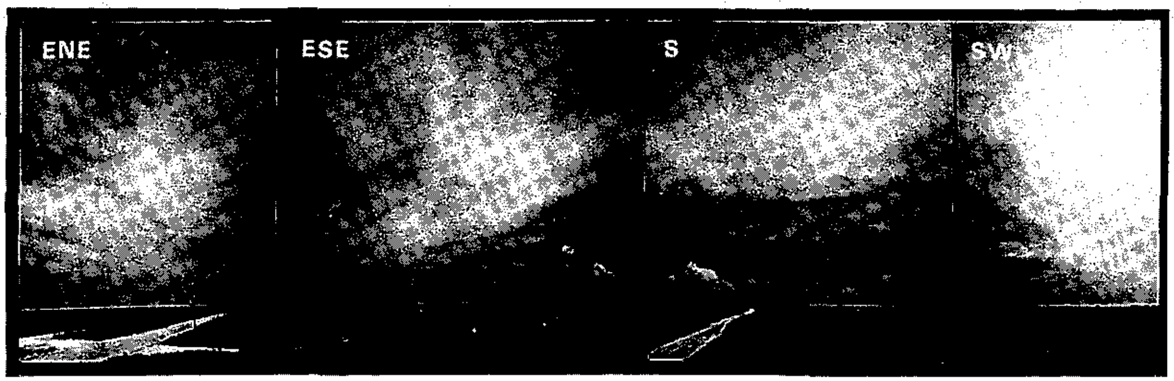
Figure 5. Interday correlations with time of the precipitation coverage in area A. The pairing of days for the purposes of correlation was made on the basis of $S\text{-}N_e$. The number of paired days used to calculate the correlation coefficient are indicated in parentheses above each plotted point.

was a disappointing 0.09. This correlation might have been somewhat higher if we had included the hours when both components of a pair had zero precipitation or if we had been able to stratify the pairs further on the basis of the $S - N_e$.

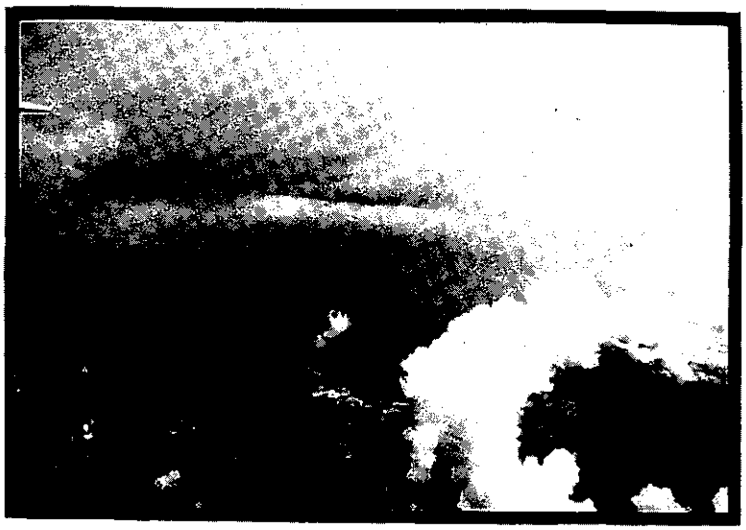
The $S - N_e$ criterion is a modest predictor of the rain coverage to be expected in the target. As an example, the correlation between the $S - N_e$ for values exceeding 1.50 and the mean target rain coverage is -0.41. This suggests that the greater the $S - N_e$, the lesser the mean target rain coverage for the day. This is exactly what was intended in the definition of $S - N_e$ but the strength of the relationship is weaker than was expected.

4. CONCLUSIONS

On days that meet the EML suitability criterion for an experimental day the percentage coverage of precipitation is better correlated for two areas on the same day than it is for a single area on paired days. This indicates that the cross-over design is superior to the random experimental design in Florida provided that contamination problems are of little consequence. Although the magnitude of this problem in Florida is not well known, we suspect that it would be very large and possibly disastrous. A common example of contamination is illustrated in figure 6. Frequently, overhanging anvils wipe out convection over major portions of the



1700Z AUGUST 5, 1973



1952Z 16 JULY 1973

Figure 6. Two photographs illustrating the effect that anvils derived from cumulonimbos can have on subsequent convection.

The upper photograph was taken from Central Site in the dense meteorological network that was operated in the summer of 1973. The lack of convection under the anvil is obvious. New cloud growth has begun in the clear region beyond the upper cloud.

The lower photograph was taken from the RFF C-130 flying at 19,000 feet p. alt. There is virtually no cumulus growth below the thick cloud anvil (left center foreground and background).

EML target area. Surface temperatures (measured by infrared thermometer) are often more than 5°C cooler under these anvils than they are with sunny conditions in the same location at the same time of day. Furthermore, our observations suggest that varieties of "dynamic" contamination are probably also operative. On some occasions, cumulus development appears enhanced, while on other occasions it is inhibited by the near neighborhood of huge merged systems. Because of the large unknown factor of cumulus interaction, we have for the time postponed serious further consideration of the crossover design in Florida.

The EML suitability criterion for a day of experimentation needs improvement. The $S - N_e$ factor does select days that are grossly similar meteorologically, but it is a relatively poor predictor of the rainfall to be expected in the EML target. The search for a more sensitive selector of experimental days is continuing, and progress is reported in Part II.

PART II

On the Use of Predictors and Covariates

Joanne Simpson, Anthony R. Olsen and Jane C. Eden

1. MOTIVATION

With the drawbacks of the cross-over design demonstrated in Part I, as well as the difficulty and expense of its execution in Florida, a remaining realistic hope of sharpening the multiple cumulus experiments lies in finding means to predict the unmodified rainfall. We believe, and shall show, that this does not involve solving the whole short-range precipitation forecast problem, but in isolating some aspects of it. The sources of guidance for this work were twofold: firstly, the Weather Service progress in use of Model Output Statistics (Klein *et al*, 1959; Klein, 1965) and secondly, the successful application by G. F. Cotton of covariate regression to the EML single cloud experiments (see Appendix - Statistical Analysis - to Woodley *et al*., 1970).

With the single clouds, Cotton (*loc. cit.*) found that the rainfall for the ten-minute period prior to the seeding run (initial wetness) was an extremely successful predictor for the subsequent rainfall for the control clouds. With the transformed (fourth root) rainfall data, the linear regression using initial wetness as a covariate reduced the variance by 32 percent and was significant at 0.5 percent. The differences of the seeded cloud rainfall observations from the control regression were significant at the 0.5 percent level. A similar linear regression for the seeded clouds was equally successful. An analysis of covariance was carried out which demonstrated that a fit of common slope and separate means was appropriate to the two populations, with the difference in means significant at 5 percent. A second predictor, namely degree of disturbance, or radar echo coverage at 2 p.m. local time, was also investigated; this predictor was found to contribute only a negligible further reduction in variance.

2. SEARCH FOR PREDICTORS IN THE AREA EXPERIMENT

Here we seek predictors for total target rainfall only. The floating target seeding factor appears nearer resolution and would present a more formidable problem in formulating predictors. Desirable criteria for predictors are that they should be readily available from routinely made observations, executed if possible prior to the time of the first seeding run. With the present pathetically small data sample (7 seed and 11 control cases from 1970, 1971, and 1972 combined) it does not make sense to utilize more than one or two predictors simultaneously. Later, with a larger data sample we hope to introduce other predictors by means of screening by stepwise regression.

We began with the hope, in view of the successful use of numerical simulation in these experiments, that one useful predictor could be found from the EML numerical cumulus modelling effort. After analyzing about six possible predictors, we have come up with three worthy of serious evaluation at the present time. These are: initial wetness of the area, model-predicted precipitation and degree of disturbance in the region.

3. DEFINITION AND EVALUATION OF THREE PREDICTORS

In the work to follow, we use transformed fourth-root rainfall amounts to eliminate extremes and to minimize heteroscedasticity. The predictand is the total target transformed rainfall for the six hours subsequent to the first seeding run on the 11 control and 7 seeded days of 1970, 1971 and 1972.

The predictors considered here are defined as follows:

1) Initial wetness W: transformed rainfall in the total target for the one-hour period ending at the first seeding run.
2) Model-predicted precipitation M: the EML one-dimensional cumulus tower model (Simpson and Wiggert, 1969; 1971) predicts, for given sounding and cloud base conditions, the precipitation production in cloud towers as a function of tower

radius. For the present study, the regular 1200 GMT Miami radiosonde observation was used, with a 915 m cloud base, which is the average cloud base. For each case, precipitation production was averaged for four tower radii, of 750, 1000, 1250 and 1500 m, the normal observed range of seedable tower sizes.

3) Degree of disturbance C: the radar echo coverage in nautical miles squared, measured within 100 n mi radius of Miami at 1800 GMT (2 p.m. local daylight time).

The general rationale for the above choices and definitions is that they express several of our physical hypotheses regarding controls on convective rainfall in terms of readily made, easily accessible observations. Specifically, predictors 1) and 3) are the same as those investigated by Cotton with the single clouds. Brier[1] has advised from long experience that for precipitation prediction, the best predictors generally involve precipitation at other places and/or times.

The model predictor was chosen for the explicit reason that with the single clouds the model-derived precipitation correlated at 0.9 with the measured precipitation from the

---

1 Personal Communication

cloud. Here we use the Miami 1200 GMT radiosonde for the model input rather than those closer in space and time to the experimental clouds, because the latter were not available in all area cases and furthermore, predictors will not be practically useful unless they can be derived from regularly made observations.

The 1970-1971-1972 data for predictand and predictors are presented in tables 1 and 2, for control and seeded clouds, respectively.

Table 1. EML Area Data - Total Target - 1970, 1971 and 1972 Control Cases

Date		$R^{1/4}$	W	M	C
		$(10^4$ acre-feet$)^{1/4}$		$g/g \times 10^2$	10^3 n. mi^2
1					
June 30		1.6237	0.8259	0.98	4.145
July 7		1.6565	0.5318	0.76	0.550
July 17	RC	1.4700	0.3720	1.17	0.775
1971					
July 1		1.1211	0.4450	0.44	0.711
July 12		1.6571	0.2300	0.93	0.365
July 15		1.1710	0.5630	0.19	0.974
July 16	RC	1.6266	0.4307	0.56	0.162
1972					
July 21	RC	0.6846	0.6623	0.71	0.147
Aug. 4	RC	0.7121	0.0000	0.89	0.000
Aug. 9	RC	1.3201	0.3246	0.60	0.084
Aug. 18	RC	1.2831	0.5563	0.91	0.126
Average		1.3025	0.4492	0.74	0.731

RC stands for "radar control", i.e., non-random control.

The rather unusual units employed with the data are used to keep all numbers involved in the same size range for convenience in programming.

Table 2. EML Area Data - Total Target - 1970, 1971 and 1972 Seeded Cases

Date	$R^{1/4}$	W $(10^4$ acre-feet$)^{1/4}$	M g/g x 10^2	C 10^3 n. mi.^{2}
1970				
June 29	1.3406	0.6751	1.02	1.000
July 2	1.1802	0.5185	0.76	0.710
July 8	1.8573	0.6766	0.96	1.275
July 18	1.7034	0.5070	0.84	0.042
1971				
June 16	0.7071	0	0	0.061
July 13	1.3150	0.1800	0.83	0.588
July 14	1.4878	0.2856	0.88	0.365
Average	1.3702	0.4061	0.75	0.577

4. DEVELOPMENT OF PREDICTOR REGRESSIONS

As previously mentioned, the small data sample sizes severely limit the number of predictors that can be considered. This limitation also applies to the evaluation of the regression relationships. Hence, only a preliminary analysis is given. In this section a description is given of all possible regressions of the three predictors on transformed total target rainfall, with control and seeded cases treated separately.

