INTER-STATION CORRELATIONS IN ANNUAL PRECIPITATION AND IN ANNUAL EFFECTIVE PRECIPITATION

By

James E. Caffey

June 1965

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

6

INTER-STATION CORRELATIONS IN ANNUAL PRECIPITATION AND IN ANNUAL EFFECTIVE PRECIPITATION

by

James E. Caffey

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO

June, 1965 No. 6

ACKNOWLEDGMENTS

The writer wishes to acknowledge the support of the U. S. National Science Foundation in the research leading to this hydrology paper. A research assistantship was awarded to the writer by Colorado State University. Research on this paper was initiated and carried out while the writer was studying towards his Doctor of Philosophy Degree at this university. The writer wishes to express his appreciation for the guidance by Dr. V. M. Yevdjevich, Professor of Civil Engineering, and for advice by Dr. M. M. Siddiqui, Professor of Statistics, in conducting research presented in this paper.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES AND TABLES (Continued)

ABSTRACT

An analysis of inter-station correlation coefficients for the two variables annual precipitation and annual effective precipitation is performed to determine their regional patterns and to relate the observed patterns to general hydrologic and meteorologic factors. The annual effective precipitation variable as used in this study is essentially the observed annual river flow.

Precipitation and river flow data from the Western United States and Southwestern Canada are used as research data. From this area, 1141 precipitation stations with an average length of record of 54 years and 446 river gaging stations with an average length of record of 37 years were selected and used in the analysis.

The regional variation in inter-station correlation coefficients is reasonably explained for both variables by the regression function

$$r_e = a_1 \exp \left[-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)\right] d$$

where r_e is the estimate of the inter-station correlation coefficient, d is the inter-station distance, and θ is the normal trigonometric angle to the radial from the block control station to the surrounding stations in a limited square area about the control station. The above function is fitted by the method of least squares to the inter-station correlation coefficients computed from the data of the block control station and every other station in the square block and plotted at each respective station location.

On the average, approximately 60 percent of the variation in inter-station correlation coefficients is explained by the selected function. Isolines of equal coefficients for the fitted surfaces approximate ellipses indicating that the meteorologic and hydrologic factors affecting the regional correlations in annual effective precipitation and annual precipitation are not isotropic. The orientations of the axes of maximal correlations for both variables are found to vary considerably over the study region. Effects of topography, general wind circulation, and frontal activity upon the axis of maximal correlation is noted for the annual precipitation variable. Other factors such as evapotranspiration and river basin characteristics tend to cause the regional patterns for the annual effective precipitation variable to differ from those observed for the precipitation variable.

INTER-STATION CORRELATIONS IN ANNUAL PRECIPITATION AND IN ANNUAL EFFECTIVE PRECIPITATION

By: James E. Caffey

CHAPTER I

INTRODUCTION

1. Significance of This Study. The casual observer of fluctuations in precipitation amounts and the resulting stream flows observes a process in nature which is probabilistic rather than deterministic. A process developing in time in a manner controlled by probabilistic laws is known as a "stochastic process." Suppose X is a variable which can take on the values

$$x_1, x_2, \ldots, x_n,$$

with the probabilities of assuming these values

$$p(x_1), p(x_2), \ldots, p(x_n), \geq 0$$

respectively. Then X is called a stochastic variable [Miller, 1963].* A stochastic process is best defined mathematically as a collection $\{X(t), t \in T\}$ of stochastic variables [Parzen, 1962]. The Greek letter ϵ is read "belongs to," and T is the set of random variables to which t belongs. There are no restrictions placed upon T. The T may be index sets such as time or distance.

The extent to which different parts of the world experience similar deviations from the long-term average experience is quite important. In attemps to control or reduce the effects of such adverse hydrologic phenomena, a detailed knowledge of the characteristics of such variables as temperature, precipitation, and stream flow are of paramount importance. There is a need for predictive capability in the area of water resources because of the changing availability of adequate supplies of water. The elusiveness of predictive capability concerning rainfall and runoff has been troubling water technologists and scientists for years.

V. M. Yevdjevich, in a paper presented in August, 1964, at Bangkok, Thailand, before the Seminar on Hydrologic Forecasting which was organized by the United Nations Economic Commission for Asia and the Far East and the World Meteorological Organization, discussed the outlook for long-range rainfall and runoff forecasting. He concluded that no predictions for the future of a deterministic nature can be obtained from the extrapolation of past rainfall and runoff experience. V. M. Yevdjevich feels that the only hope for long-range forecasting of rainfall and runoff events is to look to the main source of water vapor of hydrologic or meteorologic importance, the oceans. Only the oceans, he adds, have such physical phenomena which affect but precede sufficiently in time the deposition of moisture over large continental areas from which to draw the necessary predictive information.

Of more immediate need are means of acquiring hydrologic data for todays water resource projects. A decision must frequently be made whether or not records of a hydrologic variable taken simultaneously at a number of observing stations for relatively short periods of time can be used to replace observations taken at one station over a long interval of time. This question arises most often in areas where usable records are sparse, unavailable, or quite short in the time length of observation. In other situations, hydrologic records may be adequate in length and accuracy but quite sparsely distributed areally. Thus, it becomes increasingly important to interpolate for hydrologic variables.

The statistical concepts of correlation analysis play an important role in the estimating of hydrologic variables. The inter-station correlation coefficient r equals

$$r = \frac{\sum_{i=1}^{N} X_i Y_i - N \bar{X} \bar{Y}}{\left[\sum_{i=1}^{N} X_i^2 - N \bar{X}^2\right]^{1/2} \left[\sum_{i=1}^{N} Y_i^2 - N \bar{Y}^2\right]^{1/2}} \qquad 1.1$$

and is a measure of association between the values of the variables, X_i and Y_i, taken simultaneously in time at two different points. The autocorrelation (serial correlation) coefficient is a measure of association between successive values of a variable measured at the same point. The autocorrelation coefficient r_k equals

$$r_k = \frac{\frac{1}{N-k}\sum_{i=1}^{N-k} X_i X_{i+k} - \frac{1}{(N-k)^2}\left(\sum_{i=1}^{N-k} X_i\right)\left(\sum_{i=1}^{N-k} X_{i+k}\right)}{\sum_{i=1}^{N} X_i^2 - \frac{1}{N^2}\left(\sum_{i=1}^{N} X_i\right)} \qquad 1.2$$

*References are designated by a year in brackets associated with the author's name and may be found in the bibliography. The reference in the bibliography may appear in context as . . . Alexander [1963] or . . . [Alexander, 1963].

Both equations are for discrete time series.

In eq. 1.1, N is the number of simultaneously observed events, and in eq. 1.2, N is the number of events forming the time series. The term k gives the lag or time units between events. Autocorrelation is used as a measure of dependence or nonrandomness in time, and inter-station correlation is used as a measure of dependence or nonrandomness in space. One quite frequently encounters both simultaneously in hydrologic and meteorologic variables.

The degrees of autocorrelation and inter-station correlation in a hydrologic variable carry a significant impact when extending records of short length or interpolating for values at points where no measurements are available. Using rainfall as an example, suppose that the observations at three stations were statistically independent, then 30 years of data at each point will give us much information about some rainfall parameter as a record of 90 years at only one station, if station data describes the same rainfall parameter. But, if these three stations exhibit inter-station correlation, the information from the three stations would not necessarily be equivalent to that of 90 years for the one station. The additional gain in information resulting from additional stations in the region will be reduced by an amount depending upon the degree of inter-station correlations. Inter-station correlations between variables may be of significant magnitude over great distances. Further, autocorrelation within the time series observed at each station would tend to reduce its effective length.

When data is sparse, one often has to rely on various isopleths to estimate a hydrologic variable at a point somewhat removed from a point having observations. The accuracy of the estimate must depend upon the density of the observation points and the degree of inter-station correlation which is the measure of areal association. Areal association is the degree to which the same casual factors affect the variable at different points over a wide area. The term "degree," as used here, includes not only the actual magnitudes (high or low values of inter-station correlation coefficients) but also their consistency (small random variability of values for a given inter-station distance and direction). Knowing the degree of areal association, one is in a position to design observation networks of proper density to attain a certain degree of accuracy in interpolated estimates.

Thus, the need for a more adequate knowledge of the areal distribution of inter-station correlation coefficients in annual precipitation and in annual stream flow becomes apparent. Beyond this obvious immediate need is that of a more precise indication of the oceanic sources of atmospheric moisture. This may be possible through a trace of the directions of maximal inter-station correlations for precipitation when related to such things as general circulation, moisture flux and precipitation mechanisms, or for runoff when the influence of evapotranspiration* is also considered.

2. Objectives. The primary objective of this study is to investigate the regional (areal or spatial) distribution of inter-station correlation coefficients within the two variables: annual river flow as represented in this study by annual effective precipitation (EP), and annual precipitation (P). The term annual precipitation refers to the accumulated depth of precipitation in inches received at the place of measurement. Annual effective precipitation is the net annual water yield of the atmosphere as recognized in measured river flows. The annual effective precipitation for a drainage area or basin may be expressed as:

$$EP = P - E - X \qquad 1.3$$

where P is the accumulated volume of water due to the precipitation falling upon the drainage area, E is the volume of water lost through evapotranspiration, and X is the volume lost to all other factors.

The determination of annual effective precipitation as expressed by eq. 1.3 is merely a statement of continuity. The amounts of precipitation contributing water to a basin may be measured within the water year. The evapotranspiration and other factors removing water from the basin are assumed to be measured within the water year. The quantities of E and X are not readily determined, so the annual effective precipitation (EP) may be determined through the use of measured runoff. However, all of the water which is free to flow from the basin may not be measured as runoff within the water year in which it fell upon the basin.

The delayed appearance of a portion of the water as runoff is caused by the storage of water in the basin. Thus, the observed flows must be corrected for the change in storage from one year to the next. This change in storage occurs as change in ground water, soil moisture, snow cover, surface reservoir water and depression storage, and water temporarily existing on the surface as flowing water. The changes in carryover may be either positive or negative and in most instances they are quite small in comparison to the yearly volume of flow. The annual effective precipitation as used in this study has been corrected for changes in carryover, but essentially it is the same as the observed runoff.

