REGIONAL DISCRIMINATION OF CHANGE IN RUNOFF

by

Viboon Nimmannit and Hubert J. Morel-Seytoux

November 1969

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

37

REGIONAL DISCRIMINATION OF CHANGE IN RUNOFF

by
Viboon Nimmannit
and
Hubert J. Morel-Seytoux

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80521

November 1969

No. 37

ACKNOWLEDGMENTS

The present paper is based primarily upon Mr. Viboon Nimmannit's Ph.D. dissertation in the Department of Civil Engineering, Colorado State University: "Multivariate Analysis of Hydrologic Changes." The work was supported by the U.S. Bureau of Reclamation, Office of Atmospheric Water Resources, Contract numbered BR 14-06-D-6597, whose help is gratefully acknowledged.

RELATION OF HYDROLOGY PAPER NO. 37 TO RESEARCH PROGRAM: "HYDROLOGY OF WEATHER MODIFICATION"

The present study is part of a more comprehensive project which has as one of its objectives the development of methods of evaluation of atmospheric water resources programs. Correlatively the application of the methods to a variety of basins forms a basis for selection of suitable watersheds, basins or regions.

Several approaches were pursued. This report discusses one of them. Several other approaches were previously described in Hydrology Papers 22, 34, and 36 (see back inside cover for complete reference).

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ABSTRACT

The object of this study is to find answers to the following questions:

What is the appropriate statistical test for a regional target-control technique of evaluation?

What is a suitable method for reduction of an originally large number of variables?

Which of the Upper Basin of the Colorado River or the San Juan Mountains is a more suitable area of operations, if the effectiveness of precipitation management is to be detected as quickly as possible?

The results of this research study show:

1. The T^2-test is the appropriate test for multiple target-control technique of evaluation.

2. The canonical analysis is the suitable method for the reduction of a large number of original variables.

3. The Upper Basin of the Colorado River is preferable under the assumption of an equal percentage of increase in runoff. However, if the percentage increase in the southern area is at least 1.2 times as large as in the northern area (and recent publications suggest that this ratio is probably around 3) then the southern area is preferable.

Based on the T^2-test, the minimum number of years for detecting an increase of 10 percent in spring runoff means are three years in the Upper Basin of the Colorado River, and four years in the San Juan Mountains.

REGIONAL DISCRIMINATION OF CHANGE IN RUNOFF

by

Viboon Nimmannit* and Hubert J. Morel-Seytoux**

Chapter I

INTRODUCTION

1.1 Motivation of study. As interference with nature is accelerating [1,2,3] there is a need for early detection of direct or side effects of man's actions. Because of the rapid pace of development [3, 4,5] it is important to develop techniques that will display the effect of any given practice on water resources availability and distribution at the earliest possible time. For large scale field research, the availability of an efficient and regionally representative test would reduce the duration of experiments required to attain conclusive results and therefore costs, and provide a basis for managerial decision at an earlier stage, without additional observations. The decision may be to stop a project earlier when it becomes apparent, based on real time analysis of data, that the objectives cannot be achieved in the planned time. Better, pre-experiment data simulation would permit to assess the chances of being in that unfortunate situation as a function of a range of values of the suspected or hoped for change. Useful charts can be drawn in terms of the parameters, (magnitude of change, basin characteristics, etc.) for first stage planning.

The techniques which are described in this paper could be used for detection of the effects of watershed management of any origin upon water supply. They could be used to determine the effect of urbanization on the local hydrology, to detect when such urbanization has created a significant change that calls for reappraisal of the protective designs, e.g., flood control, etc. In other words, they are quite general. To a certain degree the techniques will indeed be discussed in a general abstract form, but their practical applicability will be demonstrated with a very special and very important application in mind.

The Bureau of Reclamation will most probably initiate in the fall of 1970 a pilot project of massive cloud seeding operations, covering some 4000 square miles within the state of Colorado. It will be the primary purpose of this paper to establish as accurately as possible how long it will take to detect a regional hydrologic change and to attribute it with little risk of error to the cloud seeding operations. To understand this practical illustration of the technique some knowledge of the geographic and hydrologic features of the region, of the water situation and of the plans of the Bureau of Reclamation is a prerequisite. The purpose of the following sections is to provide this background information.

1.2 Geographic and hydrologic setting. The Colorado River begins high in the snow-capped Rocky Mountains of north central Colorado, flows nearly 1,400 miles southwest, and empties into the Gulf of California in Mexico far to the south. It drains a vast area of 244,000 square miles, 242,000 square miles in the United States--one-twelfth of the area of Continental United States--and 2,000 square miles in northern Mexico. The basin from Wyoming to below the Mexican border is some 900 miles long and varies in width from about 300 miles in the upper section to 500 miles in the lower section. It is bounded on the north and east by the Continental Divide in the Rocky Mountains, on the west by the Wasatch Range, and on the southwest by the San Jacinto Mountains, a range of the Sierra Nevada Mountains. The area, larger than the states of New York, Pennsylvania, and New Jersey combined, above Lee Ferry, Arizona, is known as the Upper Colorado River Basin (Fig. 1). This area is the source of the greatest part of water reaching the Colorado River. The upper portion of this basin in Wyoming and Colorado is a mountainous plateau, 5,000 to 8,000 feet in altitude, marked by broad rolling valleys, deep canyons, and intersecting mountain ranges. Climatologically, the Colorado River Basin has heavy precipitation on the high peaks of the Rockies and truly desert conditions with little rain in the southern area around Yuma, Arizona. Extremes of temperatures in the basin range from 50° below zero to 130° above zero degree Fahrenheit. Development and utilization of resources in this arid land depend on the availability of water. Crops must be irrigated; cattle on the vast ranges must be partially fed from hay produced on irrigated land; towns and cities must be located within distance of dependable domestic and municipal water supplies, and mining and many other industries depend, to an extent, on the availability of hydroelectric power [1].

1.3 The water resources outlook. The U.S. Geological Survey estimates total water demand in the United States was 280 billion gallons per day (314 million acre-feet per year) in 1960. As a point of comparison let us note that the average annual flow of the biggest river in the United States, the Mississippi, is 440 maf and that of the Upper Colorado is about 14 maf. The U.S.G.S. estimates the total water demand for the U.S. will be 600 billion gallon per day (672 million acre-feet per year) by 1980. In 1960 the demand in the Western States alone was estimated at 125 billion gallons per day (140 million

*Ph.D. graduate of Colorado State University, Department of Civil Engineering, Fort Collins, Colorado, presently with Engineering Consultants Inc., Denver, Colorado.

**Associate Professor of Civil Engineering, Colorado State University, Fort Collins, Colorado.

Fig. 1 The Upper Colorado River Basin (after Upper Colorado River Commission)

acre-feet per year) and for 1980 at 190 billion gallons per day (213 million acre-feet per year). The lower percentage of demand growth for the Western States reflects different demands of industry in the East and agriculture in the West. Because rainfall is low in the Western States, the conservation use must be greater than in the East and Midwest. Municipal or domestic use has first priority in the West, with irrigation second. It is estimated the 44,000,000 population of the Western States in 1960 will expand to more than 100,000,000 by the year 2000 [2].

From the population figures given above, it is obvious much more water will be needed in the near future. So, the question one must answer is, "What can be used as sources for additional water to alleviate the shortages?" Several agencies, such as, the Bureau of Reclamation [3], the Upper Colorado River Commission [4], and the Committee on Water of the National Research Council [5], feel cloud seeding, to augment the precipitation amount in the Upper Colorado River Basin, may become a partial solution to the recurrent water shortage.

1.4 Precipitation management operations and plans. An important experimental cloud seeding operation is being conducted near Climax, Colorado, by Colorado State University under sponsorship of the National Science Foundation. These experiments are designed to show quantitative change in precipitation by cloud seeding and to determine criteria for optimum seeding conditions.

The most favorable conditions for cloud seeding are in regions where moist winds blow more or less constantly up the slopes of the mountains. Cloud seeding involves artificial introduction of tiny particles into clouds so that moisture can depose around each of the nuclei to form a crystal heavy enough to fall to the ground. Among nuclei that have been used experimentally in cloud seeding operations are solid carbon dioxide, silver iodide, water spray, and carbon black. To date, the greatest number of cloud-seeding attempts have been made by using silver iodide generators operated on the ground. However, seeding operations using aircraft flown directly over cloud layers have demonstrated that this technique may be more effective [6].

In 1968, the Bureau of Reclamation adopted a plan to start pilot programs for weather modification operations in the Upper Colorado River Basin (Fig. 1), and two regions were selected for this purpose [7]. The first was the Upper Basin of the Colorado River*, which will for brevity be referred to in this study as the Northern Project area (Fig. 2). The second area was the San Juan Mountains region referred to as the Southern Project area (Fig. 3). Since the initiation of this study, the plans of the Bureau were modified. Currently [8] only one area is considered: the Southern area. Nevertheless, because they had already been calculated, the results for the Northern area are also reported.

1.5 Objective of study and approach. The primary objective was to develop an appropriate and efficient methodology that can be used to demonstrate the effectiveness of cloud seeding in each project region. In order to achieve this, a multivariate analysis of geographically well distributed stations in each region is carried out. These stations are referred to as targets. Variables used in this study are spring runoffs. The spring runoff of a station is defined here as the average flow, in cubic feet per second, of that station during the spring months. Because this flow is substantially contributed by winter snow, it can be regarded as an indirect measure of the effect of weather modification. However, because of the lack of a precise date for the start of snow melting, two different time intervals will be used for spring months. The first interval will be composed of four months: April, May, June and July; the second of six months: March, April, May, June, July and August.

Because the use of controls, which are the stations free from the effect of weather modification, is a well proven means of making tests more effective, (9), it also will be utilized in this study. An area between the Northern and Southern Project areas has been selected (Fig. 4) to serve as the control area.

*The reader is warned for possible confusion. In this paper the expression "Upper Colorado River Basin" refers to the Colorado Basin above Lee's Ferry. On the other hand, the expression "Upper Basin of the Colorado River" refers to a much smaller drainage basin including the main stem of the Colorado close to its source and a few tributaries. The limits of that basin are shown on Fig. 2.

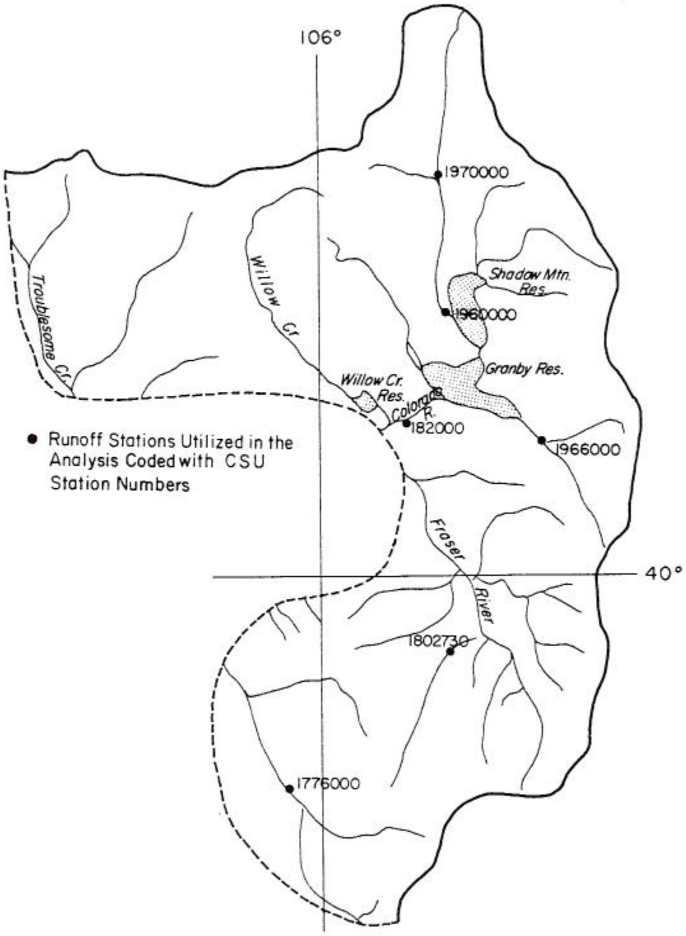


Fig. 2 General configuration of and location of gages within the Upper Basin of the Colorado River

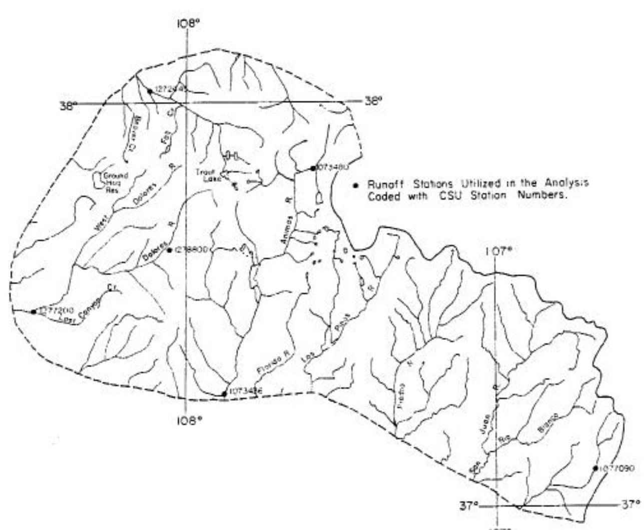


Fig. 3 General configuration of and location of gages within the Colorado River Basin Pilot Project area

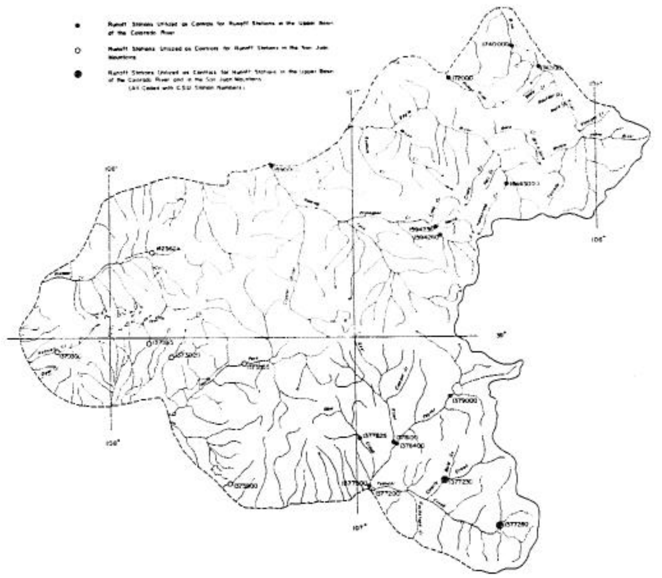


Fig. 4 General configuration of and location of gages within the Colorado River Basin Pilot Project control area

For brevity, the following symbols will be employed:

N-4:	4-month runoff series in the northern target region,
N-6:	6-month runoff series in the northern target region,
CN-4:	4-month runoff series in the northern control region,
CN-6:	6-month runoff series in the northern control region,
S-4:	4-month runoff series in the southern target region,
S-6:	6-month runoff series in the southern target region,
CS-4:	4-month runoff series in the southern control region,
CS-6:	6-month runoff series in the southern control region.
N-CN-4:	the combination of N-4 and CN-4,
N-CN-6:	the combination of N-6 and CN-6,
S-CS-4:	the combination of S-4 and CS-4,
S-CS-6:	the combination of S-6 and CS-6.

In applying theories of statistics to an engineering problem, it is necessary to assume certain properties of the variables. The assumptions made in this study are:

a) The observations of runoff follow a multivariate normal distribution.

b) The estimated means in both target and control areas from the period before seeding are essentially equal to the population values.

c) After seeding the means in the target areas will change but the means in the control areas will remain unchanged.

d) The covariance matrix of the target and control variables is the same for both periods before and after seeding.

The above assumptions are required in this study because of the difficulty in developing the theoretical distribution of the test criterion otherwise. In dealing with more than two variables, the knowledge of distributions, except that of the normal distribution, are not sufficiently developed [10]. So, even though it is rather obvious the assumptions made here will be violated to some degree in reality, they are practically as good as one can make with the present state of statistical knowledge.

From the work of Ref. [9], it is found that the χ^2-test which is based on the population values, and the conditional Student's t-test which is based on the sample values, give very closely the same results for sample sizes around 30. Thus, for convenience in handling the mathematics, the population values are assumed to be known here and this assumption appears justified. Also, all the observations of runoff station used in this study have been plotted on normal probability paper. If the runoff were exactly distributed as a normal variate, all the observations would fall exactly on a straight line. The actual observations did not in any case deviate appreciably from a straight line. The assumption of normality may therefore be entertained for these data.

Based on the above assumptions, a T^2-statistic is obtained [11,12]. The minimum number of years, N^*, to detect the expected increase can be obtained [11] from the formula,

$$N^* = \frac{\tau^2}{\underline{\mu}'\underline{V}^{-1}\underline{\mu}} , \qquad (1)$$

where τ^2 is the noncentrality parameter (it is a measure of the amount of deviation from being central which is the case when the variables under study have means zero),

$\underline{\mu} = \underline{\mu}^* - \underline{\mu}_o$, $\underline{\mu}^*$ is the runoff mean vector for the seeded period, and $\underline{\mu}_o$ is the runoff mean vector for the non-seeded period,

$\underline{\mu}'$ is the transpose of $\underline{\mu}$, and

$\underline{V}^{-1}$ is the inverse of the covariance matrix of runoff variables, $\underline{V}$.

In Chapter II, most approaches used to detect the effectiveness of weather modification by other investigators are summarized. The theoretical concepts of the principal component analysis, the canonical analysis, and the T^2-statistic are the main subjects of Chapter III. Chapters IV and V deal with data assembly, analysis of data, and results.

The study led to two major conclusions, one of general theoretical interest and the second of practical significance for the plans of the Bureau:

a) Canonical analysis coupled with the multivariate T^2-test provides an effective technique of detection of a suspected regional hydrologic change and,

b) Assuming a 10% uniform increase in runoff by precipitation management 3 and 4 years only are required for significant evaluation for the Upper Basin of the Colorado and the San Juan Mountains, respectively.

Chapter II

REVIEW OF PREVIOUSLY USED TESTS

The statistical content of this chapter is not new. The material here is provided for the sake of convenience to a reader whose statistical background is that of the average engineer. A statistician can bypass this chapter without detrimental effect to the continuity and understanding of this paper.

In this chapter the statistical tests, which have been employed by other investigators for detecting the effectiveness of weather modification, will be presented. The literature is further discussed in Ref. 12. Because all tests are concerned with the expected increase in the means of either runoff or precipitation during the seeded period, the hypotheses for all tests can be stated as:

H_o (null hypothesis) - there is no increase in the mean of the hydrologic variable during the seeded period,

H_a (alternate hypothesis) - there is an increase in the mean.

2.1 Target sample u-test. Let $q_{11}, q_{12}, \ldots, q_{1n_1}$, be n_1 observations of a hydrologic variable for the nonseeded period, and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_2 observations for the seeded period of a target watershed. When n_1 is large the mean and variance of the series $q_{11}, q_{12}, \ldots, q_{1n_1}$ can be considered to be the population mean and population variance. Assuming the variance of the seeded period is the same as the non-seeded period, the test statistic is [13]

$$u_o = \frac{\overline{q}_2 - \mu_1}{\sigma_1/\sqrt{n_2}} \quad ,$$

where u_o is normally distributed with mean o and variance 1

$$\overline{q}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} q_{2i}$$

$$\mu_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} q_{1i}$$

$$\sigma_1^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (q_{1i} - \mu_1)^2 \quad .$$

The null hypothesis, H_o , will be accepted at a 5% level of significance if u_o has a value less than 1.645. That is, there is no increase in the mean. On the contrary, if u_o is greater than 1.645 the alternative hypothesis, H_a , will be accepted at a 5% level of significance. The use of this test can be found in References [9] and [14]. South Fork San Joaquin, California, was the target basin for the study in Reference [9]. There were 15 years of seeded record, and 29 years of non-seeded record. The apparent percentage increase in the mean of the seasonal runoff for the seeded period was about 10%. By the use of the target sample u-test it was found that $u_o = 1.20$. This shows that the target sample u-test was not powerful enough to detect the increase in mean value in the order of 10% of the old mean.

2.2 Target two-sample t-test. This test does not require knowledge of population parameters. Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_1 and n_2 observations for the non-seeded and seeded periods of a target watershed.

Assuming the variances of the non-seeded and seeded periods are equal, the test statistic [15]

$$t_o = \frac{\overline{q}_2 - \overline{q}_1}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

is distributed as t-distribution with $n_1 + n_2 - 2$ degrees of freedom, where:

$$\overline{q}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} q_{1i}$$

$$\overline{q}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} q_{2i}$$

$$s^2 = \frac{\sum_{i=1}^{n_1} (q_{1i} - \overline{q}_1)^2 + \sum_{i=1}^{n_2} (q_{2i} - \overline{q}_2)^2}{(n_1 - 1) + (n_2 - 1)} \quad .$$

The use of this test can be found in References [8], [14], [16], [17], [18], [19], [20], [21], [22], [23], and [24]. The value of the t-statistic was also computed for South Fork San Joaquin [9] from the same set of data used in computing the target sample u_o . The computed t-statistic has the value of 0.89. So, again no significant increase was concluded. The target two-sample t-test, and the target sample u-test therefore can be considered to be insufficiently powerful tests for studies of this nature.

2.3 Target-control χ^2-test. The detectability of the test can be improved by the use of a control [9]. This can be done by comparing sets of hydrologic data of non-seeded and seeded periods for the target watershed with those for an unseeded control watershed located in the vicinity of the target area.

Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q'_{11}, q'_{12}, \ldots, q'_{1n_1}$ be n_1 observations for the period prior to seeding of the target and control watersheds respectively. Also, let n_2 observations for the seeded period in the target be denoted by $q_{21}, q_{22}, \ldots, q_{2n_2}$, and those in the control by $q'_{21}, q'_{22}, \ldots, q'_{2n_2}$.

When the length of record before seeding is long enough, the estimated statistics of the target and control can be assumed to be the population values. Assuming the variables in the target and control are bivariate normally distributed, then the test statistic [14]:

$$\chi_o^2 = \frac{n_2}{1-\rho^2}\left\{\left(\frac{\bar{q}_2-\mu_1}{\sigma}\right)^2 - 2\rho\,\frac{(\bar{q}_2-\mu_1)(\bar{q}'_2-\mu'_1)}{\sigma\,\sigma'} + \left(\frac{\bar{q}'_2-\mu'_1}{\sigma'}\right)^2\right\}$$

is distributed as Chi-square distribution with two degrees of freedom, where

ρ is the population coefficient of correlation between the target and control for the non-seeded period, given by

$$\rho = \frac{\sum_{i=1}^{n_1}(q_{1i}-\mu_1)(q'_{1i}-\mu'_1)}{\left[\sum_{i=1}^{n_1}(q_{1i}-\mu_1)^2\sum_{i=1}^{n_1}(q'_{1i}-\mu'_1)^2\right]^{1/2}}$$

$$\mu_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} q_{1i}$$

$$\mu'_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} q'_{1i}$$

$$\bar{q}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} q_{2i}$$

$$\bar{q}'_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} q'_{2i}$$

$$\sigma = \sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}(q_{1i}-\mu_1)^2}$$

$$\sigma' = \sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}(q'_{1i}-\mu'_1)^2}\ .$$

This test has been used in References [9] and [14].

With the use of Merced River at Pohono Bridge as a control runoff station for the target, South Fork San Joaquin, the observed χ_o^2-statistic was found to be [9] 22.2. The value of χ^2 for significance at 99% level of confidence is 9.2. Therefore, a significant increase was detected by the use of the target-control χ^2-test. This shows that for the same set of data for the target basin, the target-control χ^2-test is overwhelmingly more discriminating than the target two-sample t-test and the target two-sample u-test.

2.4 Target-control conditional Student's t-test. In this test population parameters are not known. What is tested is the normality or abnormality of the target, given the behavior of the control, normal or otherwise [9].

Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be the n_1 and n_2 observations of a hydrologic variable in the target watershed before and during seeded periods respectively. Let $q'_{11}, q'_{12}, \ldots, q'_{1n_1}$ and $q'_{21}, q'_{22}, \ldots, q'_{2n_2}$ be the corresponding observations in the control watershed.

By application of the maximum-likelihood ratio method [25], the test statistic:

$$t_o = \frac{\dfrac{\sqrt{n_1+n_2-3}\left[(\bar{q}_2-\bar{q}_1)-(\bar{q}'_2-\bar{q}'_1)\left\{\sum_{i=1}^{n_1}a_i(\Delta q_{1i})+\sum_{i=1}^{n_2}b_i(\Delta q_{2i})\right\}\right]}{\left[\dfrac{1}{n_1}+\dfrac{1}{n_2}+\left(\dfrac{\bar{q}'_2-\bar{q}'_1}{\Delta}\right)^2\right]^{1/2}}}{\left[\sum_{i=1}^{n_1}(\Delta q_{1i})^2+\sum_{i=1}^{n_2}(\Delta q_{2i})^2-\left\{\sum_{i=1}^{n_1}a_i(\Delta q_{1i})+\sum_{i=1}^{n_2}b_i(\Delta q_{2i})\right\}^2\right]^{1/2}}$$

is obtained and it is distributed as Student's t-distribution with $n_1 + n_2 - 3$ degrees of freedom, where

$$\bar{q}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} q_{1i}$$

$$\bar{q}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} q_{2i}$$

$$\bar{q}'_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} q'_{1i}$$

$$\bar{q}'_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} q'_{2i}$$

$$(\Delta q_{1i}) = q_{1i} - \bar{q}_1$$

$$(\Delta q_{2i}) = q_{2i} - \bar{q}_2$$

$$(\Delta q'_{1i}) = q'_{1i} - \bar{q}'_1$$

$$(\Delta q'_{2i}) = q'_{2i} - \bar{q}'_2$$

$$\Delta^2 = \sum_{i=1}^{n_1}(\Delta q'_{1i})^2 + \sum_{i=1}^{n_2}(\Delta q'_{2i})^2$$

$$a_i = \frac{(\Delta q'_{1i})}{\Delta}$$

$$b_i = \frac{(\Delta q'_{2i})}{\Delta} \quad .$$

The use of this test can be found in References [9] and [14].