First, linear regressions were obtained using each predictor alone on seeded and control populations separately. Results are shown in tables 3 and 4. A plot of the data against initial wetness is shown in figure 7, with the least-squares best fit linear regressions. On the radar control day of 21 July, 1972, the onset of synoptic-scale suppression caused the virtual disappearance of rain from the target after about halfway through the six-hour measurement period. When this case is omitted, the correlation between target rainfall and initial wetness is 0.51 for the control clouds. The regression reduces the variance by 0.26. For the single cloud control population, Cotton found that the correlation of total rain with initial wetness was 0.57 and that the regression reduced the variance by 0.32. The larger correlations, reductions in variance and F - values achieved by the predictors with the seeded cloud data can at present only be

Table 3. Predictor Regressions for Total Target 1970-72 Transformed Rainfall on Control Days.

Predictors Used	Correlation with $R^{\frac{1}{4}}$	Reduction in variance	Computed F-value	Intercept	Regression coefficients for W	M	C
Initial Wetness W (all cases)	0.214	0.05	0.43	1.148	0.344	-	-
Initial Wetness W (omitting 7-21-72)	0.513	0.26	2.85	1.06	0.712	-	-
Model Pred. Precip. M (all cases)	0.22	0.05	0.45	1.096	-	0.279	-
Deg. of Disturb. C (all cases)	0.36	0.13	1.30	1.224	-	-	0.107
W and C (omitting 7-21-72)	0.516	0.267	1.27	1.039	0.800	-	-0.022
W and M (omitting 7-21-72)	0.598	0.367	1.94	0.796	0.766	0.324	-
M and C (all cases)	0.386	0.149	0.70	1.084	-	0.198	0.098
W, M and C (omitting 7-21-72)	0.631	0.398	1.32	0.649	1.094	0.417	-0.079

Table 4. Predictor Regressions for Total Target 1970-72 Transformed Rainfall on Seeded Days.

Predictors Used	Correlation with $R^{\frac{1}{4}}$	Reduction in variance	Computed F-value	Intercept	Regression coefficients for W	M	C
Initial Wetness W	0.70	0.48	4.71*	0.959	1.012		
Model Pred. Precip. M	0.82	0.66	9.91**	0.700		0.890	
Deg. of Disturb. C	0.42	0.18	1.10	1.172			0.344
W and M	0.82	0.68	4.17	0.707	0.244	0.746	
W and C	0.70	0.49	1.94	0.966	1.127		-0.093
M and C	0.82	0.67	4.09	0.697		0.955	-0.084
W, M and C	0.83	0.70	2.31	0.707	0.415	0.783	-0.168

* Significant at 10% level, $F_{1,5,.90} = 4.06$
**Significant at 5% level, $F_{1,5,.95} = 6.61$

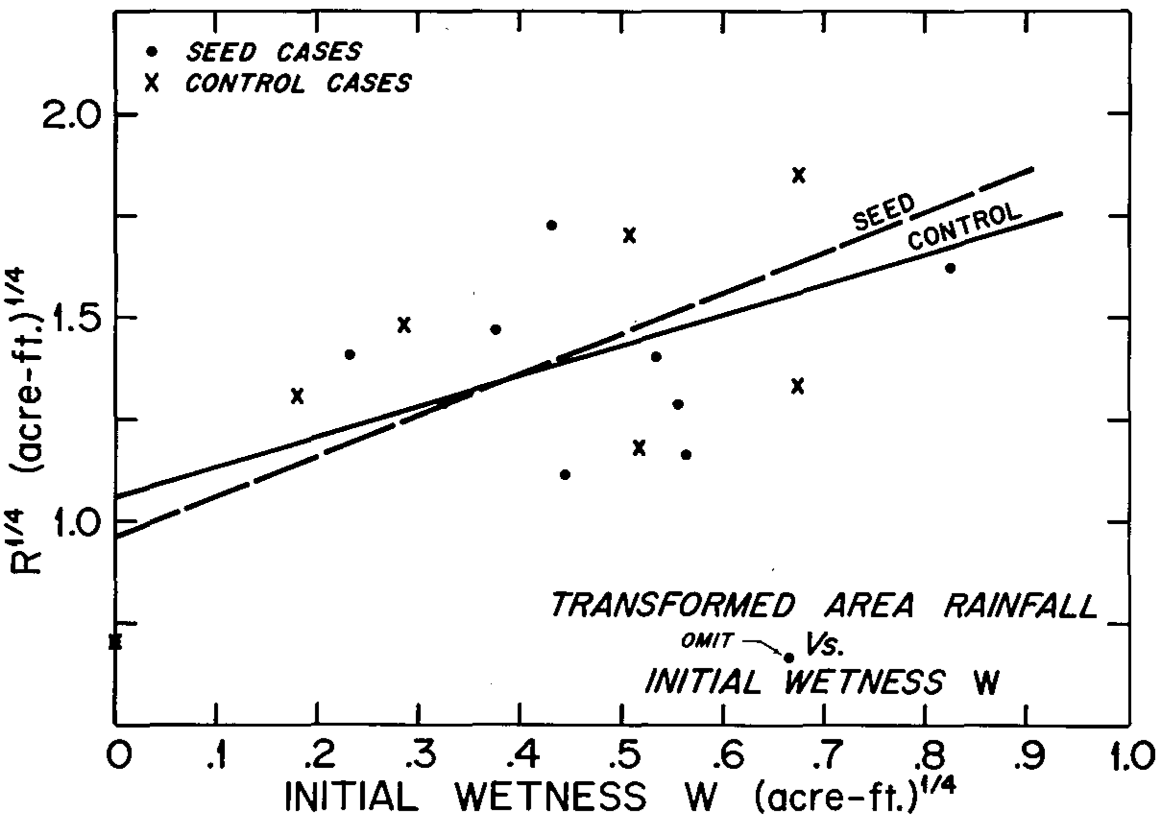


Figure 7. Plot of fourth-root transformed 1970-1971-1972 total target rainfall against initial wetness as a predictor. Lines are least squares fit to seed and control cases separately.

attributed to chance.

With the much larger sample of single cloud data, namely 26 seed and 26 control cases, Cotton (loc. cit.) found regressions which had very much larger reductions in variance and far greater significance with the F test. Table 5 presents a summary of his results for comparison. Comparison of table 5 with tables 3 and 4 forcefully brings home the handicap of a small data sample and emphasizes the urgency of obtaining more cases in the area experiment. It also suggests that even more control cases alone could be of great value, provided that bias or allegation of bias could be avoided. Table 5 further confirms our earlier deduction (Simpson, Woodley, Miller and Cotton, 1971) that on naturally rainy or disturbed days, seeded rainfall is diminished relative to fair days.

To determine if multiple linear regression relationships would be useful in predicting naturally occurring rainfall, regressions using the predictors in combination were constructed for the control cases with the results presented in table 3. For completeness, the same procedures were carried out for the seeded cases and the results given in table 4. In terms of both the variance reduction and the F test, the most successful combination of two predictors (see table 3) combines initial wetness with model predicted precipitation. This is the type of predictor combination we were seeking, in that it takes into account a model generated

Table 5. Single Cloud Regressions with Single Covariate

A. 26 Control Clouds

I. Initial Wetness

Correlations with Total Rain = 0.57

Regression: $R^{1/4} = 1.924 + 0.73\,W$

Reduction in variance = 0.324

$F = 11.55$; $F_{1,25,.995} = 9.48$; $F_{1,25,.999} = 13.88$

II. Degree of Disturbance

Correlation with Total Rain = 0.006

Regression: $R^{1/4} = 2.886 + 0.004C$

Reduction in variance = 4×10^{-5}

$F = 0.0009$; $F_{1,25,.75} = 1.39$

B. 25 Seeded Clouds*

I. Initial Wetness

Correlation with Total Rain = 0.41

Regression: $R^{1/4} = 2.788 + 0.611W$

Reduction in variance = 0.17

$F = 4.63$; $F_{1,24,.975} = 5.72$; $F_{1,24,.95} = 4.26$

II. Degree of Disturbance

Correlation with Total Rain = -0.307

Regression: $R^{1/4} = 4.16 - 0.216C$

Reduction in variance = 0,094

$F = 2.40$; $F_{1,24,.90} = 2.93$; $F_{1,24,.75} = 1.39$

* One of the 26 random seeded clouds lacked a measurement of initial wetness.

parameter and a measured precipitation variable, in this case related to time-wise persistence.

5. COVARIATE REGRESSION ANALYSIS

One method of using the predictor regression relationships established in section 4 for testing for a seeding effect is covariate regression coupled with the t-test. The basic procedure consists of using the control sample to establish a predictor regression for naturally occurring rainfall. This regression relationship is then used to predict the naturally occurring rainfall on seeded days. The difference between the predicted and observed transformed rainfall on seeded days is computed and the t-test is used to test if the difference is equal to zero.

Table 6 presents the analysis when using the best combination of two predictors for the regression relationship. The seed cases are seen to have an average departure from this regression for the controls that is not even significant at the 25 percent level.

Although a three-predictor regression is not really justified with this small data sample, we have nevertheless tried it in table 7. As expected, no improvement is obtained. The results of tables 6 and 7 show there is no significant difference between the seeded and control populations in the total target rainfall. Next we show the analysis of variance tables associated with the covariance analysis of the data, a more direct way of incorporating the control sample in the testing procedure for a possible seeding effect.

Table 6. Covariate Regression for 1970-72 Total Target Rainfall Using Initial Wetness and Model Predicted Precipitation as Predictors.*

Date	$R^{\frac{1}{4}}$ obs.	$R^{\frac{1}{4}}$ pred.	d
1970			
June 29	1.3406	1.6436	-.3030
July 2	1.1802	1.4394	-.2592
July 8	1.8573	1.6245	.2328
July 18	1.7034	1.4565	.2469
1971			
June 16	0.7071	0.7960	-.0889
July 13	1.3150	1.2028	.1122
July 14	1.4878	1.3000	.1878

$\bar{d} = .0184$
$s_d = .2341$
$t = \bar{d}/(s_d/ 7) = .2076$
$t_{6,.10} = 1.44$
$t_{6,.25} = 0.718$

*Regression equation: $R^{\frac{1}{4}} = 0.796 + .766W + .324M$

Table 7. Covariate Regression for 1970-72 Total Target Transformed Rainfall Using Initial Wetness, Model Predicted Precipitation and Degree of Disturbance as Predictors*

Date	$R^{\frac{1}{4}}$ obs.	$R^{\frac{1}{4}}$ pred.	d
1970			
June 29	1.3406	1.7338	-.3932
July 2	1.1802	1.4770	-.2968
July 8	1.8573	1.6880	.1685
July 18	1.7034	1.5507	.1527
1971			
June 16	0.7071	0.6442	.0629
July 13	1.3150	1.1456	.1694
July 14	1.4878	1.2996	.1882

$\bar{d} = .0074$
$s_d = .2446$
$t = \bar{d}/(s_d/\ 7) = .080$
$t_{6,.10} = 1.44$
$t_{6,.25} = 0.718$

*Regression equation: $R^{\frac{1}{4}} = .649 + 1.094\ W + .417\ M - .079\ C$

6. ANALYSIS OF COVARIANCE

Instead of using the control data to establish a regression between the predictors and the transformed rainfall and then applying the regression equation to predict the unmodified rainfall on seeded days, it is possible to perform an analysis of covariance utilizing both the seed and control cases to determine the appropriate regressions. The analysis is the usual covariance analysis for two treatments in a one-way classification, with the exception that more than one covariate is used. For this reason the analysis is presented in terms of the residuals after fitting various regressions.

Table 9 presents the analysis when the covariates, i.e., concomitant variables, are initial wetness and model predicted precipitation. Under the usual assumptions of the analysis of covariance, i.e., the same slope for both seed and control regressions, the test for the difference between adjusted means is not significant at even the ten percent level. This effectively tests for an additive seeding effect, since it is designed to test for additive shifts in the adjusted mean. In view of the postulated multiplicative seeding effect, a test was made jointly for a difference in slopes between the control and seeded regressions. Again, the result was not significant at even the ten percent level.

Hence, with the present data sample no difference in the seed and control regressions can be detected.

Table 9. Tests of Significance of Difference Between Regressions of Control and Seed Days With Initial Wetness and Model Predicted Precipitation as Covariates.