The variation in regional (inter-station) correlations is defined in this paper for a given station position by fitting a three-dimensional surface to values of the inter-station correlation coefficients. The statistical technique of multiple regression is employed to determine the form of the fitted surfaces.

The hypothesis assumed is that the inter-station correlation coefficients for like variables between one point taken as origin and points surrounding it can be related by a simple mathematical relation dependent only upon the distance and direction from the point of origin to the surrounding points. The parameters evaluated for the selected function are expected to vary from region to region as the character of the influencing hydrologic variables changes.

A secondary objective of this study is to furnish information to aid hydrologists in relating ungaged to gaged areas for both variables, annual effective precipitation and annual precipitation, and in estimating missing data as accurately as feasible. This can be accomplished by relating measurements at one point to those of surrounding points which contain complete sets of observations. This study should definitely aid in selecting the best correlated stations. Regional correlations have been used recently to determine homogeneous regions for runoff and precipitation studies. This study should be of benefit to the extension of this approach [Gatewood, et. al, 1964].

* The term "evapotranspiration" designates all evaporation from a river basin.

Physical phenomena that affect or explain unexpected deviations in the spatical distribution of correlation coefficients are considered and evaluated if possible. Particular interest is given to relationships between the configuration of regional correlation coefficients and moisture sources, paths of moisture flux, the influence of evapotranspiration, and the mechanism of precipitation and runoff.

3. Delineation of the Study. The study is restricted to that portion of the North American Continent including the states of the United States west of the Mississippi River and the three Canadian provinces, Alberta, Saskatchewan, and British Columbia.

The research data consists of the annual values from the historical records of 446 river gaging stations, the EP-series, and 1141 precipitation gaging stations, the P-series. The selection of these gaging stations was done upon the basis of fixed criteria rather than randomly. The reader is referred to the references by Markovic [1964], Caffey [1965], and Yevdjevich [1963a].

Each of the above gaging stations will act as the control station for a square area of approximately 90,000 - 170,000 square miles depending on the station density. Each control station is at the coordinate center of the area and is the station whose data is correlated with the data of all other stations in the block (Fig. 1).

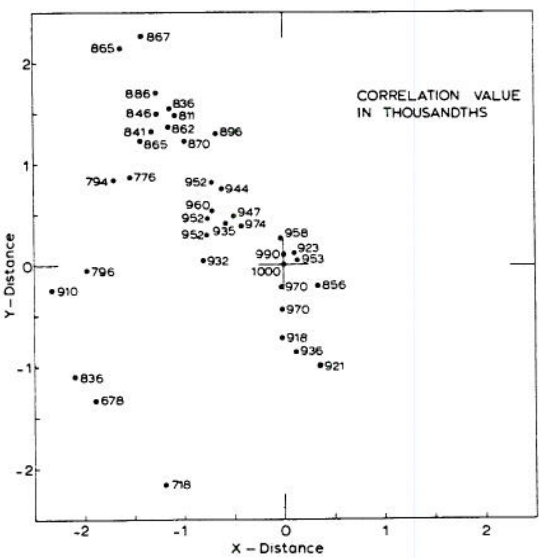

Fig. 1 Raw inter-station correlation coefficients (EP-series) plotted at their respective station locations. Each unit of distance (X or Y) is approximately equal to 60 nautical miles.

The matrix of all possible correlation coefficients between all stations within each of the control station blocks is not computed. The correlation coefficient for the data of the control station correlated with itself, of course, is 1.0. The correlation coefficients from the data of the control station with the data of all other stations in the block are the values to be plotted at each individual station map location (Fig. 1), and it is to these values that the surfaces are fitted. Every sample station, where possible, is used as a control station and, therefore, there is a surface fitted to the inter-station correlation coefficients for most stations of both annual precipitation (P),* and annual effective precipitation (EP).

The fitted surfaces are expected to be ellipsoidal in form, the degree of ellipticity varying over the study region according to the changing effects of meteorologic and hydrologic factors. The orientation of the major axis and the degree of ellipticity are two points of particular interest. More will be said in later sections about the surface fitting procedure and the parameters which describe this surface.

4. Contributions of This Study. Considering the previous work in this field, there are several items of interest presented which, to the writer's knowledge, are first attempted in this study:

a. Application of this type of study outside the continent of Australia;

b. Investigation of such a large area involving different climatic and physiographic conditions;

c. Application of this form of analysis to runoff data as expressed through the effective precipitation; and,

d. The use of inter-station correlation to trace the path of atmospheric moisture to its source region.

Another point to note is the vast amount of data being utilized through the aid of high-speed digital computers. There are 446 station time series of annual runoff values averaging 37 years in length and 1141 station time series of annual precipitation amounts averaging 54 years in length. No station record length in either set of data is less than 30 years.

5. Some Recent Results. A review of recent research which is considered to be closely related to this study may be found in the dissertation by Caffey [1965].

---

* Throughout the remainter of the text, for the sake of brevity and readability, annual effective precipitation may be expressed simply as "effective precipitation" or "EP." Also, annual precipitation may be expressed simply as "precipitation" or "P."

CHAPTER II

RESEARCH DATA ASSEMBLY

1. Introduction of Available Data. The nature of the variables sampled, annual river flow and annual precipitation, restricts the research data to the historical records obtained at widely distributed geographical points. Runoff records integrate the various processes transforming the precipitation falling upon a catchment area into streamflow. The distribution of runoff gaging stations is quite sparse, and only a small percentage of the total geographical area used for this study is sampled. Precipitation gaging stations obtain point measurements.

The sampling for this study is on a continental basis. Gaging stations were selected from the continental United States west of the Mississippi River and from the provinces of British Columbia, Alberta, and Saskatchewan in southwestern Canada (Figs. 22 and 23). The series of annual effective precipitation (EP) derived from the observed runoff series contains 446 station record series, and the series of annual precipitation (P) contains 1141 station record series. The period of record used for both variables ends with the year 1960. No station record is included that is shorter than 30 years. Thus, the inclusive period 1931-1960 is concurrent to all station records. The longest runoff record is 72 years, and the longest precipitation record is 114 years. Many gaging stations have interruptions in their series records prior to 1931.

Considering the number of stations with each of the annual precipitation and the annual effective precipitation data samples and the respective average lengths of their station records, one would, at first glance, be impressed by the apparent better quality of the annual precipitation sample. There will be, on the average, approximately two and one-half times as many inter-station correlation coefficients for the precipitation as for the effective precipitation within each block to which the surface is fitted. Remember, however, that each effective precipitation station record furnishes information for an area that may include several precipitation stations which provide only point measurements. Therefore, the effective precipitation sample may possibly contain more information than the precipitation sample.

2. Inaccuracies and Inconsistencies. There are inaccuracies and errors in the data brought about by several factors. Some of the factors contributing to this are: inaccuracies in the gaging instruments and procedures of measurement, errors in transcription and reduction of the data, changes in the observational instruments and techniques, and finally in the theoretical approaches to data reduction over the period of observation. Some of the inaccuracies are random in nature and do not introduce trends and jumps like inconsistencies and nonhomogeneities in data. However, changes in techniques and marked mistakes may affect conclusions drawn. A means of removing the latter may be disclosed by a close study of the station history.

Detailed historical studies cannot be accomplished for samples as large as those of the EP- and P-series. However, a survey was made of each station history during the gaging station selection process. Particular attention was paid to the amount of artificial regulation of streamflow, by irrigation withdrawals, storage effects and diversion from or into the stream. Stations having what were considered to be significant incorrectable changes were not included in the EP-series sample.

Precipitation gaging stations are subjected to nonhomogeneities principally by moves of the gaging instrument or shielding of the instrument. Records included in the P-series sample are assumed to be homogeneous based upon either double-mass plots and statistical tests or through a study of the station history if the former is not available. The historical study is concerned primarily with moves of the gaging instrument, and stations having no moves or moves one mile or less horizontally and 100 feet or less vertically were assumed to have a homogeneous record.

3. Sources of Data. Precipitation data used here was collected from published records of the United States Weather Bureau and the Canadian Department of Commerce, Meteorological Branch. Values extracted were the calendar year totals of precipitation from rain and snow. The values extracted were observed amounts as reported in the various publications.

Observed runoff data used here was collected from data published by the United States Geological Survey, Department of the Interior and the Dominion Water Power and Reclamation Service, Department of the Interior, Canada. The observed runoff data is reported for the water year which is from October 1 through September 30, the year being the one following that of the beginning date. For example, the water year 1960 begins October 1, 1959, and ends September 30, 1960.

4. Processing of the Observed Data to Effective Precipitation. The observed runoff data was processed through two steps prior to its use in this study. First, the measured flows were corrected for the change in carryover. Then, the corrected flows were reduced to create a "net station" series from the corrected observed values at each of the selected gaging stations. The processing applied is explained below.

Precipitation occurring during the water year because of the integrating and lagging effects of the drainage basin does not all appear as measured flow in the water year it occurred. Generally, a variable percentage of it is measured as runoff during the following water year and for larger basins with large storage within the basins the carryover effect may be observed for a period of more than one succeeding water year.

The change in carryover may influence significantly the correlation coefficients between the runoff variables at different locations. Change in carryover causes a smoothing of the runoff time series. The smoothing effect upon the time series increases as the ratio of change in carryover to the total annual volume of runoff increases. The correlation coefficient of one runoff time series with another having similar carryover characteristics would likely be higher than a coefficient from a different time series having dissimilar carryover characteristics. The change of carryover correction is applied to remove the influence of differing carryover characteristics upon the inter-station correlations.

Carryover was approximated by fitting mean recession curves to the end of the water year recession hydrographs. Curves of the form

$$Q = Q_o e^{-ct} \qquad 2.1$$

and

$$Q = Q_o e^{-ct^n} \qquad 2.2$$

were fitted to recession hydrographs plotted on semilog paper. The computation of the carryover correction is presented by Yevdjevich in a previous Hydrology Paper [1963a].

