NOTE ON PROJECT SCUD

J. M. WELLS and M. A. WELLS
UNIVERSITY OF CALIFORNIA, BERKELEY

1. General information

Information on Project SCUD stems from the article by Jerome Spar published in *Meteorological Monographs* [1].

Project SCUD originated at New York University in May 1952, as an attempt to discover quantitative effects of cloud seeding on cyclones developing in the east coastal region of the United States. A meteorological group at New York University forecast the location of the center and the zero hour of an incipient cyclone. The personnel at Naval facilities based in the coastal area were responsible for randomization, for seeding, and for collection of observations. The experiment was designed to test the hypothesis that cloud seeding in areas of cyclogenesis on the east coast of the United States has no measurable effect on the development of storms there. Precipitation data were taken because it was thought that they would reflect the effects of the seeding treatment, if any. The experiment was designed with the leading idea that seeding during an early stage of a cyclone would be more effective than during the later stages. It was hoped that the seeding of clouds would produce rain over large areas where it would not occur of its own accord and, in accordance with a suggestion of Langmuir, that the heat so generated would have a marked effect on the general circulation of the atmosphere. With the above in mind, an effort was made to seed situations in which cyclogenesis appeared imminent.

Even though the duration of the experiment was too short to detect the possible effects of seeding of an intensity that was reasonable to expect, some of the findings attained in this experiment deserve serious attention.

2. Meteorological variables

Two meteorological variables were observed: precipitation and pressure change. The present account is concerned with precipitation only.

3. Seeding

Seeding was done with silver iodide released from seventeen ground based generators and with dry ice dispensed from aircraft. The silver iodide smoke

Prepared with the partial support of the Office of Naval Research (Contract No. N00014-66-C0036-G01; NR 307-303X).

generators were installed at Coast Guard stations from Florida to New York, as indicated in figure 1. The tracks of seeding aircraft were approximately 1000 miles long and varied in location and shape depending on the predicted weather. All these tracks were within the region *Ia* (see below).

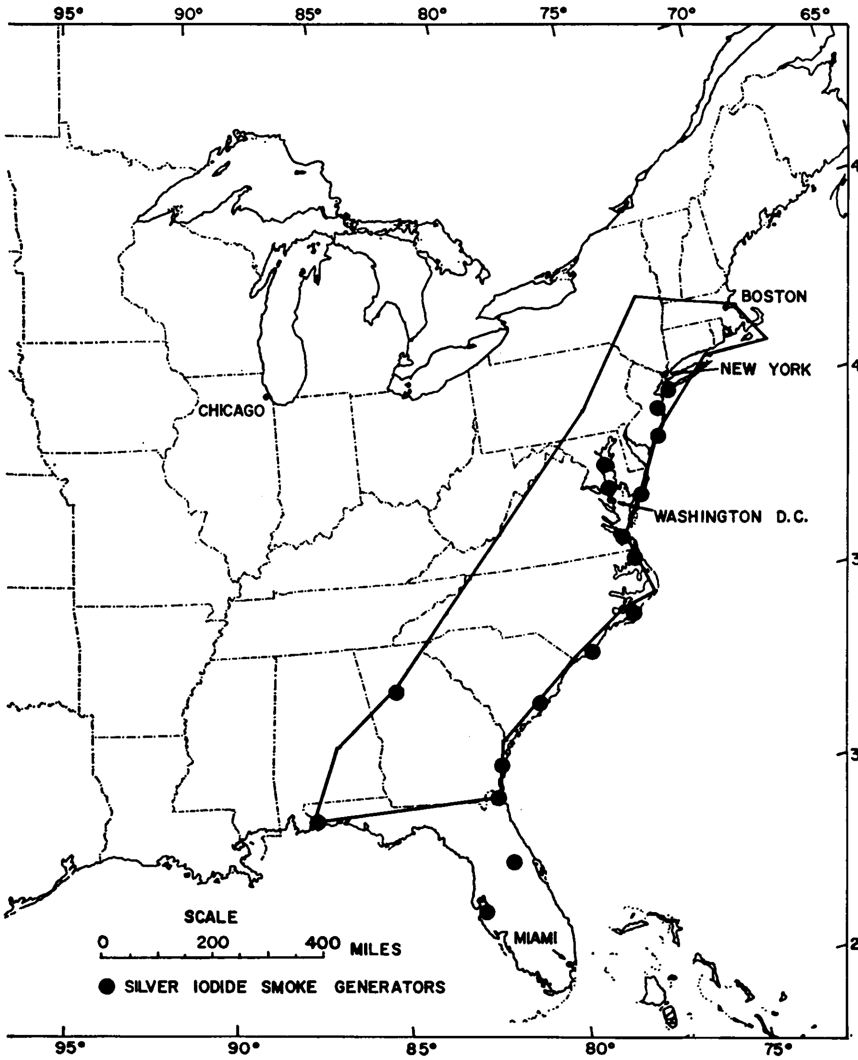

FIGURE 1

Location of silver iodide smoke generators of Project SCUD [1].

