Test of Runoff Increase Due to Precipitation Management for the Colorado River Basin Pilot Project

H. J. MOREL-SEYTOUX[1] AND F. SAHELI[2]

Dept. of Civil Engineering, Colorado State University, Fort Collins

(Manuscript received 31 March 1972, in revised form 20 November 1972)

ABSTRACT

The problem of detection of an increase in runoff due to precipitation management is considered. Classical methods of detection are reviewed and their shortcomings are discussed. The concept of grouping of observations is introduced. The paper answers two fundamental questions:

1) Given a region consisting of B basins in which changes are suspected and given that (economic, or financial, or other) constraints limit to b the number of basins where measurements can be obtained, which basins should be selected?

2) How should the measurements in individual basins be combined?

The answers are obtained by a constrained optimization procedure. When applied to the Colorado River Basin Pilot Project area the power of the test, expressed in years needed for detection, is increased by a factor of 2.

1. Introduction

a. *Water resources planning*

The increasing demand, for water has led men in positions of responsibility to be concerned with the problem of water shortage in particular and of water resources in general (Bureau of Reclamation, 1966; Gilman *et al.*, 1965; Office of Saline Water, 1963; National Academy of Sciences, 1966). The water situation is particularly critical in the Colorado River Basin. The Colorado River system is the largest in the United States that flows mainly through lands with a chronic water deficiency for cultivation of crops (National Academy of Sciences, 1968). The average specific (or unit) yield of the Lower Colorado River Basin is only 0.3 inch, the lowest yield in the United States for a drainage area of this size (National Academy of Sciences, 1968). The Upper Colorado River Basin does not fare much better, 2.2 inches. It outranks only a few basins, the Rio Grande and the Missouri, but it is far below the Mississippi's 10 inches and the Columbia's 16 inches. Population projections and the associated water demands indicate a need for actual importation of approximately 3 million acre-feet annually by the year 2080 (Smith, 1968). Development of the vast oil-shale resources alone would require an additional 1 million acre-feet by the year 2000, assuming a daily oil production of 4 million barrels (National Academy of Sciences, 1968; Smith, 1968). "This amount of water simply is not there now," (Smith, 1968) although "the Colorado Basin is closer than most other basins in the United States to utilizing the last drop of available water for man's needs" (National Academy of Sciences, 1968).

[1] Permanent affiliation: Laboratoires de Mechanique des Fluids des Solides et des Sols, Université de Grenoble, Grenoble, France.

[2] Permanent affiliation: Department of Mechanical Engineering, Aria-Mehr University, Tehran, Iran.

b. *Precipitation management: An alternative to importation*

The potential economic and quantitative significance of precipitation management by seeding winter storms, is now reasonably well established. Under full-scale operations it is estimated that an average additional 1.9 million acre-feet would appear annually in the rivers (Hurley, 1968). In the fall of 1970 The Bureau of Reclamation initiated a 4 or 5 year program of precipitation management in the San Juan Mountains area of Colorado. The area is shown in Fig. 1. The program is known as the "Colorado River Basin Pilot Project" (Bureau of Reclamation, 1968). The "target" area is divided in four zones. The original intent of the Bureau of Reclamation, Division of Atmospheric Water Resources Management, was to seed all zones. Later the decision was made to seed only zones 1 and 2. For this reason, zones 3 and 4 are considered as part of the target area but are nevertheless utilized as "control" zones. The control area refers to the area of Colorado shown in Fig. 2, for which there are currently no weather modification plans for the near future, i.e., for at least the duration of the Colorado River Basin Pilot Project.

2. Evaluation of atmospheric water resources attainments

The main difficulty in this evaluation is caused by the natural variability of hydrologic variables which

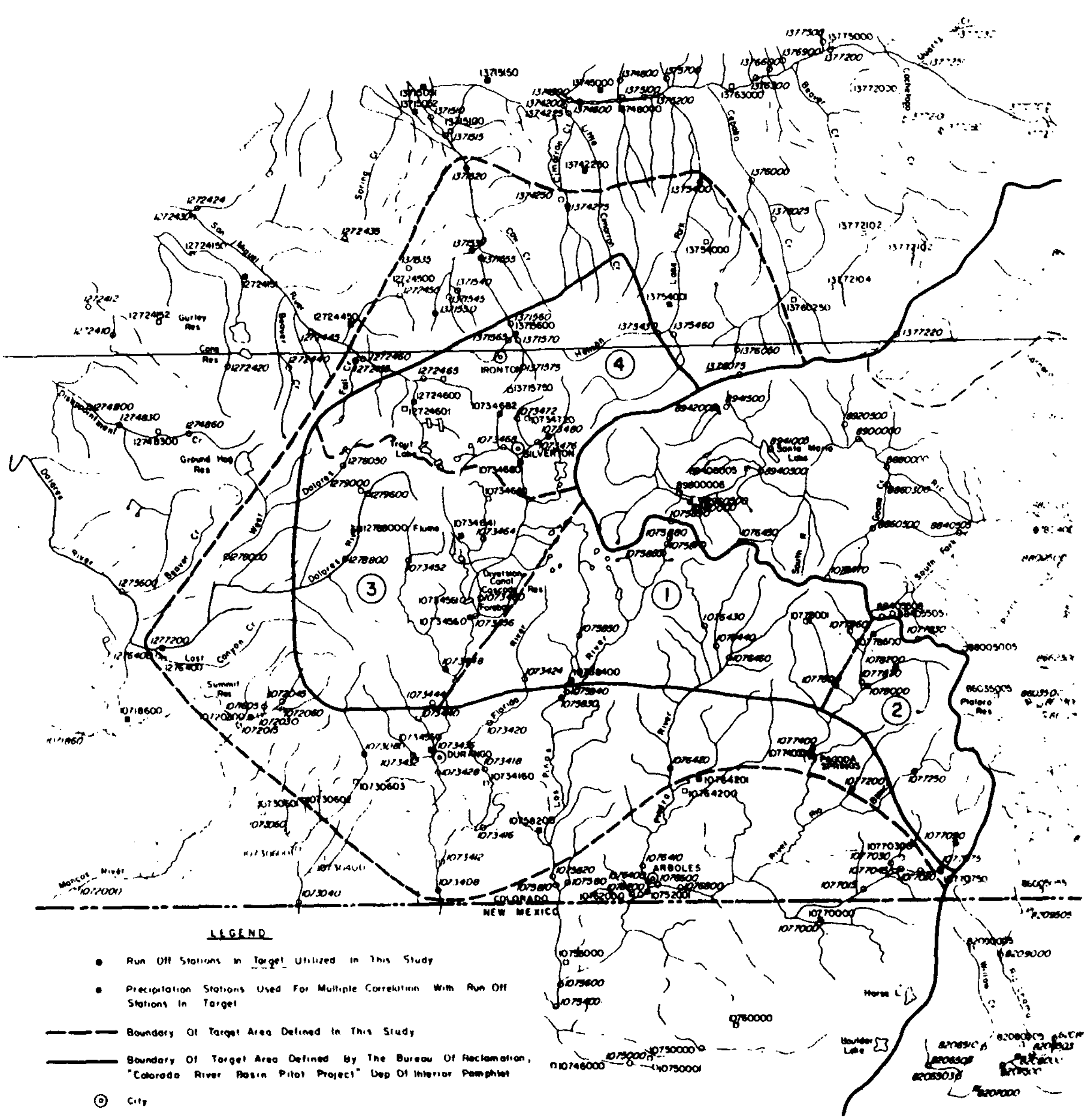

FIG. 1. Target area (San Juan Mountains of Colorado) showing zones 1, 2, 3 and 4, runoff stations, and precipitation stations.

exceeds the expected range of increases induced by man. Simple statistical tests have been developed (Markovic, 1966). They have not proven very sensitive and, as a result, require long periods of observations, prior to and during seeding operations, in order to give satisfactory test results. Furthermore, these tests are insensitive when experiments are performed during a dry period of annual stream-flow sequences. In the early days of weather modification it was thought that such simple techniques would be adequate because it was hoped that the increases would be large. By this simple technique (two-sample test) it can be shown that the number of years N needed for significance at the 95% level (two-tailed) and 50% power (Brownlee, 1961) is

$$N=(1.96)^2C_v^2/h^2\approx 4C_v^2/h^2, \tag{1}$$

where C_v is the coefficient of variation, i.e., the percent ratio of standard deviation to the mean, and h is a given (hypothesized) percentage increase in the population mean. The number N was calculated for a variety of basins in the target area under the assumption of a 10% increase in "seasonal" runoff (total runoff for the 6-month period March–August inclusive). The results are shown in Table 1. These numbers greatly exceed the 4–5 years planned duration. Even for a 20% increase the numbers would be prohibitively high. (The actual number of years would be the ones given in the