If there is more than one gaging station on a stream, flow originating upstream may be measured several times on its way to the sea. This repeated measurement creates a built in correlation between the data of gaging stations on the same channel, and this will distort regional patterns of inter-station correlation coefficients. Flows from mutually exclusive areas will not have the built in inter-station correlation. The "net station" data represents the flows from mutually exclusive areas. Uppermost gaging stations are naturally "net stations." The "net station" data was formed, after correction for the change in carryover, by subtracting the corrected flows of the upstream stations which pass directly to the station whose records are being converted. Thus, the "net station" data represents the "net flows" from the "net areas." The data of the EP-series was formed in this fashion.

In Chapter I, Section 2, the annual effective precipitation for a drainage area or basin was defined by the equation 1.3. The definition becomes more precisely defined by inclusion of the "net station" concept of this section. The factor X in eq. 1.3 is not readily evaluated and is usually negligible. Thus, eq. 1.3 becomes

$$EP = P - E = V + \Delta W, \qquad 2.3$$

where EP is the effective precipitation, V is the measured flow, and ΔW is the change in carryover. The quantities EP, V, and ΔW represent the water year volumes. The annual effective precipitation (EP) is defined as the annual contribution of moisture from the atmosphere to a river basin or a portion thereof.

5. Conversion of Data for Computations. The precipitation data required no processing of the nature required for runoff. The final operation was done alike for both variables. The series mean, $\overline{X}$, and the standard deviation, s, were computed for each station record and then used to form the series of standardized variable, x_i, with $i = 1, 2, \ldots, N$. These values were computed by the following equations:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \qquad 2.4$$

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \overline{X})^2 \qquad 2.5$$

$$x_i = \frac{(X_i - \overline{X})}{s} \qquad 2.6$$

where X_i are annual values and N is the sample size. The standardized variables were used in all computations related to this study. Standardization in no way affects the computation of correlation coefficients.

6. Coordinate System. In the study of inter-station correlation coefficients about a fixed observation station, it is necessary to relate the position of other stations to the central station of the block (Figs. 1 and 2). This is done using cartesian coordinates. The majority of blocks containing at least the established minimum of nine station locations were found to be only 5^0 wide measured from south to north along a meridian. The spacing of latitudinal lines is about 60 nautical miles per one degree. However, the meridional lines rapidly converge northward. To utilize a square control block, the longitudinal station coordinates have to be corrected for the convergence. Since the map portrayal of latitude lengths is essentially constant, the longitudinal lengths were converted to latitudinal scale. Latitudinal lines do have curvature, but the correction for curvature was omitted since the degree of error introduced over such small blocks is considered to be much less than inaccuracies in the data utilized. The assumption of a plane surface is paramount in this discussion.

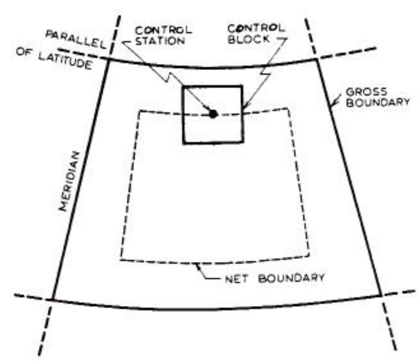

Fig. 2 Portrayal of the working subdivisional elements.

The correction for convergence is approximated using data given by Merriman [1920]. Lengths of one degree are given in meters for the Clarke Spheroid as follows:

Latitude	On Meridian Length	On Parallel Length
30^{0}N	110,850	96,489
50^{0}N	111,231	71,699
40^{0}N (Mean)	111,041 (Mean)	85,397 (Mean)

The rates of change are very nearly linear, so a linear adjustment is made on the basis of the mean values at 40^{0}N, the approximate mid-latitude of the region under study. The approximate rate of convergence is 1240 meters per degree latitude.

The cartesian value of longitude, x, for a station is given by the expression:

$$x_c = \frac{85,497 + 1240\ (40-Y)}{111,041}\ (\Delta x) \qquad 2.7$$

where x_c = cartesian coordinate distance, x, from the control block center line taken as a meridian through the control block center, coordinate (0, 0); Δx = control block center longitude minus the longitude of the station; and Y = latitude of the station.

The cartesian value of latitude, y, is given by subtracting the control block center latitude from the station latitude. One unit of x or y is equivalent to the distance between two points on a meridian separated by one degree of latitude. Thus, one unit of distance in the cartesian coordinate system is equivalent to approximately sixty nautical miles.

7. Regional Subdivisions. The number of stations selected for each of the two variables (P and EP) produced a quantity of data too great for storage in the internal memory of the Control Data Corporation 3600 digital computer which was used for all the computations of this study. The lack of memory capacity required that the study region be subdivided into areas containing a number of selected stations with a quantity of data small enough for storage.

Thirty-seven subdivisions of the region were required to accommodate the annual precipitation variable, and five were required for the annual effective precipitation variable (Figs. 2, 3 and 4). These subdivisions were selected to cover the study region without overlapping one another and are referred to as "net subdivisions."

Each station within the net subdivision is used as a "block control station" (BCS).* This means that each station becomes the center station about which a block (square area) is defined. When a station location falls at or near the net subdivision border, no stations are available to complete the square block about the BCS. To assure that there is data for stations in all quadrants of the block, it is necessary to use data from stations not included in the net subdivision. This requires the inclusion of data from a band about the net subdivision. The band is at least 3.5 units in width. The net subdivision area plus the band about it make up what is called the "gross subdivision." Each net subdivision had to be chosen so that the quantity of data from its corresponding gross subdivision did not exceed the memory capacity of the computer.

The relative positions of a net subdivision and its corresponding gross subdivisions are shown in fig. 2. The net subdivisions are mutually exclusive but the gross subdivisions are overlapping. A block whose control station is situated at the boundary of the net subdivision is also shown.

* The abbreviation "BCS" may often be used in the place of the words "block control station," but it has equivalent meaning.

CHAPTER III

THEORETICAL CONSIDERATIONS OF INTER-STATION CORRELATION AND ITS CASUAL FACTORS

High correlation coefficients indicate only a high degree of linear association and do not necessarily mean that causation is present. One would expect the temperature of water to be highly correlated with the heat added, a definite causal relation. One might find a high correlation between the temperature in Chicago and some phenomenon in India for a small sample of data, but it is doubtful that there is any causal relation even though a high association has been shown. Even when two variables have no apparent reason to be correlated, a high degree of linear correlation may be found in sample data due to chance or because both are related to some common causal factors.

This study, in one sense, is an attempt to evaluate how a variable measured at different points may be affected by the same broad scale causal factors. That is, in the case of precipitation, what effects do the meteorologic factors producing precipitation at one station have at another? Inversely, one might think of the problem as a way of determining the amount of change in the causal factors from one point to another. An analogous situation can be drawn for runoff, but the factors involved become much more complicated than those for precipitation.

One would expect to find rather high values of the correlation coefficients derived from the data of either precipitation or effective precipitation if the coefficients are from a reasonably restricted area. This is primarily because the causal factors, meteorologic and hydrologic, tend to be homogeneous over wide areas and change only gradually with distance except for regions of climatic divides.

1. Inter-station Correlation and Hydrologic Homogeneity. Thom [1940] postulated that the linear correlation (of rainfall from storms) between stations results from two causal factors: (1) "the average depths of a series of storms occurring in a record will be distributed in a frequency-function with the result that storms with greater average depths will, in general, produce greater depths at the several stations; and, (2) the centers of action of a storm system, although being random in position, have considerable area and a tendency to move in a general west-east direction so that stations in the path of a center will experience greater depths than those not in the path."

The correlation patterns should change from region to region, but within a large homogeneous region, the distribution of correlation coefficients about points therein should be basically the same. Thom [1940] defined a meteorologically homogeneous region as one in which a single type of storm produces the regional precipitation, in which there are meteorologic reasons to assume that all points will be subjected to equal frequency storm conditions, and in which a similar seasonal cycle of precipitation holds throughout. In this study, the dominant type of storm is the one most likely to produce the patterns in the areal distribution of correlation coefficients.

McDonald [1957] discussed the effects of non-normality of the distribution of correlation coefficients, and he concluded that "the kind of inferences that geophysicists wish to draw from most correlation studies can be entirely adequately drawn with the aid of the elementary sampling theory of the normal distribution." One can safely say that this holds also for the hydrologist. McDonald suggested, as an example, the use of inter-station correlation coefficients in studies of climatic homogeneity. He also suggested that a large number of inter-station correlation coefficients might be computed for several points and their geographical distribution be studied for dominant patterns of similarity and dissimilarity.

Benson [1962] reported on a study of the factors influencing the occurrence of floods. The many factors noted are not only influential in flood flows but also in the resultant mean annual runoff as well. After discussion of the interactions among meteorologic, hydrologic, and hydraulic factors, Benson stated that the topographic and meteorologic variables are not independent of each other.

The conclusion of Benson is supported in a comprehensive study of droughts in the Southwestern United States conducted in the offices of the United States Geological Survey. Thomas [1962] reporting on the study delineated, independently of other work, four meteorologic zones. Gatewood and others [1964] using inter-station correlation coefficients of runoff as a means of determining homogeneity found that the Southwest was not one homogeneous region as far as runoff patterns were concerned.

2. Meteorologic Factors Affecting Inter-station Correlation. Meteorologic factors may be divided into the two groups: climatic and synoptic. The climatic factors depend upon the major climatic controls: that is, the global location, the general circulation of the atmosphere, the proximity to geographical prominences such as oceans and orographic barriers, and other such elements. Visher [1944] indicated a similar division of factors when he wrote that the factors responsible for regional differences in the major precipitation regions of the United States were due to: (1) the position of the United States in relation to the belt of westerly winds; (2) the position of topography, especially the western mountains and the Mississippi River Basin; and, (3) the number of occurrences of cyclonic disturbances. Dorroh [1946] also concurred when he pointed out that the total annual precipitation at a given place comes about because of these dominating factors: its elevation, its location with respect to the source, and the normal directional movements of moist air masses. Runoff, since it is derived from precipitation, is affected by the meteorologic factors plus others that complicate the rainfall-runoff relationship.