4. Targets

For purposes of rainfall collection there were two very large fixed test regions and one much smaller region adjustable to weather conditions. However, rainfall observations for the adjustable region, denoted by *Rt*, are available only for the second year of operations with 16 experimental units, much too little for an effective evaluation, and the present account is limited to the two fixed targets. The rainfall variables for these regions are denoted by *RI* and *RII*, respectively; *RI* stands for the 24 hour precipitation amount, starting with "zero hour," averaged over a network of Weather Bureau raingages in the region *Ia*. This region, shown in figure 2, extended over the eastern part of the United States from 30° to 45° latitude, and from the Atlantic coast to the Appalachians; it was roughly 1000 miles long and 200 miles wide. *RII* is also the average 24 hour precipitation, but starting with zero hour plus 12. It refers to another large region, north of the 44th parallel, including Labrador and Newfoundland, as shown in figure 2. In extending the observations so far north from the region where the seeding was done, the experimenters were interested in the possibility of downwind effects.

5. Observational data and evaluation

Table I reproduces the observational data as published by Spar, except that precipitation amounts *Rt* and also pressure data are omitted. The quantities *M*, *T*, and *L* are of particular interest as covariates to predict the precipitation.

(1) *M* is the geostrophic meridional circulation index, defined as

$$M = (h_1 + h_2 + h_3) - (h_4 + h_5 + h_6), \tag{5.1}$$

where h denotes the geopotential height of the 700 mb surface at the following radiosonde stations:

1 Nantucket, Mass.	2 Hatteras, N. C.	3 Tampa, Fla.
4 Pittsburgh, Pa.	5 Nashville, Tenn.	6 Burwood, La.

(2) *T* is the sum of net water vapor influx across the boundaries of a hexagon with the six stations above as vertices (see figure 3). *T* was measured by computing geostrophic water influx at 1000, 850, 700, 500, and 400 mb levels from radiosonde data and interpolating to zero hour.

(3) *L* is a measure of the latitude of the cyclone, determined by drawing a perpendicular from the cyclone center on the zero hour map to a line between stations 1 and 6 above. *L* is then defined as the distance in tens of miles from Burwood, La. to the point of intersection, as illustrated in figure 3.

The quantity *M* was used in predicting cyclogenesis; cyclones were expected to develop only when *M* was predicted positive. In addition, the three quantities *M*, *T*, and *L* were used as predictor variables in the regression analysis of precipitation data.