Residual After Fit of	SS	d.f.	MS	F value
Common Regressions	.86515	14		
Separate Means, Common Slopes	.86408	13	.06647	
Difference Between Adjusted Means	.00107	1	.00107	.016
Common Regressions	.86515	14		
Separate Regressions	.81852	11	.0744	
Difference Between Regressions	.04663	3	.0155	.209

$F_{1,13,.90} = 3.14$ $F_{3,11,.90} = 2.56$

Table 10. Tests of Significance of Difference Between Regressions of Control and Seeded Days With Initial Wetness, Model Predicted Precipitation and Degree of Disturbance as Covariates.

Residual After Fit of	SS	d.f.	MS	F-value
Common Regressions	.83643	13		
Separate Means, Common Slopes	.83622	12	.06969	
Difference Between Adjusted Means	.00021	1	.00021	.003
Common Regressions	.83643	13		
Separate Regressions	.76542	9	.08505	
Difference Between Regressions	.07101	4	.01775	.209

$F_{1,13,.90} = 3.14$ $\quad$ $F_{4,9,.90} = 2.69$

Table 10 presents the same types of analyses as table 9, except that now three covariates are used - initial wetness, model predicted precipitation and degree of disturbance. Clearly, the addition of the third covariate did not affect the results of the previous covariance analysis. In both cases the sparse data samples really preclude making any definitive interpretation of the testing procedures. However, if further data continue to result in reasonable regression fits, as in section 5, then this analysis should demonstrate if differences exist between control and seed regressions.

7. ANALYSES USING MULTIVARIATE DATA MATRICES

One of the unfortunate aspects of using two or more predictors or covariates in a regression analysis is the inability to plot, conveniently, the data points in three or more dimensions. Gabriel (1972) has suggested the use of biplots based on the canonical decomposition of a particular form of a data matrix. The biplot is a graphical display of a two dimensional approximation to a matrix. The approximation is obtained by least squares, using the two singular value components associated with the two largest characteristic roots of the matrix of sums of squares and products.

In the process of obtaining the biplot, an inspection of the data matrix and its Moore-Penrose inverse leads to a procedure for the possible detection of data outliers either in the original variables or in the variables adjusted by the regression on all remaining variables in the data matrix. This procedure has recently been proposed by Gabriel and Haber (1973). Furthermore, it is possible to perform both parametric and non-parametric test procedures for the possible effect of a treatment, in this case a seeding effect.

Let Y denote the 18×4 matrix of deviations from the variable means including both the control and seeded data. The columns of Y correspond to the variables $R^{1/4}$, W, M and C respectively with the matrix given as

$$
Y = \begin{bmatrix}
.2950 & .3934 & .2339 & 3.474 \\
.3278 & .0993 & .0139 & -.121 \\
.1413 & -.0605 & .4239 & .104 \\
-.2076 & .0125 & -.3061 & .040 \\
.3284 & -.2025 & .1839 & -.306 \\
-.1577 & .1305 & -.5561 & .303 \\
.2979 & -.0018 & -.1861 & -.509 \\
-.6441 & .2298 & -.0361 & -.524 \\
-.6166 & -.4325 & .1439 & -.671 \\
.0086 & -.1079 & -.1661 & -.587 \\
-.0456 & .1238 & .1639 & -.545 \\
.0119 & .2426 & .2739 & .329 \\
-.1485 & .0860 & .0139 & .039 \\
.5286 & .2441 & .2139 & .604 \\
.3747 & .0745 & .0939 & -.629 \\
-.6216 & -.4325 & -.7461 & -.610 \\
-.0137 & -.2525 & .0839 & -.083 \\
.1591 & -.1469 & .1339 & -.306
\end{bmatrix}
$$

An inspection of the individual columns of Y for possible outliers, large deviations from the variable mean, suggests that the first observation on the variable degree of disturbance is much larger than the remaining deviations. This is also visually detected in the biplot of Y presented later. The inspection of Y in this manner is a subjective way of detecting univariate outliers.

Although this procedure detects univariate outliers, it does not check for departures from a multivariate pattern, in particular linear multivariate patterns. Gabriel and Haber (1973) have developed a methodology based on the computation of the Moore-Penrose inverse of Y that is useful in exploring departures from such linear multivariate patterns. In this case the Moore-Penrose inverse of Y is given by $Y^+ = (Y'Y)^{-1}Y'$ and the ith element of the jth row of Y^+ is proportional to the deviation from the value predicted for the ith observation by the regression of variable j on all the variables. Thus, an inspection of the rows of Y^+ will lead to the subjective detection of outliers from the linear multivariate pattern that may not be detectable from a similar inspection of Y.

After performing the indicated matrix manipulations, the transpose of the Moore-Penrose inverse for the example is

$$(Y^+)' = \begin{bmatrix} -.0538 & -.1300 & .0119 & .2516 \\ .1966 & .1226 & -.0958 & -.0416 \\ -.0231 & -.2034 & .3325 & .0117 \\ -.0500 & .0920 & -.2069 & .0124 \\ .2245 & -.3996 & .0828 & .0016 \\ .0258 & .2578 & -.4578 & .0170 \\ .2733 & .0472 & -.2465 & -.0540 \\ -.4725 & .6955 & .1600 & -.0861 \\ -.3395 & -.4969 & .3685 & .0379 \\ .0987 & -.0573 & -.1283 & -.0324 \\ -.1003 & .3401 & .1591 & -.0836 \\ -.1461 & .3433 & .2107 & -.0251 \\ -.1280 & .1839 & .0505 & -.0114 \\ .2172 & .1616 & -.0087 & -.0090 \\ .2219 & .1688 & -.0277 & -.0899 \\ -.0671 & -.4342 & -.4152 & .0619 \\ .0301 & -.4440 & .1010 & .0449 \\ .1108 & -.2518 & .0868 & -.0054 \end{bmatrix}$$

An inspection of the first row of Y^+ indicates that the eighth observation may be an outlier in the linear regression on the remaining variables. Similarly, the second row gives the same indication for the eighth observation. Neither of these points were detected as univariate outliers.

On the other hand the first deviation of the fourth row of Y^+ shows up as both a univariate and multivariate outlier. The first multivariate outlier corresponds to July 21, 1972, and an explanation for its deviation in terms of the total rainfall and initial wetness was given in section 4.

So far in this section no use has been made of the first 11 observations corresponding to control days and the last seven corresponding to seed days. Since the first variable is the total (fourth root) rainfall, the first column of Y can be used to test for a seeding effect. Under the assumptions of approximate normality, independence of observations, and equal variances for seed and control populations, the usual t-test for difference between two means is computed and gives a computed t-value equal to 0.377.

Since only the total rainfall is expected to be affected by the seeding treatment, the treatment means may be computed from the first row of Y^+ and will provide a comparison of total rainfall adjusted for its regression on the remaining variables. Since the remaining variables are not affected by the treatment, the adjustment should enhance the sensitivity of the comparison. Performing the t-test calculation on the first row of Y^+ gives a computed t-value of 0.548, an increase over the previous t-value indicating an increased sensitivity for the test. As explained in more detail by Gabriel and Haber (1973), the Moore-Penrose adjustment is similar to that in an analysis of covariance except that the inversion involves "total" regressions, whereas the analysis of covariance uses "within" regressions.

The basic usefulness of the biplot is the ability to have a two dimensional plot of a matrix Y. For the present example the rows of Y would have to be plotted in four dimensions, with the biplot enabling a two-dimensional approximation to be plotted. By basing the procedure on the canonical decomposition, which is related to a principal component analysis, the "best" two-dimensional approximation to Y can be written as GH' where

$$G = Y(\lambda_1^{-1} q_1, \lambda_2^{-1} q_2)$$

$$H = (\lambda_1 q_1, \lambda_2 q_2)$$

and λ_1^2 , λ_2^2 are the two largest characteristic roots and q_1 , q_2 the corresponding characteristic vectors of $Y'Y$.

The biplot for Y is contained in figure 8 where Arabic numerals are the plots of the rows of G and the vectors with Roman numerals are plots of the rows of H. Note that the first row of G when plotted shows a possible outlier that agrees with the visual inspection of Y. The clustering of points indicates similar variation of the variables within the clusters and differing variation between clusters. This characteristic can be used visually

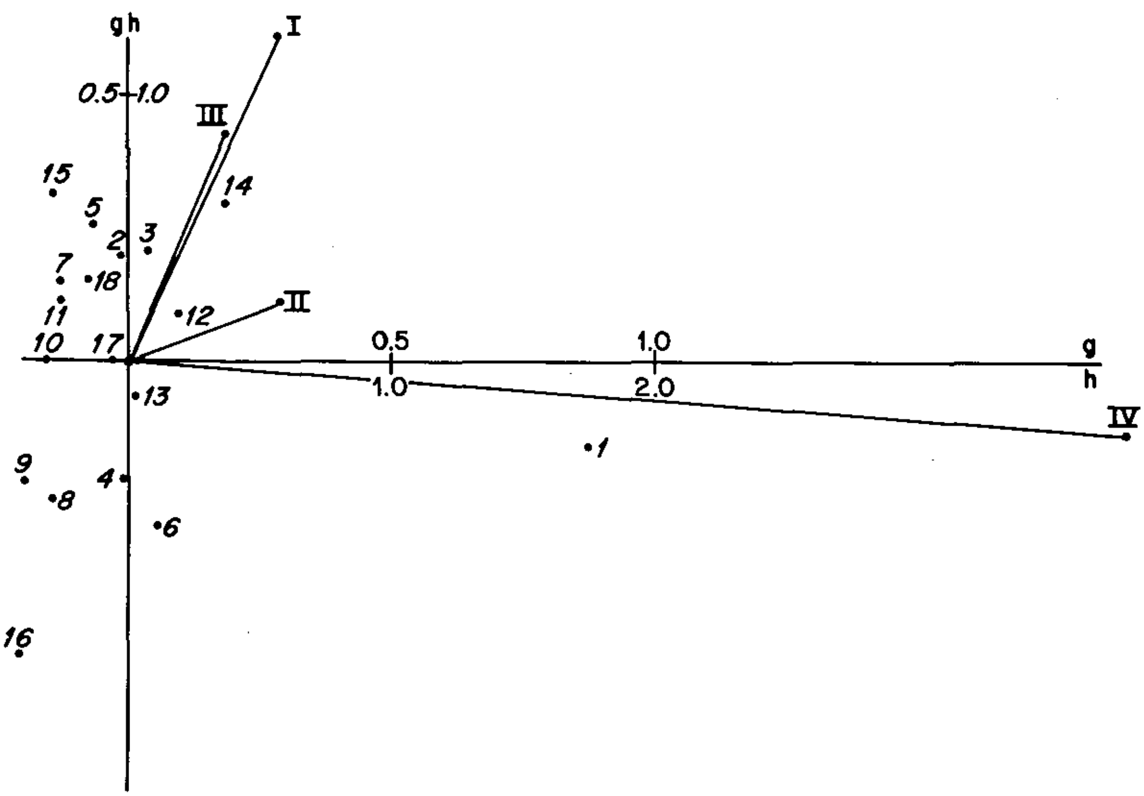


Figure 8. Biplot of total target seed and control deviations for 1970-1971-1972.

to see if the seed and control deviations form distinct clusters. In this example, no indication is shown that the control deviation (1 - 11) and the seeded deviations (12 - 18) form distinct clusters. A more detailed description of the properties of biplots is given by Gabriel (1972).

8. THE USE OF BAYES EQUATION INCORPORATING PREDICTORS

In earlier publications (Simpson and Pézier, 1971; Simpson, 1972; Simpson, Eden, Olsen and Pézier, 1973; Simpson, Woodley, Olsen and Eden, 1973) we have used Bayesian statistics in an effort to obtain a probability distribution for the seeding factor F in our dynamic cumulus seeding experiments. F is defined as the multiplicative factor by which the seeding increases the rainfall in the target in question. For example, F was found to be about 3 for the single clouds. If F were 0.5, on the other hand, this would mean that the seeded rainfall averaged half of the control rainfall. In the most meaningful portion of the earlier work, we formulated Bayes equation as follows:

$$p(F|D) = p(F)\,\frac{p(D|F)}{p(D)} \tag{1}$$

where $p(F|D)$ is the probability density distribution of F after considering the data; $p(F)$ is an assumed prior probability distribution of F; $p(D|F)$ is the probability of the data given F and $p(D)$ is the probability of the data, a normalizing factor only.