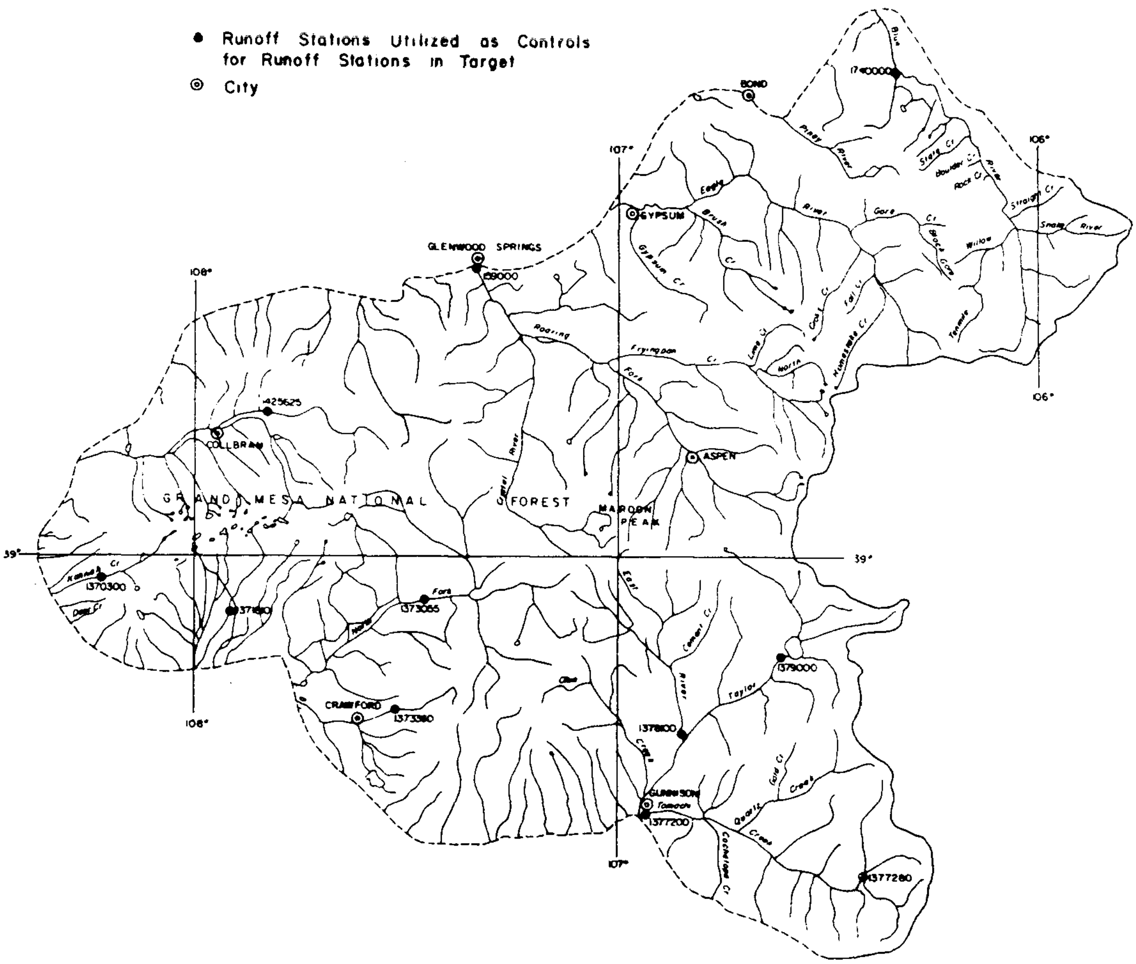

FIG. 2. Control area (Maroon Peak and Grand Mesa Region).

last column of Table 1 divided by 4, since N is inversely proportional to the square of the increase). Based on the current state of the art (Grant, 1969) it is doubtful that an increase much greater than 20% could be achieved in the Colorado River Basin Pilot Project. Under these conditions the two-sample test is inadequate.

More sophisticated techniques are needed. The target control concept was introduced, and different tests were devised, including a chi-square test and a Student t-test. In a previous study (Morel-Seytoux, 1968), a target-control chi-square test was applied to the mean annual or mean seasonal flows of some rivers and it was shown that the number of years N necessary to detect a given percentage increase in population mean h, at the 95% level of significance (two-tailed) and 50% power (Brownlee, 1961), was

$$N=(1.96)^2(1-\rho^2)C_v^2/h^2\approx 4(1-\rho^2)C_v^2/h^2, \qquad (2)$$

where ρ is the correlation coefficient between the target and the control watersheds, and C_v the coefficient of variation of the target watershed.

It is important to know before actually conducting the experiments how many years are needed to have an even chance to achieve significance. Therefore, calculations were performed for a few stations in the Upper Colorado River Basin to get an idea of what could be expected if seeding operations were conducted in the area. In particular, the number of years (needed) to detect a 10% (population) increase (at the 95% level and 50% power) was calculated. This number of years can be viewed as an average number of years needed because the power was chosen for the calculations at the 50% level. The assumed 10% increase is a conservative estimate for the region (Grant, 1969). However, because the operations of seeding are randomized on a fifty-fifty basis (Bureau of Reclamation, 1968), half the opportunities to increase the snowpack are lost. Furthermore, the operations are stopped when the snowpack exceeds 150% of the "normal" conditions. Again opportunities are lost. Under these conditions the probability that the runoff be increased by more than 10% seems rather dim. For this (plausible) assumed 10% increase the number of years needed for significance was calculated, based on Eq. (2) for basins

TABLE 1. Names and characteristics of the runoff stations in the target area and number of years needed for evaluation, as calculated from (1).

CSU number	Station name*	Drainage area (mi^2)	Elevation (ft)	Mean seasonal runoff (acre-ft)	Seasonal runoff standard deviation (acre-ft)	Coefficient of variation (%)	Number of years needed for significance at 95% level (two-tailed) and 50% power, assuming a 10% runoff increase (and no control)
1073080	La Plata River at Hesperus	37	8105	26,810	12,750	48	87
1073408	Animas River near Cedarhill	1090	5960	525,100	221,700	42	69
1073436	Animas River at Durango	692	6502	454,200	173,500	38	56
1073448	Hermosa Creek near Hermosa	172	6705	75,350	44,270	59	133
1073480	Animas River at Howardsville Park	55.9	9617	64,460	17,920	28	30
1075830	Los Pinos River near Bayfield	284	7515	210,500	84,070	40	62
1076420	Piedra River near Piedra	371	6530	191,500	96,920	51	99
1077090	Navajo River at Banded Peak Ranch near Chromo	69.8	7941	61,410	23,160	38	55
1077250	Rio Blanco near Pagosa Springs	58	7950	51,190	21,170	41	66
1077400	San Juan River at Pagosa Springs	298	7052	211,000	116,900	55	118
1078000	East Fork San Juan River near Pagosa Springs	86.9	7598	74,730	32,770	44	74
1272445	San Miguel River near Placerville	308	7096	132,400	58,210	44	75
1277200	Dolores River at Dolores	556	6925	265,100	114,100	43	72
1371520	West Fork Dallas Creek near Ridgeway	437	8400	145,600	55,620	38	56
1375400	Lake Fork at Gateview	388	7828	147,800	49,390	33	43

* U. S. Geological Survey stations in Colorado.

in zones 1 and 2, using basins in zones 3 and 4 as controls. The results are shown in Table 2. The improvement over the results of Table 1 is striking but not sufficient. Additional results can be obtained by combining data of Tables 1 and 3. For example, for station 1077250 in zone 2 without control, the number of years is 66. The highest correlation with any other station in the target area is with station 1077090 and $\rho = 0.95$ (see Table 3). With that control the value of N is $66[1-(0.95)^2] = 6.6$, which exceeds 5; moreover, that control station (Navajo River at Banded Peak Ranch) is a neighbor in zone 2 (as could be expected with such a high correlation) and therefore not acceptable as a control. What then can be done to reduce the number of years needed to obtain significance?

3. The concept of grouping of observations

Because of the many interacting complex parameters and variables involved, local weather variability, etc., the runoff of a watershed can be viewed as a random variable. This point of view is rather old in hydrology but nevertheless still gaining ground (Yevjevich, 1972). Fundamentally, a hydrologic change in the course of time is detected when the recent observations can no longer be viewed as coming from the statistical distribution that prevailed over the long past. Let us denote the (seasonal) runoff-random variable from watershed i by Q_i. An observation (realization) of this random variable (r.v.) in year j is denoted q_{ij}. There are many such r.v. in the target as well as in the control area. Up to this point all techniques involved either one target r.v. (two-sample test) or a pair of r.v. (two-sample target-control test). Of course, for the purpose of rapid detection, if the percentage increase in runoff was the same in all watersheds one would pick up the "best" pair of target control watersheds in Table 2, namely the pair with lowest value of N in the last column of Table 2 (the pair with $N = 5.9$). The criterion for the selection was the minimization of N. Looking back at (2) one realizes that there are at least two ways of minimizing N, namely minimize C_v, i.e., select the watershed with minimum natural variability in runoff, or maximize ρ, i.e., select the pair of target and control watersheds which are most highly correlated. It is well known in hydrometeorology that the natural variability tends to decrease when the hydrologic variables are averaged over a large region. Thus, another possibility for decreasing N is to consider not pairs of elemental

TABLE 2. Names and characteristics of stations in the target area and number of years needed for evaluation associated with target-control pairs as calculated from (2).