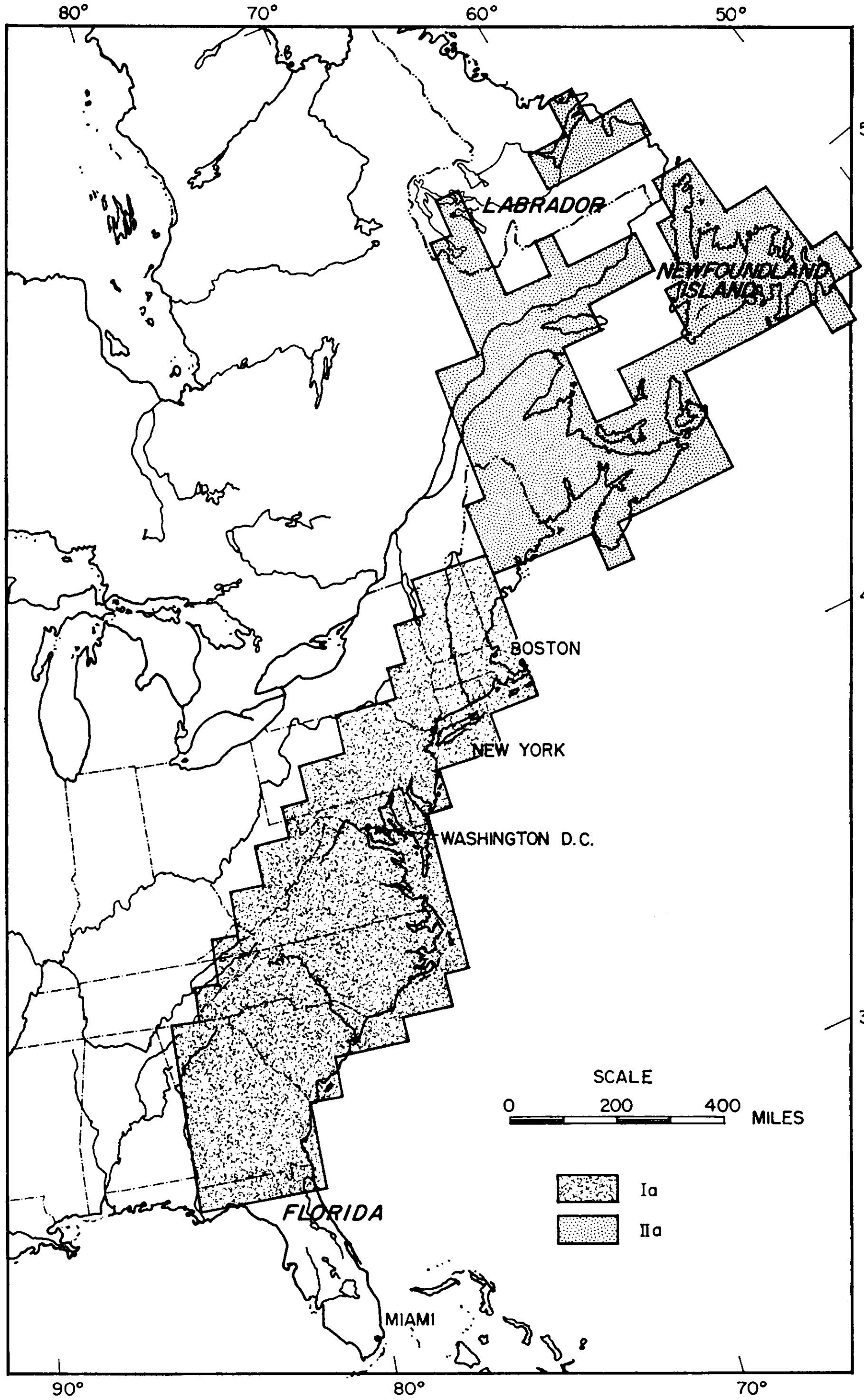

FIGURE 2

Target regions *Ia* and *IIa*
(adapted from [1]).

TABLE I

OBSERVATIONAL DATA PROJECT SCUD

No.	Zero Hour/Date	Treatment	Dry Ice Dispensed (lb)	*RI*	*RII*	*M*	*T*	*L*
			Experiment ONE					
1	1230/ 9 Jan. 1953	S	1000	.498	.011	87	961	31
2	0630/18 Jan. 53	C		.339	.016	63	269	81
3	1230/21 Jan. 53	C		.188	.063	25	−699	65
4	0030/24 Jan. 53	S	2500	.603	.721	115	1203	56
5	0630/ 1 Feb. 53	S	650	.021	.108	37	−259	135
6	0630/ 3 Feb. 53	C		.081	.086	10	−279	53
7	1830/ 6 Feb. 53	C		.519	.102	75	654	65
8	0630/12 Feb. 53	S	1370	.175	.006	28	402	105
9	0630/15 Feb. 53	C		.738	.499	90	329	38
10	1830/20 Feb. 53	S	2480	.417	.200	71	1572	77
11	1230/25 Feb. 53	C		.138	.020	−7	−97	23
12	1830/ 3 Mar. 53	S	405	.441	.439	52	1156	33
13	1230/13 Mar. 53	S	460	.180	.574	24	−122	107
14	1230/15 Mar. 53	C		.435	.137	51	312	76
15	1830/18 Mar. 53	S	1680	.178	.013	30	798	77
16	1830/23 Mar. 53	C		.423	.062	53	670	48
17	1830/ 1 Apr. 53	C		.072	.167	17	−40	88
18	1830/ 6 Apr. 53	S	1125	.715	.160	47	1173	36
19	1830/10 Apr. 53	C		.075	.113	44	−620	89
20	0630/16 Apr. 53	S	1935	.205	.066	57	−383	97
21	1230/18 Apr. 53	S	2975	.260	.031	43	709	30
Arithmetic Means:	Seeded			.336	.212	53.7	655	71
	Control			.301	.127	42.1	50	63
			Experiment TWO					
22	1830/ 4 Dec. 53	S	7425	.362	.426	99	140	65
23	1830/ 9 Dec. 53	C		.374	.757	83	320	28
24	1230/12 Dec. 53	S	7150	.454	.133	58	411	28
25	1830/10 Jan. 54	C		.364	.101	51	880	30
26	1230/15 Jan. 54	C		.598	.059	19	938	136
27	1230/21 Jan. 54	S	1350	.316	.033	32	−450	136
28	1830/27 Jan. 54	C		.089	.271	27	−1187	105
29	1830/11 Feb. 54	S	6750	.020	.198	−2	−845	87
30	1830/20 Feb. 54	S	3350	.497	.040	91	1189	36
31	1830/24 Feb. 54	C		.192	.191	44	−42	34
32	1230/26 Feb. 54	S	5825	.057	.197	48	−91	98
33	1230/ 1 Mar. 54	C		.232	.096	113	−839	90
34	1830/13 Mar. 54	S	6110	.619	.130	55	365	69
35	1830/19 Mar. 54	C		.563	.362	52	997	49
36	1230/30 Mar. 54	C		.063	.028	15	−480	104
37	0630/28 Apr. 54	S	6200	.117	.007	1	338	106
Arithmetic Means:	Seeded			.305	.146	47.7	132	78
	Control			.309	.233	50.5	73	72
Arithmetic Means:	ONE and TWO Seeded			.322	.184	51.3	435	74
	ONE and TWO Control			.305	.174	45.8	60	67