In Reference [9], the application of the target-control conditional Student's t-test was made for the target, South Fork San Joaquin, and the control, Merced River at Pohono Bridge. The observed t_o-statistic by this method was 3.80. The value of t for significance at 99% was 2.71. Therefore, a significant increase was the result of this test. Comparison of the results of the above mentioned statistic tests show that the target-control χ^2-test and the target-control conditional Student's t-test are better tests than the target two-sample t-test and the target sample u-test. Also note that for runoff data from high elevation watersheds the outcomes of the two tests are essentially the same for a sample size around 30. However, it should be noted that all these tests are applicable only when single target or single target-control technique is used. None of these tests can be applied without modification when the number of variables in the study is greater than two, which is the usual case.

2.5 Rank test. Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_1 and n_2 observations of a hydrologic variable for the non-seeded and seeded periods respectively.

Arrange the observations in a common sequence of increasing magnitude,

$$q_{11}, q_{12}, q_{21}, q_{22}, q_{13}, q_{14}, q_{15}, q_{23}, q_{16}, \ldots \ .$$

Assign ranks from 1 to n, where $n = n_1 + n_2$, to the above sequence so that rank 1 is given to the smallest observation and n to the largest.

The test statistic is now [26]:

$$Z = \frac{T_s - \overline{T}}{\sigma} \quad ,$$

where Z is approximately a standard normal variate,
T_s is the sum of ranks for seeded observations,
$\overline{T}$ is the expected mean value of T_s, given by

$$\overline{T} = \frac{n_2(n_2 + n_1 + 1)}{2} = \frac{n_2(n+1)}{2} \quad ,$$

and

$$\sigma = \sqrt{\frac{n_2 n_1 (n+1)}{12}} \quad .$$

If Z is greater than 1.645, then, one rejects the null hypothesis and concludes that at the 5% level of significance weather modification was effective.

This test has been used in References [27] and [28]. From the data in the Necaxa Watershed, Mexico, it was found that [27] the value of Z was 2.64, which is a value significant beyond the 99% level. The numbers of observations were 45 seeded days and 29 unseeded days. However, the apparent increase in the mean of the seeded period here was large. The seeded mean was about 26 percent larger than the unseeded mean. So, the use of rank test in Reference [27] does not tell much about the efficiency of the test at all. In fact, with the amount of increase of this order, one can find with any statistical test that the cloud seeding is effective. For example, when the u-test is applied the approximate number of observations needed to detect the 26 percent increase in the mean is obtained from:

$$N^* = \frac{4\sigma^2}{h^2 \mu^2}$$

where N^* is the approximate number of observations required to detect a certain amount of increase in the mean,
σ^2 is the variance of the hydrologic variable for the unseeded period,
μ is the mean of the hydrologic variable for the unseeded period, and
h is the fractional increase in mean.

Upon substituting the values of σ^2, μ, h from the data of Reference [27], it was found that

$$N^* = \frac{4 \times 600.17}{(.26)^2 (88.14)^2} \approx 5 \quad .$$

Thus, it is clear that the required number of observations to detect a 26 percent increase in the mean is much smaller than 45 which is the actual number of observations. So, with this large amount of increase any statistical test will always give the positive result.

2.6 Median test. The median of a distribution is that value which divides the distribution halfway, i.e., half the distribution have lower and half have higher values. The median test determines primarily if the medians of the populations from which the samples come are well separated or not.

Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_1 and n_2 observations of a hydrologic variable for the non-seeded and seeded periods respectively. Arrange the observations in a common sequence of increasing magnitude, e.g.,

$$q_{11}, q_{12}, q_{21}, q_{22}, q_{23}, q_{13}, q_{14}, q_{15}, q_{16}, q_{24}. \ \ldots \ .$$

If the total number of observations is even, the median is taken to be halfway between the two middle observations. If this total number is odd, the median observation is removed since it does not contribute any information to the question of whether the distribution of that sample has its median above or below the joint sample median. The case then reduces to the even case.

Let the numbers of q_{1i}'s above and below the median of the common sequence be n_{1a} and n_{1b}, and the numbers of q_{2j}'s above and below the same common sample median be n_{2a} and n_{2b}. Under the null hypothesis that the two samples come from identical distributions, the proportion of each sample lying below any point should be the same.

If the test function [29]

$$M = (|2n_{1a}-(n_{1a}+n_{1b})|-1)^2/n_1+(|2n_{2a}-(n_{2a}+n_{2b})|-1)^2/n_2$$

is greater than $\chi^2_{0.95}$ with one degree of freedom, then, one rejects, at the 95% level, the hypothesis that the samples have the same median.

This test has been used in Reference [20]. The data used in Reference [20] were obtained from an experiment on artificial stimulation of rain in three climatologically similar regions, Delhi, Agra and Jaipur in northwest India. The net increase in rainfall obtained over all three regions was 41.9%. Thus, it was found that there was a highly significant increase in the amount of rainfall. The observations were made from 1957 to 1965 (excluding 1962) in Delhi, from 1960 to 1965 in Agra, and from 1960 to 1963 in Jaipur. There was, however, no observed statistic given in this report.

2.7 The Mann-Whitney U test. Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_1 and n_2 observations of a hydrologic variable for the non-seeded and seeded periods respectively. Arrange the observations in a common sequence of increasing magnitude, e.g.,

$$q_{11}, q_{12}, q_{13}, q_{21}, q_{14}, q_{15}, q_{22}, q_{23}, q_{24}, \cdots .$$

The statistic U is defined as the number of times a q_{2j} precedes a q_{1i}. This test was used to test the null hypothesis

H_o - the q_{1i} and q_{2j} values have the same distribution against the alternative hypothesis,

H_a - the location parameter of q_{2j} is larger than the location parameter of q_{1i}, i.e., the bulk of the distribution of q_{2j}'s is to the right of the bulk of the distribution of q_{1i}'s.

If H_a is true, one expects U to be small. Mann and Whitney [30] computed tables that give probabilities associated with small (lower tail) values of U, and Auble [31] gives tables of critical values of U for significant levels of 0.001, 0.01, 0.025, and 0.05 for a one-sided test. For the one-sided alternative hypothesis that the location parameter of q_{2j} is smaller than the parameter of q_{1i}, one computes the statistic U', defined to be the number of times a q_{1i} precedes a q_{2j}, and uses Aubles's tables to test H_o.

The relationship between U and the sum of ranks for seeded observations, T_s, in the rank test can be expressed as (Wine [32]):

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - T_s .$$

The U statistic is usually computed by the above equation, since it is tedious to compute from the definition of U when n_1 and n_2 become fairly large.

The test statistic is

$$W = \frac{U-\overline{U}}{\sigma} ,$$

where W is approximately a standard normal variate, $\overline{U}$ is the expected value of U, given by

$$\overline{U} = \frac{n_1 n_2}{2} ,$$

and $$\sigma = \sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}} .$$

If W is greater than 1.65, then the null hypothesis is rejected and one can conclude the location of q_{2j} is larger than that of q_{1i}. This test has been used by many authors - [20], [21], [28], [33], and [34].

In Reference [21], the data used were collected from a five-year period experiment (1960 through 1964) in Missouri. On comparing the average rainfall (inches/hour) of the seeded days with that of the non-seeded days, it was found there was, on the average, a decrease of 67.9%. The values of W ranged from smaller than 0.01 to 0.88. Thus, it was concluded that no evidence of increases in precipitation because of cloud seeding was achieved.

2.8 Run test. Let $q_{11}, q_{12}, \ldots, q_{1n_1}$ and $q_{21}, q_{22}, \ldots, q_{2n_2}$ be n_1 and n_2 observations of a hydrologic variable for the non-seeded and seeded periods respectively.

Arrange the observations in a common sequence of increasing magnitude, e.g.,

$$q_{11}, q_{12}, q_{21}, q_{13}, q_{14}, q_{22}, q_{23}, \cdots .$$

A run is defined as an unbroken sequence of elements of the same type, i.e., a sequence of q_{1i}'s or a sequence of q_{2j}'s. Let the number of runs be denoted by η. If two samples are from the same population, the non-seeded and seeded observations will be well mixed and the number of runs, η, will be large.

The test statistic is now [14]

$$U = \frac{\eta - \overline{\eta}}{\sigma} ,$$

where U is a standard normal variate, $\overline{\eta}$ is the expected value of η, given by

$$\bar{\eta} = \frac{2n_1 n_2}{n_1 + n_2}$$

$$\sigma = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1+n_2)^2 (n_1+n_2-1)}} \quad .$$

If U is greater than 1.65, then the null hypothesis is rejected and the alternative hypothesis is accepted. This test has been used in Reference [35].

In Reference [35], the data of the King River at Piedra, California was analyzed. The observations were the annual flows from 1917 to 1954 for the non-seeded period, and 1955 to 1966 for the seeded period There was a decrease of about 3.3% in mean annual flows for the seeded period. The number of runs, η, was found to be 17, $\bar{\eta} = 19.240$, and $\sigma = 2.533$. From the above values, U was obtained as -0.88. Therefore, no significant increase in the mean annual flow was concluded.

Of all the tests stated above, it is found that none of them can be applied for testing the increase in runoff means when the number of runoff variables is greater than two. In the evaluation of weather modification effectiveness based on a multiple target-control concept the number of runoff variables involved is large. So, it is necessary to find an approach to detect the increase in means of these runoff variables.

In Chapter III, the principal components, canonical analysis, and the T^2-statistic are discussed.

Chapter III

PRINCIPAL, CANONICAL COMPONENTS AND THE T^2-STATISTIC

For small scale operations the method of evaluation of a significant change in hydrologic characteristics based on the single target-control concept is adequate. For large regions this procedure would not be very representative. Besides if the test were performed for many pairs of target and control it is not clear how one should treat the ensemble of the outcomes. On the other hand, there is no problem of interpretation when a single test is performed even though the tested statistic may itself be a complicated combination of many observations from many targets and controls. For representativity the station runoff variables should be geographically well distributed over the large area of interest. This results in a selection of a large number of variables that are usually not independent variables. Sometimes the number of variables involved may be so large that any study can hardly be made economically. In fact, this is one of the difficulties in this study since there are three big areas under investigation. It is, therefore, also an object of this study to find a suitable method for reducing the number of variables involved in the analysis.

There are several ways to reduce the number of variables. However, two methods are used here before the statistical test is carried out. One is the principal components analysis, the other the canonical analysis.

3.1 Principal component analysis. The principal components are linear combinations of random variables, which have special properties in terms of variances. Usually, the linear combination with the maximum variance is referred to as the first principal component; the second component is the one that is uncorrelated with the first and has the second largest variance, and so on. The idea of this analysis was discussed thoroughly by Hotelling [36] in 1933.

From the hydrologic point of view, these principal components can be considered as new transformed runoff variables though lacking simple physical meaning. These transformed variables have, in total, the same amount of fluctuation or variation as do the original runoff variables. But the number of the transformed variables can be smaller than that of the original variables. Also these transformed variables are independent while the original variables are not.

A priori what can be expected from the principal components analysis for the purpose of evaluation? Suppose the principal components analysis is carried for all the targets and all the controls. The first principal component for each group will be the most statistically representative single combination of targets and controls, respectively, because that combination will account for the largest fraction of the total variation. If the percentage is high (say 95%) all the other principal components can be dropped. Then the originally multivariate test reduces again to a familiar single target control t-test, even though the target variable and the control variable are each a combination of many target and control ones. The procedure will be simple and effective if the target first principal component and the control one are highly correlated. However, this need not happen because the targets and controls are treated separately and the procedure does not attempt to maximize the correlation between the two components (which canonical analysis does). It can be concluded that principal components analysis can provide the basis for a simple and highly representative test but it will not be, by far, a minimal time evaluation one. (The procedure for the actual computation of the principal components is summarized in Chapter V, Section 1).

3.2 The canonical analysis. Canonical analysis is a technique to maximize the correlations between two groups of random variables. This analysis gives new sets of transformed variables as linear combinations of the original runoff variables. The first linear combination of each group will have the highest correlation, and each is uncorrelated with the other linear combinations in its group. The second linear combinations will have the second highest correlation, the third linear combinations will have the third highest correlation and so on. These linear combinations are referred to as canonical variables or components.

In this study the first group is the group of runoff stations in the target region and the second group is made of stations in the control region. This analysis is particularly advantageous for evaluation purposes. The canonical analysis yields a smaller number of variables for the final test, and most importantly it also guarantees high correlations between the variables of the target and control regions.

3.3 Computation of canonical variables. The steps for computing the canonical variables are now described:

Step 1) Compute the covariance matrix, $\hat{\underline{V}}$, of the runoff variables of the two sets (target and control). For p_1 runoff stations in the target region and p_2 in the control region, then

$$\hat{\underline{V}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots\cdots & \sigma_{1p_1} & \sigma_{1(p_1+1)} \cdots\cdots & \sigma_{1(p_1+p_2)} \\ \sigma_{21} & \sigma_{22} \cdots\cdots & \sigma_{2p_1} & \sigma_{2(p_1+1)} \cdots\cdots & \sigma_{2(p_1+p_2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{p_1 1} & \sigma_{p_1 2} & \sigma_{p_1 p_1} & \sigma_{p_1(p_1+1)} \cdots\cdots & \sigma_{p_1(p_1+p_2)} \\ \sigma_{(p_1+1)1} & \sigma_{(p_1+1)2} \cdots & \sigma_{(p_1+1)p_1} & \sigma_{(p_1+1)(p_1+1)} \cdots & \sigma_{(p_1+1)(p_1+p_2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{(p_1+p_2)1} & \sigma_{(p_1+p_2)2} \cdots & \sigma_{(p_1+p_2)p_1} & \sigma_{(p_1+p_2)(p_1+1)} \cdots & \sigma_{(p_1+p_2)(p_1+p_2)} \end{bmatrix} \quad (2)$$

The subscripts of σ are the ordering numbers of the stations. The numbers 1 to p_1 are for the p_1 stations in the target region. The numbers p_1+1 to p_1+p_2 are for the p_2 stations in the control region. For example, the subscript 1 will refer to the first station in the target region, while the subscript p_1+1 will refer to the first station in the control region and the subscript p_1+2 the second station in the control region, etc.

σ_{ii} is the variance of the runoff series for station i , defined as,

$$\sigma_{ii} = \frac{1}{N} \sum_{s=1}^{N} (q_{is}-\overline{q}_i)^2 \quad , \tag{3}$$

where N is the number of years of recorded runoff data,

q_{is} is the s^{th} recorded runoff of station i, and

$\overline{q}_i$ is the mean of the recorded runoff of station i .

σ_{ij} is the covariance of stations i and j , defined as,

$$\sigma_{ij} = \frac{1}{N} \sum_{s=1}^{N} (q_{is}- \overline{q}_i)(q_{js}- \overline{q}_j) \tag{4}$$

$$\sigma_{ij} = \sigma_{ji} \quad .$$

Step 2) Partition the covariance matrix, $\underline{\hat{V}}$, such that,

$$\underline{\hat{V}} = \begin{bmatrix} \underline{\hat{V}}_{11} & \underline{\hat{V}}_{12} \\ \underline{\hat{V}}_{21} & \underline{\hat{V}}_{22} \end{bmatrix} \quad , \tag{5}$$

where $\underline{\hat{V}}_{11}$ is a $p_1 \times p_1$ matrix,

$$\underline{\hat{V}}_{11} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p_1} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p_1} \\ \vdots & \vdots & & \\ \sigma_{p_1 1} & \sigma_{p_1 2} & \cdots & \sigma_{p_1 p_1} \end{bmatrix} \tag{6}$$

$$\underline{\hat{V}}_{12} = \begin{bmatrix} \sigma_{1(p_1+1)} & \sigma_{1(p_1+2)} & \cdots & \sigma_{1(p_1+p_2)} \\ \sigma_{2(p_1+1)} & \sigma_{2(p_1+2)} & \cdots & \sigma_{2(p_1+p_2)} \\ \vdots & \vdots & & \vdots \\ \sigma_{p_1(p_1+1)} & \sigma_{p_1(p_1+2)} & \cdots & \sigma_{p_1(p_1+p_2)} \end{bmatrix} \tag{7}$$

$$\underline{\hat{V}}'_{12} = \underline{\hat{V}}_{21} \tag{8}$$

$$\underline{\hat{V}}_{22} = \begin{bmatrix} \sigma_{(p_1+1)(p_1+1)} & \sigma_{(p_1+1)(p_1+2)} & \cdots & \sigma_{(p_1+1)(p_1+p_2)} \\ \sigma_{(p_1+2)(p_1+1)} & \sigma_{(p_1+2)(p_1+2)} & \cdots & \sigma_{(p_1+2)(p_1+p_2)} \\ \vdots & \vdots & & \vdots \\ \sigma_{(p_1+p_2)(p_1+1)} & \sigma_{(p_1+p_2)(p_1+2)} & \cdots & \sigma_{(p_1+p_2)(p_1+p_2)} \end{bmatrix} \tag{9}$$

Step 3) Obtain the values of canonical correlations by solving the system,

$$\begin{vmatrix} -\theta\underline{\hat{V}}_{11} & \underline{\hat{V}}_{12} \\ \underline{\hat{V}}_{21} & -\theta\underline{\hat{V}}_{22} \end{vmatrix} = 0 \quad . \tag{10}$$

The values of θ are the canonical correlations.

Step 4) Let $\underline{\alpha}$ and $\underline{\gamma}$ be the column vectors of coefficients for the canonical variables of the target and control regions respectively. Then, for a given value θ_i , the vectors $\underline{\alpha}_i$ and $\underline{\gamma}_i$ can be obtained by solving the system,

$$\begin{bmatrix} -\theta_i\underline{\hat{V}}_{11} & \underline{\hat{V}}_{12} \\ \underline{\hat{V}}_{21} & -\theta_i\underline{\hat{V}}_{22} \end{bmatrix} \begin{bmatrix} \underline{\alpha}_i \\ \underline{\gamma}_i \end{bmatrix} = \underline{0} \tag{11}$$

subject to the standardization conditions:

$$\underline{\alpha}'_i \, \underline{\hat{V}}_{11} \, \underline{\alpha}_i = 1 \tag{12}$$

$$\underline{\gamma}'_i \, \underline{\hat{V}}_{22} \, \underline{\gamma}_i = 1 \quad ; \tag{13}$$

$\underline{\alpha}'_i$ and $\underline{\gamma}'_i$ are the transposes of $\underline{\alpha}_i$ and $\underline{\gamma}_i$ respectively.

Once the $\underline{\alpha}_i$ and $\underline{\gamma}_i$ are obtained, the canonical variables for the target region are obtained from the relations:

$$\zeta_i = \underline{\alpha}'_i \, \underline{Q}_1 \tag{14}$$

where ζ_i is the i^{th} canonical variable in the target region

$$\underline{\alpha}'_i = (\alpha_{i1}, \alpha_{i2}, \ldots , \alpha_{ip_1}) \tag{15}$$

$$\underline{Q}_1 = \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_{p_1} \end{bmatrix} \tag{16}$$

$Q_1, Q_2, \ldots, Q_{p_1}$ are runoff variables in the target region.

Similarly, ε_i is the i^{th} canonical variable in the control region defined by the relation:

$$\varepsilon_i = \underline{\gamma}_i' \, \underline{Q}_2 \ , \qquad (17)$$

where $$\underline{\gamma}_i' = (\gamma_{i(p_1+1)} \gamma_{i(p_1+2)} \cdots \gamma_{i(p_1+p_2)}) \qquad (18)$$

$$\underline{Q}_2 = \begin{bmatrix} Q_{p_1+1} \\ Q_{p_1+2} \\ \vdots \\ Q_{p_1+p_2} \end{bmatrix} \qquad (19)$$

$Q_{p_1+1}, Q_{p_1+2}, \ldots, Q_{p_1+p_2}$ are runoff variables in the control region.

3.4 <u>The minimum number of years for detecting an increase in runoff means</u>. In the previous sections two techniques to transform the original runoff variables were described and in the case of canonical analysis even the basic steps of the procedure were described. However, the multivariate T^2 test applies just as well for the set of original variables. The principal and canonical transformations will either simplify some of the calculations or improve the outcome of the test. Again, the transformations are not necessary to apply the test. Nevertheless in this study the test was only performed for the transformed variables.

Assuming the values of the population mean vector $\underline{\mu}^*$ and covariance matrix $\underline{V}$ for the seeded period are known, the minimum number of observations, N^*, that one needs in order to be able to reject the hypothesis $\underline{\mu}^* = \underline{\mu}_o$, where $\underline{\mu}_o$ is a given vector, is given by

$$N^* = \frac{\tau^2}{(\underline{\mu}^*-\underline{\mu}_o)' \, \underline{V}^{-1} \, (\underline{\mu}^*-\underline{\mu}_o)} \ , \qquad (20)$$

where τ^2 is the noncentrality parameter with degrees of freedom k and $N-k$,
k is the total number of runoff variables, and
N is the number of observations for the non-seeded period.

Select values of τ^2 as given by Tang [37] and Lehmer [38] are shown for convenience in Table 1.

TABLE 1 - VALUE OF τ^2

Level of significance, $\alpha = 0.05$; power $\beta = 0.50$

Degrees of freedom		
k	N-k	τ^2
2	28	5.468
4	26	7.640
5	25	8.640
6	24	9.646
8	22	11.655

In this study the value of $\underline{\mu}_o$ is assumed to be the mean vector of target and control runoff variables for the period before seeding. $\underline{\mu}^*$ is similar to $\underline{\mu}_o$ except that the means of the target runoff variables are 1.1 times greater than those in $\underline{\mu}_o$. In other words, it is assumed in this study that the effect of precipitation management over the target areas will be to increase the runoff uniformly throughout the target areas by 10%. The covariance matrix $\underline{V}$ is assumed to be the same as that of the nonseeded period.

When the principal components (or the canonical variables) are used for computing N^*, then $\underline{\mu}^*$ and $\underline{\mu}_o$ are the mean vectors of the principal components (or the canonical variables) for the seeded and non-seeded periods respectively, and $\underline{V}$ is the covariance matrix of the principal components (or the canonical variables) for the non-seeded period. The original runoff variables can also be used in computing N^*. However, because of the large number of the original runoff variables, they are not used in this study.

It should be noted here that the use of principal components in equation (20) will yield approximately the same results as the use of the original runoff variables. This is due to the fact that the amount of variation accounted for by the principal components is practically the same as the variation of the original runoff variables. Thus, the principal component analysis will merely reduce the number of original variables, but will not improve the final outcome of the test.

However, if the number of variables can be reduced to one component then the principal component analysis will be very useful because one can apply a bivariate test, such as the conditional Student's t-test which is less restrictive in its assumptions than the T^2-test. Unfortunately, this usefulness will not be known until one has completed the analysis.

In the next chapter the collection of data in the Upper Basin of the Colorado River, the San Juan Mountains area, and the Maroon Peak and Grand Mesa region is discussed.

Chapter IV

RESEARCH DATA ASSEMBLY

The data used in this study are the records of the runoff from three regions in the Colorado River Basin. These are:

1. The Upper Basin of the Colorado River,
2. The San Juan Mountains area,
3. The Maroon Peak and Grand Mesa region.

The first two areas were originally [7] proposed as sites for extensive cloud seeding operation. They are called northern and southern target regions (Figs. 2 and 3), while the third is called the control region (Fig. 4). Currently [8] only one area is considered: the southern area. The selection of the control stations is done primarily on the basis of the high correlations with those in the target regions.

It is virgin flow, which is the flow free from any man-made intervention, that is necessary for this study. So, corrections must be made for the records of runoff. The records of runoff were obtained from U.S. Geological Survey Water Supply Papers. However, only the corrections due to transmountain, transbasin diversions, and regulation can be made. The diversion for irrigation cannot be made because there is no record for the amount of water diverted for this purpose. Thus, it is assumed after making the corrections above, that virgin flows are obtained.

Out of a large number of stations, seven stations are chosen for the final analysis in the northern target region, and six stations in the southern region. There are fourteen stations used as controls for the northern region, and nine stations as controls for the southern region. These stations and their descriptions are listed in Table 2. The correlations for these stations computed from all the corresponding actually available records are shown in Tables 3, 4, 5 and 6. There are two stations used as controls for both the northern and southern regions.