Up to now our Bayesian work has been based on the assumption that the rainfall from both seeded and unseeded targets obeys a gamma distribution with the shape parameter of the distribution known and invariant under seeding. This assumption was well verified with the single cloud data (Simpson, 1972; Simpson, Eden, Olsen and Pézier, 1973), but the sample of area cases up through 1972 was not large enough to be confident of the application of the assumption to floating and total target rainfalls, particularly with regard to the magnitude and invariance of the shape parameter.

Figure 9 illustrates a probability density for the total target seeding factor, based on the existing control and seeded populations to determine the scale parameters for the respective gamma distributions. Note that the expected value of F is 1.9, with significant probability density throughout the range 0.5 to about 4. Simpson, Woodley, Olsen and Eden (1973) showed that this curve, while insensitive to widely differing choices of prior probability on F, is sensitive to extreme values in the small population samples, via the scale parameters of the distributions. The gamma population assumption enabled an estimate of the number of cases required to resolve seeding factors, as a function of F itself. With F in the range 1.5 to 2 (large for most seeding situations) it was found (Simpson, et. al., 1973, loc. cit.) that 50-100 pairs of cases might be required for a good resolution of F without some innovation in experiment design or evaluation procedure.

The Bayesian approach proposed here differs from the past analyses by taking advantage of predictor regression

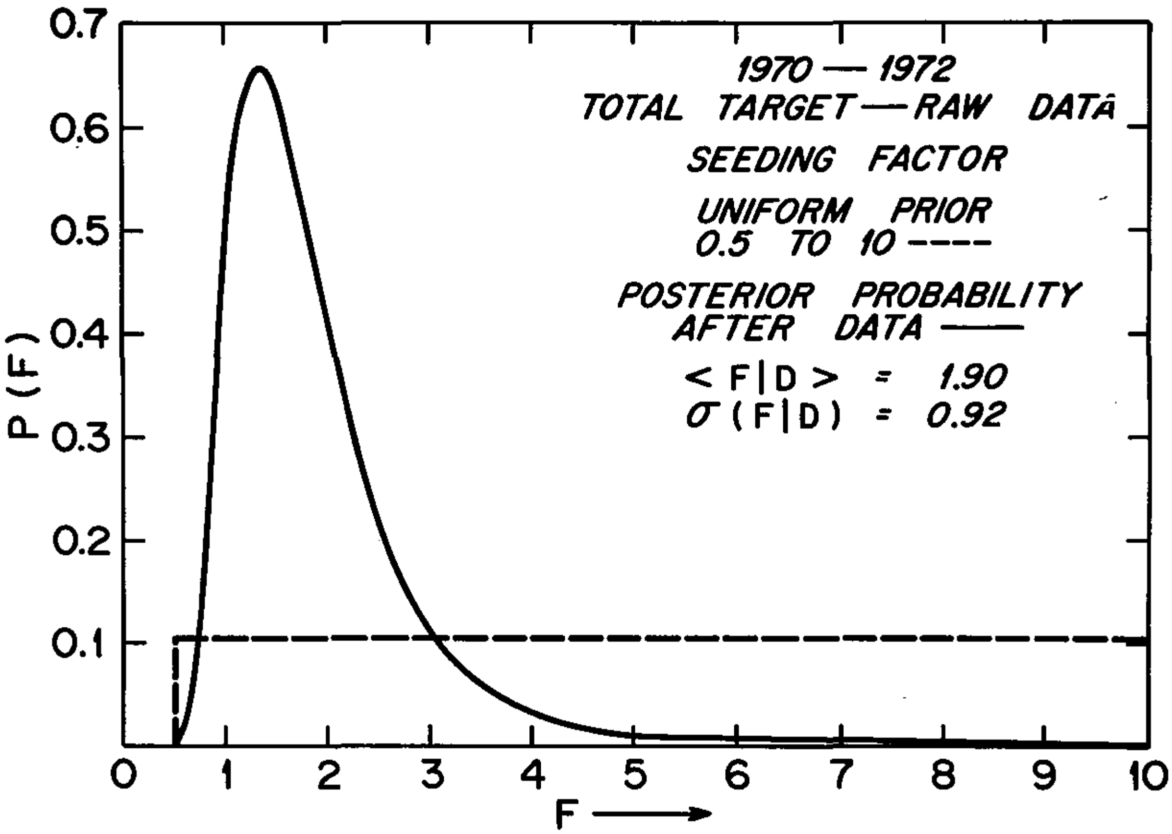


Figure 9. Plot of posterior probability density of seeding factor F with a uniform prior probability distribution in the range 0.5 to 10.

relationships. This is an initial attempt to overcome partially the natural variability problem.

The basis for the approach is the existence of a linear regression of predictors that will predict the unseeded, or naturally occurring, rainfall. It is assumed here that the naturally occurring rainfall, if it could be measured, is distributed about the regression-predicted value in a Gaussian distribution with a specified variance. The regression coefficients and the specified variance are obtained from the existing control data set with the variance obtained from the mean square error of the deviations about the regression. The Gaussian assumption requires further explanation. The assumption is based upon past experience with the usefulness of the fourth root transform in obtaining approximate normality for rainfall and the expectation that the deviations from the predicted unmodified rainfall will be mainly due to synoptic trends, variations in surface heating owing to anvil overhangs and other factors, cloud microphysical variations and measurement errors. It seems reasonable that the net sum of these effects should be a random variable that approaches a Gaussian distribution.

We again assume a multiplicative seeding factor F. Here we work with transformed (fourth root) rainfall data, so we define a transformed seeding factor F' to be evaluated by Bayes equation, namely

$$(2) \qquad p(F'|D) = p(F') \frac{p(D|F')}{p(D)}$$

where $F' \approx F^{1/4}$ (see Simpson, Eden, Olsen and Pézier, 1973, for careful analysis of this transform).

We calculate $p(D|F')$ as follows, by first calculating from the regression the anticipated unseeded rainfall. Then if the transformed seeding factor is F', the measured seeded rainfall should be distributed in a Gaussian distribution about the value F' times the unseeded predicted value, with the same standard deviation as that found for the distribution about the regression for the control cases. This assumption is logical in view of the fact that seeding factors (for raw or untransformed data) are at most 2 - 3, while natural variations range over two orders of magnitude.

With these assumptions the probability of seeing each seeded datum, given the range of F' considered, is readily calculated. Then the numerator of the right side of (2) is obtained by multiplication and the result is normalized. The procedure is further clarified by the Basic-language program entitled PREDBAYES which is described in detail by J. C. Eden (1974).

In this approach we are not making any assumptions about the seeded and control rainfall distributions, nor need they even obey the same distribution. This appears to be an advantage over our earlier Bayesian analysis, provided of course that the alternative assumptions made here prove justifiable.

As a first example, we take a uniform prior on F', for F in the range roughly 0.5 to 10 as before. Results for the posterior probabilities of F' and F are shown in figures 10 and 11.

Comparing figure 11 with figure 9, we find a promising result and a dilemma. The promising result is that the standard deviation of the probability density of F is reduced by more than a factor of three by the predictor approach, a highly desirable gain for weather modification evaluation. The dilemma, however, is that the expected value of F is reduced from 1.9 to 1.08, which if valid, might make the difference between a practical, useful modification effort and one which is not!

Therefore it is now particularly important to conduct sensitivity tests. We first investigate the effect of different prior probability assignments.

Figures 12 and 13 show the posterior probability distribution for F when the prior probability distributions are $p(F) \propto 1/F$ and the modified uniform prior

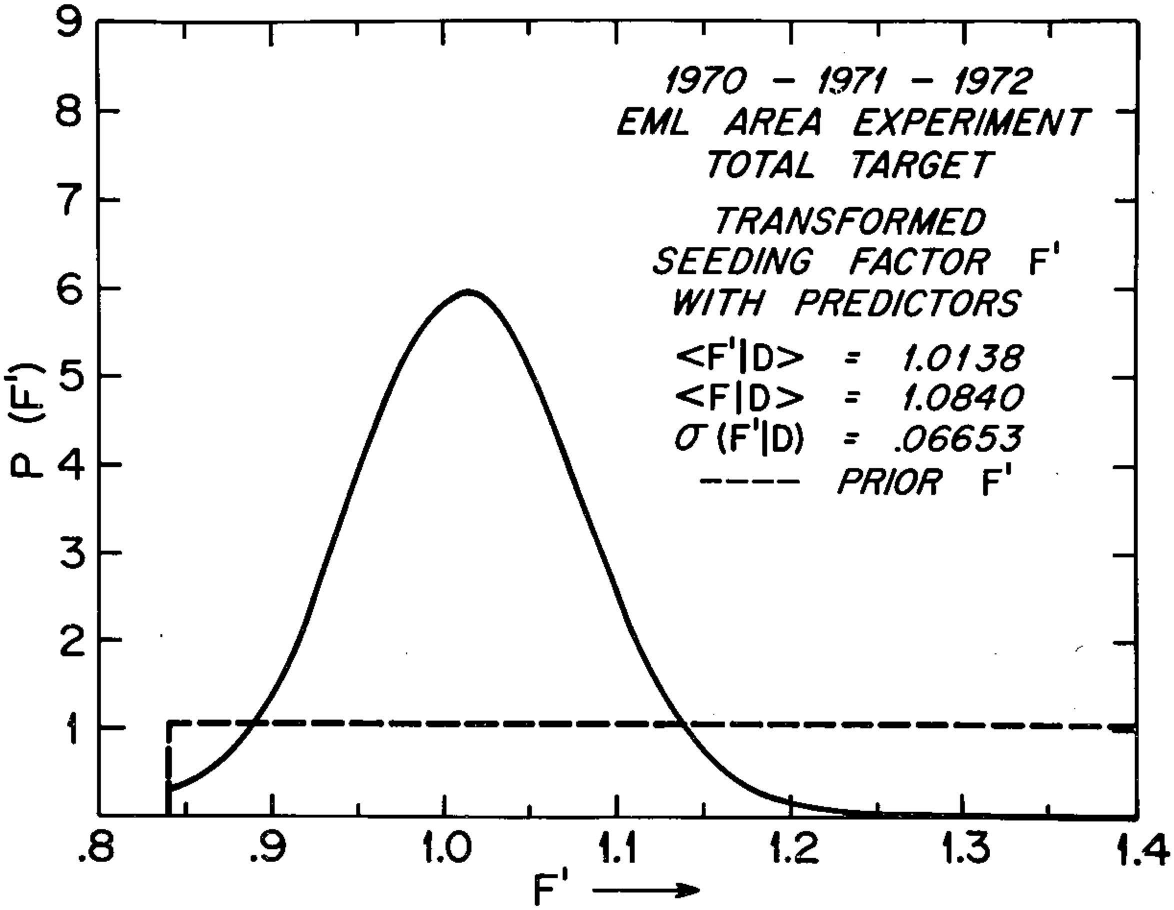


Figure 10. Posterior distribution of the transform seeding factor F' using predictors and uniform prior on F'.