The synoptic factors are those which change

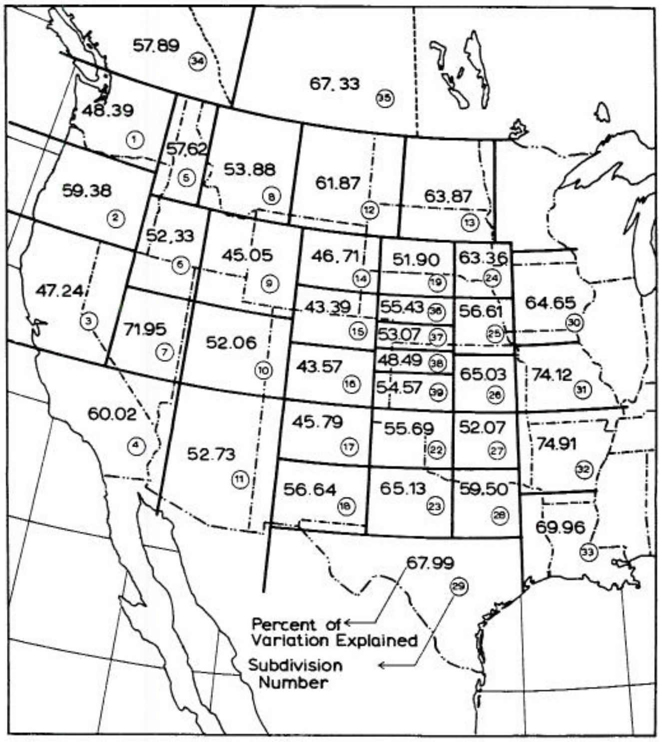

Fig. 3 Delineation of the net subdivisions and the average value of the variation in the correlation coefficients explained by regression for precipitation.

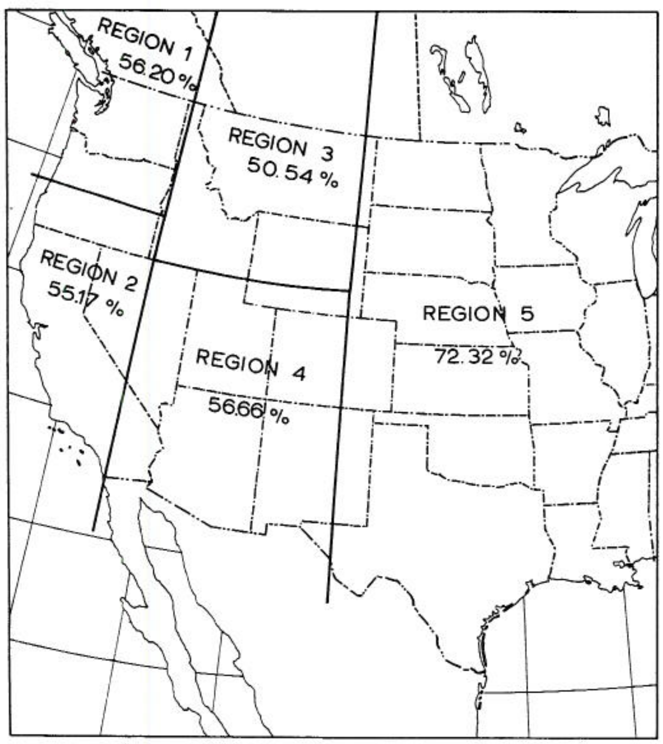

Fig. 4 Delineation of the net subdivisions and the average value of the variation in the correlation coefficients explained by regression for effective precipitation.

from day to day and season to season. The correlation between variables at different points tend to decrease as the time interval of observation decreases. Also, basically due to synoptic factors, correlation decreases with an increase of distance between points. Precipitation inter-station correlation coefficients are generally higher when produced by broad scale atmospheric disturbances rather than localized convective phenomena, and this is indirectly reflected in runoff.

For shorter periods such as a day, a week, or maybe a month, the synoptic factors may be the dominant causal factors of high or low inter-station correlation coefficients, depending on the type of disturbances. Recent integrations of the dynamic equations of the movement of an atmospheric air mass show that after a short time the state of the air mass changes and the new state becomes independent of the initial conditions. The new state may develop in a short period of time while the air mass is either stationary or in motion. If the air mass is moving while changing to a new state which is independent of initial conditions, then the inter-station correlation coefficients for the precipitation variable, for example, will become less significant as the distance increases between points where the variables are measured. The rate of decrease of inter-station correlation coefficients would be dependent upon the rate of movement and the rate of change to the new independent state. However, when the interval of observation is a year as in this study, the dominant causal factors are likely to be the climatic elements. One would expect many of the synoptic factors contributing to lower correlations to be masked and/or smoothed for larger time intervals of observation and the resultant correlation coefficients between variables at different points to be much higher than for coefficients from shorter intervals of observation.

The studies by Stenhouse and Cornish [1958] and Cornish, Hill and Evans [1961] have indicated for 6-day and monthly totals of precipitation that the degree of correlation shows a definitely greater persistence in one direction. They found that the isopleths of equal correlation coefficients were generally elliptic in shape. The same form of results are expected for annual precipitation and annual effective precipitation. In the cited studies, the authors worked with Fisher's z-transform, $z = \tanh^{-1} r$, where r is the inter-station correlation coefficient. They used the z-transform because it is normally distributed. No transformation is used in this study, and the surface is fitted to the computed values of the inter-station correlation coefficient, r.

In the region covered by this study, various factors act to produce precipitation in its several forms. Five primary factors are: (1) source-moisture paths, (2) topography, (3) convergence, (4) frontal activity, and (5) convection. The general effects these factors have upon areal patterns of correlation coefficients are considered further in the order listed.

Source of Moisture Paths. -- The principal sources of moisture that is eventually precipitated over the continents are the oceans and the seas. However, this fact was not fully accepted as late as 1943 when Horton [1943] stated that little or no vapor of truly oceanic origin may ever find its way to the small feeder streams at the headwaters of the larger rivers such as the Columbia, the Colorado, and the Mississippi. Through the analysis of the information from upper air soundings, meteorologists concluded that great amounts of water vapor move across the continents from the oceans and that vapor supplied to the atmosphere by evapotranspiration from the continent is quite small in comparison.

From a study of the Mississippi River Basin, Benton, Blackburn and Snead [1950] concluded that 100 inches of the 146 inches of water that flow over the basin as vapor can be attributed to maritime air masses. They further state that about 85 percent of all precipitation is derived from water vapor of direct oceanic origin. Bannon, Matthewman and Murray [1961] indicated that the direction of flux of water vapor over the United States and Southwestern Canada varies through the year. The winter flux is generally west to east over the entire area, but in the summer the west to east flux stagnates over the United States and gives way to a vapor flux from the Gulf of Mexico which flows northward into the midwest and then turns northeastward to flow over the Atlantic Ocean. Benton and Estoque [1954] summarized the situation by pointing out that there are two well defined paths of moisture inflow over North America: (1) a strong south to north to northeast from the Gulf of Mexico and (2) a quite diffuse westerly flow from the Pacific Ocean which is centered at approximately 50^0N.

By schematically representing the sources either as point or infinite line sources in the general circulation, one can reason about the effects upon the areal patterns of correlation coefficients. The general circulation will tend to form moisture flux lines from the point source in the fashion shown in fig. 5.

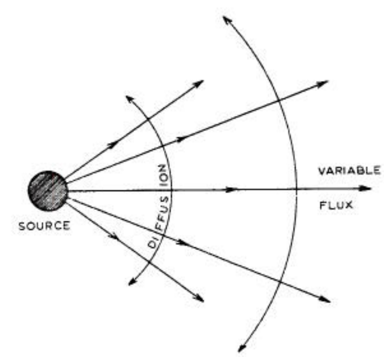

Fig. 5 Point moisture source.

The flux lines will tend to vary in direction as the general circulation shifts its direction. Furthermore, the moisture will diffuse laterally from the flux paths as the distance from the source increases. Precipitation will generally be expected to be produced over limited areas by small disturbances in the main flow of moisture, moving with the general circulation. Thus, one would expect the correlation to be greatest along a line parallel to the general circulation for two reasons: (1) the moisture available for precipitation decreases laterally to the circulation; and (2) the precipitation becomes more erratic toward the fringes of storm cells. The stations in the path of the storm cell will receive the benefits of storm center passage.

The pattern of inter-station correlation changes in the case of a line source which schematically represents the ocean. If one considers the source infinite in length and of constant strength, the correlation would be expected to be strongest laterally to the flux of moisture (Fig. 6). This is based on the assumption that storms over the period occur uniformly in the lateral direction. Correlations along the path of the flux would be less because the availability of moisture varies more with the circulation than laterally. This results from moisture being depleted irregularly by precipitation along the path of the circulation. The schematic line source comes nearest to depicting the actual situation, but the assumption of a constant line strength does not actually occur because of continental interruptions and circulation intensity variations. The strength of the circulation and the moisture content of the atmosphere in motion are constantly changing along the line. The lateral variability of this strength and moisture content, changes in the direction of the circulation, and the areal variability of storm occurrences may shift the higher correlations parallel to the moisture flux once again as in the schematic point source.

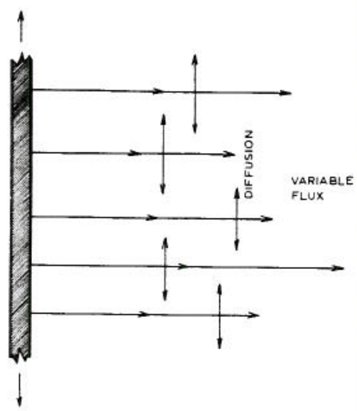

Fig. 6 Line moisture source.

Topographic Effects. -- The remaining four factors will tend to change the orientation of the axis of maximal correlations from its parallel orientation to the path of the general moisture flux. Topography will likely have a large influence upon the direction of maximum correlation. A high mountain ridge perpendicular to the general flow of water vapor presents a barrier which will lift and cool the moist air. This may result in precipitation which is quite uniform along the barrier, but less uniform as the distance from the barrier increases. Thus, orographic effects may tend to rotate the axis of strongest correlation 90^0 to the general moisture path.