TABLE 2 - DESCRIPTION OF STATIONS

Types	Seq. No.	CSU Sta. No.	USGS Sta. No.	Names	Lat. (o	'	")	Long. (o	'	")	Area (Sq. Mi.)	Elevation (ft.)
Target-stations in Northern Project	1	1970000	9.0105	Colorado River below Baker Gulch, near Grand Lake, Colorado.	40	19	33	105	51	22	53	8750
	2	1960000	9.0110	Colorado River near Grand Lake, Colo.	40	13	08	105	51	25	103	8380
	3	1866000	9.0165	Arapaho Creek at Monarch Lake outlet, Colo.	40	06	45	105	44	57	47.1	8310
	4	1830000	9.0190	Colorado River below Lake Granby, Colo.	40	08	39	105	52	00	311	8050
	5	1820000	9.0195	Colorado River near Granby, Colo.	40	07	15	105	54	00	322	7960
	6	1802730	9.0265	St. Louis Creek near Fraser, Colo.	39	54	30	105	52	45	32.8	8980
	7	1776000	9.0360	Williams Fork near Leal, Colo.	39	49	55	106	03	20	89.5	8790
Control-stations for Northern Project	1	1742100	9.0535	Blue River above Green Mountain Reservoir, Colo.	39	49	55	106	13	20	514	7947
	2	1740000	9.0575	Blue River below Green Mountain Reservoir, Colo.	39	52	50	106	20	00	599	7683
	3	1720000	9.0595	Piney River near State Bridge, Colo.	39	48	00	106	35	00	82.6	7272
	4	1666300	9.0645	Homestake Creek near Red Cliff, Colo.	39	28	25	106	22	00	58.9	8783
	5	1594260	9.0780	Fryingpan River at Norrie, Colo.	39	19	50	106	39	30	89.5	8410
	6	1594236	9.0785	North Fork Fryingpan River near Norrie, Colo.	39	20	40	106	39	50	41.2	8400
	7	1590000	9.0850	Roaring Fork River at Glenwood Springs, Colo.	39	32	50	107	19	50	1460	5721
	8	1379000	9.1090	Taylor River below Taylor Park Reservoir, Colo.	38	48	50	106	36	40	254	9170
	9	1378400	9.1100	Taylor River at Almont, Colo.	38	40	00	106	51	00	477	8011
	10	1378100	9.1125	East River at Almont, Colo.	38	40	00	106	51	00	295	8006
	11	1377825	9.1135	Ohio Creek near Baldwin, Colo.	38	42	00	107	00	00	124	8180
	12	1377500	9.1145	Gunnison River near Gunnison, Colo.	38	32	50	106	57	00	1010	7670
	13	1377280	9.1155	Tomichi Creek at Sargents, Colo.	38	24	00	106	25	00	155	8420
	14	1377230	9.1180	Quartz Creek near Ohio City, Colo.	38	33	35	106	38	10	106	8430
Target-stations in Southern Project	1	1278800	9.1650	Dolores River below Rico, Colo.	37	38	20	108	03	35	105	8422
	2	1278050	9.1665	Dolores River at Dolores, Colo.	37	28	00	108	30	00	556	6919
	3	1272445	9.1725	San Miguel River near Placerville, Colo.	38	02	05	108	07	15	308	7056
	4	1077090	9.3440	Navajo River at Banded Peak Ranch, near Chromo, Colo.	37	05	07	106	41	20	69.8	7941
	5	1073480	9.3575	Animas River at Howardsville, Colo.	37	50	00	107	36	00	55.9	9617
	6	1073436	9.3615	Animas River at Durango, Colo.	37	16	45	107	52	47	692	6502
Control-stations for Southern Project	1	1425625	9.0975	Buzzard Creek near Collbran, Colo.	39	16	20	107	51	00	139	6955
	2	1377280	9.1155	Tomichi Creek at Sargents, Colo.	38	24	00	106	25	00	155	8420
	3	1377230	9.1180	Quartz Creek near Ohio City, Colo.	38	33	35	106	38	10	106	8430
	4	1377200	9.1190	Tomichi Creek at Gunnison, Colo.	38	31	20	106	56	25	1020	7629
	5	1373900	9.1275	Crystal Creek near Maher, Colo.	38	33	05	107	30	20	42.2	8070
	6	1373055	9.1325	North Fork Gunnison River near Somerset, Colo.	38	55	45	107	26	55	521	6039
	7	1373020	9.1345	Leroux Creek near Cedaredge, Colo.	38	55	35	107	47	35	35.1	7160
	8	1371815	9.1430	Surface Creek near Cedaredge, Colo.	38	59	00	107	51	00	26.7	8180
	9	1370300	9.1520	Kannah Creek near Whitewater, Colo.	38	59	00	108	14	00	61.9	—

TABLE 3 - CORRELATION MATRIX BETWEEN N-4 AND CN-4 (as computed from all available data)

			N-4						
		CSU STA. No.	1970000	1960000	1866000	1830000	1820000	1802730	1776000
	CSU STA. No.	USGS STA. No.	9.0105	9.0105	9.0110	9.0165	9.0190	9.0265	9.0360
CN-4	1742100	9.0535	.8625	.8365	.8375	.8234	.6475	.7779	.9342
	1740000	9.0575	.6055	.7277	.6970	.7077	.4634	.8427	.8357
	1720000	9.0595	.9164	.9003	.8322	.8476	.7171	.6076	.9470
	1666300	9.0645	.6781	.7548	.8147	.8304	.6023	.6515	.8033
	1594260	9.0780	.8952	.8514	.8854	.8919	.7218	.6618	.9219
	1594236	9.0785	.8608	.8567	.9187	.9089	.7975	.6291	.8647
	1590000	9.0850	.8723	.8776	.8382	.8770	.7701	.6381	.8717
	1379000	9.0190	.8364	.8174	.7846	.8541	.6999	.4699	.9080
	1378400	9.1100	.8474	.8434	.7942	.8473	.7329	.5012	.7744
	1378100	9.1125	.8635	.8151	.7971	.8301	.6581	.6456	.7896
	1377825	9.1135	.8741	.6554	.5844	.7306	.5222	.6190	.7672
	1377500	9.1145	.8714	.8338	.7996	.8434	.6851	.5337	.8012
	1377280	9.1155	.8026	.6197	.7937	.8009	.6634	.5672	.7082
	1377230	9.1180	.8644	.6436	.7113	.7675	.6274	.5090	.7536

TABLE 4 - CORRELATION MATRIX BETWEEN N-6 AND CN-6 (as computed from all available data)

			N-6						
		CSU STA. No.	1970000	1960000	1866000	1830000	1820000	1802730	1776000
	CSU STA. No.	USGS STA. No.	9.0105	9.0110	9.0165	9.0190	9.0195	9.0265	9.0360
CN-6	1742100	9.0535	.6648	.9578	.9124	.9243	.9937	.5921	.9155
	1740000	9.0575	.7233	.4359	.8704	.8640	.9038	.4789	.8409
	1720000	9.0595	.6944	.3230	.5386	.6146	.3806	.3611	.3126
	1666300	9.0645	.7348	.6093	.5738	.5336	.3514	.3718	.6702
	1594260	9.0780	.4923	.5923	.6371	.7567	.8406	.5299	.7359
	1594236	9.0785	.7153	.5548	.8017	.6039	.2842	.3247	.4008
	1590000	9.0850	.7877	.3323	.6076	.6012	.7616	.5203	.5695
	1379000	9.0190	.6576	.5072	.6912	.2488	.4766	.5284	.7748
	1378400	9.1100	.7135	.4373	.5538	.7051	.3055	.5282	.7908
	1378100	9.1125	.3529	.4010	.7701	.1470	.4582	.4277	.4510
	1377825	9.1135	.6645	.2784	.6396	.7202	.0136	.7097	.4439
	1377500	9.1145	.7503	.4247	.5132	.5766	.5478	.4155	.8483
	1377280	9.1155	.7354	.5034	.6163	.5225	.4351	.4590	.8899
	1377230	9.1180	.8004	.5576	.1133	.6640	.6830	.6961	.4868

TABLE 5 - CORRELATION MATRIX BETWEEN S-4 AND CS-4 (as computed from all available data)

			S-4					
		CSU STA. NO.	1278800	1278050	1272445	1077090	1073480	1073436
	CSU STA. NO.	USGS STA. NO.	9.1650	9.1665	9.1725	9.3440	9.3440	9.3615
CS-4	1425625	9.0975	.9004	.8519	.8872	.7978	.8466	.8258
	1377280	9.1155	.9020	.7565	.8040	.7529	.8295	.7353
	1377230	9.1180	.9108	.7289	.5841	.6336	.7553	.6964
	1377200	9.1190	.9865	.8587	.8428	.7859	.8895	.8423
	1373900	9.1275	.8879	.8710	.9059	.7988	.8578	.8549
	1373055	9.1325	.8900	.8599	.7981	.7835	.8582	.8216
	1373020	9.1345	.8335	.8608	.7064	.8226	.8118	.8069
	1371815	9.1430	.8961	.8993	.8021	.8490	.2168	.4315
	1370300	9.1520	.9299	.8058	.8909	.8576	.8276	.7837

TABLE 6 - CORRELATION MATRIX BETWEEN S-6 AND CS-6 (as computed from all available data)

			S-6					
		CSU STA. NO.	1278800	1278050	1272445	1077090	1073480	1073436
	CSU STA. NO.	USGS STA. NO.	9.1650	9.1665	9.1725	9.3440	9.3575	9.3615
CS-6	1425625	9.0975	.8217	.6427	.7111	.6302	.8128	.3267
	1377280	9.1155	.9310	.9100	.7033	.3573	.9009	.9126
	1377230	9.1180	.8008	.8536	.7309	.7921	.7754	.8297
	1377200	9.1190	.8864	.7361	.8601	.9381	.9729	.7719
	1373900	9.1275	.9406	.7605	.8115	.7576	.7964	.7121
	1373055	9.1325	.9368	.7148	.8990	.8423	.8881	.9498
	1373020	9.1345	.8947	.6922	.8129	.5556	.8410	.6934
	1371815	9.1430	.8844	.8071	.7831	.7217	.7546	.7877
	1370300	9.1520	.7872	.7429	.6865	.7834	.8869	.7467

The major part of the spring runoff will occur because of the melting of the winter snow, which is subject to the effect of seeding during winter time. So, it is reasonable to consider whatever changes in the value of the spring runoff as an indirect indicator of the effect of cloud seeding. This is equivalent to saying a larger amount of snowfall in winter will produce a larger amount of runoff in spring. Because of the uncertainty of the start of snow melting, both the runoff during the four months of April, May, June and July, and during the six months of March, April, May, June, July and August are used. These four-month runoff and six-month runoff periods are treated separately in this analysis.

The number of years of record for all stations is fixed at 30, starting from 1938 up to 1967. To assure that these stations are still in operation, the selection has been made in such a way that only stations that have records available for 1967 are considered. It is not likely that the operation of these stations will be discontinued in the near future.

The characteristics of the data used in this study are shown in Tables 7, 8, 9, 10, 11, 12, 13, and 14. There are some data missing in the runoff record of the stations selected but they are filled in by the regression method [39] with the random component superimposed. These stations with missing data are shown in Table 15. Also shown in Table 15 are the stations used in evaluating the missing data. Graphical representations of the data used are shown in Figs. 5, 6, 7, 8, 9, 10, 11, and 12 according to the regions. The means and standard deviations computed from the year 1938 up to 1967 data are shown in Table 16; and the correlations between N-4 and CN-4, N-6 and CN-6, S-4 and CS-4, and S-6 and CS-6 are shown in Tables 17, 18, 19, and 20, respectively.

In Chapter V, the analysis of the data and the results are presented.

TABLE 7 - N-4 SERIES (CFS)

CSU USGS	1970000 9.0105	1960000 9.0110	Station Numbers 1866000 9.0165	1830000 9.0190	1820000 9.0195	1802730 9.0265	1776000 9.0360
Year							
1938	180.35	366.27	244.61	1175.06	1136.39	102.10	379.77
1939	216.18	245.75	129.61	754.37	763.81	72.41	247.57
1940	139.53	193.02	168.37	686.57	673.55	53.67	175.50
1941	163.66	251.27	318.69	994.25	805.08	70.50	224.29
1942	216.65	219.05	233.59	1192.95	850.85	83.66	282.53
1943	180.00	230.67	145.06	585.11	807.90	93.99	245.73
1944	191.33	224.75	170.59	661.96	728.79	73.77	218.40
1945	217.40	251.49	196.92	1069.87	837.71	72.80	233.42
1946	195.31	179.05	201.58	662.49	641.57	73.00	226.37
1947	200.72	317.40	246.16	1556.30	1035.63	98.51	317.49
1948	57.75	198.77	161.58	1023.12	682.34	74.92	231.42
1949	302.87	306.21	244.28	1069.03	1037.22	84.06	268.34
1950	170.28	180.74	155.98	923.67	246.07	72.07	227.35
1951	252.28	305.83	245.08	1000.46	992.64	107.37	301.82
1952	244.13	352.67	285.75	1188.94	1188.71	107.45	334.81
1953	235.87	207.48	186.89	740.89	737.48	75.53	225.97
1954	108.66	112.48	129.00	529.51	486.97	34.57	108.27
1955	145.96	150.63	171.09	749.74	438.18	52.04	163.22
1956	206.35	249.24	215.13	1046.18	874.92	42.58	219.86
1957	274.13	368.69	320.64	1625.95	1252.69	84.12	306.48
1958	226.14	249.54	220.52	929.26	995.11	80.32	247.84
1959	171.07	206.87	204.50	853.75	578.10	25.05	201.92
1960	212.75	280.40	217.77	1122.92	1384.95	43.85	252.19
1961	181.57	202.94	171.29	836.40	823.74	23.22	178.74
1962	268.49	373.30	229.44	1211.50	1108.41	48.16	313.24
1963	150.16	151.40	158.76	655.72	636.17	21.13	100.42
1964	195.19	198.73	177.40	826.43	810.51	22.48	177.83
1965	271.56	311.11	264.29	1193.47	1180.51	39.82	275.53
1966	141.71	125.82	99.35	288.08	272.07	20.50	142.67
1967	235.42	243.04	193.79	798.78	788.61	36.67	211.29

TABLE 8 - N-6 SERIES (CFS)

CSU USGS	1970000 9.0105	1960000 9.0110	Station Numbers 1866000 9.0165	1830000 9.0190	1820000 9.0195	1802730 9.0265	1776000 9.0360
Year							
1938	126.75	258.12	171.72	806.46	857.32	75.53	269.28
1939	152.42	172.51	91.96	499.87	539.45	53.70	176.43
1940	97.59	135.77	121.40	456.37	479.85	40.21	126.53
1941	114.87	176.89	228.03	683.28	571.12	53.16	162.34
1942	152.78	153.63	166.17	837.27	597.26	62.61	202.05
1943	127.65	164.24	104.99	379.79	578.21	69.92	180.20
1944	133.58	155.78	121.51	426.60	511.21	55.45	157.48
1945	159.34	188.16	149.80	773.27	637.78	61.19	183.51
1946	139.16	128.63	143.48	447.69	466.07	55.12	164.47
1947	143.83	228.14	176.76	1116.29	752.17	76.19	237.03
1948	39.11	140.56	114.31	714.13	482.15	55.95	166.26
1949	215.43	216.55	172.61	735.79	731.75	63.60	195.86
1950	119.74	127.12	109.07	638.61	178.45	53.53	163.16
1951	185.38	228.48	177.25	723.93	718.66	80.43	221.28
1952	175.29	253.57	205.26	853.98	854.73	81.65	244.36
1953	171.42	154.66	136.06	543.19	542.88	58.32	169.56
1954	77.05	82.52	91.46	377.96	359.63	26.67	80.90
1955	105.61	112.60	123.39	540.96	319.30	41.40	125.25
1956	144.56	177.11	152.30	735.35	621.87	33.12	159.23
1957	201.49	273.95	234.88	1176.83	904.12	59.48	229.80
1958	157.49	175.79	154.87	644.97	710.91	57.91	177.54
1959	123.44	151.96	148.11	618.79	422.89	20.29	147.68
1960	150.52	199.93	153.26	780.54	993.41	31.58	180.31
1961	130.37	148.48	124.20	599.67	589.04	19.44	133.18
1962	193.05	269.76	164.31	851.17	762.64	35.21	225.39
1963	112.60	119.22	120.65	493.81	477.88	16.44	79.02
1964	141.14	146.24	128.08	589.81	575.56	18.43	130.85
1965	196.22	227.93	193.31	760.75	749.85	33.52	209.79
1966	101.45	93.69	74.06	109.61	95.96	17.30	106.26
1967	166.64	175.08	135.65	409.44	399.60	27.09	153.62

TABLE 9 - CN-4 SERIES (CFS)

CSU USGS	1742100 9.0535	1740000 9.0575	1720000 9.0595	1666300 9.0645	1594260 9.0780	1594236 9.0785	Station Numbers 1590000 9.0850	1379000 9.1090	1378400 9.1100	1378100 9.1125	1377825 9.1135	1377500 9.1145	1377280 9.1155	1377230 9.1180
Year														
1938	1553.47	1425.34	268.05	201.44	335.55	67.04	3970.41	431.36	787.47	909.49	196.74	1327.83	142.26	147.15
1939	684.33	1141.80	226.68	203.33	354.16	142.06	2339.75	357.54	544.07	574.20	177.54	1514.84	111.77	92.95
1940	766.92	809.29	2.00	149.42	260.23	106.91	1747.83	213.23	302.13	427.80	102.24	1026.65	55.15	48.25
1941	1023.03	1020.70	184.80	178.80	233.90	145.17	2740.32	382.29	592.46	822.97	252.74	21.26	157.12	129.47
1942	1398.12	1191.44	100.01	197.32	243.19	114.37	3125.42	470.89	764.47	754.52	229.56	3676.62	211.14	193.06
1943	619.19	1234.40	61.42	209.79	306.11	164.67	2943.42	451.96	720.99	816.08	272.66	1167.75	148.61	142.62
1944	807.17	973.33	230.99	199.29	277.45	164.86	2897.58	372.48	587.09	774.26	274.26	1700.92	147.20	90.74
1945	865.43	1050.51	194.62	193.80	239.87	. 139.00	2763.15	333.35	490.27	700.26	220.31	1337.85	112.58	88.16
1946	867.13	983.65	169.00	199.61	342.19	121.38	2476.55	326.24	490.30	632.29	157.63	1143.95	60.34	61.08
1947	1322.76	1516.34	246.24	256.92	408.12	139.61	3720.75	501.64	755.42	902.67	259.66	1846.62	147.52	136.91
1948	1076.73	1238.84	218.43	222.78	340.15	130.56	3399.91	525.59	802.91	941.12	313.44	2022.60	153.99	128.68
1949	1133.94	1322.54	209.23	246.43	362.99	157.57	3093.04	472.44	758.16	824.38	265.90	1782.76	177.18	139.93
1950	926.02	1124.04	156.12	202.52	277.07	103.50	2489.05	354.75	543.50	729.29	216.72	1350.09	73.22	84.37
1951	1377.80	1558.42	200.49	265.24	363.02	148.72	2777.03	397.45	605.49	781.07	210.68	1436.51	120.06	86.09
1952	1298.34	1549.32	270.03	269.57	392.33	176.25	4112.90	629.16	1056.48	1228.04	381.88	2580.67	223.94	178.77
1953	982.09	1156.54	156.10	207.93	285.07	138.78	2476.33	378.85	609.77	648.70	110.59	1345.06	160.10	143.88
1954	435.33	518.35	68.95	108.61	156.27	60.74	1268.37	286.50	415.92	371.21	160.29	712.17	33.35	69.57
1955	595.77	726.05	133.83	157.44	233.58	93.63	1950.25	321.17	471.66	538.06	122.31	1027.53	53.45	100.22
1956	1019.09	1226.45	170.76	230.74	283.36	133.35	2384.36	412.39	620.21	696.76	218.34	1402.35	98.93	101.35
1957	1339.02	1588.86	290.43	323.31	477.71	246.11	5117.86	765.81	1308.94	1415.11	554.87	3221.64	278.21	216.48
1958	969.24	1167.64	229.26	200.47	298.62	131.64	2725.42	365.70	668.05	861.16	266.45	1704.05	189.82	141.25
1959	805.33	975.27	169.04	189.39	289.58	109.71	2287.30	290.53	419.22	492.07	144.33	940.66	74.47	93.10
1960	865.45	1029.95	182.68	200.25	285.94	130.83	2406.42	418.06	638.63	585.98	177.17	1278.18	122.16	103.95
1961	599.61	727.52	111.20	139.84	209.11	83.65	1768.37	294.09	436.30	477.43	129.08	913.05	90.66	68.80
1962	1084.42	1346.70	291.43	241.40	374.85	162.43	3686.80	597.80	992.49	1048.02	306.57	2343.79	163.82	148.87
1963	422.88	535.25	84.07	151.78	178.74	85.67	1413.14	253.52	355.39	417.41	81.39	791.08	57.07	53.86
1964	577.11	396.76	138.68	171.43	256.81	100.15	2240.17	298.35	473.43	563.00	136.31	1108.08	125.24	86.33
1965	1259.69	653.70	261.83	286.15	399.81	173.93	3707.54	598.89	1000.70	918.40	350.96	2574.95	193.54	168.81
1966	427.33	520.84	121.10	110.80	202.57	77.59	1833.54	317.30	483.67	524.86	134.20	1034.07	66.46	74.56
1967	624.37	588.06	162.84	59.72	263.00	107.42	2320.36	381.22	562.37	708.26	154.24	1308.58	44.29	78.78

TABLE 10 - CN-6 SERIES (CFS)

CSU USGS	1742100 9.0535	1740000 9.0575	1720000 9.0595	1666300 9.0645	1594260 9.0780	1594236 9.0785	Station Numbers 1590000 9.0850	1379000 9.1090	1378400 9.1100	1378100 9.1125	1377825 9.1135	1377500 9.1145	1377280 9.1155	1377230 9.1180
Year														
1938	1156.86	1054.07	184.23	140.66	239.33	44.61	2879.64	318.79	583.73	653.26	139.83	963.01	103.19	108.20
1939	512.72	830.52	156.10	141.35	254.00	98.28	1699.53	264.92	409.91	416.61	127.55	1112.28	82.40	71.45
1940	576.41	604.49	.46	103.25	186.31	74.01	1275.78	164.25	236.45	310.62	74.91	769.56	43.81	39.15
1941	766.62	757.80	127.32	124.66	164.67	99.70	1998.73	283.64	440.50	588.59	176.72	15.10	113.88	96.58
1942	1039.35	874.05	67.08	136.77	171.51	77.86	2255.84	350.01	565.74	539.18	161.07	2665.26	150.62	140.25
1943	492.25	933.16	47.19	151.81	225.05	116.98	2244.66	347.56	556.40	599.26	200.14	902.68	109.85	108.00
1944	595.34	711.86	172.50	138.46	198.27	114.46	2098.35	277.19	441.90	555.34	192.79	1248.25	106.73	68.86
1945	726.18	856.90	138.97	148.88	179.62	100.35	2137.63	262.68	397.07	521.22	164.32	1046.98	85.38	72.17
1946	660.20	735.90	119.43	138.42	245.65	83.88	1831.32	242.16	371.93	456.40	112.69	850.04	48.21	48.99
1947	1007.11	1145.85	171.30	182.06	295.19	97.25	2777.18	373.55	568.91	654.97	187.60	1366.27	110.36	101.76
1948	789.57	898.78	151.99	153.36	240.23	89.65	2488.47	387.42	591.17	669.06	218.91	1454.97	112.85	95.41
1949	847.38	988.77	146.09	171.20	259.18	108.31	2266.56	349.87	562.02	592.63	187.88	1295.83	130.97	104.72
1950	679.63	824.37	108.06	138.51	194.54	70.00	1810.28	261.24	401.21	517.90	151.53	977.46	55.06	63.45
1951	1051.02	1184.87	140.74	188.25	261.82	102.90	2090.26	300.69	463.77	574.60	153.15	1083.94	91.08	67.35
1952	972.44	1164.23	187.50	192.30	284.64	123.11	3045.42	466.85	778.99	879.92	268.12	1874.68	161.78	131.00
1953	744.67	883.46	111.51	147.93	212.46	99.23	1868.67	287.46	464.65	475.36	82.38	1009.12	117.86	107.21
1954	340.53	406.60	48.77	75.41	113.46	42.56	968.70	209.69	318.18	276.42	117.02	549.98	29.41	55.67
1955	492.55	592.03	94.67	113.92	171.63	66.29	1487.11	249.92	372.09	399.65	91.53	797.95	42.97	76.91
1956	753.20	898.64	118.91	158.62	199.63	91.55	1725.41	305.33	459.60	492.50	152.71	1021.19	72.53	76.05
1957	1046.28	1233.56	206.67	235.52	354.67	172.16	3812.55	589.23	999.55	1051.31	398.44	2421.41	209.08	162.85
1958	705.37	848.10	156.58	137.47	211.06	89.96	1973.10	275.53	501.10	611.32	186.40	1233.51	135.40	103.21
1959	617.20	743.62	118.89	130.26	203.00	75.55	1993.25	220.54	324.16	361.54	104.17	713.27	59.03	72.53
1960	647.29	771.17	126.52	139.80	202.69	90.29	1771.79	310.01	475.38	423.15	125.97	944.26	92.17	77.19
1961	489.88	591.31	79.11	98.25	151.59	57.94	1320.01	218.30	336.60	345.45	92.44	684.63	69.14	54.38
1962	808.98	997.67	202.09	169.55	265.87	112.78	2692.93	437.68	725.61	747.02	215.08	1694.99	118.03	108.35
1963	363.95	459.21	61.81	111.42	136.12	61.60	1123.25	202.44	285.55	307.42	62.05	630.30	48.55	44.90
1964	430.63	354.04	97.06	123.74	192.14	71.12	1994.92	232.22	370.13	415.33	99.95	858.04	91.40	66.98
1965	981.95	605.49	188.31	211.24	298.97	125.11	2797.98	463.23	769.46	684.62	251.77	1925.98	143.61	131.49
1966	332.45	410.32	86.49	81.82	149.03	55.27	1390.09	243.24	376.24	383.67	97.05	782.29	53.50	59.98
1967	455.68	467.94	114.39	48.23	188.83	75.57	1736.02	285.94	432.59	519.56	116.15	996.10	36.49	61.74