Target (zones 1 and 2)					Control (zones 3 and 4)					Target-control pair
CSU number	Station name	Drainage area (mi²)	Elevation (ft)	Coefficient of variation (%)	CSU number	Station name	Drainage area (mi²)	Elevation (ft)	Coefficient of correlation (%)	Number of years needed for significance at 95% level (two-tailed) and 50% power assuming a 10% runoff increase
1078000	East Fork San Juan River near Pagosa Springs	86.9	7598	44	1073448	Hermosa Creek near Hermosa	172	6705	92	11
					1272445	San Miguel River near Placerville	308	7096	68	40
					1277200	Dolores River at Dolores	556	6925	92	12
					1375400	Lake Fork at Gateview	388	7828	80	27
1077400	San Juan River at Pagosa Springs	295	7052	55	1073448	Hermosa Creek near Hermosa	172	6705	65	68
					1272445	San Miguel River near Placerville	308	7096	28	109
					1277200	Dolores River at Dolores	556	6925	58	79
					1375400	Lake Fork at Gateview	388	7828	56	82
1077250	Rio Blanco near Pagosa Springs	58	7950	41	1073448	Hermosa Creek near Hermosa	172	6705	87	16
					1272445	San Miguel River near Placerville	308	7096	72	32
					1277200	Dolores River at Dolores	556	6925	89	14
					1375400	Lake Fork at Gateview	388	7828	78	27
1076420	Piedra River near Piedra	371	6530	51	1073448	Hermosa Creek near Hermosa	172	6705	96	7.4
					1272445	San Miguel River near Placerville	308	7096	71	49
					1277200	Dolores River at Dolores	556	6925	92	15
					1375400	Lake Fork at Gateview	388	7828	82	33
1075830	Los Pinos River near Bayfield	284	7515	40	1073448	Hermosa Creek near Hermosa	172	6705	95	5.9
					1272445	San Miguel River near Placerville	308	7096	69	33
					1277200	Dolores River at Dolores	556	6925	93	8.8
					1375400	Lake Fork at Gateview	388	7828	89	13

TABLE 3. Coefficient of correlation between target stations (%).

CSU number	Station name	Target stations (CSU number) 1073080	1073408	1073436	1073448	1073480	1075830	1076420	1077090	1077250	1077900	1078000	1272445	1277200	1371520	1375400
1073080	La Plata River at Hesperus	100	97	95	95	89	95	94	94	93	70	92	72	93	82	84
1073408	Animas River near Cedarhill	97	100	99	97	94	98	97	94	90	66	93	75	95	87	90
1073436	Animas River at Durango	95	99	100	96	96	98	96	93	89	63	93	71	96	88	94
1073448	Hermosa Creek near Hermosa	95	97	96	100	88	95	96	92	87	65	92	66	93	84	85
1073480	Animas River at Howardsville Park	89	94	96	88	100	94	88	87	86	60	88	69	93	88	97
1075830	Los Pinos River near Bayfield	95	98	98	95	94	100	96	92	90	65	93	69	93	82	89
1076420	Piedra River near Piedra	94	97	96	96	88	96	100	96	89	63	96	71	92	83	82
1077090	Navajo River at Banded Peak Ranch near Chromo	94	94	93	92	87	92	96	100	95	59	96	75	94	86	79
1077250	Rio Blanco near Pagosa Springs	93	90	89	87	86	90	89	95	100	63	92	72	89	80	78
1077400	San Juan River at Pagosa Springs	70	66	63	65	60	65	63	59	63	100	60	28	58	50	56
1078000	East Fork San Juan River near Pagosa Springs	92	93	93	92	88	93	96	96	92	60	100	68	92	79	80
1272445	San Miguel River near Placerville	72	75	71	66	69	69	71	75	72	28	68	100	68	65	59
1277200	Dolores River at Dolores	93	95	96	93	93	93	92	94	89	58	92	68	100	93	91
1371520	Uncompahgre River at (near) Chromo	82	87	88	84	88	82	83	86	80	50	79	65	93	100	89
1375400	Lake Fork at Gateview	84	90	94	85	97	89	82	79	78	56	80	59	91	89	100

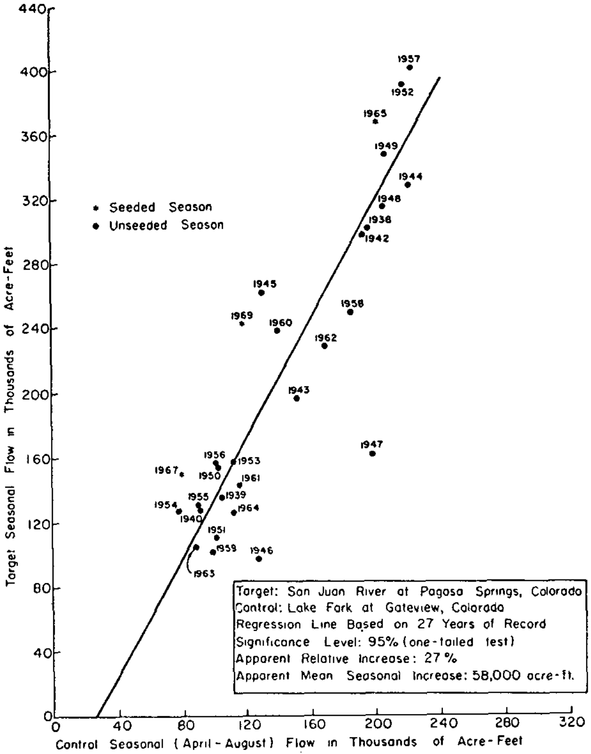

FIG. 3. Regression line between target seasonal runoff and control seasonal runoff.

target control r.v. but a global pair, i.e., total target runoff and total control runoff. That it works is demonstrated by an analysis of the effect of weather modification on runoff in the Wolf Creek Pass area. Under sponsorship of the State of Colorado and the direction of Prof. Grant of Colorado State University, the area was seeded during the years 1965, 1967 and 1969. The target-control conditional Student's t-test was applied to several pairs of target-control r.v. and to the global pair. The results are displayed graphically in Figs. 3–9. For the global pair (Fig. 9) an apparent relative increase of 20% is shown to be significant at the 99% (one-tailed) or 98% (two-tailed) level, the highest significance level shown in all the figures, whereas for an apparent relative increase of 27% (Fig. 3) the significance level is only 95% (one-tailed) or 90% (two-tailed). Note that in Prof. Grant's experiments all seedable storms were seeded. In the Colorado River Basin Pilot Project they will not be. Had the global apparent increase been only 10%, given the hydrologic characteristics of the watersheds involved, the significance level attained in three years of experiments would have been less than 95%. Nevertheless the results are encouraging.

Clearly the global (or total) target runoff is a particular linear combination of the elemental (or component) target runoffs, i.e., a combination where all the weights in the combination equal unity. This remark suggests an additional possibility of minimizing N by considering

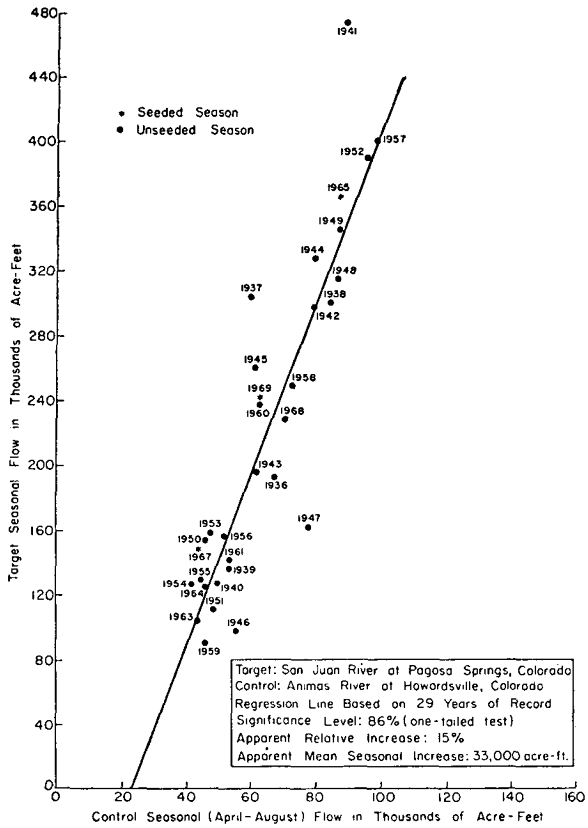

FIG. 4. Regression line between target seasonal runoff and control seasonal runoff.