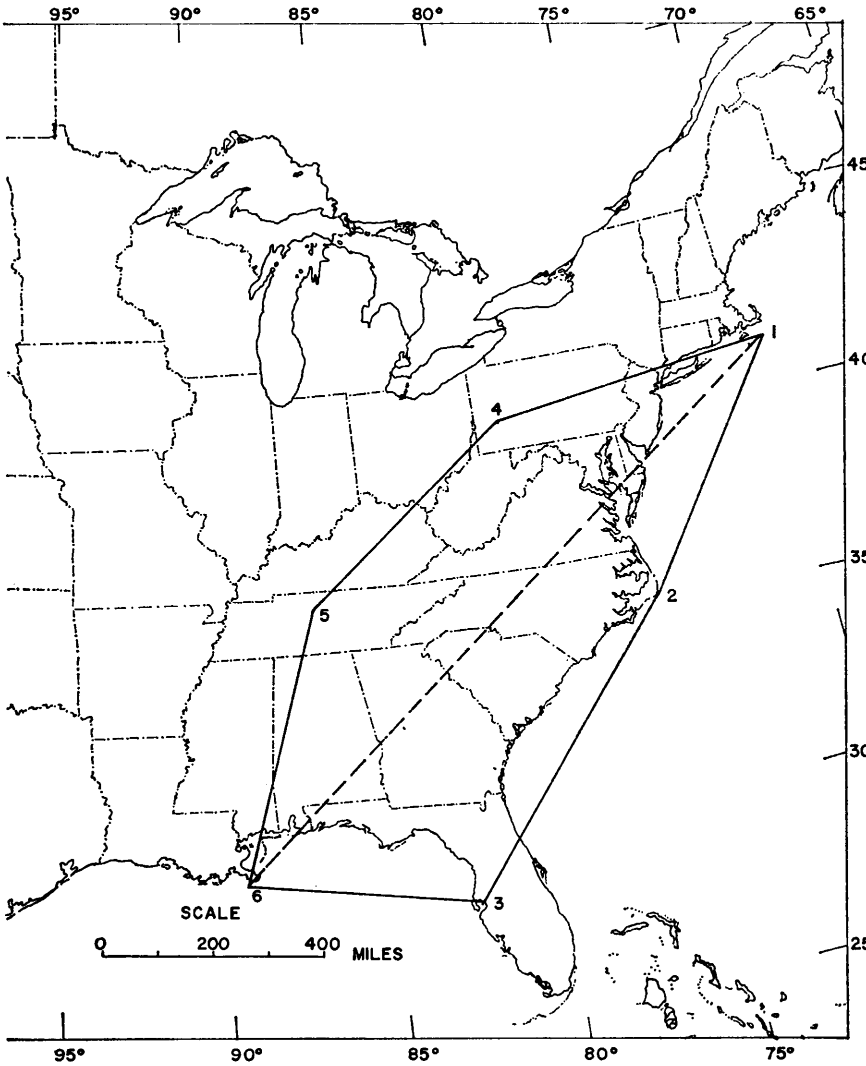

FIGURE 3

Six radiosonde stations used in determining predictors M, T, and L.

The original evaluation used not the precipitation data RI and RII as given in table I, but their logarithms, because a multiplicative effect of seeding was expected. The results of the regression analysis obtained by Spar are reproduced in table II.

It is seen that the point estimates used by Spar of the average precipitation amounts for seeded and for not seeded observational units are "adjusted"

TABLE II

MEANS AND 95 PER CENT CONFIDENCE LIMITS FOR PRECIPITATION VARIATES IN INCHES

Variate:	*RI*		*RII*		*Rt*	
Treatment:	Seeded	Control	Seeded	Control	Seeded	Control
Observed geometric mean	0.224	0.229	0.088	0.107	0.217	0.278
Adjusted geometric mean	0.220	0.233	0.079	0.118	0.225	0.268
95 per cent confidence limits	$0.66 < (S/C) < 1.35$		$0.32 < (S/C) < 1.39$		$0.20 < (S/C) < 3.52$	

geometric means, presumably obtained from the relevant regression equations. It is not clear whether a correction for bias introduced by the logarithmic transformation was applied or not. The indicated per cent change in precipitation ascribable to seeding may be obtained using these estimates as follows:

seeded/not seeded = 0.220/0.233 = .94 for *RI*;
seeded/not seeded = 0.079/0.118 = .67 for *RII*;
seeded/not seeded = 0.225/0.268 = .84 for *Rt*.