Convergence Effect. -- When convergence results in the lifting of air masses on broad scales, general precipitation may fall over a large region. For convergence precipitation to occur, lateral movement of air masses need not occur as required for orographic or frontal precipitation. Convergence precipitation tends to sustain higher correlation in precipitation in all directions for long distances. Dominance of convergence storms would override the orienting effects of the other elements.

Frontal Effects. -- Frontal activity will tend to have a strong influence on the orientation of the major axis of correlation. A front will act similarly to an orographic barrier. The effects are complicated since the barrier is moving and changing its orientation, and storm cells tend to be carried along the front by the general circulation. Fronts of Western North America are oriented generally in a southwest-northeast direction. Storm cells traverse generally from the southwest to northeast along the front. The resultant storm path is approximately easterly (Fig. 7). Stenhouse and Cornish [1958] have pointed out that "the two main features of relevance in this case are as follows: (1) slope of the frontal surface which depends upon the temperature difference at the interface of the two air masses, slope varying inversely as the temperature difference, and rainfall varying as the slope; (2) volume of air uplifted, which varies as the pressure gradient, the resultant rainfall varying as the volume of air." The slope of the frontal surface and the pressure gradient decreases southward along the front (Fig. 7). The assumption is that the properties of the air mass remain constant ahead of the front.

Thus, correlations in the amounts of precipitation will tend to be stronger along the front than perpendicular to it. Stenhouse and Cornish [1958] point out, however, that the temperature differences between the air masses and the slope of the front varies irregularly along the front as it moves over land, and that this tends to reduce the correlations along the front. The addition of the storm cell movement strengthens correlations along the front, but rotates the axis clockwise from the frontal direction.

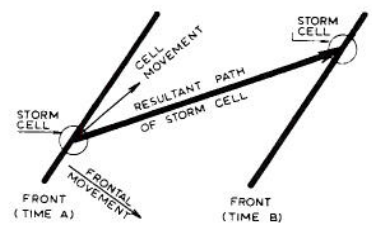

Fig. 7 Resultant frontal storm path.

Convection Effects. -- Convection storms are usually small and very localized. Their occurrence tends to be random in time and space. Thus, their general effect is to reduce the correlations in all directions.

3. Influence of Evapotranspiration upon Inter-station Correlation. Julian [1961] concluded the efficiency of the runoff process as "an important quantity defined rather simply as the percentage of precipitation falling within a basin that actually runs off." This runoff efficiency will vary with climate, the major parameter of climate being the temperature. Julian concluded that evapotranspiration processes are more efficient at higher temperatures, therefore, runoff efficiency is lower at higher temperatures. Schleusener and Crow [1961] presented some rough approximations for that portion of the observed precipitation which is lost to evapotranspiration in the Colorado River Basin. The entire watershed loses over 80 percent, the area below 5000 feet loses over 90 percent, and the area above 11,000 feet loses less than 20 percent. Other basins show similar results. Note that the percentage is less for higher elevations where the evapotranspiration processes are not so active, where the total precipitation is increased, and where the infiltration capacity is usually smaller and the basin slopes much steeper.

Some insights regarding the regional distribution of inter-station correlation coefficients about a point may be gained from a consideration of eq. 2.3, EP = P - E. Evapotranspiration has a varying influence upon the amount and timing of runoff. The effective precipitation represents the difference between the large amounts of precipitation and evapotranspiration. For the whole Colorado River Basin, the effective precipitation is relatively small, 20 percent of the total precipitation. The difference, (P - E), may be quite small in arid or semi-arid and large for humid areas.

The variance of effective precipitation is dependent upon the variances of P and E and their covariance,

$$\text{varEP} = \text{varP} + \text{varE} - 2\ \text{covPE} \qquad 3.1$$

Since

$$r_{PE} = \frac{\text{covPE}}{(\text{varP varE})^{1/2}}$$ eq. 3.1 may be

written as

$$\text{varEP} = \text{varP} + \text{varE} - 2\left[(\text{varP})\ (\text{varE})\right]^{1/2} r_{PE}. \qquad 3.2$$

The cross-correlation r_{PE} is generally positive but less than unity. The relationship of variations of EP and P is greatly dependent on the variation of E and the correlation coefficient between P and E. Therefore, the evapotranspiration factor may make the difference between the patterns in the spacial correlation coefficient of precipitation and effective precipitation to be significant.

Julian wrote [1961] that "the temporal variability of runoff exceeds either that of precipitation or evapotranspiration because it represents the relatively small differences between the latter much larger quasi-independent quantities." In this study, effective precipitation is essentially equivalent to runoff as used by Julian since, for most basins, the change in carryover is small. Thus, it is concluded that, generally, the temporal variability of annual runoff is greater than that of annual precipitation.

Assuming that the evapotranspiration rate is essentially homogeneous over areas the size of the fitting blocks of this study, the axis of maximal correlation of effective precipitation should be influenced by evapotranspiration in a way similar to the influence of the moisture flux upon precipitation. Evapotranspiration contributes moisture to the atmosphere which augments the source moisture present and should be concentrated downwind. Evapotranspiration rates are inversely proportional to the moisture present in the atmosphere, and the losses of water should be reduced downwind. Precipitation should be augmented downwind by moisture from evapotranspiration, but only to a small degree. The axis of minimal correlation should be transverse to the winds since the intensity of the winds will vary laterally and evaporated moisture will be diffused laterally. The evapotranspiration processes will generally tend to change the inter-station correlation coefficients of effective precipitation in comparison with those of precipitation.

4. <u>Combined Effects of the Causal Factors.</u> Many factors discussed in this chapter are operating simultaneously. When using a long period variable such as the annual values in this study, the various effects tend to offset or add together either to produce isotropic effects or to accentuate the ellipticity of the isolines of inter-station correlation coefficients.

As seen in Section 3, the axis of maximal correlation for the EP variable should be downwind. Evapotranspiration moisture is more closely tied to surface winds than to the general circulation. This should create some difference in the orientation of the maximal axes of correlation in precipitation and effective precipitation.

Certain sections of the study region receive moisture from predominantly one source the whole year. The Pacific Northwest receives most of its moisture from the North Pacific Ocean, and the Gulf Coast area and the lower Mississippi River Basin receive their moisture mainly from the Gulf of Mexico. However, the areas in Arizona, New Mexico, Utah, and Colorado west of the Continental Divide receive large portions of moisture from three sources (directions): the North Pacific Ocean, the South Pacific Ocean, and the Gulf of Mexico. The area in the northeast section east of the Continental Divide receives significant moisture from the North Pacific Ocean and the Gulf of Mexico.

The moisture flux from these different sources arrive over the various regions at different times of the year. Marlatt and Riehl [1963] stressed that the period of maximum precipitation for San Francisco, California, in the Pacific Coast is in the winter months, that Denver, Colorado, east of the Rocky Mountains receives its maximum precipitation in the spring and summer, and that the Upper Colorado River basin which lies between the above cities receives its precipitation rather uniformly throughout the year. The axis of maximal correlation for the precipitation variable would tend to align itself with the seasonal flux of moisture. Equally influential moisture sources are likely to create a near isotropic situation. The further investigation of inter-station correlation coefficients using the semi-annual variables (or other divisions of the year) of effective precipitation and precipitation should indicate the relative influence of the different moisture sources and give further insight into the varying effects of the factors discussed earlier in this chapter.

Should studies similar to this study be conducted using shorter period variables, it is suggested that the yearly period be first divided on the basis of the periods influenced by different moisture sources. The periods do not have to be of equal length but should begin and end with the change of moisture source. For example, the area of Arizona, Southern Utah, New Mexico, and Southern Colorado west of the Continental Divide receives near equal amounts of precipitation from the Gulf of Mexico in the summer and the North Pacific Ocean in the winter. The period of intense flow of Gulf moisture is fairly well limited to the period June through September. Each year might be divided into two periods with June through September for one and the remainder of the year's precipitation for the other. Although it would require more work, each year could be subdivided individually considering the period actually known to receive dominant moisture amounts from the different sources, or a statistically determined mean date for the change from one source to the other could be used to facilitate research data collection.

The shortening of the period of observation could show a closer relationship between the axis of maximal correlation and the path of moisture flux in the case where the region has two or more primary sources of moisture.

CHAPTER IV

MATHEMATICAL TECHNIQUE AND SELECTION OF THE REGRESSION EQUATION

The surface fitting of the type utilized in this study is a standard technique. The surfaces are fitted to the computed inter-station correlation coefficients by the method of least squares. The techniques employed are analogous to approaches used in the investigations of chemical processes and in the analysis of geophysical and meteorologic data. There is extensive literature on the subject in these areas, and the general approach seems quite adaptable to hydrologic studies.

1. The Mathematical Model. This section is devoted to a discussion of the general theory of "response surfaces," the use of the method of least squares, and the tests for goodness of fit. The discussion follows closely that given by Box [1954].

An investigator is concerned with the determination of several aspects of a functional relationship

$$Y = \phi\ (x_1, x_2, \ldots x_n) \qquad 4.1$$

which connects the "response" Y with n other quantitative variables or factors. For two factors, the response function is three-dimensional, and the term "response surface" is applied. The response surface is generally assumed to be some sort of mound or depression so that the response levels can be represented by contours as in fig. 8.

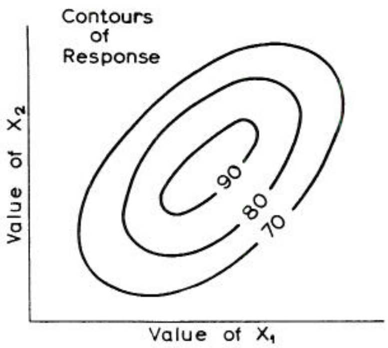

Fig. 8 Response for two factors.

In hydrologic investigations, the response may be explored by fitting some preselected function sufficiently flexible to describe a set of selected data points and the observed response values at those points. The independent factors will often be the map coordinates, and the dependent variable will be the response being studied.

A form of the response function found generally suitable in geophysical "trend surface" analysis is

$$Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1^2 + \alpha_4 x_2^2 + \alpha_5 x_1 x_2. \qquad 4.2$$

This form may be fitted by the method of least squares. Using this technique, the estimates a_o, a_1, a_2, a_3, a_4 and a_5 of α_1, α_2, α_3, α_4 and α_5 are determined in such a way that the sum of the squared differences between the observed values and those predicted by the fitted equation is the minimum.