the general combination denoted Q_t^* and defined by

$$Q_t^* = \sum_{i=1}^{n} X_i Q_i, \tag{3}$$

where the X_i are fixed but unknown parameters and n is the total number of target r.v. Similarly, a combination can be defined for the control area, denoted Q_c^*, and defined by

$$Q_c^* = \sum_{i=n+1}^{n+m} X_i Q_i, \tag{4}$$

where m is the total number of control r.v. The introduction of the variables X permits a much more thorough minimization of N if one realizes that this problem is the minimization of a conditional variance of a linear combination of several r.v., given a linear combination of other r.v. In multivariate analysis it is traditional (Graybill, 1961) to denote a set of r.v. as a random vector; we thus use $\mathbf{Q}$ in this study of runoff. Let $\mathbf{\Sigma}$ be the covariance matrix of $\mathbf{Q}$. Suppose the vector $\mathbf{Q}$ is partitioned into two vectors $\mathbf{Q}_t$ and $\mathbf{Q}_c$. The covariance matrix can be partitioned similarly into four matrices $\mathbf{\Sigma}_{tt}$, $\mathbf{\Sigma}_{tc}$, $\mathbf{\Sigma}_{ct}$ and $\mathbf{\Sigma}_{cc}$. Let Q_t^* denote the linear combination $\mathbf{X}_t\mathbf{Q}_t$, where the prime indicates the transpose operation, and Q_c^* denote $\mathbf{X}_c'\mathbf{Q}_c$, where $[\mathbf{X}_t', \mathbf{X}_c']$ is a partition of $\mathbf{X}'$, a vector of parameters, fixed but unknown weight factors. Then it is a well-known result (Graybill, 1961) that the variance of Q_t^* condi-

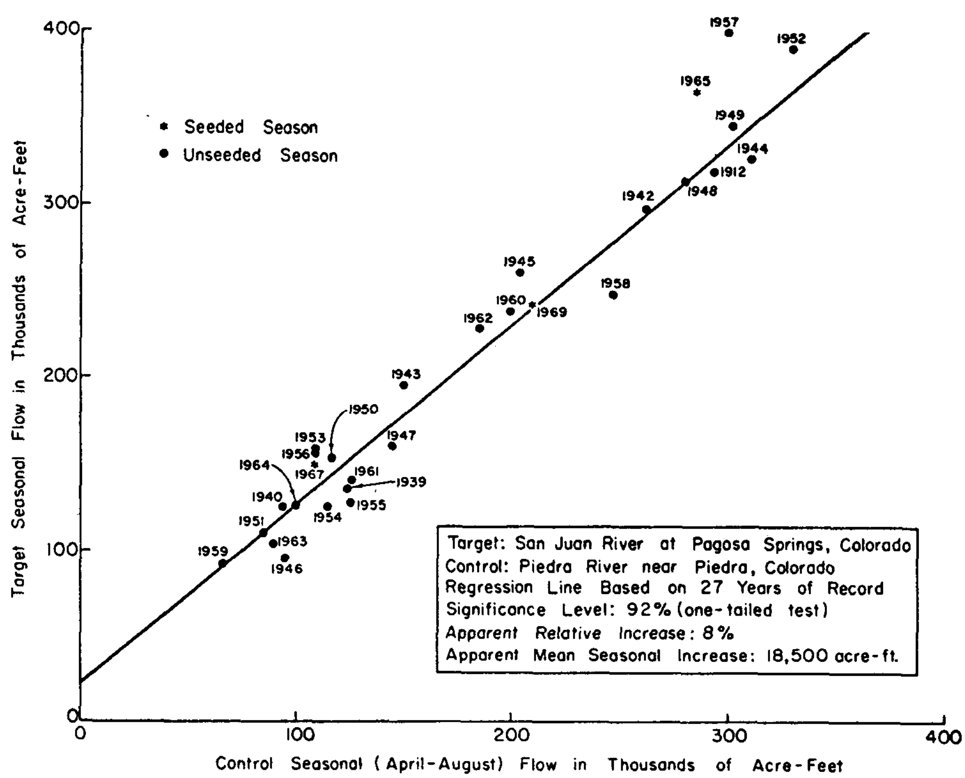

FIG. 5. Regression line between target seasonal runoff and control seasonal runoff.

tioned on Q_c^* is given by

$$\mathrm{Var}[Q_t^* | Q_c^*] = \mathbf{X}_t'\mathbf{\Sigma}_{tt}\mathbf{X}_t - (\mathbf{X}_t'\mathbf{\Sigma}_{tc}\mathbf{X}_c)^2/(\mathbf{X}_c'\mathbf{\Sigma}_{cc}\mathbf{X}_c). \quad (5)$$

Given the convariance matrix $\mathbf{\Sigma}$ (a regional characteristic) $\mathrm{Var}[Q_t^*|Q_c^*]$ can be minimized by the proper choice of the vectors $\mathbf{X}_t$ and $\mathbf{X}_c$. Given $\mathbf{X}_c$ the minimization of $\mathrm{Var}[Q_t^*|Q_c^*]$ with respect to X_t gives the obvious solution $\mathbf{X}_t = \mathbf{0}$ and the variance is zero. Clearly the minimization problem is meaningless unless restrictions are imposed on $\mathbf{X}$. Normalization conditions, however, can be imposed on $\mathbf{X}_t$ and on $\mathbf{X}_c$, for example:

$$\sum_{i=1}^{n} X_i = n, \quad (6)$$

$$\sum_{i=n+1}^{n+m} X_i = m. \quad (7)$$

In the case of the total target and control runoffs these normalization constraints are satisfied since all the X_i are equal to 1. However, the normalization constraints can be satisfied by other choices of values of the X_i than the values 1. In the problem at hand of detecting a runoff increase due to cloud seeding, one can find more physically meaningful constraints than the normalization ones.

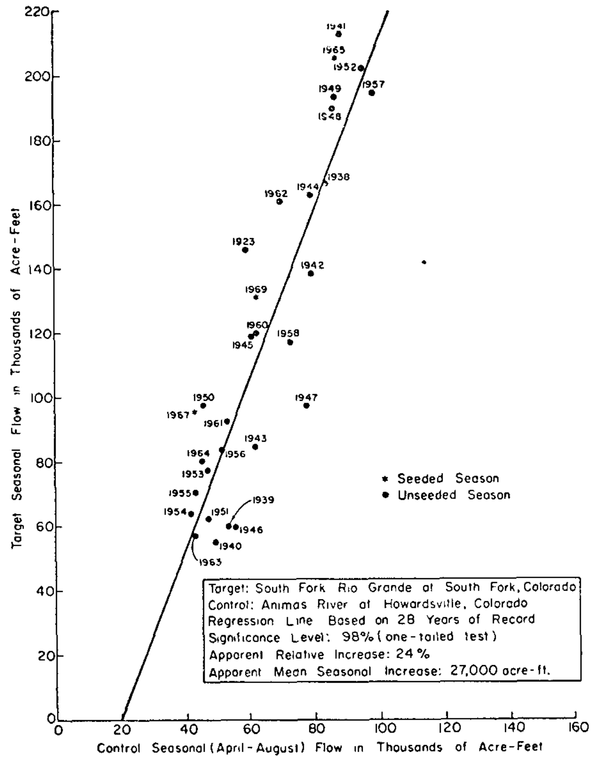

FIG. 6. Regression line between target seasonal runoff and control seasonal runoff.

4. The hydrologic constraints

Explicitly in terms of the $\mathbf{X}$ the simple-looking formula given by Eq. (2) takes the form

$$\Lambda^* = \frac{[\mathbf{X}_t'\mathbf{\Sigma}_{tt}\mathbf{X}_t - (\mathbf{X}_t'\mathbf{\Sigma}_{tc}\mathbf{X}_c)^2/(\mathbf{X}_c'\mathbf{\Sigma}_{cc}\mathbf{X}_c)]}{(\mathbf{X}_t'\mathbf{\Delta}\bar{\mathbf{Q}})^2}, \quad (8)$$

where $\Delta\bar{Q}_i$ is the expected change of mean runoff in basin i due to cloud seeding and $\mathbf{\Delta}\bar{\mathbf{Q}}$ is the vector with components $\Delta\bar{Q}_i$. The first question regarding these $\Delta\bar{Q}_i$ is: how are they known? If they were precisely known what is the need for a statistical evaluation?