TABLE 11 - S-4 SERIES (CFS)

	Station Numbers					
CSU	1278800	1278050	1272445	1077090	1073480	1073436
USGS	9.1650	9.1665	9.1725	9.3440	9.3575	9.3615
Year						
1938	376.81	1562.90	575.39	325.62	327.16	2336.76
1939	38.94	606.58	425.55	182.92	209.50	1098.61
1940	84.90	745.99	298.12	164.90	197.06	1028.62
1941	622.30	1851.80	1088.38	478.13	336.10	3077.72
1942	318.67	1715.27	530.48	338.20	304.86	2230.25
1943	449.01	1137.48	422.13	213.89	224.83	1595.65
1944	424.04	1639.32	759.53	311.61	308.98	2621.76
1945	540.66	1162.65	509.43	289.46	232.80	1706.87
1946	110.87	714.89	408.29	127.24	215.96	1217.53
1947	329.00	1042.36	517.55	176.18	291.08	1815.22
1948	358.02	1347.34	726.76	260.75	337.26	2426.32
1949	456.36	1380.07	638.53	290.50	340.17	2677.44
1950	92.95	793.92	354.92	148.56	180.39	1206.87
1951	203.16	462.49	260.02	127.78	185.28	928.70
1952	563.69	1859.43	697.05	380.26	365.02	2825.00
1953	241.09	647.71	397.84	169.55	180.92	1152.80
1954	194.39	494.48	258.19	144.24	162.09	1055.70
1955	232.12	638.44	349.38	149.37	164.89	1118.89
1956	248.65	668.54	347.71	161.12	205.02	1146.48
1957	568.05	1769.26	765.01	379.73	363.13	2551.41
1958	506.66	1494.09	888.34	310.33	285.13	2401.89
1959	143.14	352.93	282.65	123.73	176.30	861.30
1960	337.76	1020.05	516.91	253.53	248.82	1735.15
1961	265.36	770.40	482.98	187.66	208.59	1372.06
1962	361.31	1059.16	511.90	246.10	273.25	1855.92
1963	190.51	522.54	270.80	129.95	165.65	957.57
1964	230.50	630.49	398.27	142.47	175.87	1013.54
1965	503.80	1387.75	669.28	325.56	323.32	2487.34
1966	269.79	852.00	364.29	219.23	216.47	1419.96
1967	185.41	510.44	285.76	170.36	165.16	973.56

TABLE 12 - S-6 SERIES (CFS)

	Station Numbers					
CSU	1278800	1278050	1272445	1077090	1073480	1073436
USGS	9.1650	9.1665	9.1725	9.3440	9.3575	9.3615
Year						
1938	264.70	1087.32	428.99	233.04	234.72	1679.61
1939	32.79	441.02	323.29	136.78	149.07	832.04
1940	64.40	335.57	232.45	126.39	138.39	760.03
1941	438.72	1294.88	799.02	344.08	247.28	2228.01
1942	228.85	1207.55	403.54	243.35	220.95	1618.76
1943	323.33	820.58	351.01	158.26	172.19	1237.58
1944	298.37	1142.23	555.73	221.03	222.45	1864.03
1945	383.68	826.26	402.87	208.82	171.12	1270.55
1946	85.90	524.96	320.06	97.90	154.90	915.27
1947	243.36	767.28	410.77	137.72	215.76	1396.48
1948	254.48	948.10	530.04	185.66	241.83	1749.94
1949	323.65	972.30	469.79	211.16	242.48	1905.66
1950	70.94	571.55	263.88	109.55	128.96	886.80
1951	145.70	336.17	201.96	96.99	134.88	707.76
1952	389.84	1288.06	520.53	272.82	264.68	2010.24
1953	173.03	479.38	304.84	126.15	131.97	869.32
1954	138.63	361.00	204.30	108.44	117.40	788.60
1955	172.42	483.31	270.27	114.10	121.39	857.79
1956	176.90	494.15	258.79	123.56	144.73	853.34
1957	422.62	1291.87	587.51	285.37	273.42	1939.02
1958	349.47	1046.41	628.61	221.47	203.19	1701.05
1959	111.50	268.50	225.53	95.02	129.51	671.83
1960	237.14	723.01	376.56	182.67	175.81	1265.89
1961	189.47	555.96	361.68	141.88	149.11	1024.97
1962	251.22	754.19	378.03	176.50	196.04	1348.15
1963	144.45	405.77	229.42	103.45	122.02	764.80
1964	172.04	474.99	305.47	114.14	127.19	777.66
1965	367.10	1007.67	517.49	234.53	241.37	1863.86
1966	194.47	633.65	277.44	163.20	154.73	1080.99
1967	143.22	406.01	235.85	140.37	121.59	771.36

TABLE 13 - CS-4 SERIES (CFS)

	Station Numbers								
CSU	1425625	1377280	1377230	1377200	1373900	1373055	1373020	1371815	1370300
USGS	9.0975	9.1155	9.1180	9.1190	9.1275	9.1325	9.1345	9.1430	9.1520
Year									
1938	211.98	142.26	147.15	405.19	114.53	1642.21	205.27	90.56	111.74
1939	74.54	111.77	92.95	262.73	72.35	894.18	100.73	65.04	67.78
1940	65.77	55.15	48.25	89.25	105.32	841.89	98.93	75.51	97.24
1941	182.76	157.12	129.47	416.78	86.48	1395.74	175.09	132.19	148.55
1942	289.53	211.14	193.06	647.95	152.50	1579.42	172.69	123.79	143.21
1943	73.88	148.61	142.62	363.90	75.68	1059.44	111.46	75.26	60.63
1944	185.98	147.20	90.74	412.16	120.69	1480.44	139.55	116.49	132.25
1945	130.83	112.58	88.16	261.20	32.12	1249.94	126.28	101.08	87.00
1946	92.40	60.34	61.08	124.30	60.75	847.02	67.73	55.91	56.20
1947	136.23	147.52	136.91	401.64	75.77	1194.08	136.04	113.53	81.28
1948	170.25	153.99	128.68	485.11	115.13	1392.71	118.72	97.64	92.22
1949	134.23	177.18	139.93	582.28	86.42	1239.87	142.84	95.66	96.21
1950	80.95	73.22	84.37	189.60	57.45	1222.23	119.98	84.06	53.78
1951	42.20	120.06	86.09	239.52	48.28	900.74	83.58	62.22	44.54
1952	181.21	223.94	178.77	636.87	120.89	1760.47	195.21	166.94	112.23
1953	74.08	160.10	143.88	330.64	77.19	857.61	74.72	64.32	56.70
1954	34.92	33.35	69.57	68.52	21.66	445.60	70.88	57.12	43.29
1955	59.96	53.45	100.22	114.93	39.45	819.78	103.11	85.36	61.89
1956	47.87	98.93	101.35	227.28	51.96	868.61	38.79	61.85	42.92
1957	221.73	278.21	216.48	800.97	147.08	2200.99	259.23	128.75	128.55
1958	185.21	189.82	141.25	496.77	139.64	1297.41	116.86	132.44	125.35
1959	45.48	74.47	93.10	130.29	66.34	660.52	129.36	52.46	42.69
1960	86.97	122.16	103.95	271.42	45.37	882.82	172.63	71.83	66.45
1961	53.26	90.66	68.80	150.39	46.02	651.57	66.94	69.35	43.75
1962	190.55	163.82	148.87	407.72	95.54	1651.73	161.40	111.98	76.97
1963	23.34	57.07	53.86	107.67	30.68	558.43	66.95	49.42	19.60
1964	106.35	123.24	86.33	270.83	77.64	989.99	95.81	61.17	49.85
1965	142.90	193.54	168.81	524.61	112.12	1573.51	131.16	97.77	82.81
1966	62.55	66.46	74.56	156.68	56.62	739.28	85.61	84.05	45.48
1967	41.83	44.29	78.78	104.93	30.83	829.28	105.95	71.72	29.71

TABLE 14 - CS-6 SERIES (CFS)

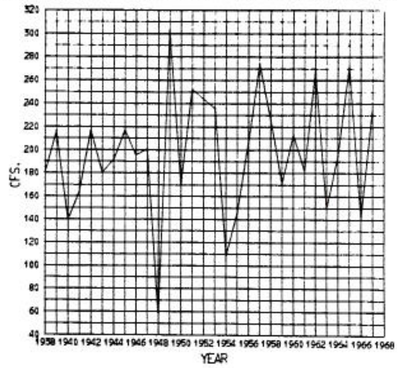
	Station Numbers								
CSU	1425625	1377280	1377230	1377200	1373900	1373055	1373020	1371815	1370300
USGS	9.0975	9.1155	9.1180	9.1190	9.1275	9.1325	9.1345	9.1430	9.1520
Year									
1938	144.14	105.19	108.20	302.53	78.66	1136.71	142.02	69.48	80.24
1939	53.13	82.40	71.45	238.90	52.07	636.67	70.95	50.37	49.77
1940	46.75	43.81	39.15	91.35	72.57	592.49	69.85	56.37	69.68
1941	123.01	113.88	96.58	326.14	60.40	969.91	122.12	97.97	105.32
1942	194.33	150.62	140.25	474.42	103.08	1090.52	119.28	92.60	101.63
1943	54.77	109.85	108.00	291.91	53.81	758.22	78.86	59.76	45.54
1944	124.48	106.73	68.86	308.07	81.75	1018.14	97.68	89.57	96.05
1945	90.12	85.38	72.17	229.69	24.91	884.45	90.13	80.43	64.59
1946	64.12	48.21	48.99	123.36	43.47	600.70	49.37	45.81	42.73
1947	92.85	110.36	101.76	320.93	53.40	846.48	96.35	85.87	60.44
1948	115.45	112.35	95.41	363.62	79.24	966.05	84.57	75.55	66.79
1949	90.92	130.97	104.72	435.31	60.49	860.50	101.69	75.59	71.48
1950	55.82	55.06	63.45	152.94	40.33	843.82	86.45	67.46	40.53
1951	29.41	91.08	67.35	196.48	34.44	634.29	61.60	50.05	34.28
1952	122.63	161.78	131.00	468.72	82.59	1220.74	136.39	122.21	80.63
1953	51.39	117.86	107.21	271.75	54.44	613.11	55.97	50.46	42.77
1954	24.46	29.41	55.67	71.27	15.99	323.08	52.32	45.07	33.25
1955	40.93	42.97	76.91	109.68	28.76	584.86	74.74	66.55	47.04
1956	33.63	72.53	76.05	174.49	37.67	615.69	64.30	48.27	33.49
1957	149.42	209.08	162.35	619.51	100.67	1553.36	181.89	100.35	92.67
1958	126.36	135.40	103.21	363.15	94.29	900.00	82.66	99.83	90.44
1959	31.24	59.03	72.53	123.33	46.51	468.01	91.14	42.57	33.60
1960	60.31	92.17	77.19	219.38	33.19	623.26	120.63	57.48	49.85
1961	37.06	69.14	54.38	133.32	32.17	460.72	50.14	47.68	30.89
1962	129.10	118.03	108.33	304.76	65.28	1146.48	114.67	85.67	54.38
1963	20.40	48.55	44.90	114.68	25.22	405.58	48.99	39.41	13.95
1964	72.26	91.40	66.98	217.55	53.85	693.73	68.87	49.22	35.18
1965	96.84	143.61	131.49	419.56	77.91	1102.29	94.71	76.68	59.73
1966	46.73	53.50	59.98	143.56	41.21	538.26	62.46	65.36	32.40
1967	31.96	36.49	61.74	101.06	23.33	613.83	76.43	58.01	21.54

TABLE 15 - STATIONS WITH MISSING DATA

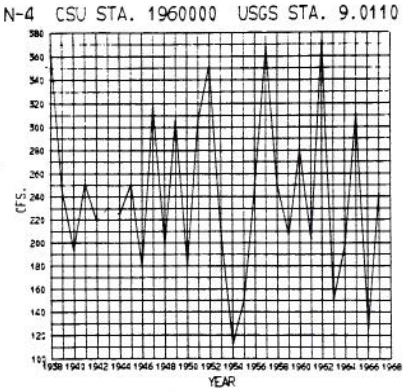
Station with missing data		Filling in of missing data is made with station		Year of missing data
CSU Sta. No.	USGS Sta. No.	CSU Sta. No.	USGS Sta. No.	
1820000	9.0195	1830000	9.0190	54-60
1830000	9.0190	1960000	9.0110	38-50
1970000	9.0105	1960000	9.0110	38-53
1272445	9.1725	1277200	9.1665	38-42
1278800	9.1650	1277200	9.1665	38-51
1371815	9.1430	1370300	9.1520	38-39
1373020	9.1345	1373055	9.1325	57-60
1373900	9.1275	1373360	9.1285	38-45; 55-60
1377230	9.1180	1377280	9.1155	51-60
1377825	9.1135	1378100	9.1125	38-40; 51-58
1594236	9.0785	1378400	9.1100	38-47
1720000	9.0595	1590000	9.0850	38-44
1377500	9.1145	1378400	9.1100	38-44
1379000	9.1090	1378400	9.1100	38
1594260	9.0780	1378400	9.1100	38-47

Fig. 5 N-4 series

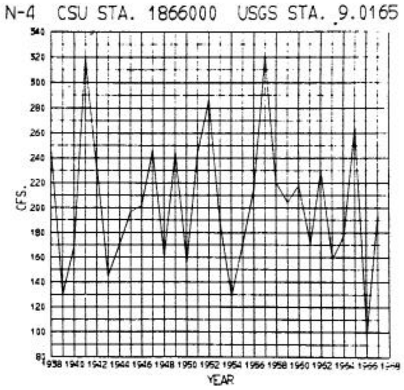
N-4 CSU STA. 1970000 USGS STA. 9.0105



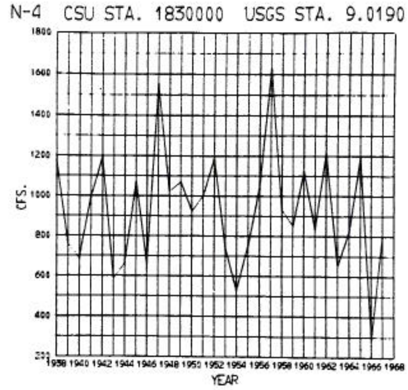
N-4 CSU STA. 1960000 USGS STA. 9.0110



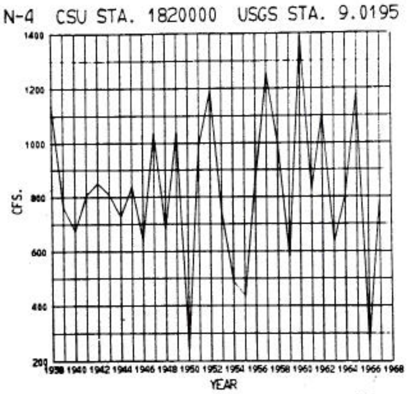
N-4 CSU STA. 1866000 USGS STA. 9.0165



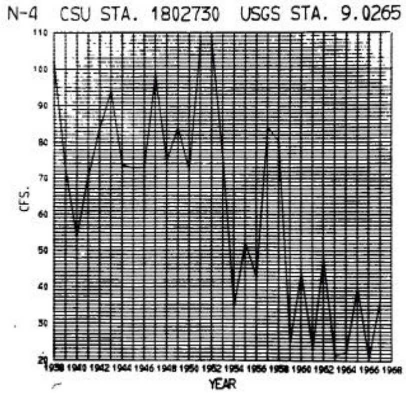
N-4 CSU STA. 1830000 USGS STA. 9.0190



N-4 CSU STA. 1820000 USGS STA. 9.0195



N-4 CSU STA. 1802730 USGS STA. 9.0265



N-4 CSU STA. 1776000 USGS STA. 9.0360

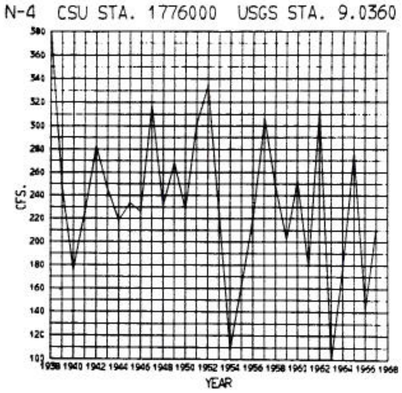


Fig. 6 N-6 series

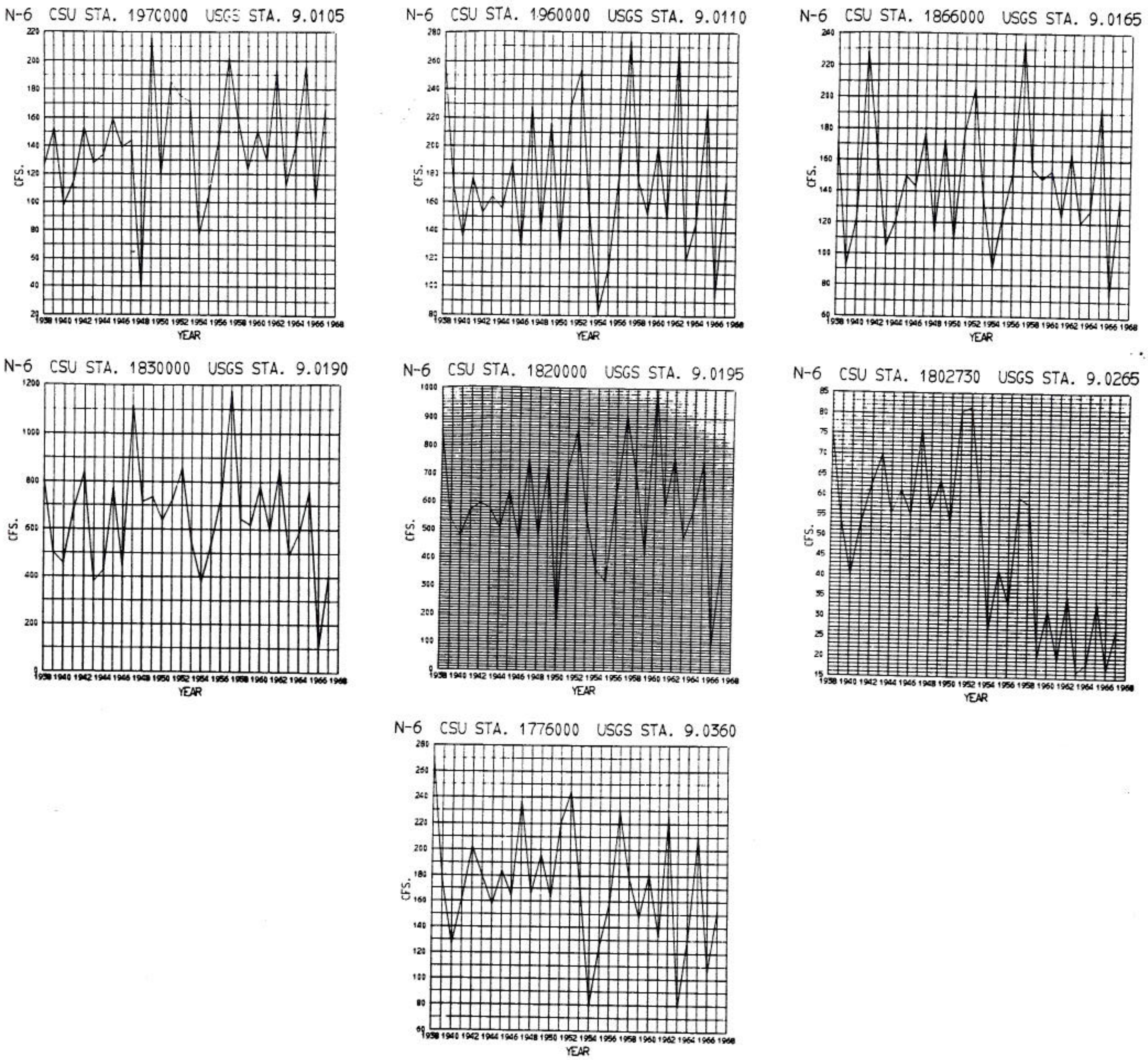


Fig. 7 CN-4 series

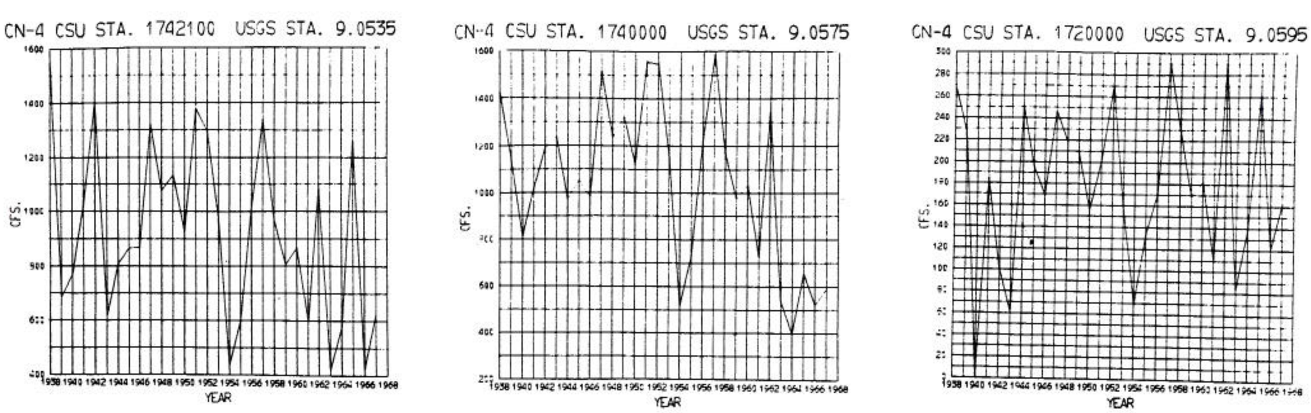
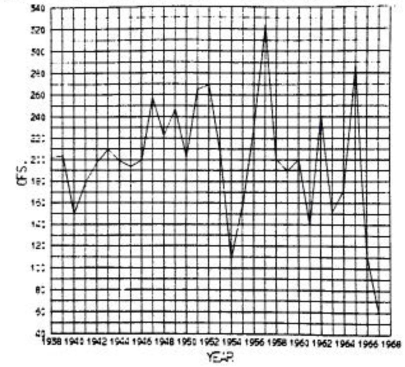
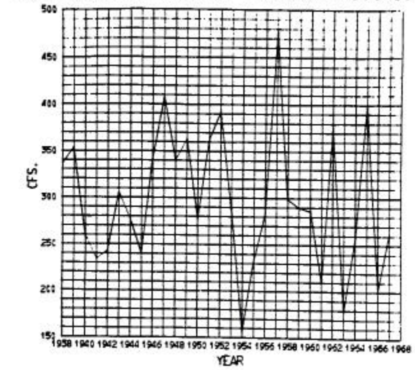