Thus, the indicated effects of seeding are decreases in precipitation of 6, 33, and 16 per cent, respectively. As seen from the confidence intervals in table II, none of these effects was found significant.

6. Reevaluation of SCUD

In this section we indicate the method used in the reevaluation of SCUD data which yielded the relevant entries in table I in [2] and also in table A-IV of [3]. The reevaluation has two phases. Phase (i) consists of the use of the optimal $C(\alpha)$ criterion to test the hypothesis that seeding has no effect on precipitation. The criterion used is asymptotically optimal with regard to the alternative that with nonzero precipitation the effect of seeding is multiplicative and that it may be either an increase or a decrease. The multiplicativity of the effect of seeding is understood to mean that the conditionally expected seeded precipitation, given any values of the predictors, is equal to the corresponding expected not seeded precipitation multiplied by a fixed factor ρ independent of the predictors. The second phase of the evaluation, phase (ii), consists in obtaining a point estimate of the factor ρ. The two phases of evaluation will be illustrated on precipitation data *RI*, only.

The method of evaluation adopted is based on figures 4, 5, and 6, representing scatter diagrams of precipitation amounts *RI* plotted against each of the predictors.

Inspection of these three figures suggests the adoption of the working hypothesis that, given the values of the predictors, either any one of them, or any two, or all three, the conditional distribution of *RI* is approximately normal with a fixed variance and with linear regression on the predictors. The method of

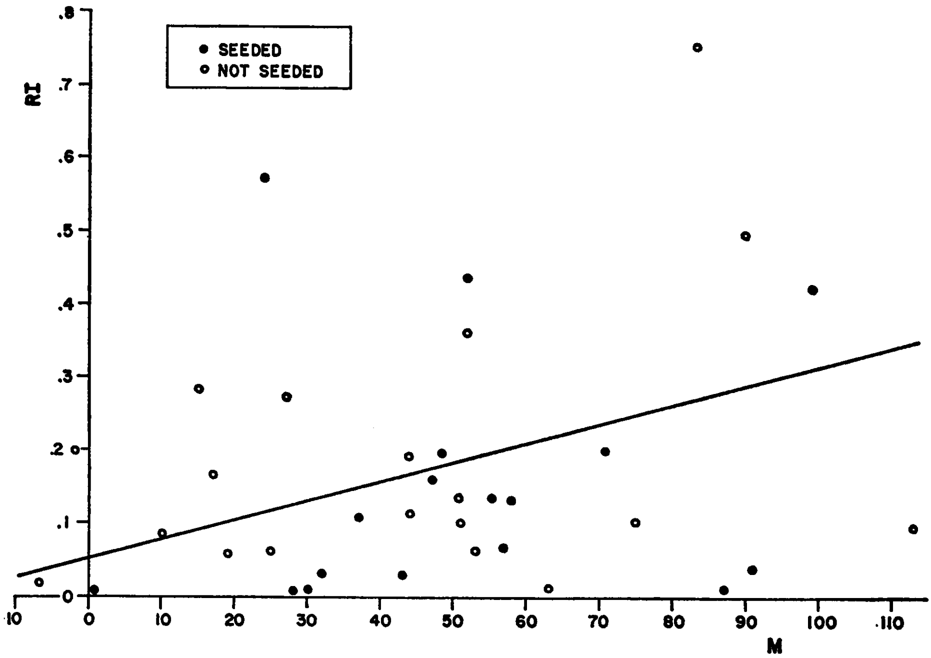

FIGURE 4

Precipitation amounts RI versus circulation index M.
Line fitted by least squares: $Y = a_0 + a_1 x$, $a_0 = 0.122$, $a_1 = 0.00396$; correlation coefficient $r = 0.588$.

evaluation used is based on this hypothesis. This method is the same whether only one of the predictors is used, any two of them, or all three, and the following description is intended to apply to any of these cases.

6.1. *Phase* (*i*). In order to simplify the notation, the precipitation RI delivered by the ith storm and the corresponding value of the jth predictor will be denoted by y_i and x_{ji}, respectively. Bold face letter $\mathbf{x}_i$ will denote the values that all the predictors used in the given evaluation assumed for the ith storm.

On the hypothesis tested, that seeding has no effect, the conditional expectation of the target precipitation, given the predictors used, is equal to, say

$$Y(\mathbf{x}_i) = a_0 + \sum_j a_j x_{ji}, \tag{6.1}$$

where the summation for j extends over all the predictors used.