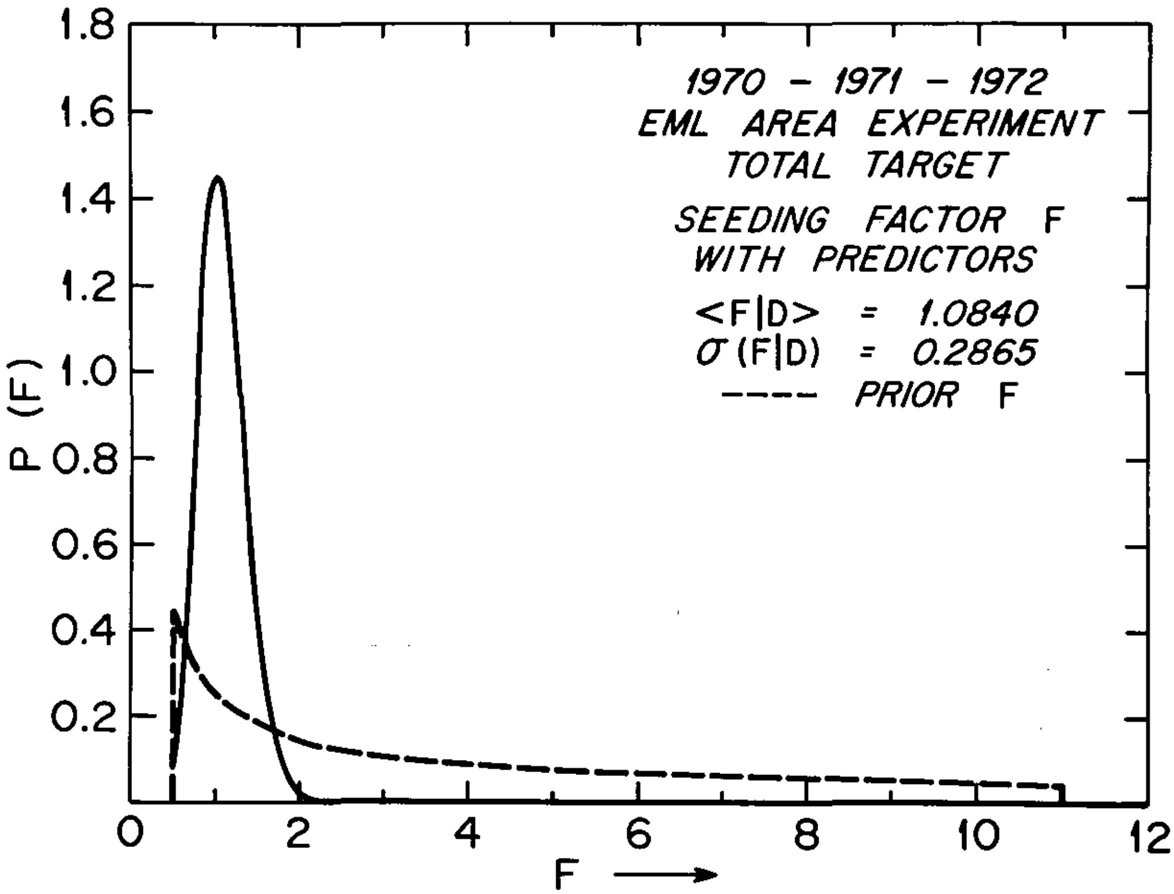


Figure 11. Posterior distribution of seeding factor F using predictors and prior on F corresponding to prior for F' in figure 10.

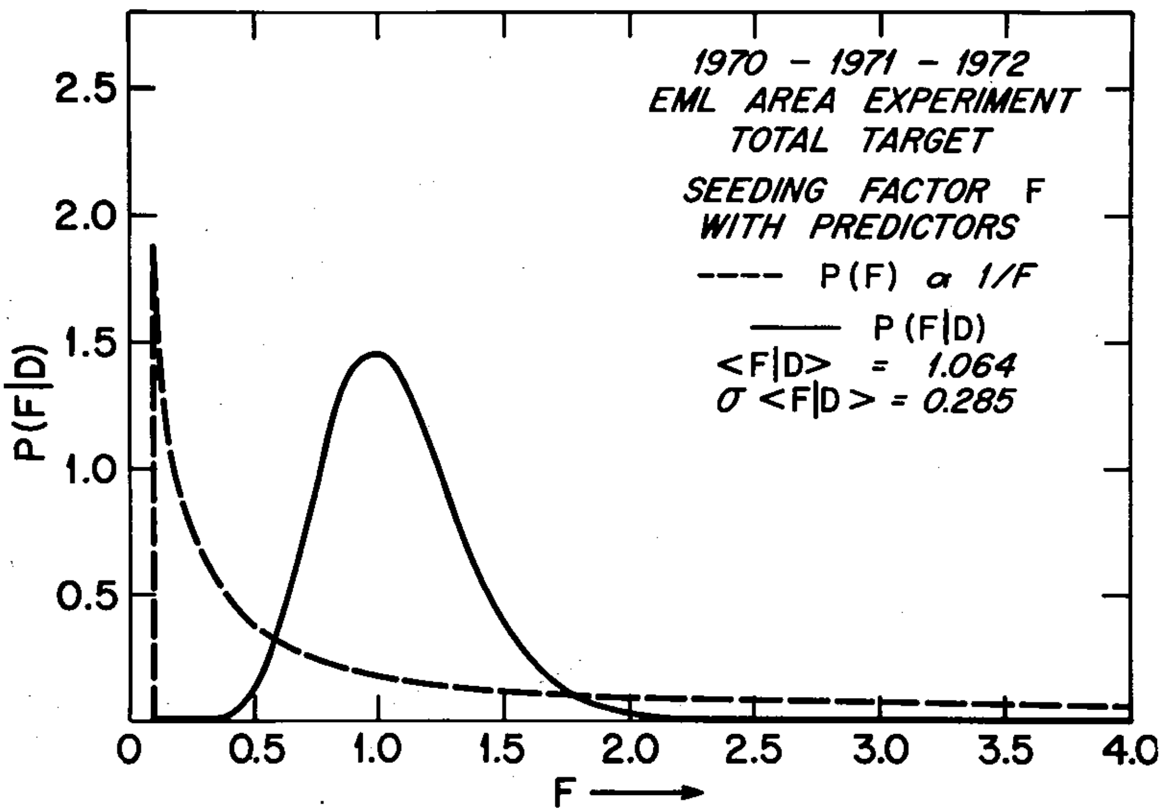


Figure 12. Posterior distribution for seeding factor F using predictors and prior proportional to 1/F.

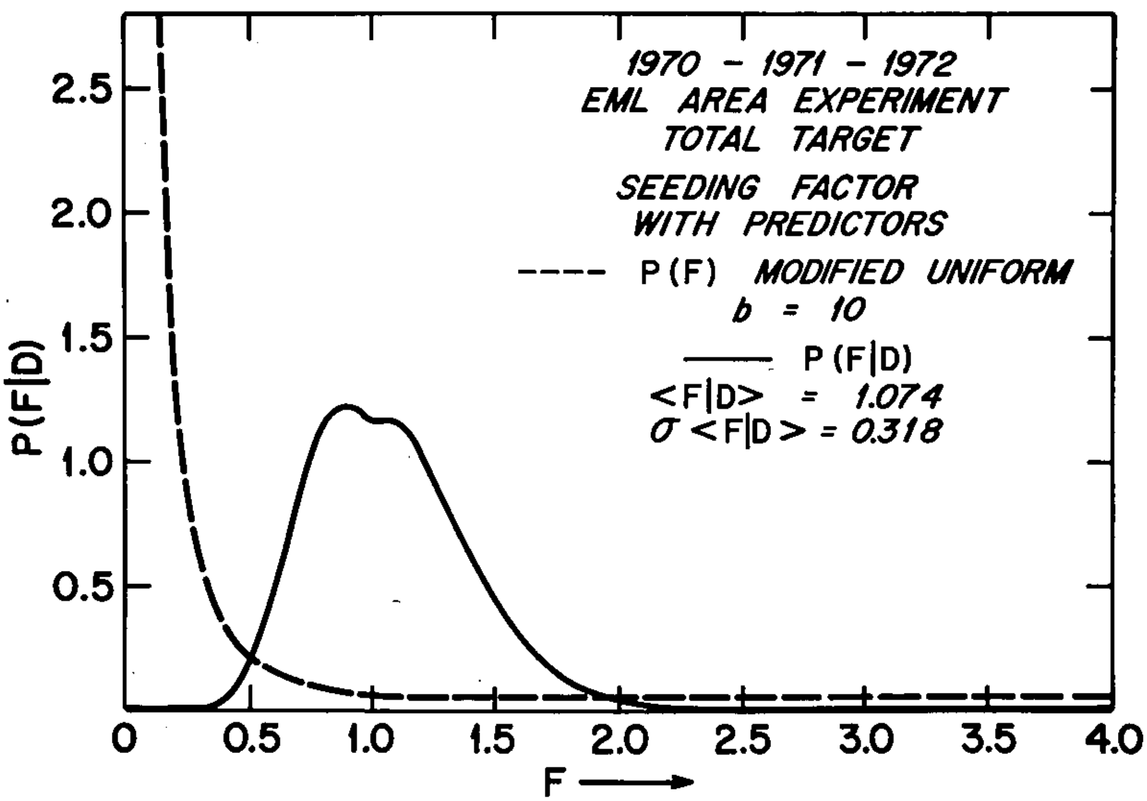


Figure 13. Posterior distribution for seeding factor F using predictors and modified uniform prior on F.

$$p(F) = \begin{cases} \frac{1}{18} F^{-2} & .1 \leq F \leq 1 \\ \frac{1}{18} & 1 \leq F < 10 \end{cases}$$

Both of the priors are attempts to present a non-informative or diffuse prior. In both cases the posterior expected value is slightly decreased and the posterior standard deviation remains approximately the same when compared to the results of figure 11. However, the change in the functional form of the modified uniform prior at F = 1, in conjunction with the probability of the data given F being concentrated at 1, results in a slight bimodal feature in the posterior distribution. An awareness of this possibility should be included in an assessment of the validity of the modified uniform to describe one's prior beliefs.

As a further indication of the sensitivity of the posterior distribution to the choice of priors, figures 14 to 20 show the shifts in the posterior that occur with the same inverse gamma priors used by Simpson, Eden, Olsen and Pézier (1973).

It is clear from these figures that a great deal of care must be taken in the selection of the prior for the encoding of the prior information. This sensitivity will, however, diminish as the number of sample cases increases and dominates the contribution of the prior.

A basic assumption in the present formulation is that

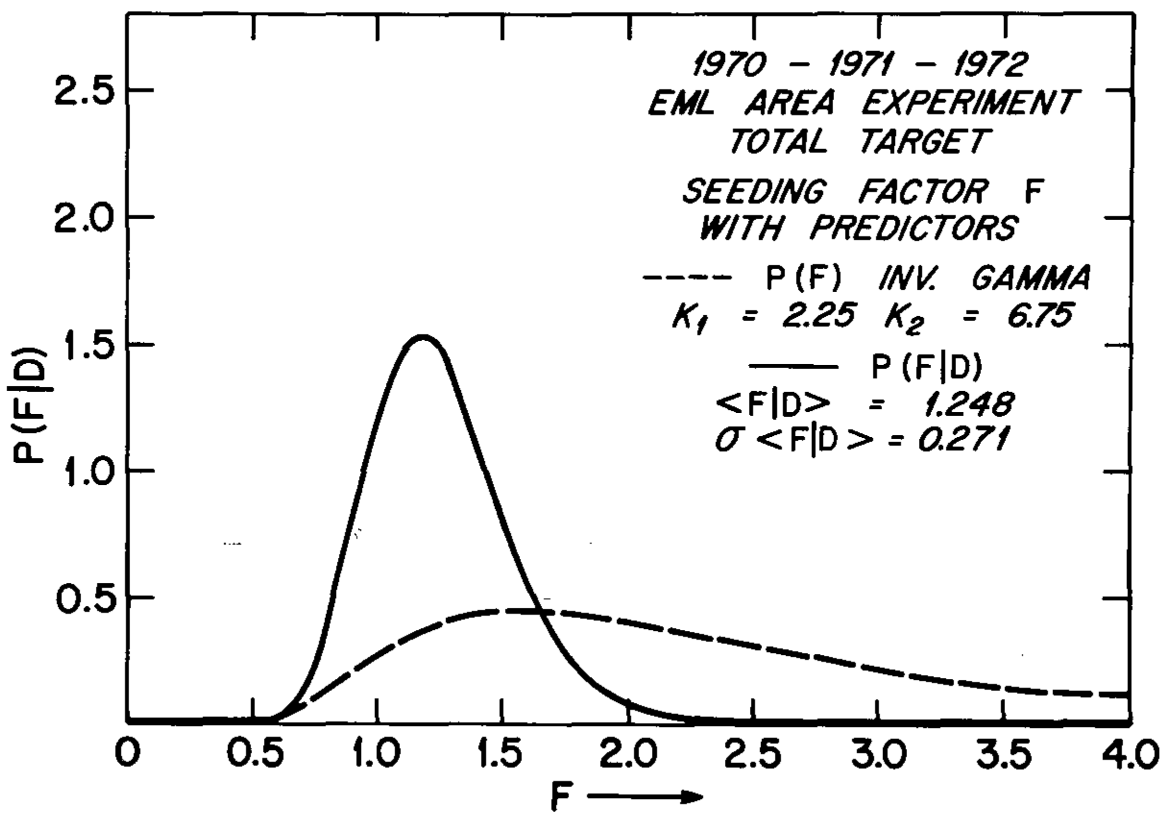


Figure 14. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 2.25$ and $K_2 = 6.75$