Fig. 7 CN-4 series - Continued

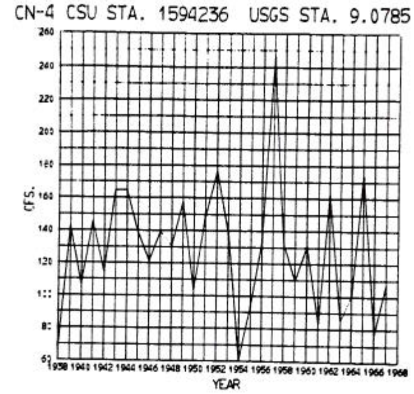
CN-4 CSU STA. 1666300 USGS STA. 9.0645



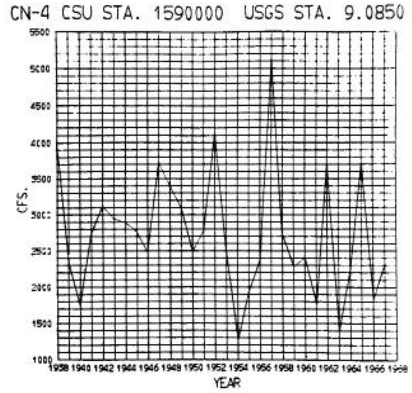
CN-4 CSU STA. 1594260 USGS STA. 9.0780



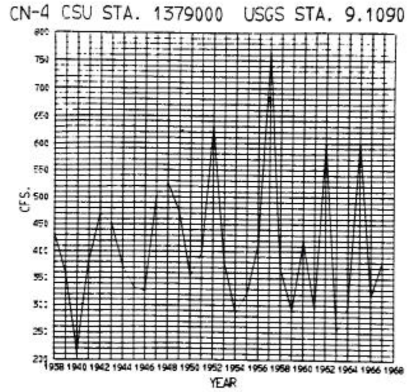
CN-4 CSU STA. 1594236 USGS STA. 9.0785



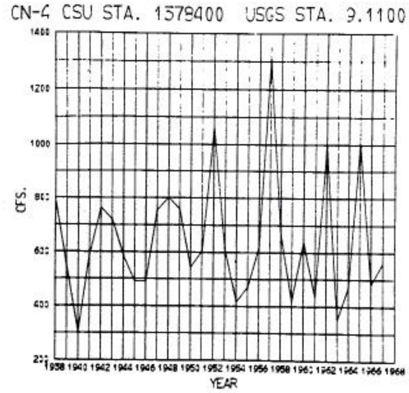
CN-4 CSU STA. 1590000 USGS STA. 9.0850



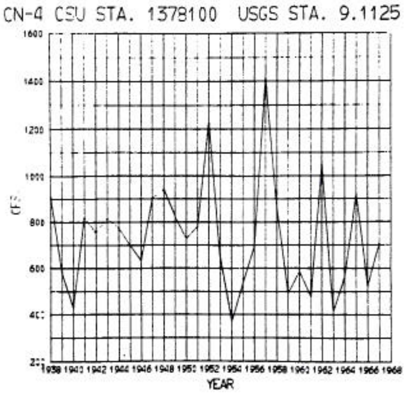
CN-4 CSU STA. 1379000 USGS STA. 9.1090



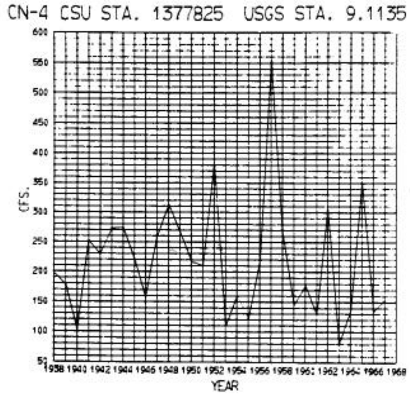
CN-4 CSU STA. 1378400 USGS STA. 9.1100



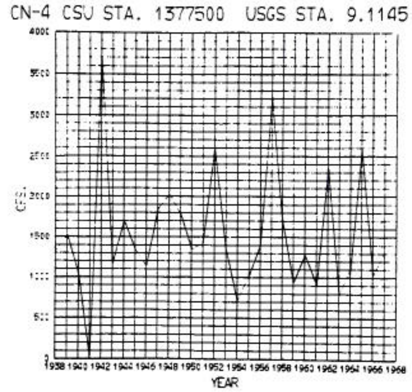
CN-4 CSU STA. 1378100 USGS STA. 9.1125



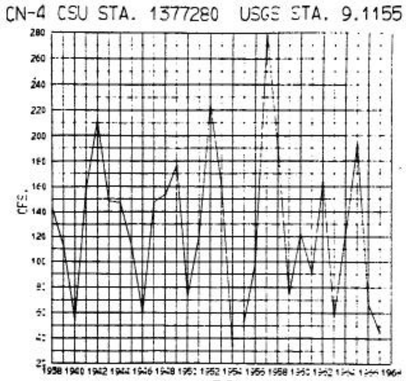
CN-4 CSU STA. 1377825 USGS STA. 9.1135



CN-4 CSU STA. 1377500 USGS STA. 9.1145



CN-4 CSU STA. 1377280 USGS STA. 9.1155



CN-4 CSU STA. 1377230 USGS STA. 9.1180

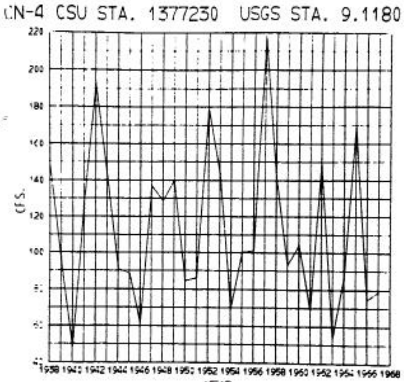


Fig. 8 CN-6 series

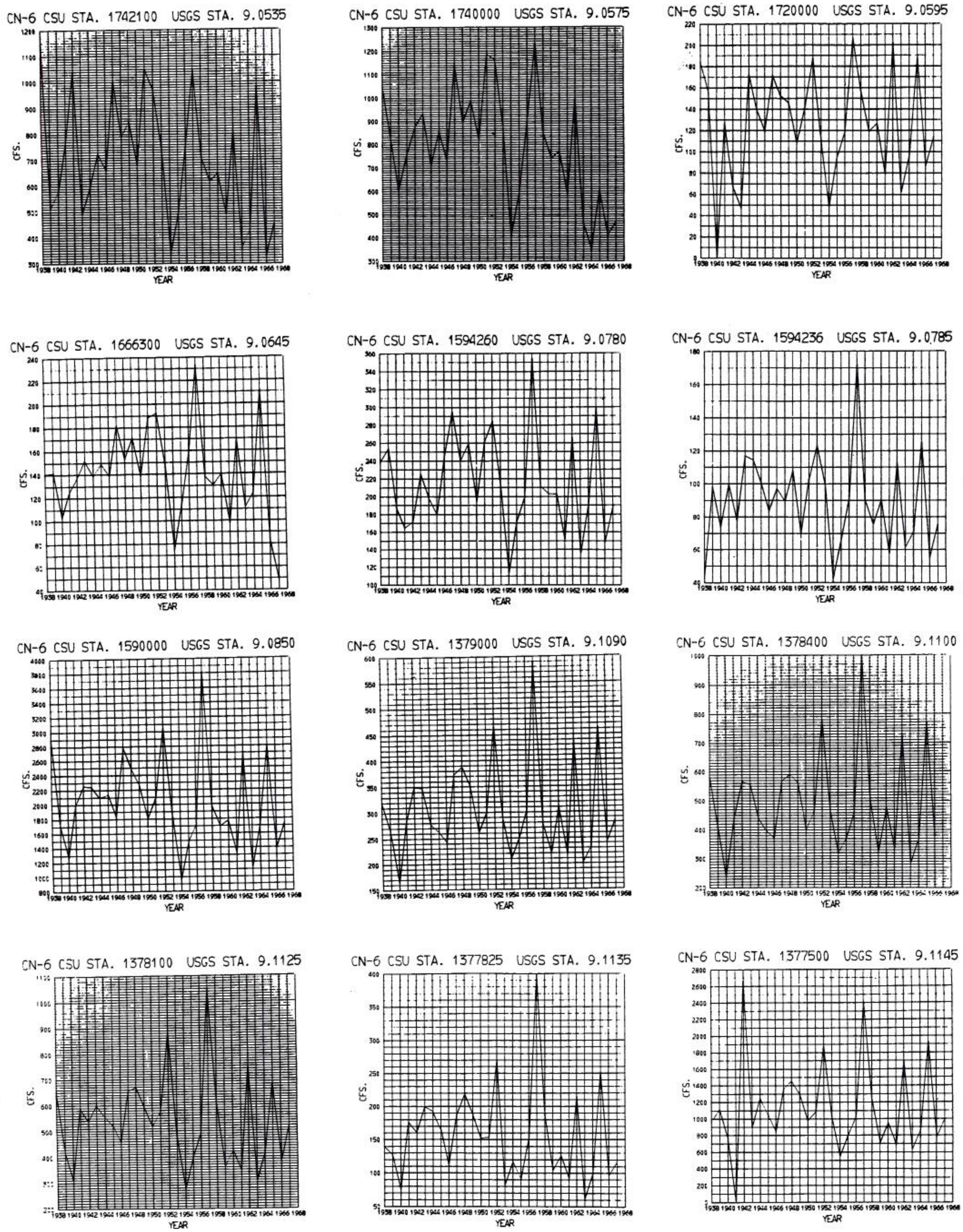


Fig. 8 CN-6 series - Continued

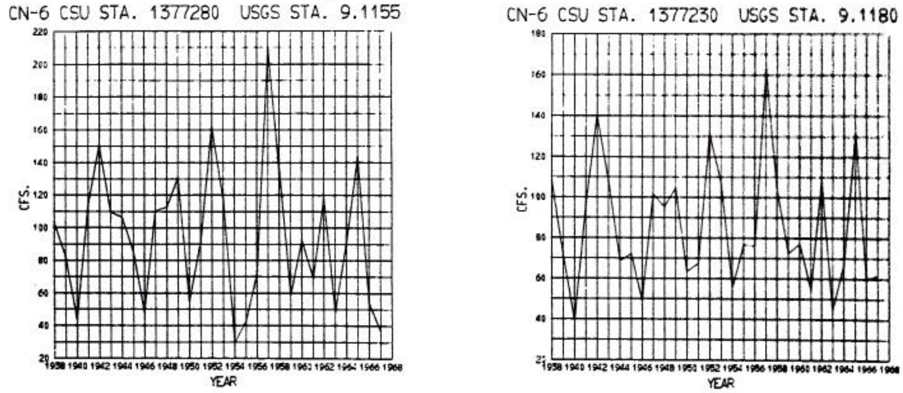


Fig. 9 S-4 series

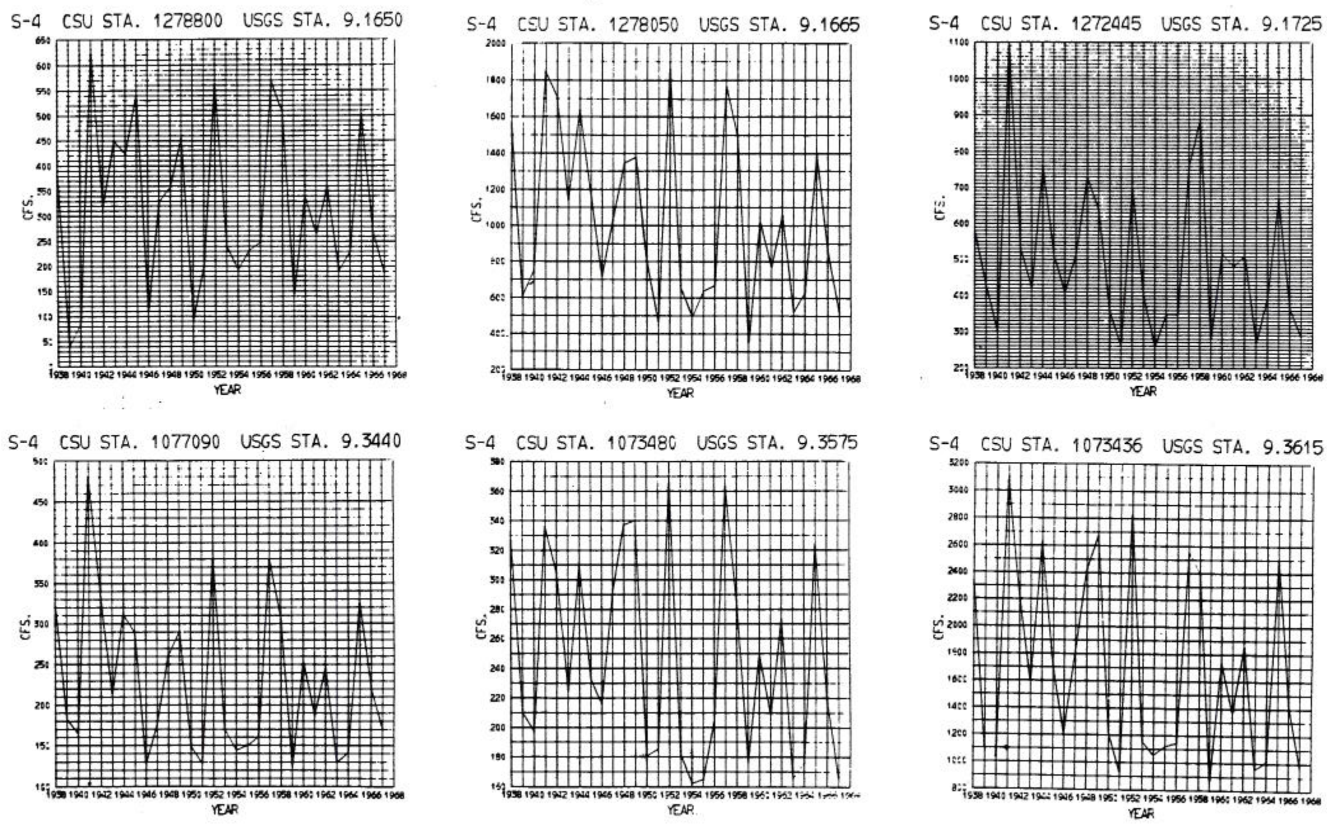


Fig. 10 S-6 series

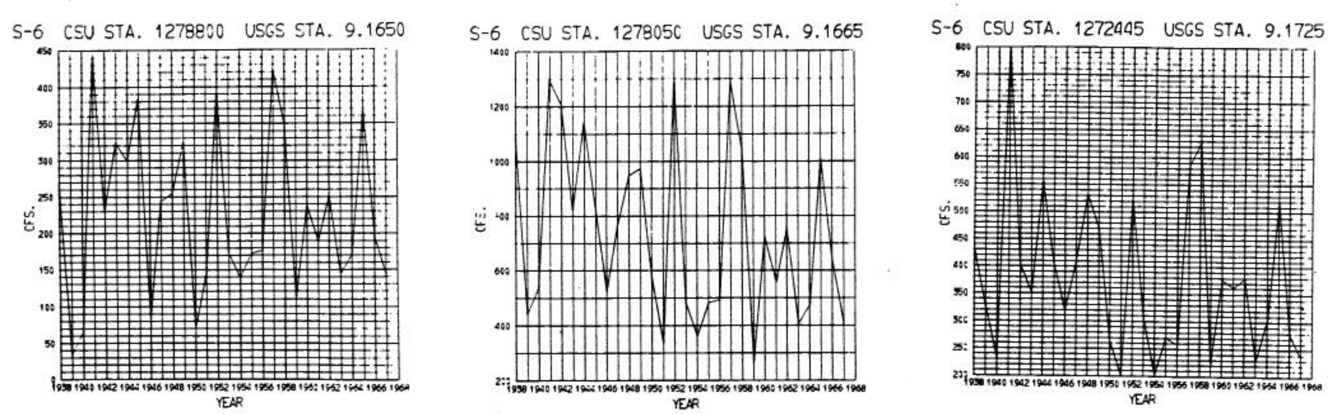
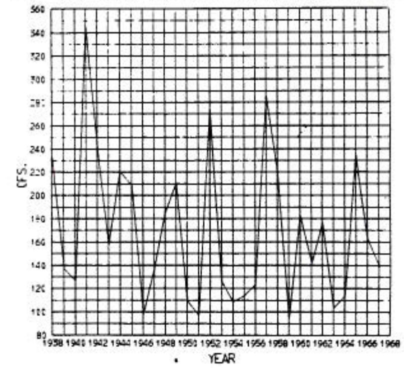
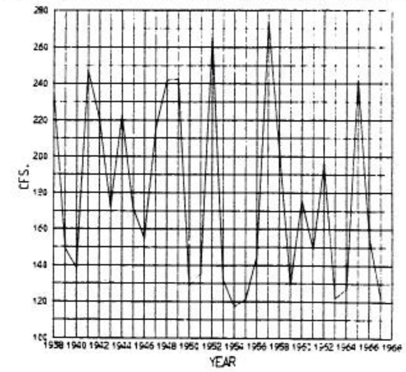


Fig. 10 S-6 series - Continued

S-6 CSU STA. 1077090 USGS STA. 9.3440



S-6 CSU STA. 1073480 USGS STA. 9.3575



S-6 CSU STA. 1073436 USGS STA. 9.3615

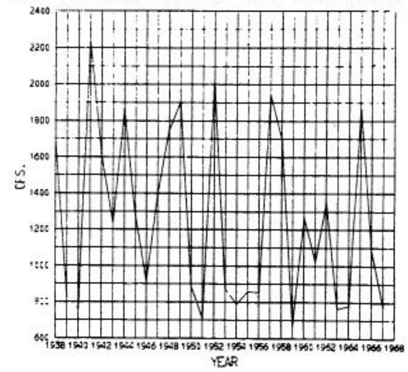
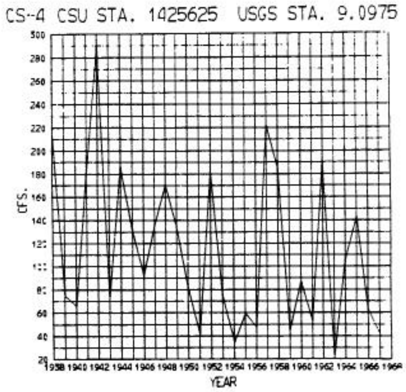
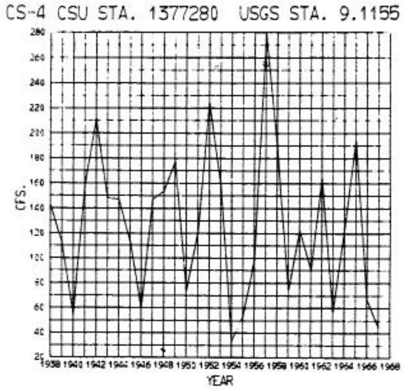


Fig. 11 CS-4 series

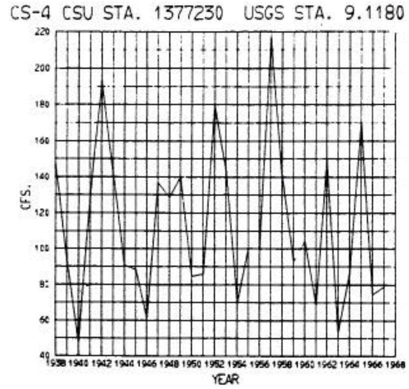
CS-4 CSU STA. 1425625 USGS STA. 9.0975



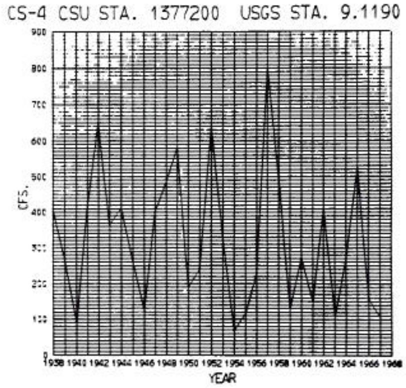
CS-4 CSU STA. 1377280 USGS STA. 9.1155



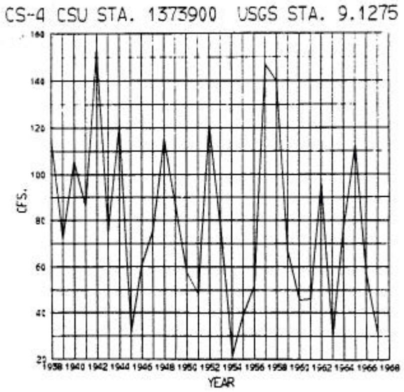
CS-4 CSU STA. 1377230 USGS STA. 9.1180



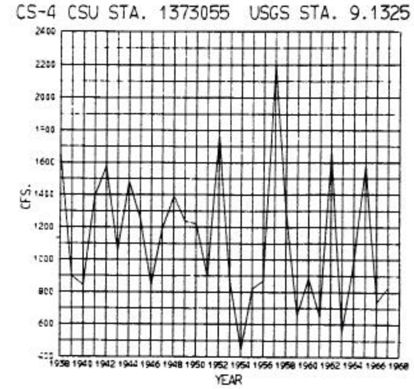
CS-4 CSU STA. 1377200 USGS STA. 9.1190



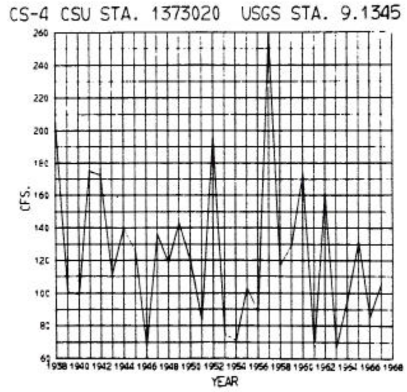
CS-4 CSU STA. 1373900 USGS STA. 9.1275



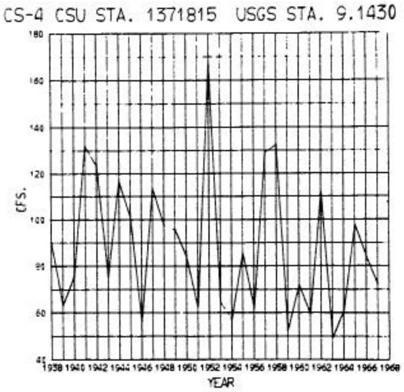
CS-4 CSU STA. 1373055 USGS STA. 9.1325



CS-4 CSU STA. 1373020 USGS STA. 9.1345



CS-4 CSU STA. 1371815 USGS STA. 9.1430



CS-4 CSU STA. 1370300 USGS STA. 9.1520

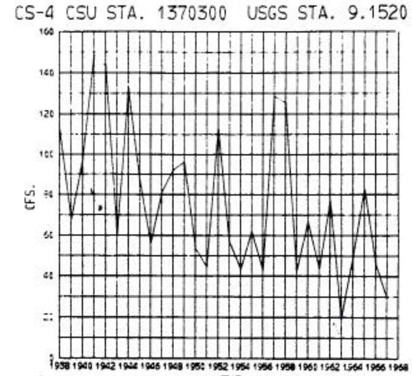


Fig. 12 CS-6 series

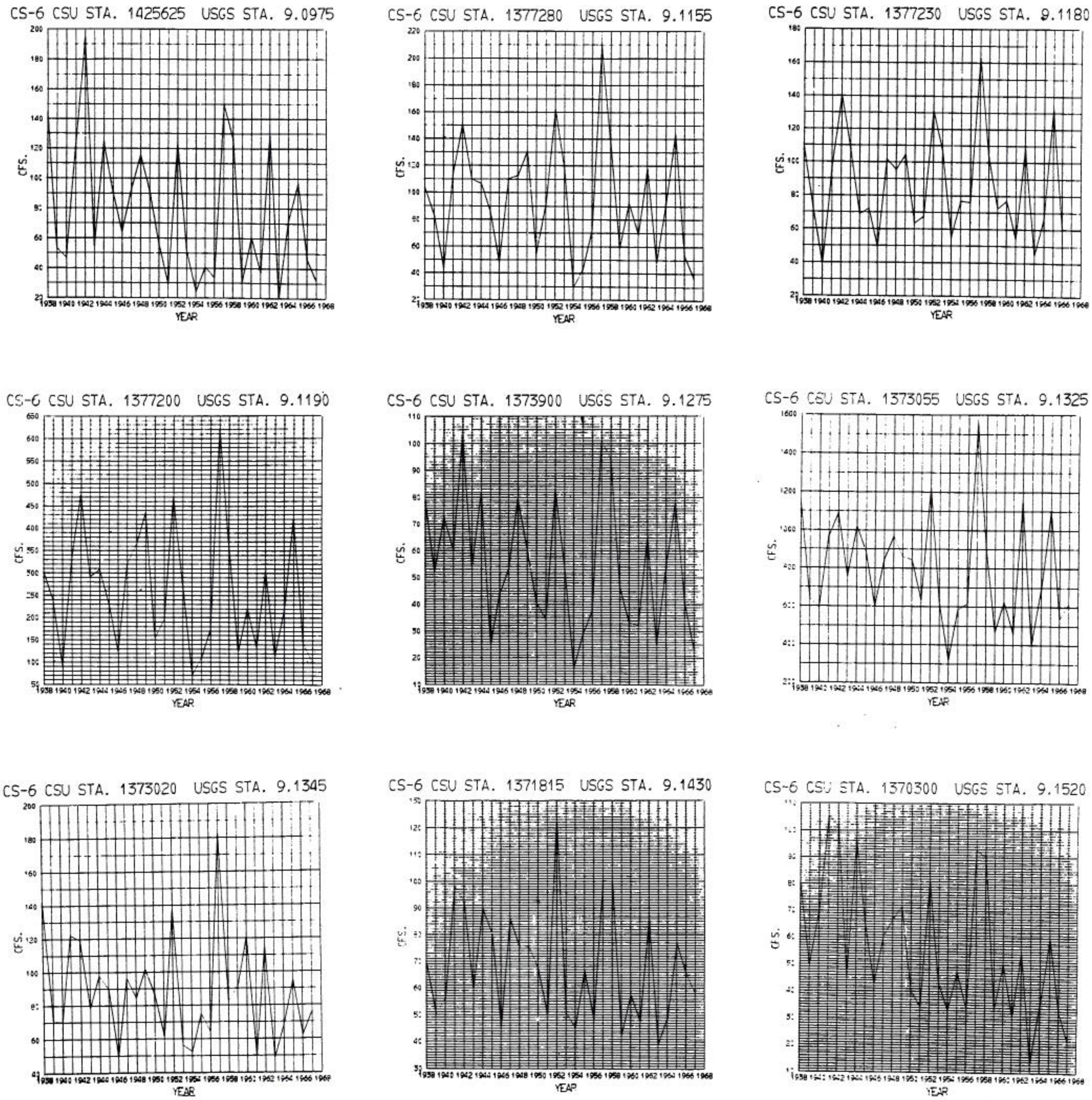


TABLE 16 - MEANS AND STANDARD DEVIATIONS OF 30 YEAR RAW DATA

CSU Sta. No.	USGS Sta. No.	Mean of 4-month averages (cfs)	Std. Dev. of 4-month averages (cfs)	Mean of 6-month averages (cfs)	Std. Dev. of 6-month averages (cfs)
1970000	9.0105	198.449	55.552	141.865	38.338
1960000	9.0110	241.821	71.196	174.569	51.065
1866000	9.0165	203.590	53.612	146.297	38.545
1830000	9.0190	931.757	290.050	644.206	220.359
1820000	9.0195	826.556	274.385	582.724	203.842
1802730	9.0265	63.007	27.307	47.814	20.085
1776000	9.0360	234.679	64.188	171.954	45.884
1742100	9.0535	924.237	316.280	702.783	233.361
1740000	9.0575	1043.263	341.708	794.292	241.819
1720000	9.0595	177.674	72.281	124.358	50.072
1666300	9.0645	199.184	55.142	141.030	39.140
1594260	9.0780	297.711	73.851	215.045	53.847
1594236	9.0785	128.576	38.515	89.611	27.138
1590000	9.0850	2739.444	854.102	2031.847	624.926
137900	9.1090	406.685	122.728	306.039	91.987
1378400	9.1100	641.932	224.886	486.119	165.435
1378100	9.1125	736.162	236.020	534.129	170.870
1377825	9.1135	219.301	98.861	157.011	69.485
1377500	9.1145	1521.405	754.366	1129.644	546.188
1377280	9.1155	126.388	59.684	94.178	42.221
1377230	9.1180	113.268	42.698	85.892	30.376
1278800	9.1650	314.930	159.674	226.413	111.344
1278050	9.1665	1028.025	467.197	738.323	319.454
1272445	9.1725	500.048	204.184	379.190	144.622
1077090	9.3440	230.964	93.315	170.480	64.553
1073480	9.3575	245.702	68.310	178.304	50.207
1073436	9.3615	1696.563	688.607	1254.713	481.412
1425625	9.0975	114.324	68.267	78.467	45.159
1377280	9.1155	126.388	59.684	94.178	42.221
1377230	9.1180	113.268	42.698	85.892	30.376
1377200	9.1190	322.737	192.349	257.047	136.341
1373900	9.1275	78.756	36.701	55.056	24.150
1373055	9.1325	1124.250	410.613	790.065	280.861
1373020	9.1345	124.166	46.158	88.247	31.349
1371815	9.1430	88.115	29.509	68.390	21.124
1370300	9.1520	76.695	35.297	56.029	24.863