The sampling error variance s_e^2, is given by the equation

$$s_e^2 = \frac{\sum_{i=1}^{N} (y_i - Y_i)^2}{(N - L)} \qquad 4.3$$

where y_i is the observed value, Y_i is the predicted value, $(N - L)$ is the number of degrees of freedom, and L is the number of parameters in eq. 4.2. If an "a priori" value of s_e^2 is known, the sample estimate from eq. 4.3 allows the analyst to obtain some knowledge of the goodness of fit of the postulated equation.

In the case where there is no "a priori" knowledge of the sampling error variance, the percentage of the total variation explained by the regression function may be utilized in conjunction with observed deviations from the regression function to evaluate the goodness of fit:

$$E_{var} = \frac{\sum_{i=1}^{N} (Y_i - \bar{y})^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \times 100 \qquad 4.4$$

where E_{var} is the value in percent, $\bar{y}$ is the mean of y_i, and the other terms are the same as defined above. An explanation of 50 percent of the variation (a coefficient of determination of 0.5) seems a reasonable acceptance level for a regression model. The corresponding correlation coefficient is 0.707. However, this alone should not be used for final acceptance. The deviations from the regression function should be studied for systematic patterns or features not explainable by natural processes.

The total response must also be considered. Regressions approximating small portions of a larger medium may give very high coefficients of determination, but when the continuous medium beyond the bounds of the fitted region are considered, the resulting regression becomes completely unrealistic. Figure 12 is an example of this, and fig. 18 shows the surface produced by the selected function of this study.

2. The Selection of the Regression Function. Very little is known about the areal or regional variation of the inter-station correlation coefficients within the variables of annual precipitation and annual

effective precipitation. Experience has shown that values of correlation coefficients generally diminish with distances [Gatewood and others, 1964; Julian, 1961]. However, nothing was found in the literature about the physical processes of precipitation and runoff that pointed to a mathematical function which explains the regional variation of inter-station correlation coefficients about a central station. Further studies of the factors discussed in Chapter III which act to vary the degree of the correlation in the variables studied might lead to the determination of a physical mathematical model, but that was beyond the scope of this study. The physical mathematical function should be used when available, although an empirical function might explain a greater percentage of the regional variation. Not having the physical model in this case meant that a suitable empirical function had to be selected.

Previous studies using correlation coefficients of runoff and precipitation data have shown considerable variation in coefficient values caused by localized hydrologic factors and inaccuracies and inconsistencies in the data. The need of a simple function to act as a smoothing function more representative of the gradual ever-changing atmospheric and hydrologic processes is assumed. Excessive deviations from the smoothing function should be explainable by localized phenomena. This is not always possible. Furthermore, one should not try for an accuracy that cannot be achieved or justified considering the accuracy of his research data.

Some preliminary studies of correlation coefficients from station data along selected meridians and parallels of latitude were particularly helpful in giving clues to the form of the function (Figs. 9 and 10). Three or four stations which are indicated by the double circles were selected along each line, and the inter-station correlation coefficients were computed from the data of each of the four stations with the data of all other stations along the line. The stations did not fall exactly upon the line, but the largest portion fell within a distance of $\pm\ 0.5^0$ of latitude or longitude from the line. Where the stations are sparse, some selections were made that fell as far as $\pm\ 1.0$ from the line. The resulting sets of coefficients were plotted at their respective station locations on each line. In addition, the profile of the topography was plotted for each line. These plots indicate that the coefficients decrease with distance and approach zero, fluctuating randomly about zero beyond some distance. Some exceptions are noted, but when they are related to the topography, most of the exceptions can be explained. The line plots definitely show the same form of change in both the west-east and the south-north directions. The form of the change appears to be essentially symmetric in opposite directions from a station.

Isolines of inter-station correlation coefficients about a station show a definite elliptic form in most cursory plots. Therefore, there are indications that the rate of decrease of the correlation coefficients with distance from the station varies with the direction and that this variation changes across the study region. Thus, a function is needed whose slope changes with direction and which is approximately symmetrical about the central station on a given axis. A comparison of the rates of decrease of the coefficients along the perpendicular parallels of latitude and meridians shows pronounced differences in the rates. This contrast in the rates of decrease and a consideration of the factors discussed in Chapter III indicate that the axes of the maximal and minimal correlations are likely orthogonal.

In summary the surface should:

a. Be simple to effect a smoothing operation;

b. Be symmetric about its origin on any axis;

c. Have orthogonal major and minor axes; and;

d. Produce surface contour lines that may vary from circular to elliptic and even to a ridge form if the local surface fit requires it.

From the line studies alone, it is not possible to determine whether a section through the surface looks bell-shaped (normal curve), triangular (linear change with distance), or cuspal (negative exponential), but the additional more specific criteria must be added at this point:

a. The maximum response should be at $Y(0,\ 0),\ Y(x,\ y) \leq |1.0|$;

b. Within the limits of the control block, the function should be positive and decreasing; and,

c. $Y\ (0,\ \pm\ \infty) = Y\ (\pm\ \infty,\ 0) = 0$.

That the response be zero at its infinite bounds is not out of order if the small positive and negative correlations (not significant) at extreme distances are due to chance.

Functions similar to the bell-shaped normal curve exhibiting the general and specific criteria outlined are:

$$r_e = a_1 e^{-(a_2 x^2 + a_3 y^2 + a_4\, x\, y)} \qquad 4.5$$

$$r_e = (a_1 + a_2\, x^2 + a_3\, y^2 + a_4\, x\, y)^{-1} \qquad 4.6$$

where "r_e" is the response or the estimate of the inter-station correlation coefficient for a given distance and direction and where x and y are the cartesian distances for the control block as defined in Chapter II, Section 6 (Fig. 11). The terms a_1, a_2, a_3, and a_4 are the estimates of the regression coefficients. The specific criteria are satisfied if $a_1 = 1$; $a_2,\ a_3\ > 0$; and $a_4{}^2 - a_2\, a_3 < 0$.

Preliminary fittings of various functions to the inter-station correlation coefficients within the blocks about the block control stations indicated that the specific criterion that a_1 as unity causes very poor explanations of the variation in the correlation coefficients. By allowing a_1 to differ from unity, a greater percentage of the variation was explained. However, the average percentages of the variation explained in test fittings for both variables in this study were less than or very near the 50 percent minimum desired for both of the eqs. 4.5 and 4.6. Aside from the fact that the specific criteria are not satisfied, inclusion of higher degree terms in x and y and linear terms through the absolute values of x and y did give an increase in the variation explained, but many of the resulting fitted surfaces assumed unrealistic shapes (Fig. 12).

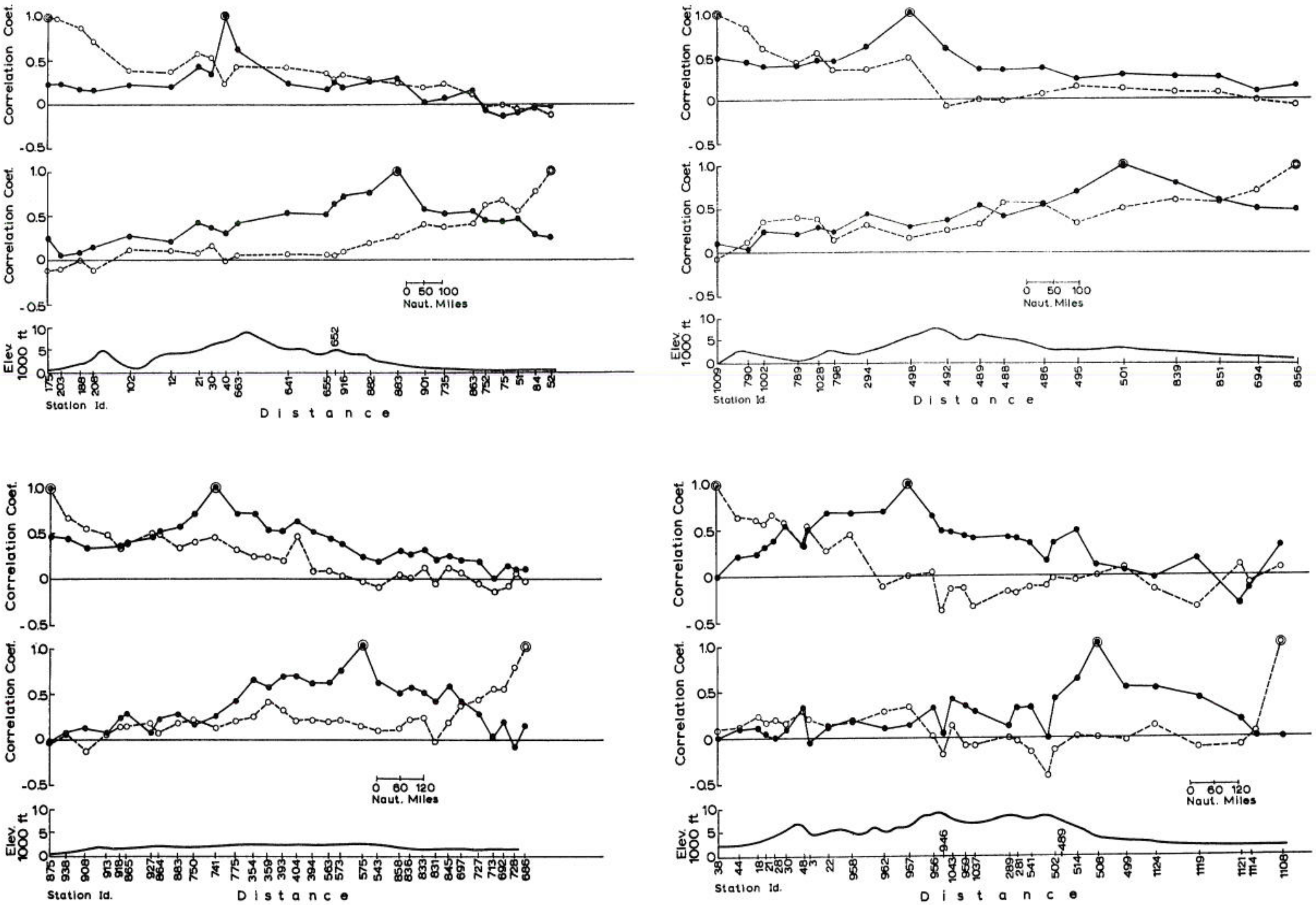

Fig. 9 Inter-station correlation coefficients for the annual precipitation variable (P) at stations along or near meridional or latitudinal lines. The station identification number is plotted at the station positions along the abscissa (distance). Double circles indicate stations whose variables were correlated with those at other station locations. The lowest plot represents the general profile of the topography along the line of study. Coefficients are plotted for latitudinal lines at 34^{0}N (Top, left) and 45^{0}N (Top, right) and for meridional lines at 100^{0}W (Lower, left) and 111^{0}W (Lower, right).