The values of the parameters a in (6.1) are unknown and symbols $\hat{a}$ will be used to denote their least squares estimates obtained by minimizing the sum

$$\sum_{i=1}^{n} \left(y_i - a_0 - \sum_j a_j x_{ji}\right)^2 \tag{6.2}$$

with respect to the unrestricted variation of the a. Here the summation extends over all the storms, say n, whether seeded or not. The symbol $y(\mathbf{x}_i)$ will stand for the estimate of $Y(\mathbf{x}_i)$ in (6.1), that is to say,

$$y(\mathbf{x}_i) = \hat{a}_0 + \sum_j \hat{a}_j x_{ji}. \tag{6.3}$$

Finally, S^2 will denote the estimate of the conditional variance, that is, the

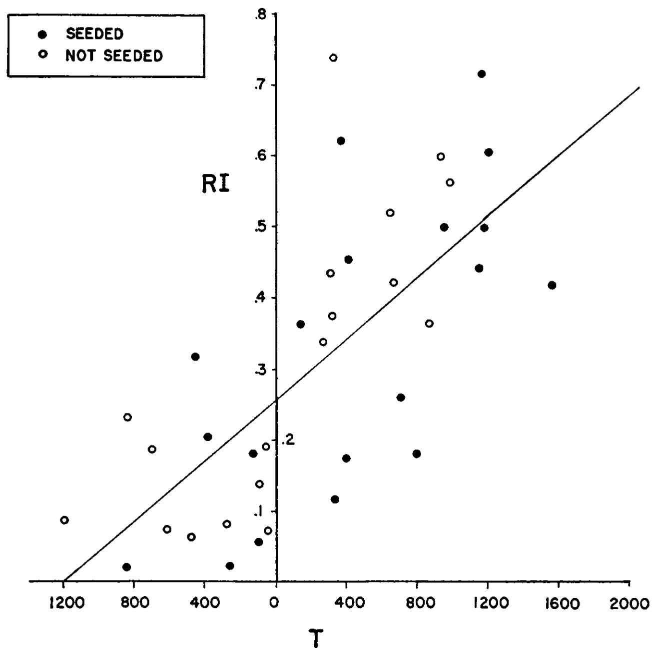

FIGURE 5

Precipitation amounts RI versus net water vapor influx T; $a_0 = 0.260$, $a_1 = 0.000215$; $r = 0.705$.

minimum of (6.2) divided by the number of degrees of freedom, say m, equal to the total number of observations minus one and minus the number of predictors used,

$$S^2 = \sum_{i=1}^{n} [y_i - y(\mathbf{x}_i)]^2/m. \tag{6.4}$$

With this notation, and on the assumption of randomization either in pairs or unrestricted with probability of seeding equal to 1/2, the optimal $C(\alpha)$ criterion can be written as, say

$$Z_n = \frac{\sum_s[y_i - y(\mathbf{x}_i)]y(\mathbf{x}_i) - \sum_c[y_i - y(\mathbf{x}_i)]y(\mathbf{x}_i)}{S\,[\sum y^2(\mathbf{x}_i)]^{1/2}}, \tag{6.5}$$

where symbols $\sum_s$ and $\sum_c$ denote summations for i extending over seeded and

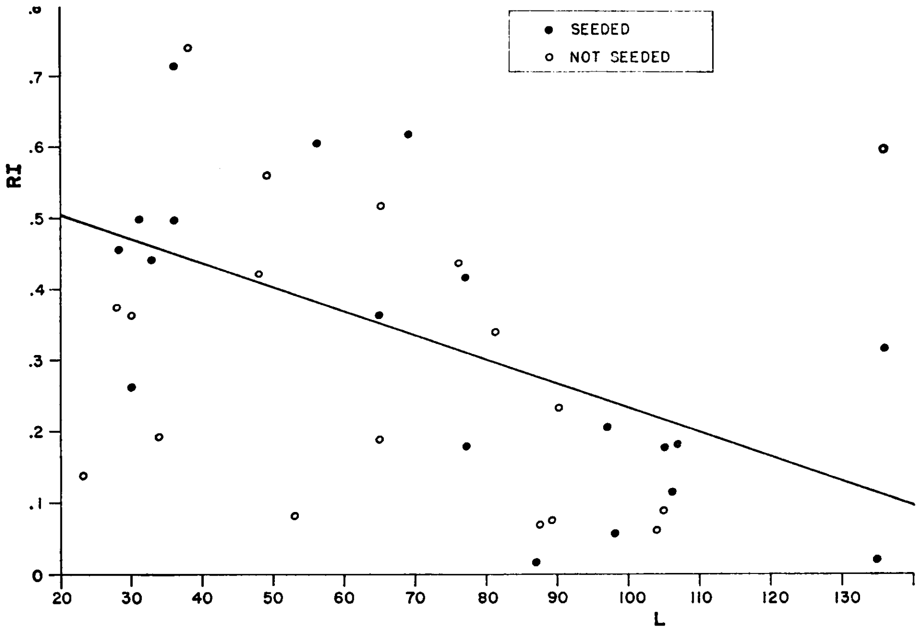

FIGURE 6

Precipitation amounts RI versus measure of latitude L;
$a_0 = 0.503$, $a_1 = -0.00268$; $r = -0.432$.

over control storms, respectively. The summation in the denominator extends over all the storms, seeded and not seeded.