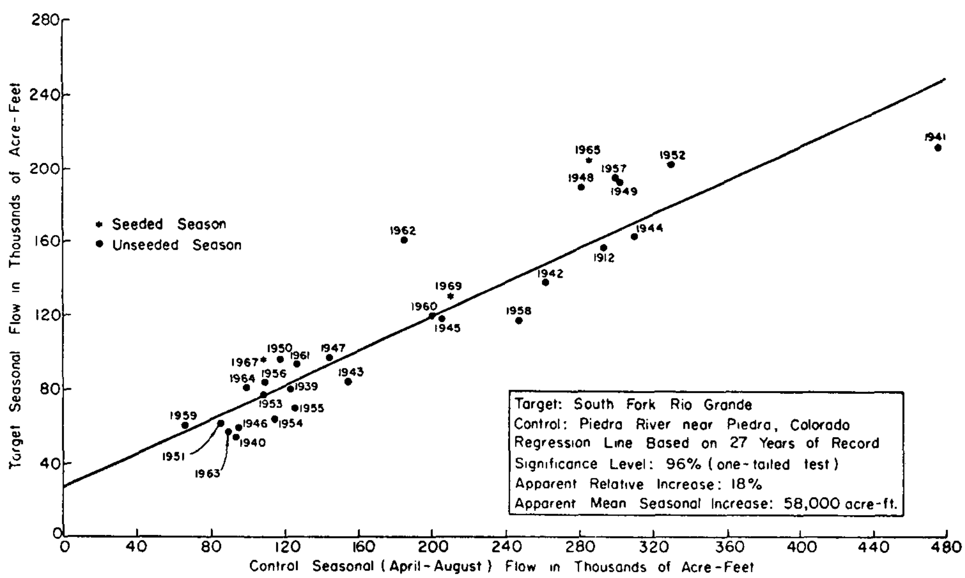

FIG. 7. Regression line between target seasonal runoff and control seasonal runoff.

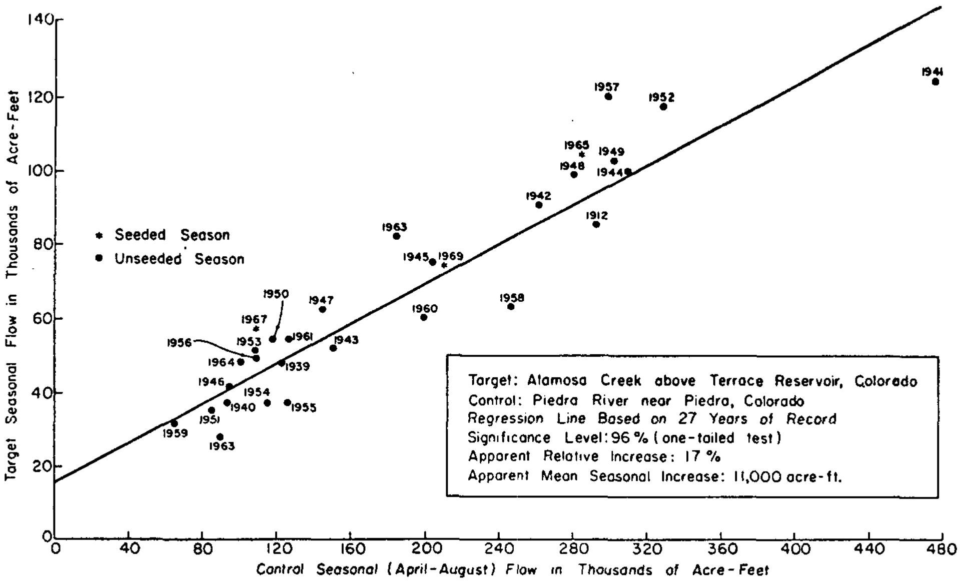

FIG. 8. Regression line between target seasonal runoff and control seasonal runoff.

a. Estimation of the mean runoff changes

Though various aspects of research on cloud modification have been conducted successfully, it is still difficult to determine accurately its quantitative effect. Indeed, one of the purposes of the Colorado River Basin Pilot Project is to determine the exact magnitude of the increase in precipitation on a large areal scale. Following this experiment, it may be possible to isolate the major factors that determine the magnitude of the increase of precipitation. At present it is a somewhat accepted opinion that the mean increase of precipitation by cloud seeding is proportional to the natural mean

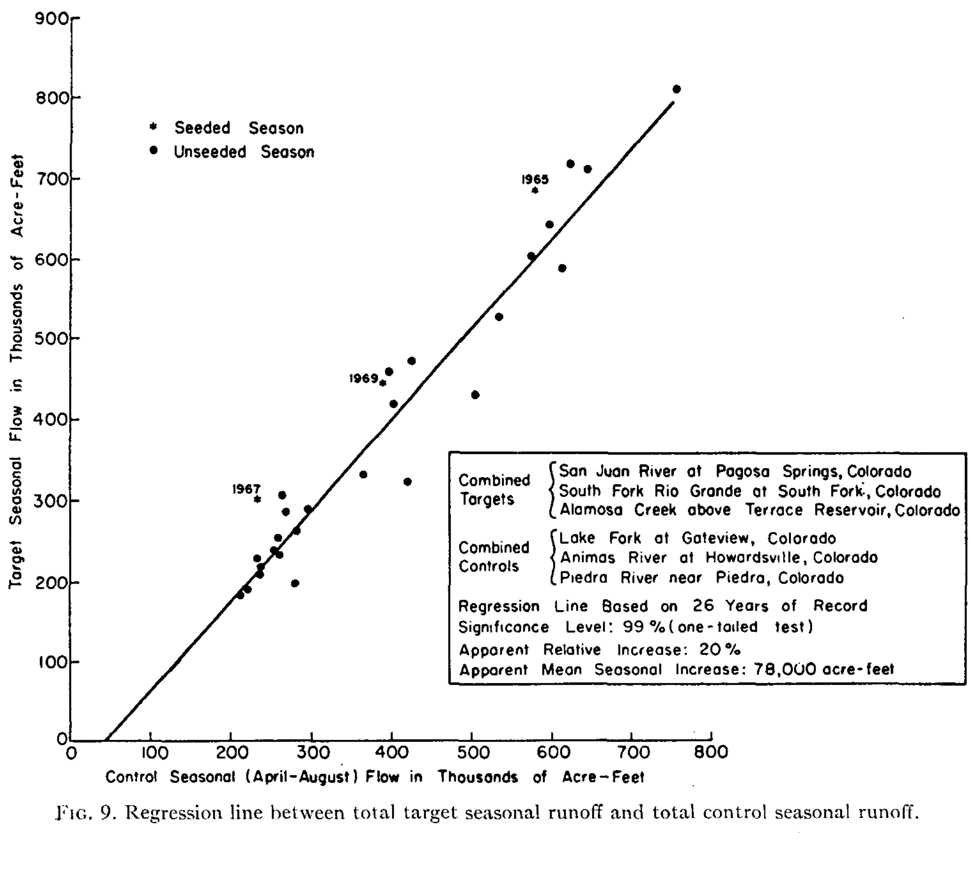

FIG. 9. Regression line between total target seasonal runoff and total control seasonal runoff.

TABLE 4. Estimation of increase in spring runoff due to an increase in winter precipitation.

CSU number	Station name	Mean seasonal runoff (acre-ft)	Mean seasonal runoff (inches)	Mean winter precipitation (inches) at gages No. 1	No. 2	No. 3	No. 4	Mean seasonal runoff increase (inches)	Mean seasonal runoff increase (acre-ft)	Percentage increase in seasonal runoff assuming a 10% increase in mean winter precipitation
1375400	Lake Fork at Gateview	150,673	7.29	14.84	7.79	6.97		0.821	17,000	11.00
1371520	Uncompahgre River at (near) Chromo	149,610	6.40	6.97	5.35	5.22		0.61	14,200	9.55
1277200	Dolores River at Dolores	270,647	9.13	11.41	9.74	9.06	7.70	1.10	32,800	12.10
1272445	San Miguel River near Placerville	135,282	8.33	7.83	7.67			0.791	13,000	9.50
1078000	East Fork San Juan River near Pagosa Springs	75,028	16.10	30.23	11.02			2.25	10,040	14.00
1077400	San Juan River at Pagosa Springs	225,865	13.25	30.23	11.02	10.66		1.46	23,200	11.00
1077250	Rio Blanco near Pagosa Springs	51,398	16.50	30.23	11.02	9.14		1.31	4,050	7.95
1077090	Navajo River at Banded Peak Ranch near Chromo	61,878	16.50	11.02	9.14			1.58	5,900	9.60
1076420	Piedra River near Piedra	193,524	9.68	10.68	8.22			0.88	17,410	9.10
1075830	Los Pinos River near Bayfield	212,086	13.90	15.44	8.22			1.01	15,300	7.26
1073480	Animas River at Howardsville	65,397	21.60	16.94	11.61			1.98	5,860	9.20
1073448	Hermosa Creek near Hermosa	76,161	8.22	19.7	15.44	10.73		0.98	9,000	11.90
1073436	Animas River at Durango	479,575	12.30	10.73	8.22			1.01	37,200	8.20
1073408	Animas River near Cedarhill	536,923	9.00	8.34	5.56			0.11	6,660	1.22
1073080	La Plata River at Hesperus	27,186	13.58	10.73	9.74			1.15	2,280	8.50