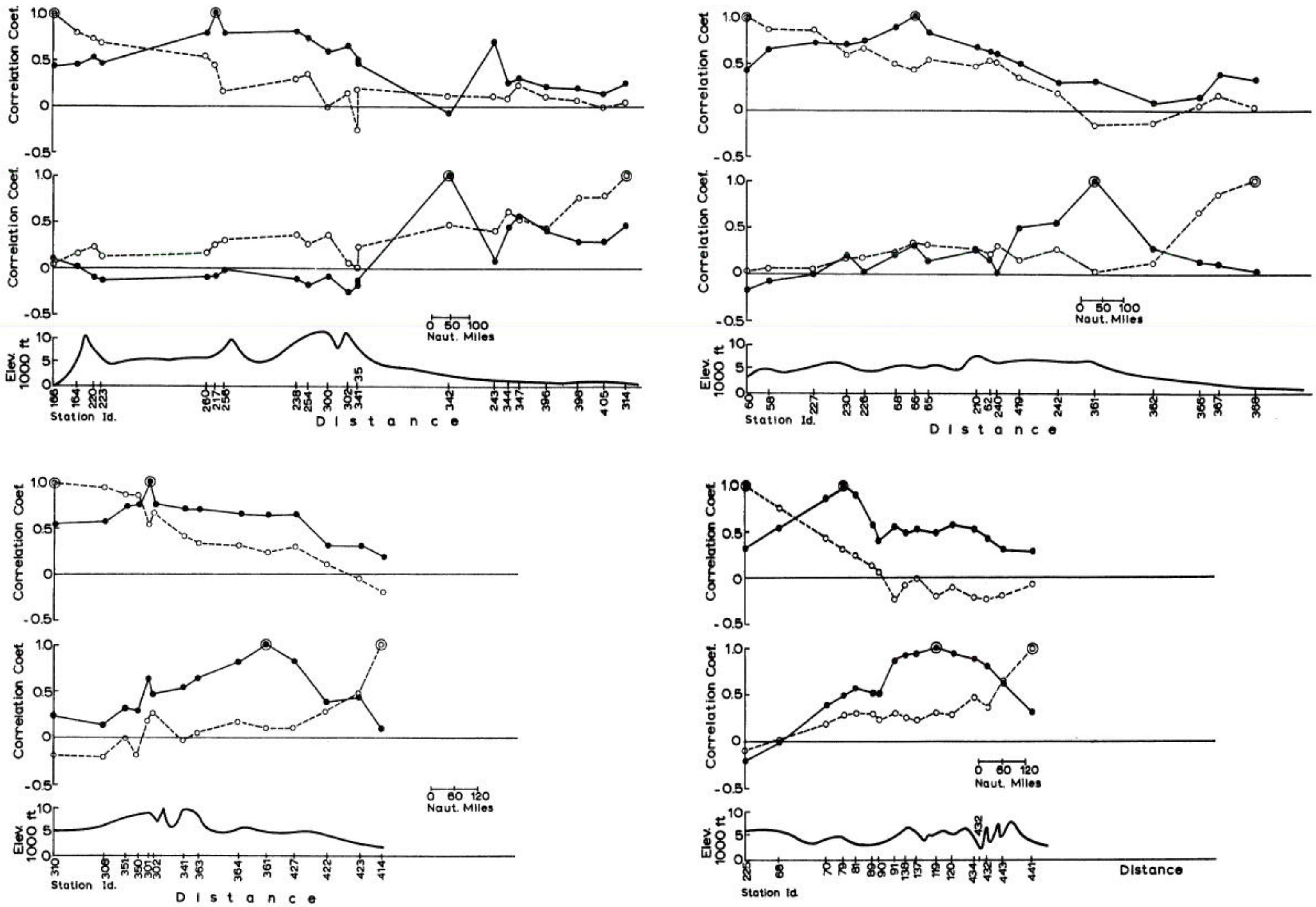

Fig. 10 Inter-station correlation coefficients for the annual effective precipitation (EP) at stations along or near meridional or latitudinal lines. The station identification number is plotted at the station position along the abscissa (distance). Double circles indicate stations whose variables were correlated with those at other station locations. The lowest plot represents the general profile of the topography along the line of study. Coefficients are plotted for latitudinal lines at 38° N (Top, left) and 42° N (Top, right) and for meridional lines at 105° W (Lower, left) and 116° W (Lower, right).

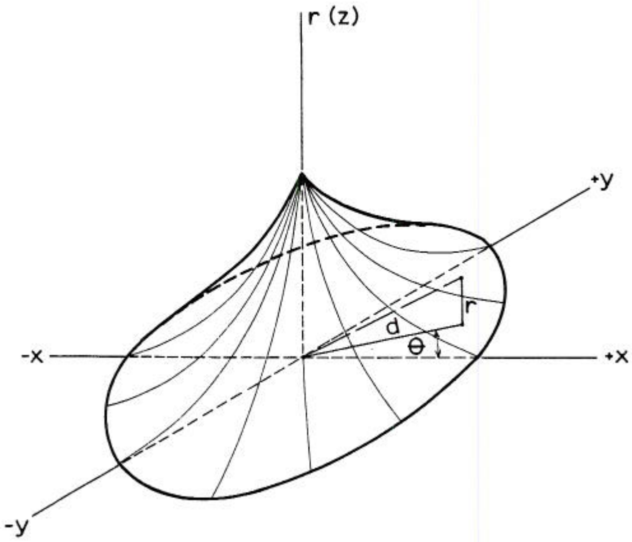

Fig. 11 Diagram of surface elements.

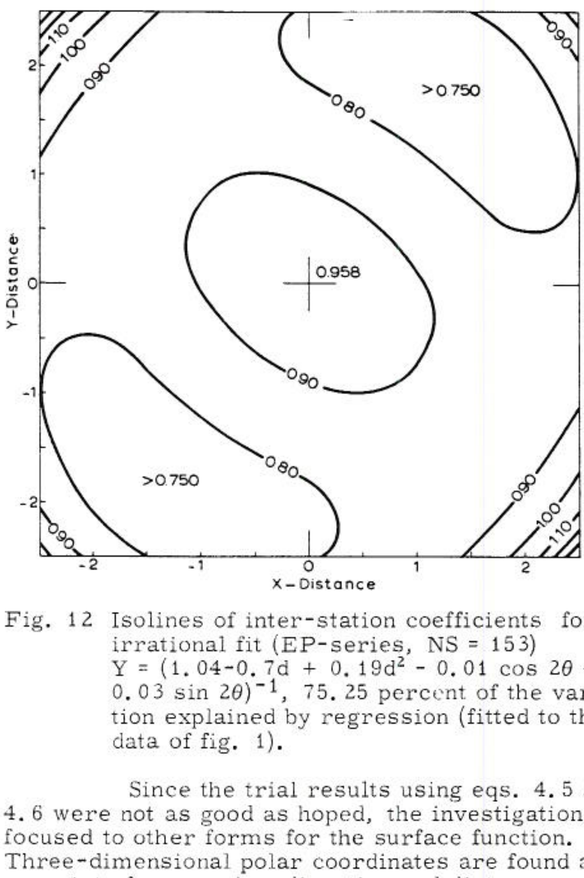

Fig. 12 Isolines of inter-station coefficients for an irrational fit (EP-series, NS = 153) $Y = (1.04 - 0.7d + 0.19d^2 - 0.01 \cos 2\theta + 0.03 \sin 2\theta)^{-1}$, 75.25 percent of the variation explained by regression (fitted to the data of fig. 1).

Since the trial results using eqs. 4.5 and 4.6 were not as good as hoped, the investigation was focused to other forms for the surface function. Three-dimensional polar coordinates are found appropriate for use when direction and distance are prime variables, and functions investigated were in the form:

$$r_e = f(d, \theta) \qquad 4.7$$

where r_e is the response value (z-axis); d is the radial distance from the control block center; and, θ is the normal trigometric angle (angle from due east line, fig. 11). The equation

$$r_e = a_1 + a_2 d + a_3 d^2 + a_4 d \cos 2\theta + a_5 d \sin 2\theta \qquad 4.8$$

was found to be suitable by Stenhouse and Cornish [1958] and Cornish, Hill and Evans [1961] in their work similar to this study.

Several variations of eq. 4.8 were applied to twenty test blocks selected from each of the P and EP variables, since it showed great promise of meeting the objectives. Some of the variations tested are:

$$r_e = a_1 + a_2 d + a_3 d^2 + a_4 d^3 + a_5 d \cos 2\theta + a_6 d \sin 2\theta \qquad 4.9$$

$$r_e = (a_1 + a_2 d + a_3 d^2 + a_4 d \cos 2\theta + a_5 d \sin 2\theta)^{-1} \qquad 4.10$$

$$r_e = (a_1 + a_2 d + a_3 d^2 + a_4 \cos 2\theta + a_5 \sin 2\theta)^{-1} \qquad 4.11$$

$$r_e = (a_1 + a_2 d + a_3 d^2 + a_4 d^3 + a_5 d \cos 2\theta + a_6 d \sin 2\theta)^{-1} \qquad 4.12$$

$$r_e = (a_1 + a_2 d^2 + a_3 d \cos 2\theta + a_4 d \sin 2\theta)^{-1} \qquad 4.13$$

$$r_e = (a_1 + a_2 d + a_3 d \cos 2\theta + a_4 d \sin 2\theta)^{-1} \qquad 4.14$$

$$r_e = a_1 e^{-(a_2 d + a_3 d \cos 2\theta + a_4 d \sin 2\theta)} \qquad 4.15$$

Equations 4.9 through 4.12 were rejected because the shapes of the surfaces near and outside the bounds of the control block were often unrealistic even though they explained more of the variation in the correlation coefficients. Equation 4.13 showed a smaller percentage of explained variation than eqs. 4.14 and 4.15.