It is seen that, in order to compute Z_n, all that is needed from the observations is the arithmetic means of all the variables and the covariance matrix,

	M	T	L	RI
M	950.6366	7765.549	−356.8746	3.762277
T	—	464063.1	−10252.32	99.74423
L	—	—	1118.974	−2.997889
RI	—	—	—	0.04307822

All the rest is straightforward. Because of the particular interest in the pair of predictors (M, T) and in its relation to L discussed in [3] we give the estimated regression formulas of RI on M and T and on M, T, and L.

$$\text{Regression of } RI \text{ on } (M, T) = 0.147 + 0.00255M + 0.000172T, \tag{6.6}$$

$$\text{Regression of } RI \text{ on } (M, T, L) = 0.179 + 0.00246M + 0.000165T - 0.000378L.$$

The variance of L in the above matrix is of the order of 1119 with a standard deviation of about 33. Multiplying this number by 0.0004, the approximate partial regression coefficient of *RI* on *L*, we obtain 0.0132. Thus, it should not be surprising that the inclusion of *L* into a regression equation in terms of *M* and *T* will result in a change in the expectation of *RI* expressed by no more than a couple of units in the second decimal. As a result, the estimates of the residual variance of *RI* obtained alternatively using only the couple of predictors (*M*, *T*) and using all three predictors (*M*, *T*, *L*), are very much the same, 0.017 and 0.018, respectively. The same applies to the square root of the sum $\sum y^2(\mathbf{x}_i)$ appearing in the denominator of (6.5). This explains the circumstance noted in [3] that, when the predictors *M* and *T* are used, the addition of *L* does not increase the precision of the experiment.

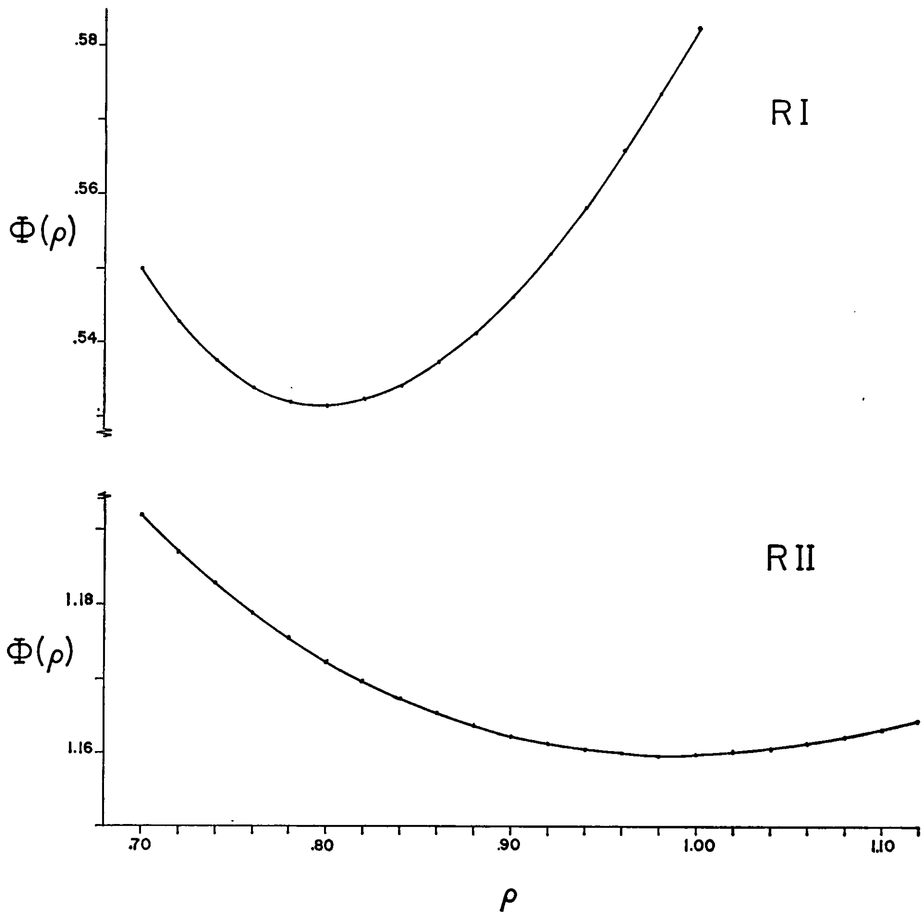

FIGURE 7

Illustration of procedure for obtaining least squares estimates of factor ρ for rainfall variables *RI* and *RII*.

6.2. *Phase (ii).* Because of the frequency of cases where a large cloud seeding experiment fails to yield significant results, the appraisal of the overall experiment must be based in part, on point estimates of the effects of seeding and it is important that these point estimates be efficient.