precipitation, i.e., that

$$\Delta\bar{P}_w = k\bar{P}_w, \tag{9}$$

in which $\Delta\bar{P}_w$ is the mean increase of winter precipitation by cloud seeding, $\bar{P}_w$ the natural winter precipitation, and k the ratio of mean increase of precipitation to the natural value or relative increase. The effect of cloud seeding is measured by the increase of usable runoff. It is assumed that runoff Q is a function of a representative precipitation P. Then, in the general form,

$$Q = f(P). \tag{10}$$

However, it is hard to find an integrated precipitation that represents the whole basin. Suppose that precipitation data P_j corresponding to Q are collected, as many as possible, in the basin in question. Eq. (10) is then modified as

$$Q = f(P_1, P_2, \cdots). \tag{11}$$

In the case of precipitation management in the San Juan Mountains area, it is the seasonal runoff Q, caused mainly by winter precipitation, P_{wj}, and partially by spring precipitation, P_{sj}, which is of concern. The relationship is represented more precisely by

$$Q = f(P_{w1}, P_{s1}, P_{w2}, P_{s2}, \cdots). \tag{12}$$

Multiple linear regression analysis is applied to find the approximate relationship. Finally,

$$Q = a + b_1P_{w1} + c_1P_{s1} + b_2P_{w2} + c_2P_{s2} + \cdots, \tag{13}$$

in which a, b_j, c_j are coefficients determined from available data.

The increase of spring runoff, ΔQ, caused by the increase of winter precipitation, ΔP_w, is given by

$$\Delta Q = b_1\Delta P_{w1} + b_2\Delta P_{w2} + \cdots. \tag{14}$$

Averaging over the period of record and substituting (9) into (14) yields

$$\overline{\Delta Q} = b_1k_1\bar{P}_{w1} + b_2k_2\bar{P}_{w2} + \cdots. \tag{15}$$

By this procedure all $\overline{\Delta Q}_i$ can be estimated in the target area. Because it is not possible to predict at present the average increase of precipitation due to cloud seeding, it was assumed that the k_i did not vary from basin to basin. For the purpose of this study and for reasons previously discussed, a uniform k value of 10% was assumed. Table 4 shows the results of the estimation of the $\Delta\bar{Q}_i$. Note that the percentage increase in runoff varies significantly from the assumed uniform 10% increase in mean winter precipitation. It is recognized that these values of the $\overline{\Delta Q}_i$ are only very rough estimates.

b. Hydrologic significance of Q^*

From its definition [Eq. (3)] Q^* is a combination of runoff r.v. However, Q^* *per se* has no hydrologic meaning. It is an "artificial" runoff, since the Q_i are both runoff quantities and random variables but Q^* is only a random variable. Hydrologic meaning may be given to one parameter in its distribution, namely its mean, by requiring of the X_i that they satisfy the constraints:

$$\bar{Q}_t^* = \sum_{i=1}^{n} X_i\bar{Q}_i = \sum_{i=1}^{n} \bar{Q}_i, \tag{16}$$

$$\bar{Q}_c^* = \sum_{i=n+1}^{n+m} X_i\bar{Q}_i = \sum_{i=n+1}^{n+m} \bar{Q}_i. \tag{17}$$

The hydrologic interpretation of (16) and (17) is that the expectation of the random variable Q_t^* is the mean of the total runoff for the group of n basins in the target area and similarly in the control area. The above equations are more meaningful constraints than the normalization ones defined by (6) and (7). Similarly, in the target area, one can impose the constraint that

$$\Delta\bar{Q}_t^* = \sum_{i=1}^{n} X_i\Delta\bar{Q}_i = \sum_{i=1}^{n} \Delta\bar{Q}_i. \tag{18}$$

The hydrologic interpretation of (18) is that the mean increase of $\bar{Q}_t^*$ is that of the total runoff for the group of n basins. This constraint is crucial because there would be no point in reducing the variance of Q_t^* given Q_c^* if at the same time the quantity $\overline{\Delta Q_t^*}$ was also reduced considerably below that of the mean increase of the total runoff. In addition, because the $\overline{\Delta Q_i}$ used in (18) are only rough estimates, some may be underestimated and some overestimated. The calculated $\overline{\Delta Q_t^*}$ from (18) may be somewhat different than the observed total increase in runoff during the years of cloud seeding. That chance becomes particularly serious if some of the X_i, solutions of the minimization of N^* subject to the constraints defined by (16)–(18), turn out to be negative. In the extreme case the $\overline{\Delta Q_t^*}$ calculated on the basis of the observed $\overline{\Delta Q_i}$ during the years of seeding rather than the estimated ones could even be negative. For this reason the X_i in the target area are subjected to a non-negativity condition, namely:

$$X_i \geqslant 0, \quad \text{for} \quad i=1, 2, \cdots n. \tag{19}$$

5. Minimal time detection test design

In summary, the most powerful test of the effect of precipitation management on runoff in the Colorado River Basin Pilot Project can be deduced from the solution to a minimization problem. The formulation of the minimization problem is the following.

We minimize the objective function N^*, defined as

$$N^* = \frac{3.84\{\sum_{i=1}^{n}\sum_{j=1}^{n} X_iX_ja_{ij} - [(\sum_{i=1}^{n}\sum_{k=n+1}^{n+m} X_iX_kb_{ik})^2/(\sum_{k=n+1}^{n+m}\sum_{l=n+1}^{n+m} X_kX_lC_{kl})]\}}{(\sum_{i=1}^{n} X_i\Delta\bar{Q}_i)^2}, \tag{20}$$

(where a_{ij} is an element of the covariance matrix in the target area, b_{ij} an element of the covariance matrix between target and control areas, and c_{kl} an element of the covariance matrix in the control area) with respect to the X_i, where $i=1, 2, \cdots n, n+1, \cdots n+m$, subject to the constraints:

$$\sum_{i=1}^{n} X_i\delta_i\bar{Q}_i + \sum_{i=1}^{n} (1-\delta_i)\bar{Q}_i = \sum_{i=1}^{n} \bar{Q}_i, \tag{21}$$

$$\sum_{i=1}^{n} X_i\delta_i\Delta\bar{Q}_i + \sum_{i=1}^{n} (1-\delta_i)\Delta\bar{Q}_i = \sum_{i=1}^{n} \Delta\bar{Q}_i, \tag{22}$$

$$\sum_{i=n+1}^{n+m} X_i\delta_i\bar{Q}_i + \sum_{i=n+1}^{n+m} (1-\delta_i)\bar{Q}_i = \sum_{i=n+1}^{n+m} \bar{Q}_i, \tag{23}$$

$$X_i \geqslant 0, \quad \text{for} \quad i=1, 2, \cdots n, \tag{24}$$

$$\sum_{i=1}^{n} \delta_i = \nu_t \tag{25}$$

$$\sum_{i=n+1}^{n+m} \delta_i = \nu_c, \tag{26}$$

where δ_i is a variable that takes only two values, zero or one, depending upon whether the corresponding variable X_i is zero or non-zero, respectively, and where $\nu_t(\leqslant n)$ and $\nu_c(\leqslant m)$ are the maximum number of target and control runoff variables to be used in the combinations Q_t^* and Q_c^*, respectively. The solution of this minimization problem will systematically select which basins should be used in the significance test and with which weights.

The mathematical programming problem mentioned here is not a standard one. It was solved successfully by a modified version of the Jacobi differential algorithm (Wilde and Beightler, 1967). The details of the method

TABLE 5. Covariance matrix for stations in target. Units: 10^5 $(mi^2 \times inches)^2$.

Target stations CSU numbers	Target stations CSU numbers 1375400	1371520	1277200	1272445	1078000	1077400	1077250	1077090	1076420	1075830	1073480	1073448	1073436	1073408	1073080
1375400	8.289	8.334	17.475	5.790	4.419	10.936	2.755	3.084	13.329	12.625	2.906	6.312	27.345	33.478	1.792
1371520	8.334	10.512	19.967	7.118	4.914	10.957	3.197	3.754	15.114	13.099	2.969	7.017	29.009	36.261	1.966
1277200	17.475	19.967	4.422	15.303	11.655	26.193	7.289	8.440	34.615	30.160	6.491	16.028	64.759	81.955	4.578
1272445	5.790	7.118	15.303	11.516	4.423	6.374	3.001	3.423	13.584	11.444	2.462	5.781	24.234	32.693	1.805
1078000	4.419	4.914	11.655	4.423	3.650	7.874	2.169	2.475	10.390	8.731	1.763	4.550	17.948	23.008	1.311
1077400	10.936	10.957	26.193	6.374	7.874	46.464	5.316	5.416	24.228	21.667	4.251	11.456	43.229	58.077	3.530
1077250	2.755	3.197	7.289	3.001	2.169	5.316	1.523	1.583	6.218	5.444	1.111	2.774	11.101	14.420	0.852
1077090	3.084	3.754	8.440	3.423	2.475	5.416	1.583	1.823	7.305	6.057	1.229	3.216	12.648	16.449	0.944
1076420	13.329	15.114	34.615	13.584	10.390	24.228	6.218	7.305	31.922	26.589	5.207	14.024	54.707	70.799	3.962
1075830	12.625	13.099	30.160	11.444	8.731	21.667	5.444	6.057	26.589	24.020	4.826	12.026	48.601	61.939	3.448
1073480	2.906	2.969	6.491	2.462	1.763	4.251	1.111	1.229	5.207	4.826	1.090	2.374	10.188	12.731	0.693
1073448	6.312	7.017	16.028	5.781	4.550	11.456	2.774	3.216	14.024	12.026	2.374	6.661	25.094	32.312	1.830
1073436	27.345	29.009	64.759	24.234	17.948	43.229	11.101	12.648	54.707	48.601	10.188	25.094	102.273	129.401	7.143
1073408	33.478	36.261	81.955	32.693	23.008	58.077	14.420	16.449	70.799	61.939	12.731	32.312	129.401	166.990	9.329
1073080	1.792	1.966	4.578	1.805	1.311	3.530	0.852	0.944	3.962	3.498	0.693	1.830	7.143	9.329	0.552

TABLE 6. Covariance matrix for stations in target and control. Units: 10^5 $(mi^2 \times inches)^2$.