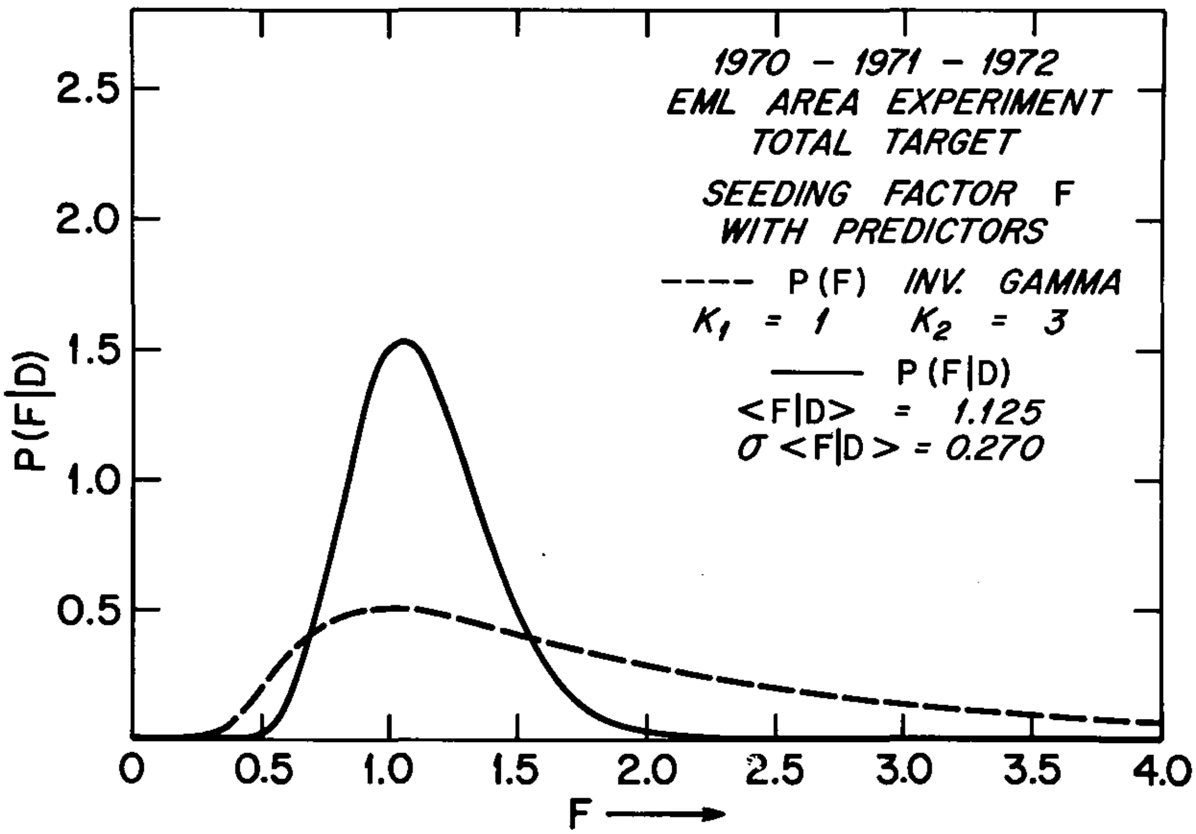


Figure 15. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 1$ and $K_2 = 3$.

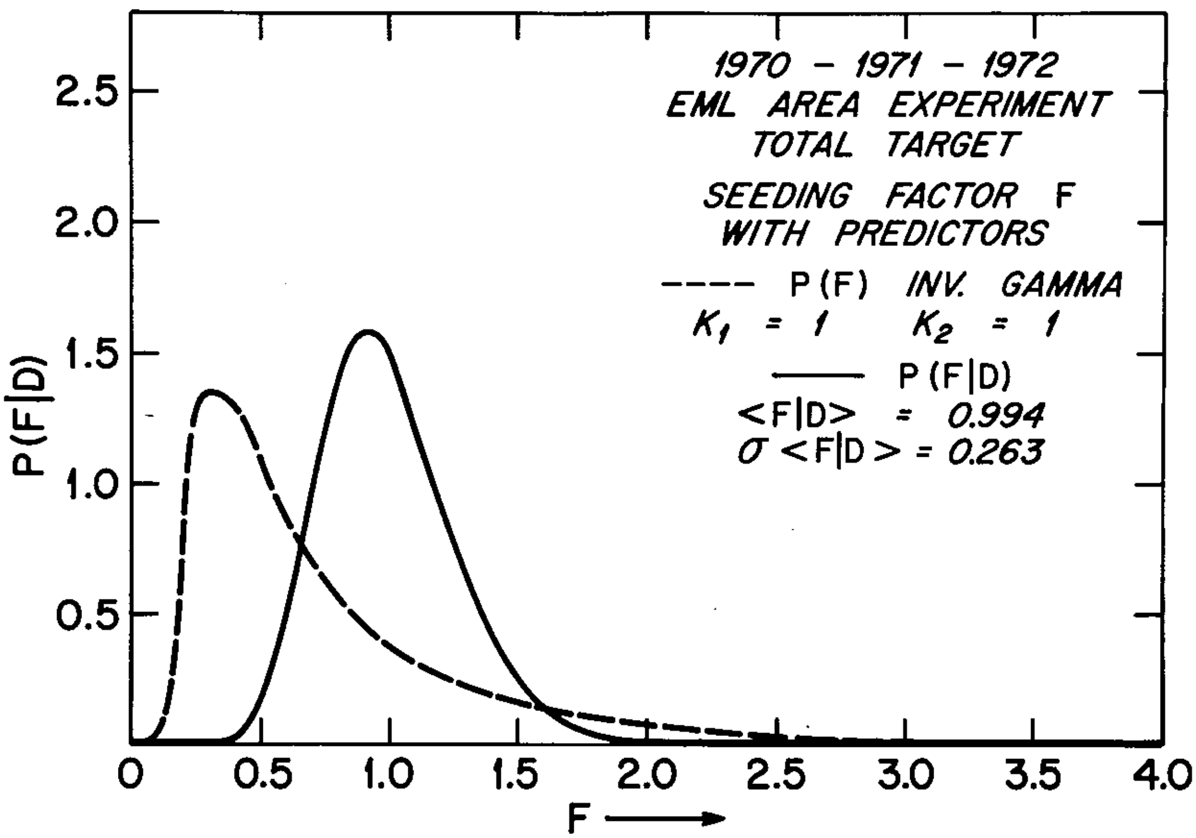


Figure 16. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 1$ and $K_2 = 1$

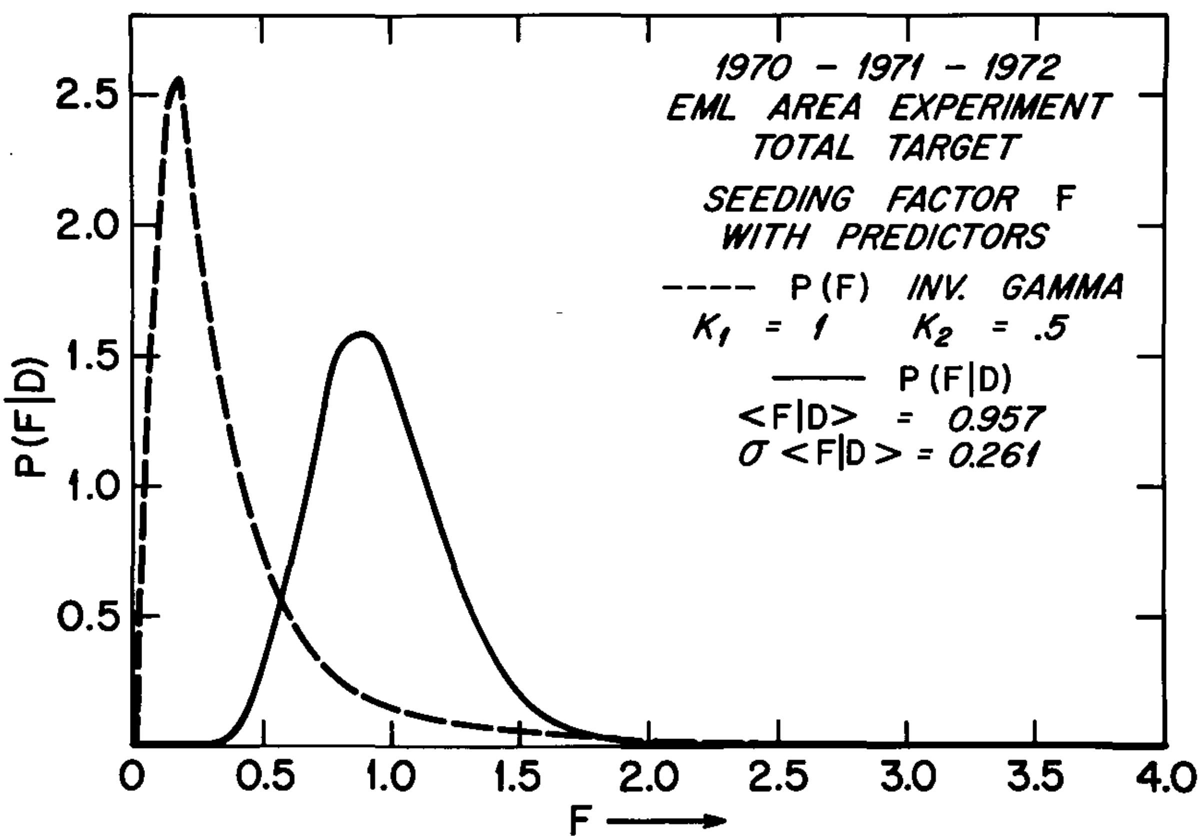


Figure 17. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 1$ and $K_2 = 0.5$.

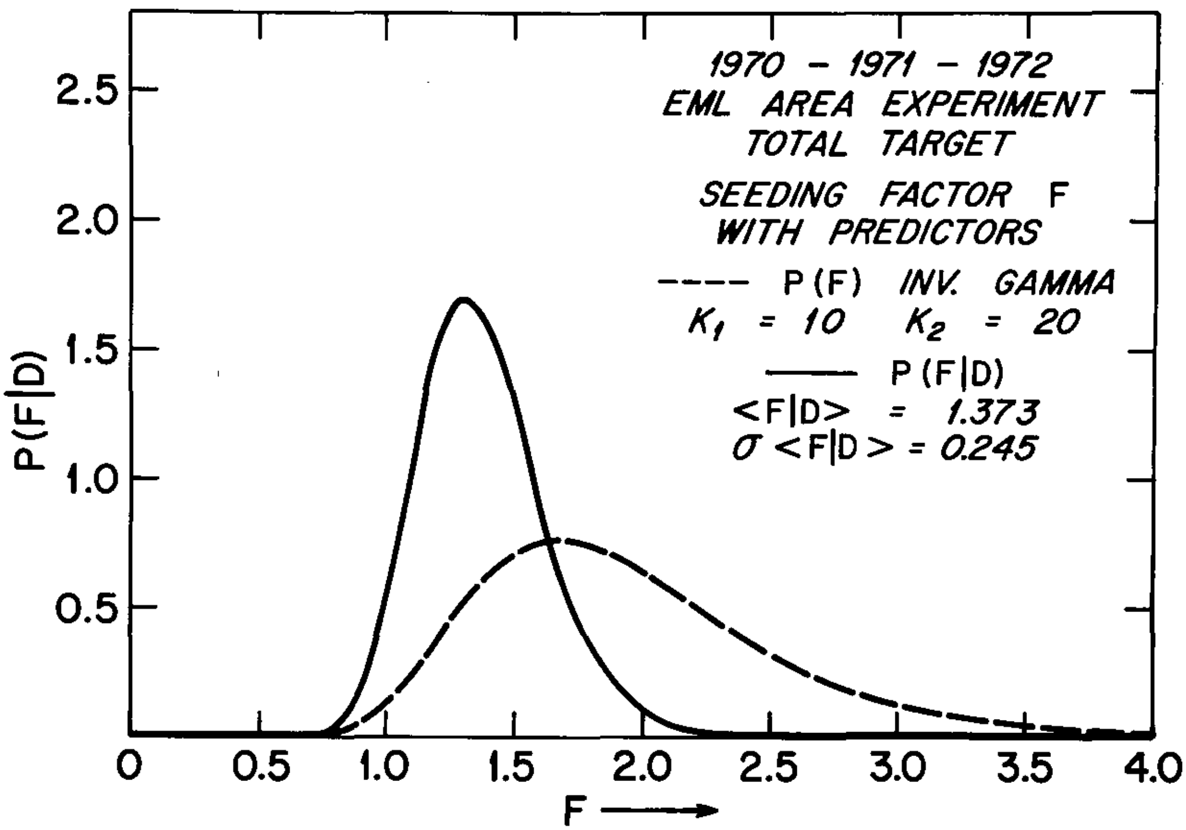


Figure 18. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 10$ and $K_2 = 20$.