Under the hypothesis explained above, the maximum likelihood estimate of the factor ρ is obtained by minimizing

$$\sum_c [y_i - (b_0 + \sum b_j x_{ji})]^2 + \sum_s [y_i - \rho(b_0 + \sum b_j x_{ji})]^2 \tag{6.7}$$

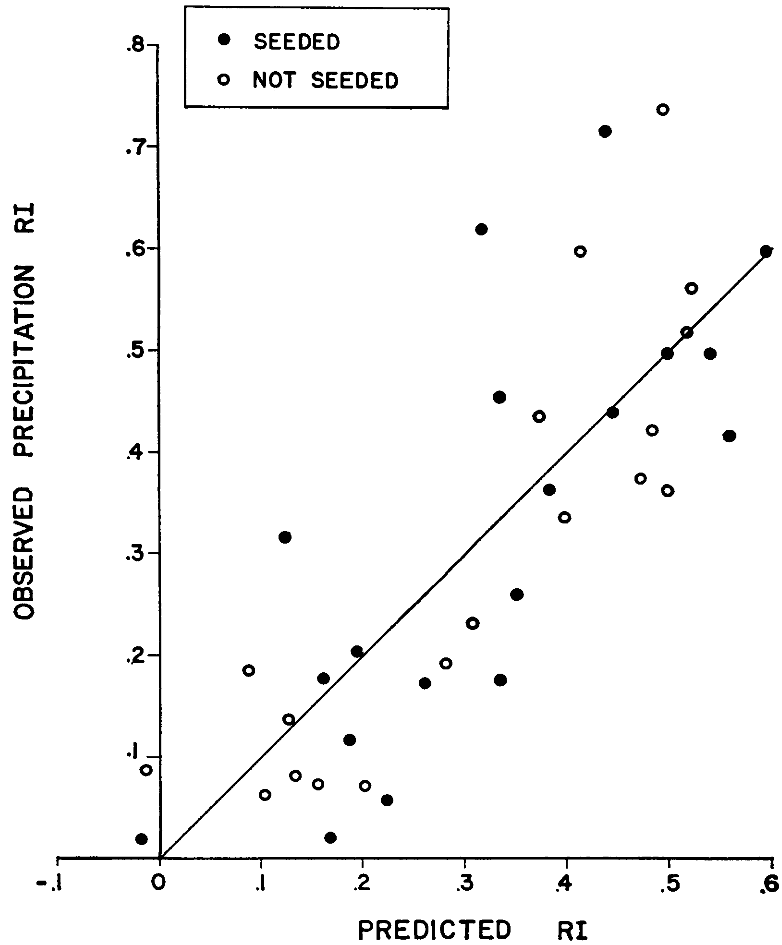

FIGURE 8

Comparison of observed and predicted value of precipitation *RI* using assumptions: (i) effect of seeding, if any, is multiplicative, and (ii) the regression of *RI* on (M, T, L) is linear; $Y = \rho(b_0 + b_1M + b_2T + b_3L)$ with $b_0 = 0.168$, $b_1 = 0.00286$, $b_2 = 0.000212$, $b_3 = -0.0000722$, and $\rho = 1.00$ for control and 0.80 for seeded. Correlation coefficient $r = 0.811$.

with respect to the unrestricted variation of parameters b and ρ. Here again the summation for j extends over the predictors used and the summations $\sum_c$ and $\sum_s$ over all control and over all seeded storms.

With the use of a digital computer, the most convenient method of minimizing (6.7) seems to be as follows. We begin by selecting a set of trial values of ρ, perhaps unity ± 0.1, ± 0.2, ± 0.3, and so forth. Next, with each such trial value of ρ, the sum of squares (6.7) is minimized with respect to the unrestricted variation of the b, which requires only the solution of a system of linear equations. Let $\Phi(\rho)$ denote the minimum of (6.7) so obtained for a given value of ρ. The final stage consists in plotting $\Phi(\rho)$ against ρ and interpolating that value $\hat{\rho}$ that yields the minimum. Figure 7 illustrates the procedure performed for *RI* and *RII*, respectively, using all three predictors. For *RI*, the maximum likelihood estimate of ρ happens to be $\hat{\rho} = 0.80$ indicating a twenty per cent decrease in precipitation due to seeding. For *RII* the same process yields $\hat{\rho} = 0.98$. Here, then, the estimated decrease in rain due to seeding is only 2 per cent, essentially zero. Because of the great distance between the target area and the area where seeding was performed, this latter estimate is quite convincing.

Figure 8 shows the resulting scatter diagram of observed precipitation versus predicted precipitation *RI*, based on (6.7).

REFERENCES

[1] J. Spar, "Project SCUD," *Meteor. Monogr.*, Vol. 2 (1957), pp. 5–23.

[2] J. Neyman and E. L. Scott, "Some outstanding problems relating to rain modification,' *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 5, pp. 293–325.

[3] ———, "Planning an experiment with cloud seeding," Appendix to [2], *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 5, pp. 327–350.