Control stations CSU numbers	Target stations CSU numbers 1375400	1371520	1277200	1272445	1078000	1077400	1077250	1077090	1076420	1075830	1073480	1073448	1073436	1073408	1073080
1590000	29.752	27.377	60.232	23.048	15.536	36.364	11.186	10.718	39.887	41.943	10.911	18.476	90.521	108.529	6.062
1371810	0.662	0.770	1.717	0.597	0.469	1.102	0.308	0.346	1.368	1.198	0.249	0.628	2.503	3.198	0.174
1373055	14.183	15.332	33.154	11.316	8.643	21.610	5.849	6.122	23.566	22.122	5.218	10.776	47.483	58.292	3.230
1373360	1.422	1.731	3.505	1.218	0.851	2.034	0.611	0.642	2.321	2.176	0.529	1.070	4.728	5.881	0.328
1377200	7.493	7.986	16.571	6.540	4.246	10.803	2.896	3.075	12.003	11.426	2.724	5.510	24.584	30.935	1.723
1377280	2.195	2.374	4.860	1.969	1.237	3.177	0.870	0.912	3.421	3.252	0.795	1.607	7.128	9.050	0.516
1378100	8.504	7.585	18.195	7.231	5.000	12.835	3.530	3.381	12.872	13.297	3.197	6.001	27.862	34.160	1.954
1379000	4.162	3.487	8.563	3.466	2.409	5.969	1.687	1.581	5.906	6.417	1.585	2.747	13.151	16.078	0.906
1425625	2.295	2.786	5.453	1.914	1.342	2.537	0.836	0.994	3.956	3.474	0.810	1.827	7.708	9.445	0.485
1370300	1.196	1.561	3.085	1.037	0.766	1.936	0.482	0.591	2.421	1.991	0.440	1.083	4.396	5.615	0.300
1740000	8.411	6.998	15.281	7.122	3.939	11.798	2.883	2.644	9.361	10.221	3.139	4.303	23.928	28.885	1.608

TABLE 7. Covariance matrix for stations in control. Units: 10^5 (mi^2 × inches)^2.

Control stations CSU numbers	Control stations CSU numbers										
	1590000	1371810	1373055	1373360	1377200	1377280	1378100	1379000	1425625	1370300	1740000
1590000	165.185	2.539	68.019	6.473	32.503	10.034	46.842	23.754	8.448	3.796	53.764
1371810	2.539	0.848	1.423	0.140	0.640	0.191	0.790	0.370	0.228	0.124	0.588
1373055	68.019	1.423	33.821	3.368	15.045	4.540	20.004	9.668	4.724	2.316	19.689
1373360	6.473	0.140	3.368	0.393	1.583	0.479	1.865	0.905	0.501	0.263	1.795
1377200	32.503	0.640	15.045	1.583	8.411	2.555	9.307	4.826	2.258	1.137	10.269
1377280	10.034	0.191	4.540	0.479	2.555	0.811	2.871	1.482	0.664	0.333	3.329
1378100	46.842	0.790	20.004	1.865	9.307	2.871	14.250	7.015	2.372	1.074	15.239
1379000	23.754	0.370	9.668	0.905	4.826	1.482	7.015	3.827	1.148	0.461	7.730
1425625	8.448	0.228	4.724	0.501	2.258	0.664	2.372	1.148	0.856	0.420	2.108
1370300	3.796	0.124	2.316	0.263	1.137	0.333	1.074	0.461	0.420	0.270	1.003
1740000	53.764	0.588	19.689	1.795	10.269	3.329	15.239	7.730	2.108	1.003	22.824

will be presented in a separate paper (Saheli and Morel-Seytoux, 1973). Because the problem is highly nonlinear several local minima exist and it is necessary to initiate the algorithm at different initial feasible points.

6. Results for the Colorado River Basin Pilot Project

Tables 5–7 show the calculated covariance matrices based on 30 years of available records for the period 1939–68 inclusive. Records were corrected for regulation and diversion when needed. The needed $\overline{\Delta Q_i}$ values have already been given in Table 4. Three sets of calculations were performed for the two possible situations (i.e., the entire target area is seeded and the control area is the one shown in Fig. 2, or only zone 1 or zones 1 and 2 of the target area are seeded and zones 3 and 4 are used as controls (see Fig. 1).

a. Entire target area is seeded

Tables 8 and 9 summarize two combinations which are essentially as good. The improvement over the use of weights all equal to 1 is drastic. Fig. 10 shows the regression line between the total target runoff (stations

TABLE 8. Optimal combination of stations in the target (San Juan Mountains) and the control (Maroon Peak).

CSU number	Station name	Location (Basin)	Number of existing years of record within (1939–68)	Gage elevation (ft)	Drainage area (mi^2)	Weight factor X
		Stations in target				
1371520	Uncompahgre River at (near) Colona	Uncompahgre	30	6318.80	437.0	5.142
1077250	Rio Blanco near Pagosa Springs	Rio Blanco	30	7950.00	58.0	2.585
1073480	Animas River at Howardsville	Animas	30	9616.00	55.9	17.249
1073408	Animas River near Cedarhill	Animas	30	5960.00	1090.0	1.217
		Stations in control				
1590000	Roaring Fork at Glenwood Springs	Crystal	30	5720.70	1460.0	0.298
1373055	North Fork Gunnison River near Crawford	North Fork Gunnison	30	6038.60	521.0	−2.608
1373360	Smith Fork near Crawford	North Fork Gunnison	30	7200.00	42.3	−13.927
1377200	Tomichi Creek at Sargents	Tomichi Creek	30	5720.70	1020.0	11.920
1378100	East River at Almont	East River	30	8008.29	295.0	9.628
1379000	Taylor River below Taylor Park Reservoir	Taylor	30	9168.67	254.0	−3.597
1425625	Buzzard Creek near Collbran	Plateau Creek	30	6920.00	139.0	4.068
1370300	Kannah Creek near White River	Kannah River	30		61.9	62.538
1740000	Blue River below Green Mountain Reservoir	Blue River	30	7510.00	623.0	−4.056

Minimum number of years with optimal X's is $3.8 \approx 4$ years.
Minimum number of years with all X's$=1$ is 99 years.

TABLE 9. Optimal combinations of stations in the target (San Juan Mountains) and the control (Maroon Peak).

CSU number	Station name	Location (Basins)	Number of existing years of record within (1939–68)	Gage elevation (ft)	Drainage area (mi²)	Weight factor X
		Stations in target				
1078000	East Fork San Juan River above Sand Creek near Pagosa Springs	San Juan	30	7597.63	86.9	1.0
1073448	Hermosa Creek near Hermosa	Animas	29	6705.0	172.0	1.0
		Stations in control				
1377200	Tomichi Creek at Sargents	Tomichi Creek	30	8420.0	155.0	9.755
1590000	Roaring Fork at Glenwood Springs	Roaring Fork	30	5720.7	1460.0	−0.110

Number of years with optimal weights in the combination is $3.8 \approx 4$ years.
Number of years with all X's$=1$ is 33 years.

listed in Table 8) and total control runoff, i.e., for the case where all the X_i are *a priori* given the value 1. Fig. 11 shows the corresponding regression line when optimal weights are used.