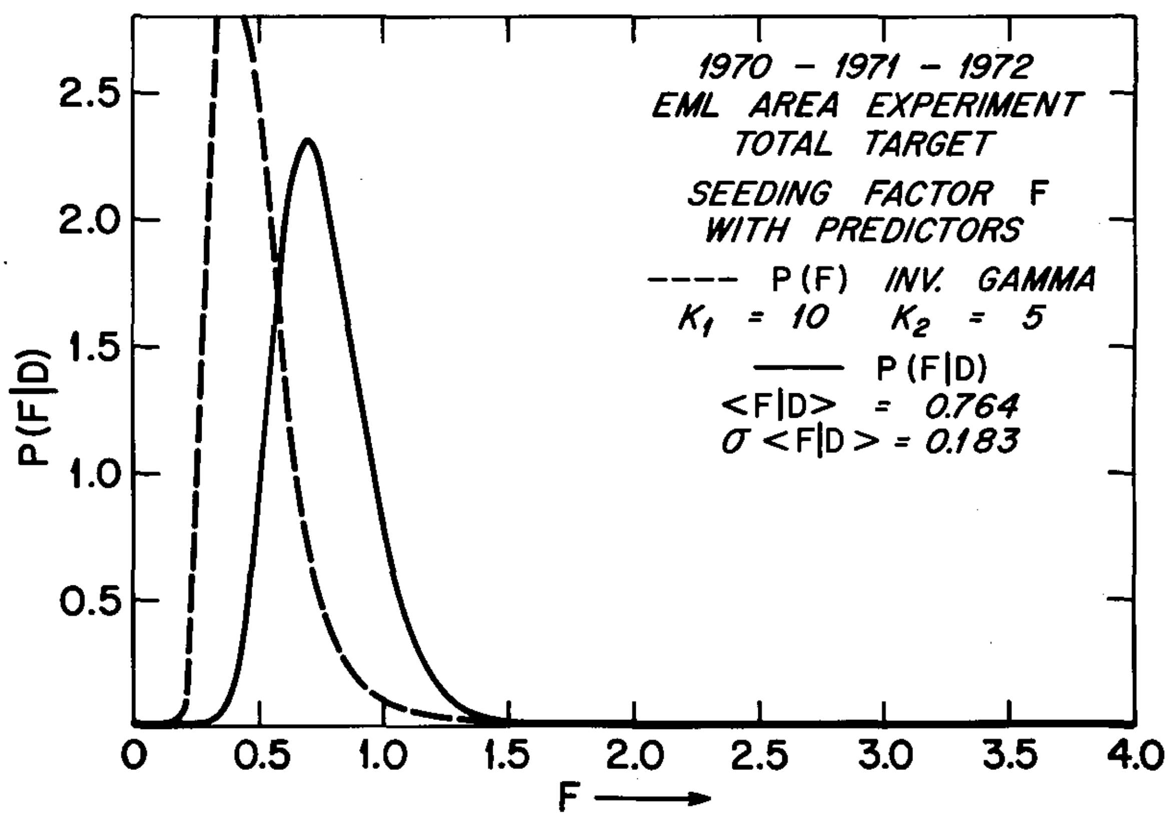


Figure 19. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 10$ and $K_2 = 5$.

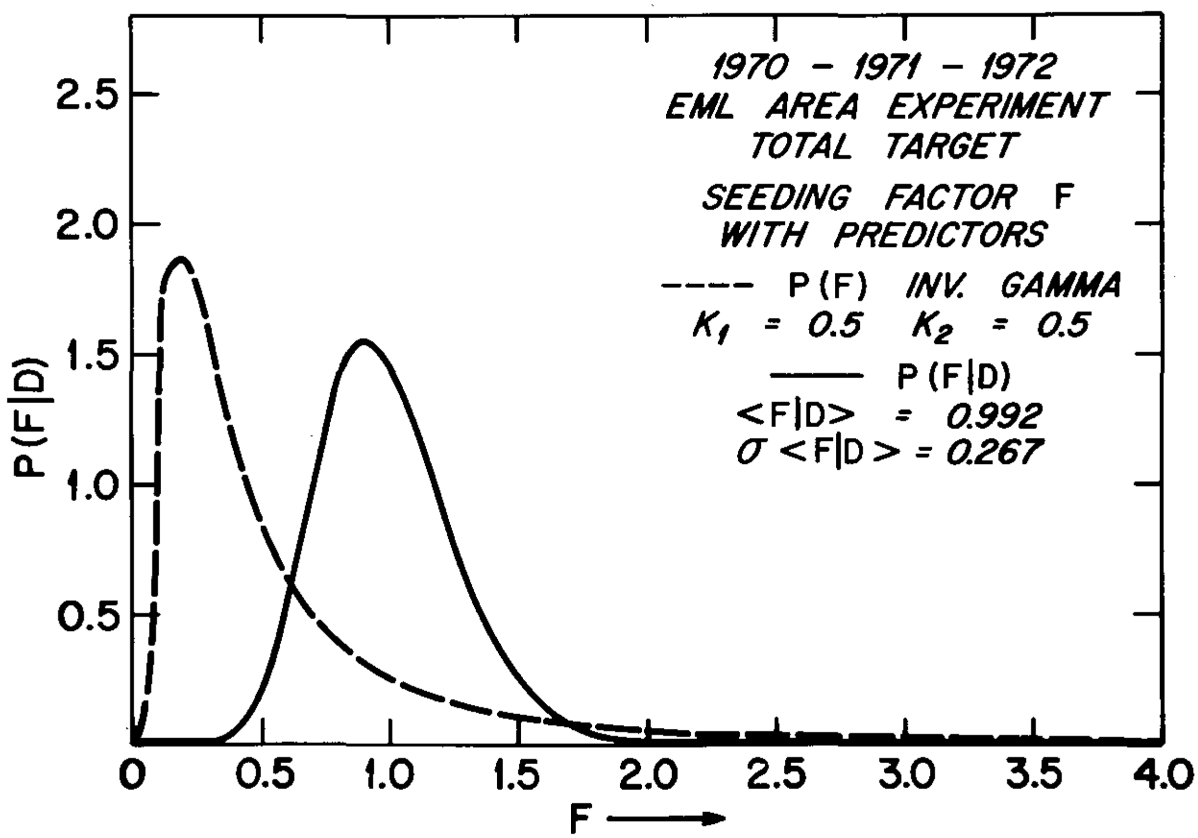


Figure 20. Posterior distribution for seeding factor F using predictors and inverse gamma prior with $K_1 = 0.5$ and $K_2 = 0.5$.

the standard deviation entering the calculations for the normal distribution is constant for all levels of the seeding factor and is given by the deviations about the regression for the control rainfall. If the transformed seeded rainfall for a given seeding factor F is predicted as $F^{1/4}\ \hat{R}_p^{\ 1/4}$, then an alternative model would give the variation as

$$\mathrm{Var}(F^{1/4}\ \hat{R}_p^{\ 1/4}) = F^{1/2}\mathrm{Var}(\hat{R}_p^{\ 1/4})$$

with $\mathrm{Var}(\hat{R}_p^{\ 1/4})$ again obtained from the regression mean square error. Figure 21 compares the posterior distributions resulting from the two different assumptions, but using the same prior. The posterior is shifted to the right with a slight increase in variability as a result of the change in the assumption concerning the variance. A detailed assessment of this model assumption is clearly indicated.

From the preceding analyses of the sensitivity of the posterior to various priors and modifications in the model assumptions further clarification and verification of the model is indicated as being necessary. While carrying out this program, consideration should also be given to a possible modification in the parameterization of the multiplicative seeding factor F. One of the unpleasant features of F is that a positive or negative seeding effect is reflected through F being greater than or less than one. By defining a new parameter θ, whose relationship to F

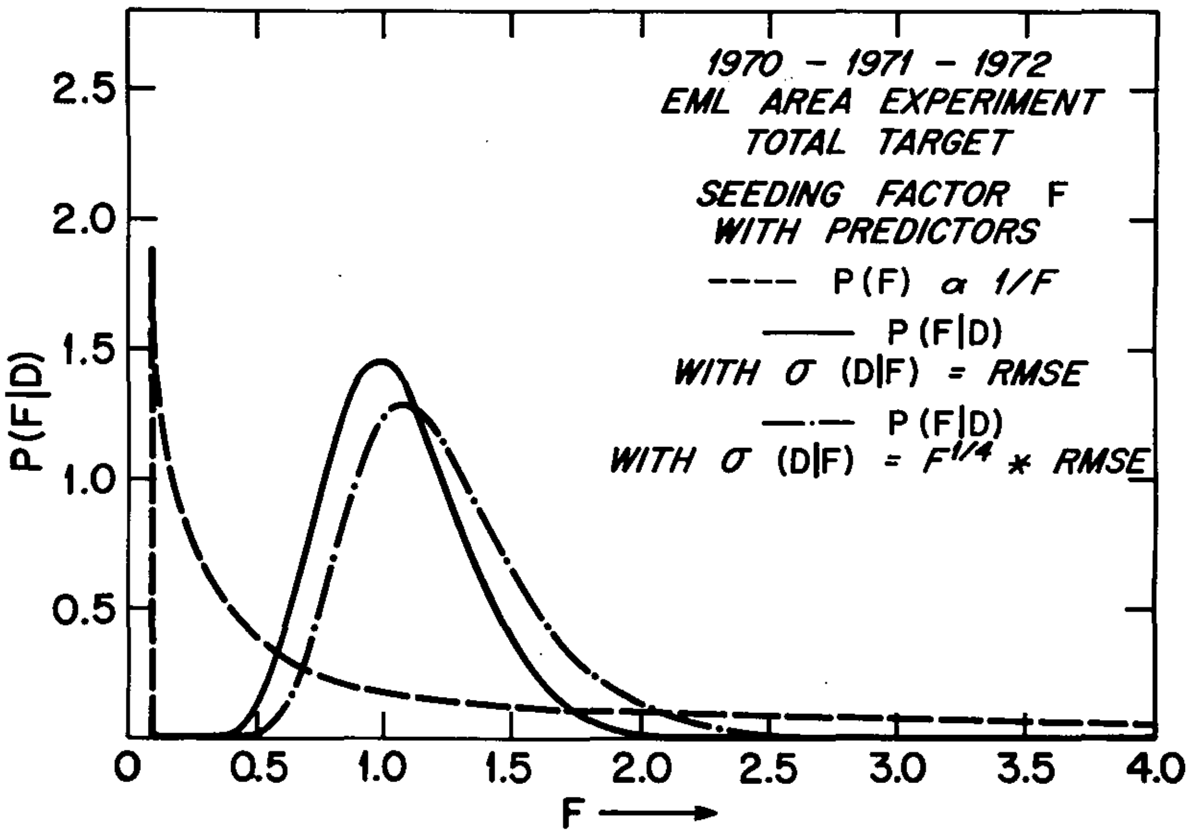


Figure 21. Comparison of posterior distributions for seeding factor F when the variance of the data distribution given F is assumed to be constant or assumed to be a function of F.

is $F = e^{\theta}$, a positive or negative effect is reflected in θ being positive or negative with no effect indicated when θ equals zero.