TABLE 17 - CORRELATION MATRIX BETWEEN N-4 AND CN-4 (computed from 30-year data)

			N-4						
		CSU STA. NO.	1970000	1960000	1866000	1830000	1820000	18-2730	1776000
	CSU STA. NO.	USGS STA. NO.	9.0105	9.0110	9.0165	9.0190	9.0195	9.0265	9.0360
CN-4	1742100	9.0535	.477	.785	.778	.815	.641	.710	.894
	1740000	9.0575	.411	.728	.566	.669	.531	.843	.836
	1720000	9.0595	.524	.771	.592	.660	.593	.421	.730
	1666300	9.0645	.535	.717	.640	.728	.619	.592	.721
	1594260	9.0780	.582	.802	.597	.702	.640	.585	.805
	1594236	9.0785	.621	.652	.602	.565	.584	.477	.549
	1590000	9.0850	.502	.845	.722	.806	.673	.641	.870
	1379000	9.1090	.517	.767	.655	.754	.662	.469	.730
	1378400	9.1100	.541	.789	.671	.743	.685	.489	.757
	1378100	9.1125	.497	.801	.714	.740	.621	.627	.791
	1377825	9.1135	.426	.680	.639	.686	.564	.498	.627
	1377500	9.1145	.497	.533	.419	.632	.491	.395	.602
	1377280	9.1155	.553	.710	.692	.694	.694	.567	.708
	1377230	9.1180	.490	.674	.674	.704	.612	.537	.719

TABLE 18 - CORRELATION MATRIX BETWEEN N-6 AND CN-6 (computed from 30-year data)

			N-6						
		CSU STA. NO.	1970000	1960000	1866000	1830000	1820000	1802730	1776000
	CSU STA. NO.	USGS STA. NO.	9.0105	9.0110	9.0165	9.0190	9.0195	9.0265	9.0360
CN-6	1742100	9.0535	.486	.783	.780	.813	.671	.706	.896
	1740000	9.0575	.457	.756	.598	.731	.631	.825	.855
	1720000	9.0595	.529	.770	.587	.620	.569	.408	.736
	1666300	9.0645	.570	.740	.659	.750	.668	.578	.745
	1594260	9.0780	.598	.798	.592	.659	.621	.555	.803
	1594236	9.0785	.627	.650	.603	.529	.552	.448	.562
	1590000	9.0850	.509	.840	.720	.764	.659	.623	.876
	1379000	9.1090	.529	.766	.658	.704	.613	.438	.734
	1378400	9.1100	.551	.788	.669	.693	.639	.456	.761
	1378100	9.1125	.507	.802	.716	.700	.597	.599	.799
	1377825	9.1135	.433	.681	.640	.650	.541	.477	.644
	1377500	9.1145	.503	.532	.420	.596	.450	.373	.604
	1377280	9.1155	.555	.708	.701	.689	.695	.542	.715
	1377230	9.1180	.497	.668	.673	.674	.593	.509	.723

TABLE 19 - CORRELATION MATRIX BETWEEN S-4 AND CS-4 (computed from 30-year data)

			S-4					
		CSU STA. NO.	1278800	1278050	1272445	1077090	1073480	1073436
	CSU STA. NO.	USGS STA. NO.	9.1650	9.1665	9.1725	9.3440	9.3575	9.3615
CS-4	1425625	9.0975	.656	.890	.752	.811	.849	.830
	1377280	9.1155	.748	.807	.718	.753	.830	.782
	1377230	9.1180	.682	.742	.593	.695	.760	.709
	1377200	9.1190	.748	.859	.737	.786	.889	.842
	1373900	9.1275	.491	.776	.661	.663	.770	.714
	1373055	9.1325	.707	.861	.711	.792	.858	.810
	1373020	9.1345	.643	.776	.612	.773	.777	.723
	1371815	9.1430	.772	.884	.784	.836	.819	.850
	1370300	9.1520	.675	.908	.835	.866	.828	.856

TABLE 20 - CORRELATION MATRIX BETWEEN S-6 AND CS-6 (computed from 30-year data)

			S-6					
		CSU STA. NO.	1278800	1278050	1272445	1077090	1073480	1073436
	CSU STA. NO.	USGS STA. NO.	9.1650	9.1665	9.1725	9.3440	9.3575	9.3615
CS-6	1425625	9.0975	.653	.892	.754	.804	.846	.826
	1377280	9.1155	.753	.805	.722	.747	.840	.786
	1377230	9.1180	.702	.747	.608	.699	.779	.722
	1377200	9.1190	.753	.854	.746	.783	.894	.841
	1373900	9.1275	.492	.778	.659	.660	.770	.712
	1373055	9.1325	.716	.863	.720	.791	.866	.813
	1373020	9.1345	.656	.779	.624	.780	.787	.733
	1371815	9.1430	.777	.888	.793	.837	.828	.854
	1370300	9.1520	.666	.896	.829	.848	.821	.845

Chapter V

DATA ANALYSIS AND RESULTS

In this chapter the data described in Chapter IV are analyzed according to the procedures discussed in Chapter III. The approaches used for reducing the number of runoff variables are the principal component analysis and the canonical analysis. The minimum numbers of years to detect the increase in the runoff means are obtained by application of equation (1).

In the principal component analysis and the canonical analysis, the coefficients for the principal components and the canonical variables are obtained basically from the analysis of the covariance matrix. Therefore, because the covariance matrix is assumed to be the same for both periods, it follows that the coefficients obtained for the non-seeded period apply for the seeded period as well. The suspected change in the means of the runoff leave the coefficients of the components invariant.

5.1 <u>The application of principal component analysis</u>. The numerical procedures for the reduction of the number of runoff variables by the principal components method were executed separately in each region on the CDC 6400 digital computer of Colorado State University. The program BMDO1M from the University of California Press was modified to accommodate nonstandardized variables. The zero mean is not desirable here because a certain percent increase in the mean will be postulated later.

The steps in obtaining the principal components in each region may be summarized as follows:

1) Compute the covariance matrix of the runoff variables in that region, $\hat{\underline{V}}$, as defined in equation (2).

2) Solve the system,

$$|\hat{\underline{V}} - \lambda I| = 0 \quad , \tag{21}$$

to obtain $\lambda_1, \lambda_2, \ldots, \lambda_p$, the characteristic roots, which are the amounts of variances of components 1, 2, ..., p .

3) Solve the system,

$$(\hat{\underline{V}} - \lambda_i \; I)\underline{\beta}_i = \underline{0} \tag{22}$$

subject to the normalization condition,

$$\underline{\beta}_i' \; \underline{\beta}_i = 1 \tag{23}$$

to obtain $\underline{\beta}_i$ which is the vector of the coefficients for the i^{th} component in that region.

For example, when N-4, which is the four-month runoff of the northern region, is used the coefficients for the first principal component are found to be (Table 21),

$$\beta_{1,1} = 0.0859$$

$$\beta_{1,2} = 0.1679$$

$$\beta_{1,3} = 0.1151$$

$$\beta_{1,4} = 0.7065$$

$$\beta_{1,5} = 0.6576$$

$$\beta_{1,6} = 0.0332$$

$$\beta_{1,7} = 0.1359$$

where the first subscript of β indicates the ordering number of the principal component, the second one indicates the sequential number of the station as shown in Table 2.

Let ξ_i be the i^{th} principal component in the target region before seeding, then for N-4,

$$\xi_1 = \sum_{j=1}^{7} \beta_{1,j} \; Q_j$$

$$= 0.0859Q_1 + 0.1679Q_2 + 0.1151Q_3 + 0.7065Q_4$$

$$+ 0.6576Q_5 + 0.0332Q_6 + 0.1359Q_7$$

where Q_1, Q_2, Q_3, Q_4, Q_5, Q_6 and Q_7 are runoff variables listed in order corresponding to the numbers in the 'Seq. No.' column in Table 2. This first principal component will account for the largest percentage of the total variation in this whole region based on the four-month spring runoff.

The coefficients for the principal components in N-4, N-6, CN-4, CN-6, S-4, S-6, CS-4 and CS-6 are shown in Tables 21, 22, 23, 24, 25, 26, 27, and 28, respectively. The cumulative percentages of total variation accounted for by the principal components in each region are shown in Table 29. A 99 cumulative percentage was used to limit the number of the principal components to be retained for the study, because it was found that beyond this percentage of total variation, the rate of increase of the cumulative percentage was very slow.

After the coefficients of the principal components in each region have been found, then the series of the principal components can be simply obtained from the original series [12].

TABLE 21 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF N-4

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.	4th Comp.
1970000	9.0105	.0859	-.0894	-.4339	-.8081
1960000	9.0110	.1679	-.0529	-.4719	.0637
1866000	9.0165	.1151	.0334	-.1221	-.2757
1830000	9.0190	.7065	.6848	.1407	-.0308
1820000	9.0195	.6576	-.7201	.1966	.0688
1802730	9.0265	.0332	.0191	-.3072	.2822
1776000	9.0360	.1359	.0132	-.6491	.4262

TABLE 22 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF N-6

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.	4th Comp.
1970000	9.0105	.0767	-.0806	-.5604	-.6494
1960000	9.0110	.1549	-.0680	-.5037	.0808
1866000	9.0165	.1084	.0256	-.2105	-.2926
1830000	9.0190	.7191	.6784	.1135	-.0377
1820000	9.0195	.6510	-.7266	.2122	.0154
1802730	9.0265	.0339	.0048	-.1892	.4046
1776000	9.0360	.1279	-.0079	-.5424	.5664

TABLE 23 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF CN-4

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.	4th Comp.
1742100	9.0535	.2250	-.1496	-.4782	-.7580
1740000	9.0575	.2155	-.2714	-.7505	.5419
1720000	9.0599	.0444	-.0508	.0329	-.0134
1666300	9.0645	.0378	-.0229	-.0434	.0232
1594260	9.0780	.0524	-.0388	-.0108	.0455
1594236	9.0785	.0240	-.0114	.0178	.0891
1590000	9.0850	.7025	-.4654	.3580	-.0949
1379000	9.1090	.0971	-.0158	.0912	.1170
1378400	9.1100	.1803	-.0355	.1781	.1584
1378100	9.1125	.1862	-.1309	.1338	.2226
1377825	9.1135	.0733	-.0277	.0926	.1475
1377500	9.1145	.5637	.8144	-.0985	.0404
1377280	9.1155	.0444	-.0050	.0128	.0178
1377230	9.1180	.0320	-.0011	.0114	.0005

TABLE 24 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF CN-6

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.	4th Comp.
1742100	9.0535	.2268	-.1513	-.5634	-.7013
1740000	9.0575	.2156	-.2566	-.6879	.6239
1720000	9.0595	.0422	-.0478	.0304	-.0128
166300	9.0645	.0375	-.0220	-.0413	.0318
1594260	9.0780	.0528	-.0362	-.0035	.0583
1594236	9.0785	.0233	-.0100	.0220	.0898
1590000	9.0850	.7062	-.4663	.3442	-.1245
1379000	9.1090	.1005	-.0147	.1038	.1050
1378400	9.1100	.1825	-.0337	.1947	.1356
1378100	9.1125	.1858	-.1261	.1495	.2038
1377825	9.1135	.0710	-.0262	.0944	.1331
1377500	9.1145	.5577	.8196	-.0869	.0420
1377280	9.1155	.0433	-.0047	.0120	.0192
1377230	9.1180	.0313	-.0008	.0125	-.0051

TABLE 25 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF S-4

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.
1278800	9.1650	-.1608	-.0738	-.8889
1278050	9.1665	-.5304	.8066	-.0525
1272445	9.1725	-.2180	-.4039	-.2817
1077090	9.3440	-.1027	.0634	-.1532
1073480	9.3575	-.0754	-.0045	.1153
1073436	9.3615	-.7931	-.4205	.3017

TABLE 26 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF S-6

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.	3rd Comp.
1278800	9.1650	-.1622	-.1421	-.8496
1278050	9.1665	-.5207	.8186	-.1102
1272445	9.1725	-.2240	-.3730	-.3252
1077090	9.3440	-.1013	.0618	-.1660
1073480	9.3575	-.0802	-.0038	.1155
1073436	9.3615	-.7973	-.4084	.3456

TABLE 27 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF CS-4

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.
1425625	9.0975	-.1341	.0453
1377280	9.1155	-.1167	.2714
1377230	9.1180	-.0799	.1717
1377200	9.1190	-.3879	.8378
1373900	9.1275	-.0658	.0819
1373055	9.1325	-.8906	-.4286
1373020	9.1345	-.0859	-.0353
1371815	9.1430	-.0537	.0138
1370300	9.1520	-.0616	.0331

TABLE 28 - COEFFICIENTS FOR THE PRINCIPAL COMPONENTS OF CS-6

CSU Sta. No.	USGS Sta. No.	1st Comp.	2nd Comp.
1425625	9.0975	-.1278	-.0209
1377280	9.1155	-.1201	-.2711
1377230	9.1180	-.0831	-.1661
1377200	9.1190	-.3994	-.8374
1373900	9.1275	-.0625	-.0688
1373055	9.1325	-.8857	.4363
1373020	9.1345	-.0853	.0376
1371815	9.1430	-.0567	-.0003
1370300	9.1520	-.0621	-.0283

TABLE 29 - CUMULATIVE PERCENTAGE OF TOTAL VARIATION ACCOUNTED FOR BY THE PRINCIPAL COMPONENTS

Type	Principal component	Cumulative percentage of total variation accounted for
N-4 series	ξ_1	85
	ξ_1 and ξ_2	97
	ξ_1, ξ_2 and ξ_3	98
	ξ_1, ξ_2, ξ_3 and ξ_4	99
N-6 series	ξ_1	85
	ξ_1 and ξ_2	97
	ξ_1, ξ_2 and ξ_3	98
	ξ_1, ξ_2, ξ_3 and ξ_4	99
CN-4 series	η_1	82
	η_1 and η_2	94
	η_1, η_2 and η_3	98
	η_1, η_2, η_3 and η_4	99
CN-6 series	η_1	83
	η_1 and η_2	95
	η_1, η_2 and η_3	98
	η_1, η_2, η_3 and η_4	99
S-4 series	ξ_1	97
	ξ_1 and ξ_2	98
	ξ_1, ξ_2 and ξ_3	99
S-6 series	ξ_1	97
	ξ_1 and ξ_2	98
	ξ_1, ξ_2 and ξ_3	99
CS-4 series	η_1	95
	η_1 and η_2	99
CS-6 series	η_1	95
	η_1 and η_2	99

The means and standard deviations of the series of the principal components for N-4, N-6, CN-4, CN-6, S-4, S-6, CS-4 and CS-6 are given in Table 30.

It is simply proven [12] that if all the means in the target areas during the seeded period have been increased by a certain fraction of the old means, say h, that is, the increase of Q_1 is hQ_1, of Q_2 is hQ_2, and so on, then the increase in the means of the principal components will also be h. If h is assigned a value of 0.10, then

$$E\{\xi_i^*\} = 1.1\ E\{\xi_i\} \quad ,$$

where $E\{\}$ denotes the expected value of $\{\}$, which is the cloud seeding effect assumed in this study.

TABLE 30 - MEANS AND STANDARD DEVIATIONS OF THE PRINCIPAL COMPONENTS

Type	Principal component	Mean (cfs)	Std. Dev. (cfs)
N-4 series	ξ_1	1316.896	385.728
	ξ_2	23.431	144.526
	ξ_3	-103.167	50.427
	ξ_4	-55.122	37.622
N-6 series	ξ_1	919.996	289.498
	ξ_2	-7.066	106.279
	ξ_3	-103.770	39.834
	ξ_4	-19.400	27.483
CN-4 series	η_1	3566.570	1171.757
	η_2	-616.325	446.487
	η_3	-124.669	238.558
	η_4	41.293	146.087
CN-6 series	η_1	2656.142	853.439
	η_2	-442.328	326.332
	η_3	-118.678	169.248
	η_4	49.830	101.846
S-4 series	ξ_1	-2092.706	865.153
	ξ_2	-95.873	108.688
	ξ_3	30.022	81.462
S-6 series	ξ_1	-1538.060	601.913
	ξ_2	-71.786	74.436
	ξ_3	28.886	55.548
CS-4 series	η_1	-1190.877	459.172
	η_2	-146.711	86.254
CS-6 series	η_1	-849.227	315.594
	η_2	85.939	61.805

For the control region, it is obvious that following the assumption that the means of the runoff stations in the control region remain unchanged,

$$E\{\eta_i^*\} = E\{\eta_i\}$$

where η_i^* is the i^{th} principal component of the control region during the seeded period.

After the principal components in each separate region have been obtained, they are gathered into four major target-control groups as N-4 and CN-4, N-6 and CN-6, S-4 and CS-4, and S-6 and CS-6. For brevity, after the principal components in the target are combined with those in the control, the following symbols will be used:

N-CN-4 - the combination of N-4 and CN-4

N-CN-6 - the combination of N-6 and CN-6

S-CS-4 - the combination of S-4 and CS-4

S-CS-6 - the combination of S-6 and CS-6.

Since it is the principal components that will be utilized in the final test, the computations of the covariance matrices are carried out for these principal components. These are as shown in Tables 31, 32, 33, and 34; also shown are the correlations matrices in Tables 35, 36, 37, and 38.

TABLE 31 - COVARIANCE MATRIX OF N-CN-4 PRINCIPAL COMPONENT SERIES

		N-4				CN-4			
		ξ_1	ξ_2	ξ_3	ξ_4	η_1	η_2	η_3	η_4
N-4	ξ_1	148786.103	-5.194	-1.641	-4.525	361585.644	-32549.149	-7106.856	-8545.409
	ξ_2	-5.194	20887.828	-.361	-.075	31294.727	1427.066	-4668.399	-2250.743
	ξ_3	-1.641	-.361	2542.946	.155	-18240.198	5629.908	2877.776	-679.743
	ξ_4	-4.525	-.075	.155	1415.475	4402.075	-4938.812	-1612.864	149.454
CN-4	η_1	361585.644	31294.727	-18240.198	4402.075	1373015.745	-71.56	57.238	-49.204
	η_2	-32549.149	1427.066	5629.908	-4938.812	-71.561	199350.870	.050	-3.719
	η_3	-7106.856	-4668.399	2877.776	-1612.864	57.238	.050	56910.290	2.404
	η_4	-8545.409	-2250.743	-679.743	149.454	-49.204	-3.719	2.404	21341.511

TABLE 32 - COVARIANCE MATRIX OF N-CN-6 PRINCIPAL COMPONENT SERIES

		N-6				CN-6			
		ξ_1	ξ_2	ξ_3	ξ_4	η_1	η_2	η_3	η_4
N-6	ξ_1	83809.330	3.376	15.820	1.313	191098.859	-20577.629	-10540.852	-2678.418
	ξ_2	3.376	11295.421	2.644	.066	13712.454	3162.736	-2003.334	-963.144
	ξ_3	15.820	2.644	1586.754	.068	-12182.229	1678.320	24.742	-126.575
	ξ_4	1.313	.066	.068	755.354	3884.894	-2718.278	-993.925	121.930
CN-6	η_1	191098.859	13712.454	-12182.229	3884.894	728358.925	20.301	20.187	-14.259
	η_2	-20577.629	3162.736	1678.320	-2718.278	20.301	106493.191	-5.314	-.917
	η_3	-10540.852	-2003.334	24.742	-993.925	20.187	-5.314	28645.205	-.067
	η_4	-2678.418	-963.144	-126.575	121.930	-14.259	-.917	-.067	10372.757

TABLE 33 - COVARIANCE MATRIX OF S-CS-4 PRINCIPAL COMPONENT SERIES

		S-4			CS-4	
		ξ_1	ξ_2	ξ_3	η_1	η_2
S-4	ξ_1	748491.385	-45.282	41.666	338072.405	-12524.935
	ξ_2	-45.282	11813.224	-.485	-11808.907	-1953.006
	ξ_3	-41.666	-.485	6636.209	-636.724	-167.057
CS-4	η_1	338072.405	-11808.907	-636.724	210839.108	7.238
	η_2	-12524.935	-1953.006	-167.057	7.238	7439.863

TABLE 34 - COVARIANCE MATRIX OF S-CS-6 PRINCIPAL COMPONENT SERIES

		S-6			CS-6	
		ξ_1	ξ_2	ξ_3	η_1	η_2
S-6	ξ_1	362299.490	9.116	18.085	162180.856	5816.690
	ξ_2	9.116	5540.858	-.073	-5460.831	1113.406
	ξ_3	18.085	-.073	3085.627	10.105	11.309
CS-6	η_1	162180.856	-5460.831	10.105	99600.481	-.854
	η_2	5816.690	1113.406	11.309	-.854	3819.884

TABLE 35 - CORRELATION MATRIX OF N-CN-4 PRINCIPAL COMPONENT SERIES

		N-4				CN-4			
		ξ_1	ξ_2	ξ_3	ξ_4	η_1	η_2	η_3	η_4
N-4	ξ_1	1.000	-.000	-.000	-.000	.800	-.189	-.077	-.152
	ξ_2	-.000	1.000	-.000	-.000	.185	.022	-.135	-.107
	ξ_3	-.000	-.000	1.000	.000	-.309	.250	.239	-.092
	ξ_4	-.000	-.000	.000	1.000	.100	-.294	-.180	.027
CN-4	η_1	.800	.185	-.309	.100	1.000	-.000	.000	-.000
	η_2	-.189	.022	.250	-.294	-.000	1.000	.000	-.000
	η_3	-.077	-.135	.239	-.180	.000	.000	1.000	.000
	η_4	-.152	-.107	-.092	.027	-.000	-.000	.000	1.000

TABLE 36 - CORRELATION MATRIX OF N-CN-6 PRINCIPAL COMPONENT SERIES

		N-6				CN-6			
		ξ_1	ξ_2	ξ_3	ξ_4	η_1	η_2	η_3	η_4
N-6	ξ_1	1.000	.000	.001	.000	.773	-.218	-.215	-.091
	ξ_2	.000	1.000	.001	.000	.151	.091	-.111	-.089
	ξ_3	.001	.001	1.000	.000	-.358	.129	.004	-.031
	ξ_4	.000	.000	.000	1.000	.166	-.303	-.214	.044
CN-6	η_1	.773	.151	-.358	.166	1.000	.000	.000	-.000
	η_2	-.218	.091	.129	-.303	.000	1.000	-.000	-.000
	η_3	-.215	-.111	.004	-.214	.000	-.000	1.000	-.000
	η_4	-.091	-.089	-.031	.044	-.000	-.000	-.000	1.000

TABLE 37 - CORRELATION MATRIX OF S-CS-4 PRINCIPAL COMPONENT SERIES

		S-4			CS-4	
		ξ_1	ξ_2	ξ_3	η_1	η_2
S-4	ξ_1	1.000	-.000	-.001	.851	-.168
	ξ_2	-.000	1.00	-.000	-.237	-.208
	ξ_3	-.001	-.000	1.000	-.017	-.024
CS-4	η_1	.851	-.237	-.017	1.000	.000
	η_2	-.168	-.208	-.024	.000	1.000

TABLE 38 - CORRELATION MATRIX OF S-CS-6 PRINCIPAL COMPONENT SERIES

		ξ_1	ξ_2	ξ_3	η_1	η_2
		S-6			CS-6	
S-6	ξ_1	1.000	.000	.001	.854	.156
	ξ_2	.000	1.000	-.000	-.232	.242
	ξ_3	.001	-.000	1.000	.001	.003
CS-6	η_1	.854	-.232	.001	1.000	-.000
	η_2	.156	.242	.003	-.000	1.000

5.2 The minimum number of years needed to detect a 10% increase in runoff based on the principal components. The minimum number of years, N^*, for detecting the increase of one-tenth of the old runoff means can be computed from equation (1) again,

$$N^* = \frac{\tau^2}{\underline{\mu}'\underline{V}^{-1}\underline{\mu}} \tag{24}$$

where τ^2 = the noncentrality parameter,

$\underline{\mu}$ = $\underline{\mu}^* - \underline{\mu}_o$,

$\underline{\mu}^*$ = the mean vector of the runoff variables for the seeded period,

$\underline{\mu}_o$ = the mean vector of the runoff variables for the period before seeding, and

$\underline{V}^{-1}$ = the inverse of covariance matrix $\underline{V}$.