Comparison of the two figures explains the reason for the efficiency of the technique.

b. Only zone 1 is seeded (zones 3 and 4 are used as controls)

Table 10 summarizes the results for this case. It is somewhat disappointing that in this case the minimum number of years is 6 which exceeds the planned 4–5 years duration of the project. Note that an 11% increase in precipitation (rather than 10%) would suffice to return the power of the test to 50% for the 5-year period.

c. Zones 1 and 2 are both seeded (zones 3 and 4 are used as controls)

Table 11 summarizes the results for this case. In this instance for the 5-year period the power of the test is slightly better than 50%. For this case the optimal X's were rounded off to simpler values to see if the calculated number of years would change appreciably. For the set of X's; 1.6, 0.1, 2.0, −0.3, 0.7, 0.9, 4.5 (compare with the last column of Table 11), the calculated value was again 4.9, indicating that this number of years is not overly sensitive to small changes in the X's around the optimal point.

7. Conclusions

From a theoretical point of view this paper shows the value of grouping stations both in the target and the control areas in an optimized manner to maximize the power of the detection test for a given significance level. From a practical point of view, the study shows that this new technique of testing yields a power for the test over a 5-year period of experimentation which is slightly better than 50% if only zones 1 and 2 in the target are seeded. It is better than 50% if the entire target area were seeded (which it is not, currently).

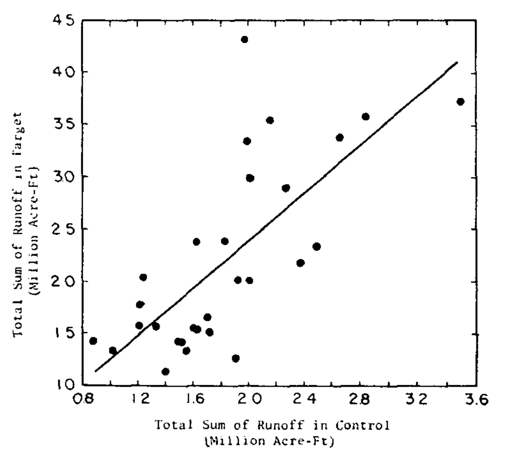

FIG. 10. Regression line between total target runoff and total control runoff for stations listed in Table 8.

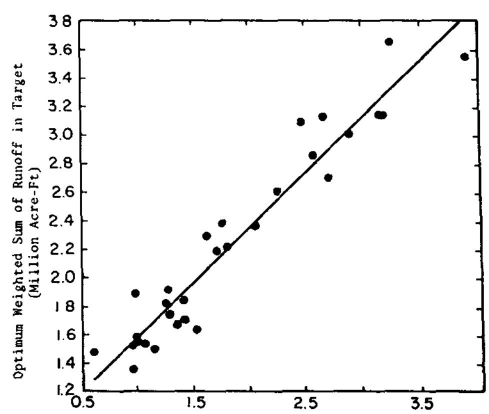

FIG. 11. Regression line between optimal grouped target runoff and optimal grouped control runoff for stations listed in Table 8.

TABLE 10. Optimal combination of stations in the target (San Juan Mountains zone 1) and the control (San Juan zones 3 and 4).

CSU number	Station name	Location (Basin)	Number of existing years of record within (1939–68)	Gage elevation (ft)	Drainage area (mi²)	Weight factor X
		Stations in target (area 1 in Fig. 1)				
1076920	Piedra River near Piedra	Piedra	30	6530.00	371.00	1.0
1075830	Los Pinos River near Bayfield	Los Pinos	30	6501.51	692.00	1.0
		Stations in control (areas 3 and 4 in Fig. 1)				
1371520	Uncompahgre River at (near) Colona	Uncompahgre	30	6318.8	437.00	0.810
1272445	San Miguel River near Placerville	San Miguel	26	7096.44	308.00	0.705
1277200	Dolores River at Dolores	Dolores	30	6924.91	556.00	1.161
1073448	Hermosa Creek near Hermosa	Animas	29	6705.08	172.00	4.543

Minimum number of years with optimal X's is 6 years.
Minimum number of years with all X's$=1$ is 13 years.

The percentage increase in precipitation was conservatively assumed in the calculations of this paper to be 10%. Thus, one can say with some confidence and optimism that the chance of significant evaluation (based on runoff) for the pilot project is at least 50%. That probability is most likely to be greater as it increases rapidly as the percentage increase in precipitation exceeds 10%. For example, for the case summarized in Table 9, the power of the test which is essentially 50% for a 4-year period increases to 84% for the same period if the precipitation increase is 15%. On the other hand, if the precipitation increase is only 10% as assumed in this paper, and because the test procedure was optimized, one must unfortunately add that this is the best that can be done under the circumstances.

Acknowledgments. This work was supported by the Bureau of Reclamation, Division of Atmospheric Water Resources Management, under Contract 14-06-D-6597, in connection with the SKYWATER program. The help of the Bureau is gratefully acknowledged.

The optimization technique was supported by funds provided by the U. S. Department of Interior, Office of Water Resources Research, as authorized under the Water Resources Research Act of 1964; and pursuant to Grant Agreement 14-31-0001-3506 (OWRR Project A-011-Colorado).

TABLE 11. Optimal combination of stations in the target (San Juan Mountains zones 1 and 2) and the control (San Juan Mountains zones 3 and 4).

CSU number	Station name	Location (Basin)	Number of existing years of record within (1939–68)	Gage elevation (ft)	Drainage area (mi²)	Weight factor X
		Target stations areas 1 and 2				
1076420	Piedra River near Piedra	Piedra	30	6530.00	371	1.613
1077400	San Juan River at Pagosa Springs	San Juan	30	7052.04	295.00	0.086
1078000	East Fork San Juan River above Sand Creek near Pagosa Springs	San Juan	30	7597.63	86.9	2.010
		Control stations areas 3 and 4				
1371520	Uncompahgre River at (near) Colona	Uncompahgre	30	6318.18	437	−0.313
1272445	San Miguel River near Placerville	San Miguel	26	7096.44	308	0.689
1277200	Dolores River at Dolores	Dolores	30	6924.9	556	0.895
1073448	Hermosa Creek near Hermosa	Animas	29	6705.08	172	4.543

Number of years with optimal X's$=4.9\approx5$ years.
Number of years with all X's$=1$ is 22 years.

REFERENCES

Brownlee, K. A., 1961: *Statistical Theory and Methodology in Science and Engineering*. New York, Wiley, 570 pp.

Bureau of Reclamation, 1966: Plan to develop technology for increasing water yield from atmospheric resources. Denver, Dept. of Interior, Division of Atmospheric Water Resources Management.

——, 1968: Colorado River Basin Pilot Project—Cloud seeding research in the San Juan Mountains. Denver, Dept. of Interior, Division of Atmospheric Water Resources Management.

Cochran, W. G., and G. M. Cox, 1957: *Experimental Designs*. New York, Wiley, 611 pp.

Gilman, D. L., J. R. Hibbs, and P. L. Laskin, 1965: Weather and climate modification. Washington, D. C., U. S. Weather Bureau, 41 pp.

Golze, A. R., 1961: *Reclamation in the United States*. Caldwell, Idaho, Caxton Printers, 270 pp.

Grant, L., 1969; Weather Modification: An operational adaptation program for the Colorado River Basin. Fort Collins, Colorado State University, Interim Report to the Bureau of Reclamation, July 1968–June 1969, 98 pp.

Graybill, F. A., 1961: *Introduction to Linear Statistical Models*. New York, McGraw–Hill, 500 pp.

Hurley, P. A., 1968: Augmenting Colorado River by weather modification. *J. Irrig. Drain. Div., ASCE*, **94**, 363–380.

Markovic, R. D., 1966: Statistical evaluation of weather modification attainments. Hydrology Paper No. 16, Fort Collins, Colorado State University, 44 pp.

Marlatt, W., and H. Riehl, 1963: Precipitation regimes over the upper Colorado River. *J. Geophys. Res.*, **68**, 6447–6458.

Morel-Seytoux, H. J., 1968; Suitability of basins to weather modification and statistical evaluation of attainment. Fort Collins, Colorado State University, Final Report to Bureau of Reclamation, 1966 and 1967, Part I, 65 pp.

Nakamichi, H., and H. J. Morel-Seytoux, 1971: Runoff grouping for detection of change in runoff. *J. Irrig. Drain. Div., ASCE*, **97**, 275–290.

National Academy of Sciences, 1966: *Weather and Climate Modification*. Washington, D. C., National Research Council.

——, 1968: *Water and Choice in the Colorado Basin*. Washington, D. C., Report by Committee on Water of the National Research Council, Publ. 1869, 107 pp.

Office of Saline Water, 1963: *Saline Water Conversion*. Washington, D. C., Govt. Printing Office.

Saheli, F., and H. J. Morel-Seytoux, 1973: An optimization technique for detection of a hydrologic change. To be submitted to *J. Water Resources Res.*

Smith, L. W., 1968: Western States water augmentation concept. Denver, Federation of Rocky Mountain States Inc., 37 pp.

Wilde, D. J., and C. S. Beightler, 1967: *Foundations of Optimization*. New York, Prentice-Hall, 480 pp.

Yevjevich, V., 1972: *Probability and Statistics in Hydrology*. Fort Collins, Water Resources Publ., 302 pp.