The use of θ in a Bayesian analysis is easily accomplished by replacing F by e^{θ} and then placing a prior on θ instead of F. The only problem that may arise is in the ability of the experimenter to encode his prior beliefs in a probability distribution for θ. As an example of the posterior distributions arising from this new parameterization, figures 22 and 23 present the results from two different choices of priors. The prior on θ in figure 22 is equivalent to a uniform prior on F from .1 to 10. The uniform prior on θ in figure 23 is equivalent to the prior on F being proportional to $1/F$. Although the figures 12 to 23 present the same information under different parameterizations, one presentation may be more meaningful than the other for different experimenters.

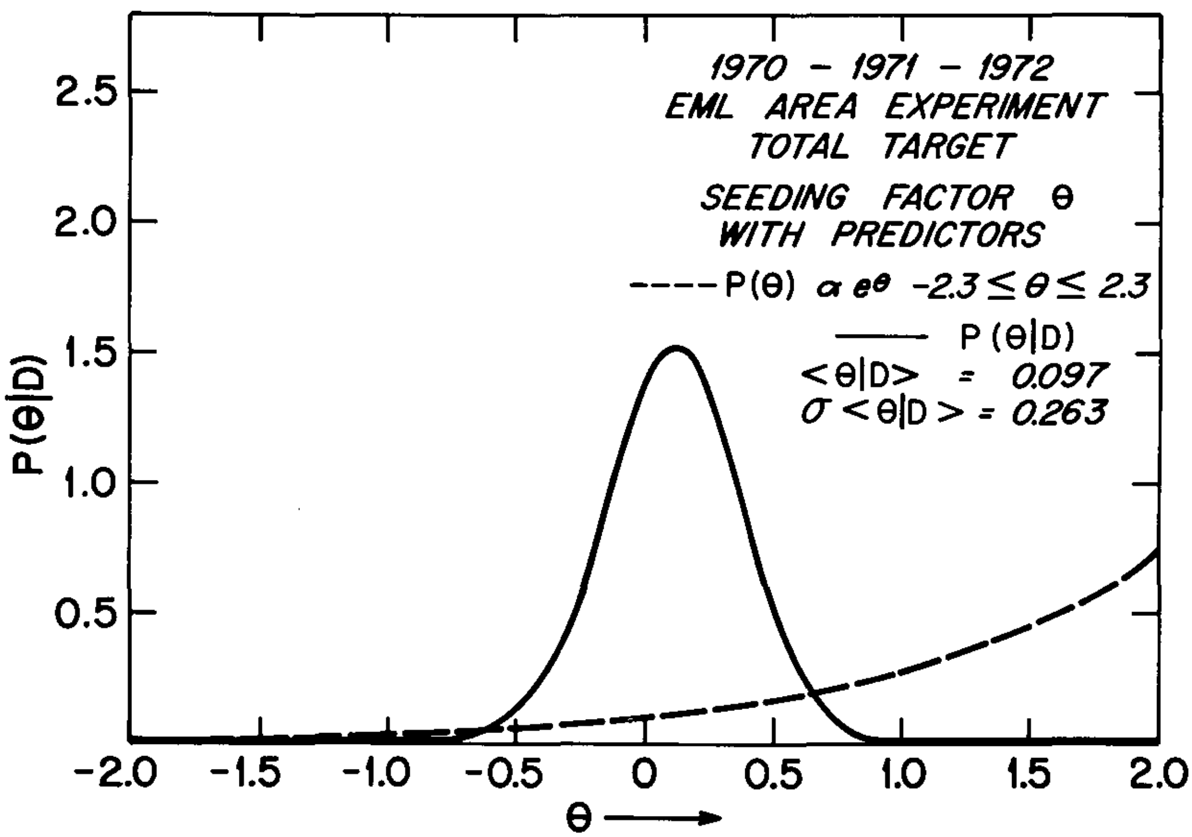


Figure 22. Posterior distribution for seeding factor θ under same assumptions used in analyses for F. Priors on θ is equivalent to a uniform prior on F from .1 to 10.

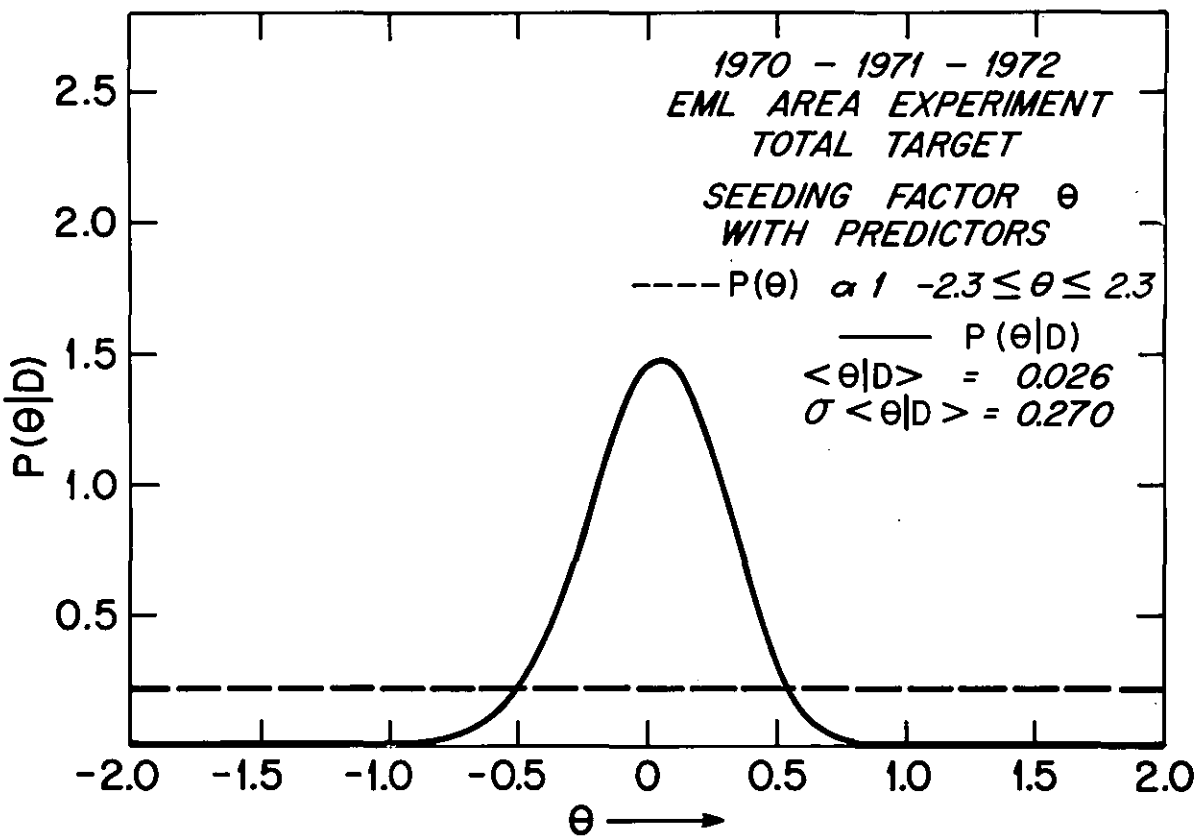


Figure 23. Posterior distribution for seeding factor θ under same assumptions used in analyses for F. Uniform prior on θ is equivalent to prior 1/F for F.

9. SUMMARY

In Part II several different types of analyses were presented which incorporated the use of predictors directly into the analysis of area seeding experiments. After finding what appear to be useful predictors, an analysis was completed using covariate regressions and analysis of covariance, which further indicated that the predictors could be useful in a classical statistical analysis of an area experiment. This methodology is promising, but indicates that larger sample sizes will be required before a definite conclusion could be reached on whether seeding has an effect over a large area.

Previously, the area experiment has been analyzed using the philosophically different Bayesian approach, but no predictors were included in the analysis. Since the predictors were shown to reduce the natural variability of the rainfall data, in section 8, an attempt was made to modify the previous Bayesian framework to include the use of predictors. The basic procedure involved using the control data sample to establish a prediction equation for the naturally occurring rainfall and then using this prediction equation to partly specify the distribution of the seeded data sample given the seeding factor. Preliminary results indicate that the posterior distribution on F has smaller variability than the previous Bayesian analyses but that the size of the seeding factor is also reduced. Hence, although the approach is promising, no firm conclusions can yet be made. When the additional data from the 1973 experiments are available, the methodology developed here will be used to complete the analysis. Hopefully, more definitive conclusions can then be drawn.

10. REFERENCES

Eden, J. C. (1974), Guide to computer programs used in the statistical analysis of Florida cumulus seeding experiments, NOAA Tech. Memo. ERL WMPO, in process.

Gabriel, K. R. (1972), Analysis of meteorological data by means of canonical decomposition and biplots, J. Appl. Meteorol. 11: 1071-1077.

Gabriel, K. R. and M. Haber (1973), The Moore-Penrose inverse of a data matrix - a statistical tool with some meteorological applications. Preprint of Third Conference on Probability and Statistics in Atmospheric Science, June 19-22, 1973, Boulder, Col., 110-117.

Herndon, A., W. L. Woodley, A. H. Miller, A. Samet and H. Senn (1973), Comparison of gage and radar methods of convective precipitation measurement, NOAA Tech. Memo. ERL OD-18, 67 pp.

Klein, W. H. (1965), Application of synoptic climatology and short-range numerical prediction to five-day forecasting, Research Paper No. 46, June, 1965, U. S. Weather Bureau, Washington, D. C., 109 pp.

Klein, W. H., B. M. Lewis and I. Enger (1959), Objective prediction of five-day mean temperatures during winter, J. Meteorol. 16: 672-682.

Schickedanz, P. T. and F. A. Huff (1971), The design and evaluation of rainfall modification experiments, J. Appl. Meteorol. 10: 502-514.

Simpson, J. (1972), Use of the gamma distribution in single cloud rainfall analysis, Monthly Wea. Rev. 100: 309-312.

Simpson, J. and J. Pézier (1971), Outline of a Bayesian approach to the EML multiple cloud seeding experiments, NOAA Tech. Memo. ERL OD-8, 43 pp.

Simpson, J. and V. Wiggert (1969), Models of precipitating cumulus towers, Monthly Wea. Rev. 97: 471-489.

Simpson, J. and V. Wiggert (1971), 1968 Florida cumulus seeding experiment: Numerical model results, Monthly Wea. Rev. 99: 87-118.

Simpson, J. and W. L. Woodley (1971), Seeding cumulus in Florida: New 1970 results, Science 172: 117-126.

Simpson, J., W. L. Woodley, A. H. Miller and G. F. Cotton (1971), Precipitation results from two randomized pyrotechnic cumulus seeding experiments, J. Appl. Meteorol. 10: 526-544.

Simpson, J., J. C. Eden, A. R. Olsen and J. Pezier (1973), On the use of gamma functions and Bayesian analysis in evaluating Florida cumulus seeding results, NOAA Tech. Memo. ERL OD-15, 86 pp.

Simpson, J., W. L. Woodley, G. F. Cotton and J. C. Eden (1973), Statistical analysis of EML multiple cumulus experiments in 1970, 1971, and 1972, NOAA Tech. Memo. ERL OD-17, 80 pp.

Simpson, J., W. L. Woodley, A. Olsen and J. C. Eden (1973), Bayesian statistics applied to dynamic modification experiments on Florida cumulus clouds, J. Atmos. Sci. 30: 1178-1190.

Woodley, W. L., J. Simpson, A. H. Miller, S. MacKay, R. Williamson and G. F. Cotton (1970), Some results of single cloud pyrotechnic seeding in Florida, 1970, NOAA Tech. Memo. ERLTM AOML 10, 85 pp.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1974 — 784-576 / 1260 REGION NO. 8