The values of τ^2 are given in Table 1.

With this table the number of years needed to detect the increase can be computed easily. The values of N^* are shown in Table 39.

5.3 The application of canonical analysis. In this analysis the set of the runoff variables in the target region is first combined with the set of those in the control region. As for the principal component analysis, the computation of the canonical variables were performed on the CDC 6400 digital computer of the University Computer Center at Colorado State University. The steps in finding the coefficients for the canonical variables were described in Chapter III Section 3.

After the coefficients of the canonical variables for N-4, N-6, CN-4, CN-6, S-4, S-6, CS-4 and CS-6 are all computed and tabulated in Tables 40-47, the canonical series of each region are easily calculated [12].

TABLE 39 - MINIMUM NUMBER OF YEARS TO DETECT THE INCREASE OF 10 PERCENT IN RUNOFF MEAN USING PRINCIPAL COMPONENTS

Type	No. of principal components in target	No. of principal components in control	Value of $\underline{\mu}'\underline{V}^{-1}\underline{\mu}$	τ^2	Minimum number of years to detect the increase, N*	Remarks
N-CN-4	4	4	1.066	11.655	11	The minimum value of N* is obtained from the larger of $N^* = \tau^2/\underline{\mu}'\underline{V}^{-1}\underline{\mu}$ or $N^* = k + 1$ where k is the total number of components in both target and control regions
N-CN-6	4	4	0.813	11.655	15	
S-CS-4	3	2	0.243	8.640	36	
S-CS-6	3	2	0.273	8.640	32	

TABLE 40 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF N-4

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1970000	9.0105	-.003956	-.006628	-.003592	.018543
1960000	9.0110	.003128	-.009783	.011935	-.042535
1866000	9.0165	.005767	-.004685	.026278	.009310
1830000	9.0190	.000796	.003972	-.002342	.002199
1820000	9.0195	-.001320	-.002450	-.001804	.001937
1802730	9.0265	.008752	.008348	.024461	-.012694
1776000	9.0360	.008385	.002618	-.023413	.019209

TABLE 41 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF N-6

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1970000	9.0105	-.006033	-.009805	-.005114	.032451
1960000	9.0110	.004802	-.007516	.011799	-.033462
1866000	9.0165	.009597	.005297	.041092	-.001293
1830000	9.0190	.001003	.003991	-.004721	.002069
1820000	9.0195	-.001910	-.003885	-.001016	.002457
1802730	9.0265	.013825	.025201	.014417	.035857
1776000	9.0360	.010705	-.008078	-.021553	-.008330

TABLE 42 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF CN-4

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1742100	9.0535	.001564	.001900	.003294	.003207
1740000	9.0575	.000620	-.000087	-.002110	-.002931
1720000	9.0595	.000086	-.001640	-.004363	.002621
1666300	9.0645	-.001139	.015690	.001530	.004480
1594260	9.0780	.001374	.001985	-.002575	.003694
1594236	9.0785	-.003596	-.040136	.019047	.013573
1590000	9.0850	.000525	.000354	-.001849	-.001949
1379000	9.1090	.002959	.029446	-.003503	.007790
1378400	9.1100	-.004647	-.030526	.005096	-.010398
1378100	9.1125	.001847	.006202	.009424	.002551
1377825	9.1135	.001334	.011723	.005682	-.002092
1377500	9.1145	-.000174	.000685	-.001047	.000878
1377280	9.1155	-.003380	-.010015	-.008777	-.003425
1377230	9.1180	.008358	.033933	.021453	.033986

TABLE 43 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF CN-6

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1742100	9.0535	.002365	-.001262	.006282	.000515
1740000	9.0575	.000555	-.000926	-.004814	.000489
1720000	9.0595	.000399	-.003421	-.005800	-.000807
1666300	9.0645	-.002081	.019407	-.009901	.012097
1594260	9.0780	.002055	.004230	-.001362	.003426
1594236	9.0785	-.003831	-.049581	.040428	.029201
1590000	9.0850	.000478	-.000866	-.002470	-.002815
1379000	9.1090	.006095	.041344	-.006513	-.013108
1378400	9.1100	-.008394	-.045125	.001749	-.003420
1378100	9.1125	.004031	.013690	.010848	.005814
1377825	9.1135	.001566	.017428	-.000181	-.005258
1377500	9.1145	-.000219	.000494	-.001734	.001375
1377280	9.1155	-.005293	-.011148	-.003944	.002038
1377230	9.1180	.013811	.053074	.011299	.031200

TABLE 44 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF S-4

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1278800	9.1650	-.000949	.010797	-.004734	.000415
1278050	9.1665	.002273	-.002086	.003651	-.002148
1272445	9.1725	.000895	.002056	.008422	-.004012
1077090	9.3440	.000256	.009180	.003705	-.009945
1073480	9.3575	.007460	.009551	-.023825	-.047496
1073436	9.3615	-.003435	-.003435	-.002076	.008598

TABLE 45 - COEFFICIENTS FOR THE CANONICAL VARIABLES OF S-6

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1278800	9.1650	-.001790	.017228	.001080	-.003766
1278050	9.1665	.003301	-.004374	.004401	-.006282
1272445	9.1725	.001264	-.000937	.010721	-.014541
1077090	9.3440	.000707	.007274	-.011061	.011509
1073480	9.3575	.014087	.018233	-.052315	-.056559
1073436	9.3615	-.001675	-.002813	.000808	.013618

TABLE 46.- COEFFICIENTS FOR THE CANONICAL VARIABLES OF CS-4

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1425625	9.0975	.001734	-.004611	.004333	-.005155
1377280	9.1155	-.003347	.055553	.035314	-.034470
1377230	9.1180	-.005054	.005968	.000726	.005480
1377200	9.1190	.003365	-.014545	-.017608	.013770
1373900	9.1275	.000054	-.003457	.002488	-.02948
1373055	9.1325	.000225	.000372	-.000186	-.000378
1373020	9.1345	.002328	.007410	-.007410	-.022485
1371815	9.1430	.004076	.010501	.000507	.023824
1370300	9.1520	.010696	.012852	.036629	.024040

TABLE 47.- COEFFICIENTS FOR THE CANONICAL VARIABLES OF CS-6

CSU Sta. No.	USGS Sta. No.	1st Variable	2nd Variable	3rd Variable	4th Variable
1425625	9.0975	.003565	-.010123	.004390	-.000819
1377280	9.1155	-.007498	.056512	.054781	-.081378
1377230	9.1180	-.006890	.014796	-.000303	.016373
1377200	9.1190	.005324	-.012051	-.020296	.025777
1373900	9.1275	.004037	-.053190	-.028107	-.030019
1373055	9.1325	.000329	.001040	.000236	-.003660
1373020	9.1345	.004450	-.000957	-.031094	.005098
1371815	9.1430	.005299	.012212	.024251	.036120
1370300	9.1520	.010325	.000287	.053491	.006772

The series of the canonical variables are tabulated in Tables 48-55 for N-4, N-6, CN-4, CN-6, S-4, S-6, CS-4, and CS-6, respectively. The means and standard deviations of the canonical series are shown in Table 56.

TABLE 48 - N-4 CANONICAL SERIES (cfs)

Year	ζ_1	ζ_2	ζ_3	ζ_4
1938	5.356	-2.195	-1.045	.827
1939	2.963	-2.067	-1.608	1.737
1940	2.621	-1.618	.602	1.452
1941	4.203	-1.884	3.477	2.474
1942	4.103	- .582	- .923	5.512
1943	3.128	-2.357	- .364	1.256
1944	2.971	-2.235	.303	1.703
1945	3.402	-1.408	- .306	2.703
1946	3.167	-1.729	.509	4.003
1947	5.015	- .291	-1.002	2.790
1948	3.835	.538	- .802	1.186
1949	3.636	-3.039	.383	3.314
1950	3.739	.635	- .523	2.882
1951	4.328	-2.584	.609	2.509
1952	4.910	-2.824	.700	2.171
1953	2.966	-2.111	.030	3.728
1954	1.655	- .943	.533	2.180
1955	2.723	- .476	.674	2.865
1956	3.098	-1.871	- .248	2.906
1957	4.865	-2.033	.654	3.208
1958	3.365	-2.401	.150	3.345
1959	2.979	-1.403	.070	2.835
1960	2.855	-3.081	-1.656	3.486
1961	2.184	-2.026	- .791	2.903
1962	3.978	-3.188	-1.473	1.454
1963	1.504	-1.735	.921	2.158
1964	2.148	-2.119	- .679	3.336
1965	3.473	-3.180	- .720	3.962
1966	1.651	-1.613	- .401	1.842
1967	2.634	-2.746	- .196	2.709

TABLE 49 - N-6 CANONICAL SERIES (cfs)

Year	ζ_1	ζ_2	ζ_3	ζ_4
1938	5.220	-2.656	.059	- .505
1939	2.893	-2.476	- .902	1.870
1940	2.679	-1.385	1.301	.978
1941	4.411	- .711	4.330	.885
1942	4.138	- .805	- .152	3.364
1943	3.198	-2.354	.342	1.723
1944	3.012	-1.995	1.019	1.780
1945	3.747	-1.514	.188	2.513
1946	3.236	-1.534	1.364	2.704
1947	5.197	- .650	- .825	1.721
1948	3.884	.209	- .482	- .298
1949	3.712	-2.711	1.024	3.491
1950	3.721	.334	- .572	1.811
1951	4.513	-2.259	1.273	2.446
1952	5.098	-2.365	1.539	1.697
1953	3.143	-1.963	.609	3.348
1954	1.736	- .761	.828	1.569
1955	2.933	- .278	.877	1.845
1956	3.152	-1.874	.550	1.478
1957	5.089	-1.963	1.283	1.943
1958	3.370	-2.207	.873	2.707
1959	3.080	-1.422	1.006	.546
1960	2.774	-3.571	- .238	1.683
1961	2.289	-2.217	.168	1.377
1962	4.004	-3.549	- .196	.046
1963	1.706	-1.470	1.505	1.636
1964	2.227	-2.279	.342	1.726
1965	3.805	-3.339	1.237	1.362
1966	1.851	-1.664	.974	.259
1967	2.803	-2.707	1.528	.894

TABLE 50 - CN-4 CANONICAL SERIES (cfs)

Year	ε_1	ε_2	ε_3	ε_4
1938	5.491	.115	-1.004	- .420
1939	2.734	- .116	-1.548	1.356
1940	2.615	.354	.971	1.517
1941	4.068	.150	3.984	1.335
1942	4.074	2.026	- .968	4.727
1943	3.220	.655	.902	- .005
1944	3.106	- .109	.309	1.172
1945	3.502	1.417	.684	1.071
1946	3.261	.466	- .008	.919
1947	5.064	2.293	- .718	1.325
1948	4.185	2.310	- .353	.906
1949	3.901	- .062	.090	1.921
1950	3.628	2.314	.803	.889
1951	4.410	.050	.907	1.639
1952	4.792	.306	.297	1.187
1953	3.284	-1.026	.337	2.767
1954	1.706	1.475	.351	.867
1955	2.617	1.851	.501	2.010
1956	3.454	.852	.812	1.785
1957	5.042	- .562	.234	1.013
1958	3.567	-0.	1.252	1.950
1959	3.244	1.619	- .080	1.708
1960	2.788	-1.017	- .883	1.251
1961	2.161	.360	- .346	.701
1962	3.934	- .921	-1.099	.704
1963	1.605	.767	.207	1.260
1964	2.151	.491	.053	1.798
1965	3.539	- .812	.163	3.907
1966	1.896	.599	- .422	.600
1967	2.791	- .149	.551	1.160

TABLE 51 - CN-6 CANONICAL SERIES (cfs)

Year	ε_1	ε_2	ε_3	ε_4
1938	5.435	- .743	- .753	-1.522
1939	2.761	- .672	-1.612	2.013
1940	2.720	.376	.907	1.868
1941	4.347	1.186	3.955	.553
1942	4.185	1.426	-1.182	3.186
1943	3.364	.413	.095	1.183
1944	3.191	- .404	.263	1.457
1945	3.889	.973	- .041	1.240
1946	3.395	- .020	- .033	.823
1947	5.229	1.489	-1.339	.433
1948	4.290	1.883	- .778	- .139
1949	4.025	- .170	- .034	1.808
1950	3.739	2.145	- .209	.809
1951	4.659	.076	.437	2.087
1952	5.008	.190	- .012	1.158
1953	3.472	-1.173	.497	2.826
1954	1.825	1.054	.052	- .140
1955	2.919	1.474	- .004	1.060
1956	3.554	.719	.244	1.561
1957	5.326	- .878	.242	1.018
1958	3.617	.154	.726	2.551
1959	3.364	1.098	- .704	1.214
1960	2.811	-1.515	- .516	.645
1961	2.313	- .274	- .274	.450
1962	3.998	-1.686	- .719	.351
1963	1.888	.542	- .128	1.010
1964	2.322	.302	.169	.812
1965	3.921	-1.187	1.097	1.002
1966	2.035	.174	- .273	- .195
1967	3.040	- .127	1.315	- .522

TABLE 52 - S-4 CANONICAL SERIES (cfs)

Year	ζ_1	ζ_2	ζ_3	ζ_4
1938	4.095	.364	-2.671	-4.194
1939	2.327	.148	- .980	-5.317
1940	2.453	- .015	-1.387	-4.918
1941	4.406	2.663	.354	-2.341
1942	4.391	- .427	-1.419	-4.347
1943	2.809	2.182	-2.294	-3.036
1944	3.989	- .096	-1.276	-1.624
1945	2.834	3.730	-2.042	-3.576
1946	2.414	- .203	-1.678	-4.181
1947	3.075	.861	-3.444	-4.148
1948	3.735	.191	-2.762	-3.410
1949	3.437	.397	-4.331	-1.361
1950	2.313	- .805	- .805	-2.759
1951	1.656	1.645	-2.952	-4.037
1952	4.550	1.260	-3.162	-3.385
1953	1.937	1.592	-1.501	-3.254
1954	1.450	.972	-2.460	-2.073
1955	1.786	1.169	-1.523	-2.373
1956	2.116	1.677	-2.476	-4.210
1957	4.638	2.586	-2.328	-5.720
1958	3.719	1.944	- .092	-2.539
1959	1.478	1.392	-2.539	-4.031
1960	2.794	1.582	-2.113	-3.544
1961	2.280	1.493	-1.499	-3.458
1962	2.926	1.492	-2.984	-3.647
1963	1.642	1.144	-2.166	-3.056
1964	1.992	1.696	-1.201	-3.911
1965	3.494	1.786	-3.343	-2.664
1966	2.378	1.268	-2.392	-3.432
1967	1.625	1.464	-1.932	-3.333

TABLE 53 - S-6 CANONICAL SERIES (cfs)

Year	ζ_1	ζ_2	ζ_3	ζ_4
1938	4.315	.650	-3.827	-1.785
1939	2.608	- .295	-3.195	-3.121
1940	2.712	- .147	-3.104	-3.009
1941	4.493	1.887	- .200	-1.089
1942	4.659	- .473	-3.052	-1.967
1943	3.037	2.460	-2.033	-2.540
1944	4.106	.041	-1.266	-1.033
1945	2.979	3.682	-1.863	-2.466
1946	2.701	- .155	-2.611	-3.445
1947	3.413	1.457	-3.637	-3.310
1948	3.950	.576	-3.158	-2.331
1949	3.596	1.477	-3.813	.509
1950	2.501	- .872	-1.819	-1.651
1951	1.887	2.023	-3.754	-2.471
1952	4.765	1.748	-3.567	-1.583
1953	2.149	1.476	-2.031	-2.269
1954	1.611	1.328	-2.774	- .413
1955	1.982	1.232	-1.707	-1.486
1956	2.338	1.780	-3.107	-2.676
1957	5.055	2.684	-3.450	-4.024
1958	3.792	1.383	.019	-2.808
1959	1.737	1.697	-3.562	-2.468
1960	2.923	1.542	-2.718	-1.512
1961	2.436	1.360	-2.011	-2.307
1962	3.145	1.739	-3.473	-1.878
1963	1.881	1.324	-2.507	-1.724
1964	2.215	1.561	-1.735	-3.363
1965	3.766	2.294	-3.333	- .807
1966	2.578	1.285	-3.051	- .899
1967	1.901	1.538	-2.819	-1.276

TABLE 54 - CS-4 CANONICAL SERIES (cfs)

Year	ε_1	ε_2	ε_3	ε_4
1938	2.930	1.215	1.510	-3.400
1939	1.591	2.091	1.492	-1.727
1940	1.759	.584	3.668	-2.275
1941	3.394	3.652	3.261	- .218
1942	3.803	.792	1.554	-1.400
1943	1.592	2.963	.687	-1.611
1944	3.313	1.341	2.705	-1.182
1945	2.203	4.083	2.156	- .240
1946	1.248	1.090	1.964	-1.471
1947	2.324	2.780	.819	- .605
1948	2.743	.375	.301	- .867
1949	2.928	1.488	- .815	.334
1950	1.582	1.640	.697	-1.436
1951	1.172	3.505	1.263	-1.684
1952	3.543	3.331	.329	- .868
1953	1.218	3.747	1.844	-1.956
1954	.790	1.778	1.233	- .007
1955	1.242	2.389	1.687	- .296
1956	1.121	2.417	.685	-1.332
1957	4.062	3.149	- .353	-3.382
1958	3.096	2.103	2.758	- .973
1959	.921	1.808	1.273	-3.346
1960	1.737	3.652	1.126	-2.610
1961	.962	2.936	1.963	-1.602
1962	2.434	2.865	1.149	-2.768
1963	.634	1.824	.478	-1.287
1964	1.480	2.019	1.262	-2.769
1965	2.462	2.433	.434	-1.807
1966	1.234	1.464	.990	- .835
1967	.924	1.732	.216	-1.053

TABLE 55 - CS-6 CANONICAL SERIES (cfs)

Year	ε_1	ε_2	ε_3	ε_4
1938	3.125	.666	- .268	-1.690
1939	1.867	1.119	.241	- .796
1940	1.871	- .579	1.762	-2.090
1941	3.368	2.396	2.874	.131
1942	3.969	.358	.766	- .335
1943	1.786	2.652	.400	-1.213
1944	3.375	.546	2.335	-1.515
1945	2.292	3.207	2.501	- .102
1946	1.463	.475	1.185	-1.301
1947	2.513	2.392	.927	.182
1948	2.948	.382	.056	- .645
1949	3.056	1.832	- .294	.970
1950	1.764	1.431	.319	- .692
1951	1.245	3.109	1.424	-2.270
1952	3.640	3.039	.844	- .329
1953	1.387	3.068	1.520	-2.353
1954	.848	1.630	1.126	.792
1955	1.357	2.024	1.420	.655
1956	1.223	2.196	.605	-1.273
1957	4.357	3.251	-1.240	-2.029
1958	3.165	1.284	2.821	-1.556
1959	1.145	1.309	- .355	-1.342
1960	1.855	2.945	.362	- .880
1961	1.025	2.263	1.682	-1.794
1962	2.618	2.358	.671	-2.190
1963	.816	1.432	- .026	- .750
1964	1.645	1.454	.473	-2.581
1965	2.718	2.316	- .082	-1.514
1966	1.418	1.063	.549	- .018
1967	1.119	1.632	- .258	.303

TABLE 56 - MEANS AND STANDARD DEVIATIONS OF CANONICAL VARIABLES

Type	Canonical Variable	Mean (cfs)	Std. Dev. (cfs)
N-4 series	ζ_1	3.315	1.000
	ζ_2	-1.819	1.000
	ζ_3	-0.104	1.000
	ζ_4	2.648	1.000
N-6 series	ζ_1	3.421	1.000
	ζ_2	-1.804	1.000
	ζ_3	.695	1.000
	ζ_4	1.620	1.000
CN-4 series	ε_1	3.394	1.000
	ε_2	0.523	1.000
	ε_3	0.199	1.000
	ε_4	1.434	1.000
CN-6 series	ε_1	3.555	1.000
	ε_2	.227	1.000
	ε_3	.046	1.000
	ε_4	1.020	1.000
S-4 series	ζ_1	2.825	1.000
	ζ_2	1.172	1.000
	ζ_3	-2.047	1.000
	ζ_4	-3.463	1.000
S-6 series	ζ_1	3.041	1.000
	ζ_2	1.276	1.000
	ζ_3	-2.639	1.000
	ζ_4	-2.040	1.000
CS-4 series	ε_1	2.015	1.000
	ε_2	2.241	1.000
	ε_3	1.278	1.000
	ε_4	-1.489	1.000
CS-6 series	ε_1	2.166	1.000
	ε_2	1.775	1.000
	ε_3	.811	1.000
	ε_4	-0.941	1.000

Similar to the principal component analysis, it is clear now that,

$$E\{\zeta_i^*\} = (1+h)\ E\{\zeta_i\}$$

where 100h is the percent increase of the runoff means in the target region. If $h = 0.10$, then,

$$E\{\zeta_i^*\} = 1.1\ E\{\zeta_i\}$$

and

$$E\{\varepsilon_i^*\} = E\{\varepsilon_i\}$$

where ε_i^* is the i^{th} canonical variable of the control region for the seeded period.

The covariance matrices of N-CN-4, N-CN-6, S-CS-4, and S-CS-6 are shown in Tables 57-60, respectively. In this analysis the correlation matrices are the same as the covariance matrices since all the canonical variables have unit variances.

5.4 The minimum number of years needed to detect a 10% increase in runoff based on the canonical variables. As discussed before in Section 5.2, the minimum number of years needed to detect the increase can be obtained with the use of Table 1, which gives the value of τ^2 . After the canonical analysis has been performed because the high correlation between target and control variables are desirable here, only the highly correlated canonical variables will be retained for further study.

For example, consider the case of S-CS-4. The correlation between the first canonical variable in S-4 and the first canonical variable in CS-4 is found to be 0.968, which is the maximum of all the correlations between the canonical variables for S-CS-4. If it is decided to use only these two canonical variables in the test, then all one needs to do is the following. From Table 56, obtain

$$\underline{\mu}_o = \begin{bmatrix} 2.825 \\ 2.015 \end{bmatrix} .$$

TABLE 57 - COVARIANCE MATRIX OF N-CN-4 CANONICAL SERIES

		N-4				CN-4			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
N-4	ζ_1	1.000	0.	0.	0.	.989	0.	0.	0.
	ζ_2	0.	1.000	0.	0.	0.	.890	0.	0.
	ζ_3	0.	0.	1.000	0.	0.	0.	.847	0.
	ζ_4	0.	0.	0.	1.000	0.	0.	0.	.767
CN-4	ε_1	.989	0.	0.	0.	1.000	0.	0.	0.
	ε_2	0.	.890	0.	0.	0.	1.000	0.	0.
	ε_3	0.	0.	.847	0.	0.	0.	1.000	0.
	ε_4	0.	0.	0.	.767	0.	0.	0.	1.000

TABLE 58 - COVARIANCE MATRIX OF N-CN-6 CANONICAL SERIES

		N-6				CN-4			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
N-6	ζ_1	1.000	0.	0.	0.	.990	0.	0.	0.
	ζ_2	0.	1.000	0.	0.	0.	.894	0.	0.
	ζ_3	0.	0.	1.000	0.	0.	0.	.869	0.
	ζ_4	0.	0.	0.	1.000	0.	0.	0.	.768
CN-6	ε_1	.990	0.	0.	0.	1.000	0.	0.	0.
	ε_2	0.	.894	0.	0.	0.	1.000	0.	0.
	ε_3	0.	0.	.869	0.	0.	0.	1.000	0.
	ε_4	0.	0.	0.	.768	0.	0.	0.	1.000

TABLE 59 - COVARIANCE MATRIX OF S-CS-4 CANONICAL SERIES

		S-4				CS-4			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
S-4	ζ_1	1.000	0.	0.	0.	.968	0.	0.	0.
	ζ_2	0.	1.000	0.	0.	0.	.771	0.	0.
	ζ_3	0.	0.	1.000	0.	0.	0.	.703	0.
	ζ_4	0.	0.	0.	1.000	0.	0.	0.	.617
CS-4	ε_1	.968	0.	0.	0.	1.000	0.	0.	0.
	ε_2	0.	.771	0.	0.	0.	1.000	0.	0.
	ε_3	0.	0.	.703	0.	0.	0.	1.000	0.
	ε_4	0.	0.	0.	.617	0.	0.	0.	1.000

TABLE 60 - COVARIANCE MATRIX OF S-CS-6 CANONICAL SERIES

		S-6				CS-6			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
S-6	ζ_1	1.000	0.	0.	0.	.969	0.	0.	0.
	ζ_2	0.	1.000	0.	0.	0.	.777	0.	0.
	ζ_3	0.	0.	1.000	0.	0.	0.	.696	0.
	ζ_4	0.	0.	0.	1.000	0.	0.	0.	.568
CS-6	ε_1	.969	0.	0.	0.	1.000	0.	0.	0.
	ε_2	0.	.777	0.	0.	0.	1.000	0.	0.
	ε_3	0.	0.	.696	0.	0.	0.	1.000	0.
	ε_4	0.	0.	0.	.568	0.	0.	0.	1.000

Assuming that there is an increase of 10% in the means of the target region and the means in the control region remain unchanged, then, the mean vector for the seeding period can be obtained as

$$\underline{\mu}^* = \begin{bmatrix} 3.107 \\ 2.015 \end{bmatrix} .$$

Now $\underline{\mu} = (\underline{\mu}^* - \underline{\mu}_o)$, that is,

$$\underline{\mu} = \begin{bmatrix} 3.107 \\ 2.015 \end{bmatrix} - \begin{bmatrix} 2.825 \\ 2.015 \end{bmatrix}$$

$$\underline{\mu} = \begin{bmatrix} 0.282 \\ 0.0 \end{bmatrix} .$$

Compute the inverse of the covariance matrix of the first canonical variables in the target and control regions, $\underline{V}^{-1}$. In this case,

$$\underline{V}^{-1} = \begin{bmatrix} 15.879 & -15.371 \\ -15.371 & 15.879 \end{bmatrix}$$

and then compute,

$$\underline{\mu}'\underline{V}^{-1}\underline{\mu} = [0.282 \quad 0.0] \begin{bmatrix} 15.879 & -15.371 \\ -15.371 & 15.879 \end{bmatrix} \begin{bmatrix} 0.282 \\ 0.0 \end{bmatrix}$$

$$= 1.271 \quad .$$

TABLE 61 - INVERSE OF COVARIANCE MATRIX OF N-CN-4 CANONICAL SERIES

		N-4				CN-4			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
N-4	ζ_1	45.706	0.	0.	0.	-45.203	0.	0.	0.
	ζ_2	0.	4.810	0.	0.	0.	-4.281	0.	0.
	ζ_3	0.	0.	3.539	0.	0.	0.	-2.997	0.
	ζ_4	0.	0.	0.	2.429	0.	0.	0.	-1.863
CN-4	ε_1	-45.203	0.	0.	0.	45.706	0.	0.	0.
	ε_2	0.	-4.281	0.	0.	0.	4.810	0.	0.
	ε_3	0.	0.	-2.997	0.	0.	0.	3.539	0.
	ε_4	0.	0.	0.	-1.863	0.	0.	0.	2.429

TABLE 62 - INVERSE OF COVARIANCE MATRIX OF N-CN-6 CANONICAL SERIES

		N-6				CN-6			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
N-6	ζ_1	50.251	0.	0.	0.	-49.749	0.	0.	0.
	ζ_2	0.	4.981	0.	0.	0.	-4.453	0.	0.
	ζ_3	0.	0.	4.084	0.	0.	0.	-3.549	0.
	ζ_4	0.	0.	0.	2.438	0.	0.	0.	-1.872
CN-6	ε_1	-49.749	0.	0.	0.	50.251	0.	0.	0.
	ε_2	0.	-4.453	0.	0.	0.	4.981	0.	0.
	ε_3	0.	0.	-3.549	0.	0.	0.	4.084	0.
	ε_4	0.	0.	0.	-1.872	0.	0.	0.	2.438

TABLE 63 - INVERSE OF COVARIANCE MATRIX OF S-CS-4 CANONICAL SERIES

		S-4				CS-4			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
S-4	ζ_1	15.879	0.	0.	0.	-15.371	0.	0.	0.
	ζ_2	0.	2.466	0.	0.	0.	-1.901	0.	0.
	ζ_3	0.	0.	1.977	0.	0.	0.	-1.390	0.
	ζ_4	0.	0.	0.	1.615	0.	0.	0.	-.996
CS-4	ε_1	-15.371	0.	0.	0.	15.879	0.	0.	0.
	ε_2	0.	-1.901	0.	0.	0.	2.466	0.	0.
	ε_3	0.	0.	-1.390	0.	0.	0.	1.977	0.
	ε_4	0.	0.	0.	-.996	0.	0.	0.	1.615

TABLE 64 - INVERSE OF COVARIANCE MATRIX OF S-CS-6 CANONICAL SERIES

		S-6				CS-6			
		ζ_1	ζ_2	ζ_3	ζ_4	ε_1	ε_2	ε_3	ε_4
S-6	ζ_1	16.383	0.	0.	0.	-15.875	0.	0.	0.
	ζ_2	0.	2.524	0.	0.	0.	-1.961	0.	0.
	ζ_3	0.	0.	1.940	0.	0.	0.	-1.350	0.
	ζ_4	0.	0.	0.	1.476	0.	0.	0.	-.839
CS-6	ε_1	-15.875	0.	0.	0.	16.383	0.	0.	0.
	ε_2	0.	-1.961	0.	0.	0.	2.524	0.	0.
	ε_3	0.	0.	-1.350	0.	0.	0.	1.940	0.
	ε_4	0.	0.	0.	-.839	0.	0.	0.	1.476

The degrees of freedom here are 2 and 28, which are the number of canonical variables and the number of observations less the number of canonical variables, respectively. With these degrees of freedom, the value of τ^2 is found to be 5.468, at the level of significance $\alpha = 0.05$ and power $\beta = 0.50$. Now from

$$N^* = \frac{\tau^2}{\underline{\mu}'\underline{V}^{-1}\underline{\mu}}$$

the value of N^* is obtained as

$$N^* = \frac{5.468}{1.271} = 4.3 = 5 \text{ years} ,$$

since N^* must be an integer. These values of N^* are shown in Table 65.

The previous results are based on the assumption that the sample mean is the same as the population mean during the non-seeded period. Now consider what effect a violation of this assumption would have on the results.

Suppose the true population mean is not equal to the sample mean. Instead it lies at the upper extremity of the 50% confidence interval established for the sample mean of the non-seeded period. Then a 10% increase in the true population mean results in a larger absolute increase than does a 10% increase in the assumed population mean (simply because the actual population mean is larger than the assumed population mean).

In the northern region, an actual 10% increase in the true population mean yields a 14.2% increase in the assumed population mean. This results in a reduction in the number of observations required to detect a change. The number of observations would be reduced to 50% of the previously determined number of observations. Similarly, in the southern region an

TABLE 65 - MINIMUM NUMBER OF YEARS TO DETECT THE INCREASE OF 10 PERCENT IN RUNOFF MEANS USING CANONICAL VARIABLES

Type	No. of canonical variables in target	No. of canonical variables in control	Value of $\underline{\mu}'\underline{V}^{-1}\underline{\mu}$	τ^2	Minimum number of years to detect the increase, N*	Remarks
N-CN-4	1	1	5.037	5.468	3	The minimum value of N* is obtained from the larger of $N^* = \tau^2/\underline{\mu}'\underline{V}^{-1}\underline{\mu}$ or $N^* = k + 1$ where k is the total number of variables in both target and control
	2	2	5.197	7.640	5	
	3	3	5.198	9.646	7	
	4	4	5.368	11.655	9	
N-CN-6	1	1	5.877	5.468	3	
	2	2	6.040	7.640	5	
	3	3	6.060	9.646	7	
	4	4	6.124	11.655	9	
S-CS-4	1	1	1.271	5.468	5	
	2	2	1.305	7.640	6	
	3	3	1.388	9.646	7	
	4	4	1.581	11.655	9	
S-CS-6	1	1	1.423	5.468	4	
	2	2	1.465	7.640	6	
	3	3	1.690	9.646	7	
	4	4	1.752	11.655	9	

actual 10% increase in the true population mean yields a 15.6% increase in the assumed population mean, and a corresponding reduction in the required number of observations by 60 percent.

Now, suppose that the true population mean lies at the lower end of the 50% confidence interval. Then a 10% increase in the true population mean results in a smaller absolute increase than does a 10% increase in the assumed population mean.

In the northern region, an actual 10% increase in the true population mean yields a 5.8% increase in the assumed population mean. This results in an increase in the number of observations required to detect a change. The number of observations would be increased by a factor of three. Similarly, in the southern region an actual 10% increase in the true population mean yields a 4.4% increase in the assumed population mean, and the number of observations required would be increased by a factor of 5.2.

In view of the above discussion, it is seen that if the number of observations is calculated by assuming different values for the population mean a distribution is obtained. The median number of observations will be the same as that number obtained by using the sample mean of the non-seeded period.

Chapter VI

CONCLUSIONS

It was the objective of this study to develop a technique for detection of a geographically widespread change in a minimum amount of time.

It was found that a combination of techniques, namely canonical analysis and multivariate T^2 test was the most effective means to provide positive results in the least time. Assuming a 10% increase in runoff, 3 and 4 years are the minimum number of years needed for significance in the Upper Basin of the Colorado and the San Juan Mountains, respectively.

A word of caution is needed at this point. If the effect of precipitation management is to produce exactly a uniform 10% increase in runoff the use of only one set of canonical components is very efficient. However, if the increase is not uniform, it is safer to use several canonical components. With more canonical components, however, the number of years needed for significance increases.

It is apparent that there exists a trade-off between power of the test and representativity of the tested variables. This is well illustrated by the combined use of principal components analysis and the T^2 test. The first three or four principal components account for 99% of the total variation in the target regions. These sets of components so to speak, are 99% representative. The number of years calculated from the T^2 test is much higher than the corresponding figure for the same number of canonical components. This number of years could be decreased by using only one principal component, which already accounts on the average for 90% of the total variation. (This number was not actually calculated but the validity of the statement can be inferred from examination of the covariance matrices).

Note that when the χ^2-test is applied to each target station with the best correlated control station, the lowest minimum number of years is found to be seven in both northern and southern regions. Again, a single station is, of course, poorly representative of the entire region. The technique (canonical components - T^2 test) improves both the power of the test and the regional representativity of the tested variable, over what it would have been even with the best single target control pair.

The results from the use of four-months or six-months spring runoff are very similar. Nevertheless, better results are obtained with the six-months runoff series, particularly in the southern area.

LIST OF SYMBOLS

Symbol	Meaning
Q_i	Runoff at station i (i is the number in the 'Seq. No.' column in Table 2)
q_i	Observation of Q_i
$\overline{q}_i$	The mean of Q_i
$q_{i,m}$	The m^{th} observation of Q_i
$\underline{Q}$	Column vector of runoff at all stations
$\underline{q}_i$	Column vector of the i^{th} observation of $\underline{Q}$
$\overline{\underline{q}}$	Mean vector of observations of $\underline{Q}$
N	Number of observations of non-seeded period
N*	Minimum number of years for detecting a 10% increase in the runoff means of seeded period
N-4	Four-month runoff series in the northern target region (the 4 months are: April, May, June, and July)
N-6	Six-month runoff series in the northern target region (the 6 months are: March, April, May, June, July and August)
CN-4	Four-month runoff series in the northern control region
CN-6	Six-month runoff series in the northern control region
S-4	Four-month runoff series in the southern target region
S-6	Six-month runoff series in the southern target region
CS-4	Four-month runoff series in the southern control region
CS-6	Six-month runoff series in the southern control region
N-CN-4	The combination of N-4 and CN-4
N-CN-6	The combination of N-6 and CN-6
S-CS-4	The combination of S-4 and CS-4
S-CS-6	The combination of S-6 and CS-6
k	Total number of runoff variables, i.e., the number of all target and control variables
h	The fractional increase in the runoff mean
E{}	The expected value of {}
p	The number of runoff variables in target (or control) region in the principal component analysis
p_1	The number of runoff variables in target region
p_2	The number of runoff variables in control region
$\underline{\beta}_i$	Column vector of coefficients for computing the i^{th} principal component

Symbol	Meaning
β_{ij}	Coefficient of runoff at station j in the computation of the i^{th} principal component
I	Identity matrix
$\underline{V}$	Covariance matrix of runoff variables
$\underline{V}^{-1}$	Inverse of $\underline{V}$
$\underline{W}$	$\underline{V}^{-1}$
ξ_i	The i^{th} principal component of target region before seeding
ξ_i^*	The i^{th} principal component of target region for seeded period
η_i	The i^{th} principal component of control region before seeding
η_i^*	The i^{th} principal component of control region for seeded period
$\xi_{i,m}$	The m^{th} data point of ξ_i
$\eta_{i,m}$	The m^{th} data point of η_i
λ_i	The amount of variance accounted for by the i^{th} principal component
ζ_i	The i^{th} canonical variable of target region before seeding
ζ_i^*	The i^{th} canonical variable of target region for seeded period
ε_i	The i^{th} canonical variable of control region before seeding
ε_i^*	The i^{th} canonical variable of control region for the seeded period
$\zeta_{i,m}$	The m^{th} data point of ζ_i
$\varepsilon_{i,m}$	The m^{th} data point of ε_i
θ_i	Correlation between ζ_i and ε_i
$\underline{\alpha}_i$	Vector of coefficients for computing ζ_i
$\underline{\gamma}_i$	Vector of coefficients for computing ε_i
$\alpha_{i,j}$	Coefficient of runoff at station j (target region) in the computation of ζ_i
$\gamma_{i,j}$	Coefficient of runoff at station j (control region) in the computation of ε_i
$\underline{\mu}^*$	Runoff mean vector for the seeded period
$\underline{\mu}_o$	Runoff mean vector for the non-seeded period
$\underline{\mu}$	$\underline{\mu}^* - \underline{\mu}_o$
$\underline{\mu}'$	Transpose of $\underline{\mu}$
$\sum_{i=1}^{N}$	Summation from i=1 to i=N
$\prod_{i=1}^{N}$	Product from i=1 to i=N
τ^2	Noncentrality parameter
$\hat{}$	Estimated value

LIST OF SYMBOLS - Continued

Symbol	Meaning
'	Transpose of a matrix
σ_{ii}	Variance of runoff variable Q_i
σ_{ij}	Covariance of runoff variables Q_i and Q_j
*	Of seeded period
cfs	Cubic feet per second

REFERENCES

1. Krug, J. A., "The Colorado River," Bureau of Reclamation, U.S. Dept. of the Interior, March 1966, 295 p.

2. Golzé, A. R., "Reclamation in the United States," the Caxton Printers, Ltd., Caldwell, Idaho, 1961, 486 p.

3. Bureau of Reclamation, U.S. Dept. of the Interior, "Plan to Develop Technology for Increasing Water Yield from Atmospheric Sources," November 1966, 54 p.

4. Upper Colorado River Commission, Salt Lake City, Utah, the Nineteenth Annual Report, September 30, 1967, 152 p.

5. Committee on Water of the National Research Council, National Academy of Sciences, Washington, D.C., 1968, "Water and Choice in the Colorado Basin," 107 p.

6. Miller, D. W., J. J. Geraghty and R. S. Collins, "Water Atlas of the United States," Water Information Center, Inc., Port Washington, L. I., N. Y., 80 p.

7. Bureau of Reclamation, U.S. Dept. of the Interior, Reclamation to conduct Project Skywater Program in Colorado Rockies this winter, September 22, 1968, pamphlet.

8. Bureau of Reclamation, U.S. Dept. of the Interior, "Scientific Review Plans and Design of the Colorado River Basin Pilot Project with Appendix: Design Changes and Decisions" Proceedings: Skywater Conference V, February 1969, Denver, Colo.

9. Morel-Seytoux, H. J., "Suitability of Basins to Weather Modification of Statistical Evaluation of Attainment," Final report for the Bureau of Reclamation, Hydrology Program, Colorado State University, Fort Collins, Colorado, 1967, Part 1, 62 p.

10. Kendall, M. G. and A. Stuart, "The Advanced Theory of Statistics," Vol. 1-2-3, Charles Griffin and Co., Ltd., London.

11. Anderson, T. W., "An Introduction to Multivariate Statistical Analysis," Wiley Publications in Statistics, 374 p.

12. Nimmannit, V., "Multivariate Analysis of Hydrologic Changes," Ph.D. dissertation, August 1969, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, CED68-69VN45.

13. Li, J. C. R., "Statistical Inference," Vol. I, Edwards Brothers, Inc., Ann Arbor, Michigan, 658 p.

14. Markovic, R. D., "Statistical Evaluation of Weather Modification Attainments," Hydrology Paper, Colorado State University, Fort Collins, Colorado, 1966, No. 16.

15. Brownlee, K. A., "Statistical Theory and Methodology in Science and Engineering," John Wiley and Sons, Inc., 570 p.

16. Brier, G. W. and I. Enger, "An Analysis of the Results of the 1951 Cloud Seeding Operations in Central Arizona," Bulletin of the American Meteorological Society, 1952, Vol. 33, No. 5, pp. 208-210.

17. Buell, E. C., "An Evaluation of the Results of Cloud Seeding in Western New Mexico and Southern Arizona During July and August 1951 and 1952," Bulletin of the American Meteorological Society, 1955, Vol. 36, No. 1, pp. 6-15.

18. Dennis, A. S., "Cloud Seeding Results in Santa Clara County," Journal of Irrigation and Drainage, March 1960, pp. 93-99.

19. Thom, H. C. S., "An Evaluation of a Series of Orographic Cloud Seeding Operations," Advisory Committee on Weather Control, Technical Report No. 2, 1957, p. 25.

20. Biswas, K. R., R. K. Kapoor and K. K. Kanuga, "Cloud Seeding Experiment Using Common Salt," Journal of Applied Meteorology, 1967, pp. 914-923.

21. Decker, L. W. and P. T. Schickedanz, "The Evaluation of Rainfall Records from a 5-Year Cloud Seeding Experiment in Mississippi," 5th Berkeley Symposium of Mathematical Statistics and Probability, 1967, pp. 55-63.

22. U.S. Advisory Committee on Weather Control, Final Report, 1957, Vol. 1, 32 p.

23. Howell, W. E., "Twelve Years of Cloud Seeding in the Andes of Northern Peru," Journal of Applied Meteorology, 1965, Vol. 4, pp. 693-700.

24. Howell, W. E. and M. E. Lopez, "Cloud Seeding in Southern Puerto Rico, April-July 1965," Journal of Applied Meteorology, 1966, Vol. 5, pp. 692-696.

25. Mood, A. M. and F. A. Graybill, "Introduction to the Theory of Statistics," McGraw-Hill Book Company, Inc., 443 p.

26. Edwards, A. L., "Statistical Methods for the Behavioral Sciences," Rinehart and Company, Inc., N. Y., 1956, 420 p.

27. Siliceo, E. P., A. A. Ahumada and P. A. Mosino, "Twelve Years of Cloud Seeding in the Necaxa Watershed, Mexico," Journal of Applied Meteorology, 1963, Vol. 2, No. 3, pp. 311-323.

28. Battan, L. J. and A. R. Kassander, Jr., "Summary of Results of a Randomized Cloud Seeding Project in Arizona," 5th Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 29-33.

29. Mack, C., "Essentials of Statistics for Scientists and Technologists," Plenum Press, N.Y., 174 p.

30. Mann, H. B. and D. R. Whitney, "On a Test Whether One of Two Random Variables is Stochastically Larger than the Other," Annals of Mathematical Statistics, Vol. 18, 1947, pp. 50-60.

31. Auble, D., "Extended Tables for the Mann-Whitney Statistics," Bulletin of Institution of Educational Research, Indiana University, Vol. 1, No. 2, 1953.

32. Wine, R. L., "Statistics for Scientists and Engineers," 1964, Prentice-Hall, Inc., 671 p.

33. Elliott, R. D., "Note on Cloud Seeding Evaluation with Hourly Precipitation Data, 1962," Journal of Applied Meteorology, Vol. 1, pp. 578-580.

34. Battan, L. J., "Silver-Iodide Seeding and Precipitation Initiation in Convective Clouds," Journal of Applied Meteorology, 1967, Vol. 6, p. 317.

35. Markovic, R. D., "Statistical Evaluation of Weather Modification: Target Two-Sample Run Method," Report to the Bureau of Reclamation, Civil Engineering Department, Colorado State University, Fort Collins, 1967, 17 p.

36. Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components," J. Educ. Psych., 24, 1933, pp. 417-441, 498-520.

37. Tang, P. C., "The Power Function of the Analysis of Variance Tests with Tables and Illustrations of Their Use," Statistic Research Memo, 2, pp. 126-157, 1938.

38. Lehmer, E., "Inverse Tables of Probabilities of Errors of the Second Kind," Ann. Math. Stat., 15, pp. 388-398, 1944.

39. Chow, V. T., editor, "Handbook of Applied Hydrology" Section 8-II "Statistical and Probability Analysis of Hydrologic Data," by V. M. Yevjevich, McGraw-Hill, 1964.

Key Words: Statistical discrimination, regional hydrologic change, seasonal runoff, precipitation management, evaluation

Abstract: The object of this study is to find answers to the following questions: What is the appropriate statistical test for a regional target-control technique of evaluation? What is a suitable method for reduction of an originally large number of variables? Which of the Upper Basin of the Colorado River or the San Juan Mountains is a more suitable area of operations, if the effectiveness of precipitation management is to be detected as quickly as possible? The results of this research study show: 1. The T^2-test is the appropriate test for multiple target-control technique of evaluation. 2. The canonical anlysis is the suitable method for the reduction of a large number of original variables. 3. The Upper Basin of the Colorado River is preferable under the assumption of an equal percentage of increase in runoff. However, if the percentage increase in the southern area is at least 1.2 times as large as in the northern area (and recent publications suggest that this ratio is probably around 3) then the southern area is preferable. Based on the T^2-test, the minimum number of years for detecting an increase of 10 percent in spring runoff means are three years in the Upper Basin of the Colorado River, and four years in the San Juan Mountains.

References: Viboon Nimmannit and Hubert J. Morel-Seytoux, Colorado State University Hydrology Paper No. 37 (November 1969) "Regional Discrimination of Change in Runoff."

Key Words: Statistical discrimination, regional hydrologic change, seasonal runoff, precipitation management, evaluation

Abstract: The object of this study is to find answers to the following questions: What is the appropriate statistical test for a regional target-control technique of evaluation? What is a suitable method for reduction of an originally large number of variables? Which of the Upper Basin of the Colorado River or the San Juan Mountains is a more suitable area of operations, if the effectiveness of precipitation management is to be detected as quickly as possible? The results of this research study show: 1. The T^2-test is the appropriate test for multiple target-control technique of evaluation. 2. The canonical anlysis is the suitable method for the reduction of a large number of original variables. 3. The Upper Basin of the Colorado River is preferable under the assumption of an equal percentage of increase in runoff. However, if the percentage increase in the southern area is at least 1.2 times as large as in the northern area (and recent publications suggest that this ratio is probably around 3) then the southern area is preferable. Based on the T^2-test, the minimum number of years for detecting an increase of 10 percent in spring runoff means are three years in the Upper Basin of the Colorado River, and four years in the San Juan Mountains.

References: Viboon Nimmannit and Hubert J. Morel-Seytoux, Colorado State University Hydrology Paper No. 37 (November 1969) "Regional Discrimination of Change in Runoff."

Key Words: Statistical discrimination, regional hydrologic change, seasonal runoff, precipitation management, evaluation

Abstract: The object of this study is to find answers to the following questions: What is the appropriate statistical test for a regional target-control technique of evaluation? What is a suitable method for reduction of an originally large number of variables? Which of the Upper Basin of the Colorado River or the San Juan Mountains is a more suitable area of operations, if the effectiveness of precipitation management is to be detected as quickly as possible? The results of this research study show: 1. The T^2-test is the appropriate test for multiple target-control technique of evaluation. 2. The canonical anlysis is the suitable method for the reduction of a large number of original variables. 3. The Upper Basin of the Colorado River is preferable under the assumption of an equal percentage of increase in runoff. However, if the percentage increase in the southern area is at least 1.2 times as large as in the northern area (and recent publications suggest that this ratio is probably around 3) then the southern area is preferable. Based on the T^2-test, the minimum number of years for detecting an increase of 10 percent in spring runoff means are three years in the Upper Basin of the Colorado River, and four years in the San Juan Mountains.

References: Viboon Nimmannit and Hubert J. Morel-Seytoux, Colorado State University Hydrology Paper No. 37 (November 1969) "Regional Discrimination of Change in Runoff."

Key Words: Statistical discrimination, regional hydrologic change, seasonal runoff, precipitation management, evaluation

Abstract: The object of this study is to find answers to the following questions: What is the appropriate statistical test for a regional target-control technique of evaluation? What is a suitable method for reduction of an originally large number of variables? Which of the Upper Basin of the Colorado River or the San Juan Mountains is a more suitable area of operations, if the effectiveness of precipitation management is to be detected as quickly as possible? The results of this research study show: 1. The T^2-test is the appropriate test for multiple target-control technique of evaluation. 2. The canonical anlysis is the suitable method for the reduction of a large number of original variables. 3. The Upper Basin of the Colorado River is preferable under the assumption of an equal percentage of increase in runoff. However, if the percentage increase in the southern area is at least 1.2 times as large as in the northern area (and recent publications suggest that this ratio is probably around 3) then the southern area is preferable. Based on the T^2-test, the minimum number of years for detecting an increase of 10 percent in spring runoff means are three years in the Upper Basin of the Colorado River, and four years in the San Juan Mountains.

References: Viboon Nimmannit and Hubert J. Morel-Seytoux, Colorado State University Hydrology Paper No. 37 (November 1969) "Regional Discrimination of Change in Runoff."