Equations 4.6, 4.14, and 4.15 showed similar applicability. Comparative figures are given in Table 1. Overall, it is apparent that eqs. 4.14 and 4.15 are better suited than eq. 4.6. The explained variation is lowest for eq. 4.6, and its maximum values do not approach the theoretical value of unity within reasonable limits. In the case of effective precipitation, eq. 4.15 seems better than eq. 4.14. This choice for the precipitation variable is not that apparent. The explained variation of eq. 4.14 is 0.89 percentage points greater. The number of maximum values greater than 1.0 favors 4.15 three to one. But, eq. 4.15 thereby sacrifices a higher average maximum value.

The choice of either eq. 4.14 or eq. 4.15 would serve the purpose of the inquiry concerning the correlation coefficients of precipitation. Equation 4.15 was chosen for the study of precipitation correlation coefficients because of easier direct comparison between the two variables, P and EP, and because of its mathematical simplicity.

3. Selected Function and Its Characteristics. The function utilized is eq. 4.15. This function is linear in d, symmetric about its major and minor axes, and always positive for any values of a_2, a_3, and a_4 if $a_1 > 0$. The major axis may be found by maximizing the right hand side of the equation in θ. First, by converting to logarithmic form, eq. 4.15 becomes

$$\ln r_e = \ln a_1 - (a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d \qquad 4.16$$

which reduces to

$$\frac{\ln a_1 - \ln r_e}{d} = a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta. \qquad 4.17$$

The maximization of the right hand side of eq. 4.17 gives

$$\frac{d\left[\dfrac{\ln a_1 - \ln r_e}{d}\right]}{d\theta} = -2a_3 \sin 2\theta + 2a_4 \cos 2\theta = 0. \qquad 4.18$$

Solving for θ, one obtains

$$\theta = \frac{1}{2} \arctan a_4/a_3. \qquad 4.19$$

Because of the selection of terms and the manner of fitting, the axes of maximal and minimal correlation will be at right angles.

The angle θ is measured from the east radial and may be readily converted to an azimuth angle, ϕ. The rate of change of the correlation coefficient with distance at any point on the surface is

$$\frac{\partial r_e}{\partial d} = -\mu e^{-\mu d} \qquad 4.20$$

where $\mu = a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta$. The term μ evaluated for $\theta = \theta_{max}$ and $\theta = (\theta_{max} + \pi/2)$ will give the coefficients C_1 and C_2 which are constant for the cross-sections along the major and minor axes. The coefficients C_1 and C_2 serve as surface parameters showing the rates of change of the correlation coefficients along the major and minor axes, respectively. The ratio of C_1 divided by C_2 serves to show the degree of ellipticity. The correlation along the maximal axis increases with distance rather than decreases if $C_1 < 0$, and the isolines of correlation coefficients are no longer closed contours. When both C_1 and C_2 are negative, the surface becomes a closed depression which will have closed contours. The equations for C_1 and C_2 are

$$C_1 = a_2 + a_3 \cos 2\theta_{max} + a_4 \sin 2\theta_{max} \qquad 4.21$$

$$C_2 = a_2 + a_3 \cos 2(\theta_{max} + \pi/2) + a_4 \sin 2(\theta_{max} + \pi/2) \qquad 4.22$$

The four regression coefficients, a_1, a_2, a_3, and a_4 are determined by the method of least squares and from these the four descriptive parameters of the surface are obtained. The parameters are:

a. ϕ, the azimuth to the major axis;

b. C_1, parameter for the rate of decrease for the major axis;

c. C_2, parameter for the rate of decrease for the minor axis; and,

d. $R_c = C_1/C_2$, parameter for the degree of ellipticity.

CHAPTER V

COMPUTATIONS AND PRESENTATION OF RESULTS

The procedure and techniques used in computing the various coefficients and surface parameters are detailed in this chapter. The manner of their presentation in the several figures is also indicated.

1. Control Block Inter-station Correlation Coefficients. Each station in both sets of variables, annual precipitation (P) and annual effective precipitation (EP), is used as a "block control station" (BCS), which is at the center of the square area surrounding it. The inter-station correlation coefficients are computed between the data of the BCS and the data of all stations within the bounds of the control station block using eq. 1.1. The variables used for these computations are more specifically defined here than was done in Chapter I, Section 1:

X_i represents the observed annual values at the BCS;

Y_i represents the observed annual values at the other stations as their station series are correlated one at a time;

N is the number of paired observations (N is less than or equal to the minimum length of the time series records of X_i and Y_i depending on the chronological coincidences of annual observations);

$\sum_{i=1}^{N}$ is summed over the N number of chronologically matched pairs of X_i and Y_i.

The BCS correlated with itself gives an inter-station correlation coefficient of 1.0. The number of coefficients in a block is equal to the total number of stations in the block.

A minimum number of nine stations per block was established to assure an adequate number of values to maintain a smoothing effect. The width of the block was allowed to increase until this number was included. Block sizes range from 5 to 7 units (a unit equals 60 nautical miles) for effective precipitation and from 5 to 6 units for precipitation. Only 1139 out of the 1141 precipitation station control blocks included enough stations for a surface fitting. All 446 of the effective precipitation station control blocks contained the minimum number of stations. The computations of the inter-station correlation coefficients for each selected BCS were done by digital computer and stored on magnetic tape for use by the surface fitting program which was utilized in the next stage of the computations.

2. Subdivision Inter-station Correlation Coefficients. A secondary computation, performed at the time the BCS coefficients were determined, evaluated all possible inter-station correlation coefficients for certain gross subdivisions. At the same time, the distribution (and its parameters) were calculated for the coefficients of each subdivision.

Precipitation Gross Subdivisions. -- The gross subdivisions for which these computations were performed are shown in fig. 13 for precipitation. The statistics for the resulting frequency distributions of the coefficients are also given on these figures. Two other samples of precipitation data were also used: (1) the set of 235 stations that was selected for use in the line studies of inter-station correlation coefficients (Chap. IV, Sect. 2); and (2) a group of 472 precipitation stations for which data was collected but whose stations were not included in the study sample. The two latter samples are from the study region. The 235-station set is a subsample of the precipitation sample selected for this study. The 472-station sample represents those stations meeting all the selection criteria imposed upon the precipitation stations selected for this study except that their records were considered to be inconsistent. The stations of the 235-station subsample were selected from a band along the parallels of latitude and meridians of the line study, and therefore, the stations are not uniformly distributed over the region. However, the 472-station sample has its stations reasonably uniformly distributed over the study region.

Effective Precipitation Subdivisions. -- The matrices of inter-station correlation coefficients are computed for all the gross subdivisions and the total sample for the effective precipitation variable. The gross subdivisions are shown in fig. 16. The statistics of the frequency distributions of the coefficients from each subdivision are also shown on these figures.

Distribution of the Inter-station Correlation Coefficients. -- The frequency distributions of the coefficients resulting from the computations for each subdivisions and the other samples are given in figs. 14 and 15 for the precipitation variable and in fig. 17 for the effective precipitation variable. The 472-station sample distribution is referred to as "P2" in fig. 15. The statistics of the frequency distributions for the 235-station and the 472-station precipitation samples and the total effective precipitation sample are also given in these figures.

3. Surface Fitting Procedure. The basic program used to fit the selected function to the previously computed coefficients of each BCS in the IBM 1620 Users Group Library Program, Number 6.0.134, entitled "Non-linear Least Squares Curve Fitting Program," by George Struble. The program fits a finite number of points with a function of arbitrary non-linear form in the least squares sense. The numerical method used is that of Newton iteration from a first guess at the coefficients in the function. Each iteration linearizes the function at the current estimate of the coefficients. The program is capable of handling an unlimited number of observations for up to eight independent variables and one dependent variable. Functions may be used with, at most, six coefficients. The Newton iteration method does not always converge for all functions and all initial guesses.

According to the IBM library write-up of the program, the accuracy of the surface fit is highly data-dependent, and an "a priori" analysis of errors would be quite difficult. However, a round-off errors will, in general, be much less than the errors in the raw data. Deviations from a perfect fit of the function will likewise be much less than the deviations of the fitted data. This program was utilized throughout the exploratory and final fittings because of its generalized nature. The exact solutions for the selected function would have required the writing of a new program.

A complete fitting of the surface function is made to the points of each BCS of both variables (P and EP). The independent variables of the fitted function (Eq. 4.15) are d and θ as defined in Chapter IV, Section 3. The dependent variable is the inter-station correlation coefficient. The following information is also produced by the basic program for the function fitted:

a. The inverse of the matrix of the function coefficients;

b. The variance of the deviations from the fitted function;

c. The final estimates of the coefficients of the fitted function (a_1, a_2, a_3 and a_4), and

d. The standard error of estimate of each of the coefficients of the fitted function.

A typical example of the resulting fitted surface is shown in fig. 18.

The coefficients and parameters ϕ, a_1, a_2, C_1, C_2, and R_c were summarized for display of their regional variation. The first step was to find the average values of the estimated regression coefficients, a_1, a_2, a_3, and a_4, for block control stations falling in the non-overlapping areas having dimensions of 2^0 latitude by 2^0 longitude. The summarized coefficients and parameters were then computed by equations

$$\theta'_{max} = \frac{1}{2} \arctan a'_4/a'_3 \qquad 5.1$$

$$C'_1 = a'_2 + a'_3 \cos 2\theta'_{max} + a'_4 \sin 2\theta'_{max} \qquad 5.2$$

$$C'_2 = a'_2 + a'_3 \cos 2(\theta'_{max} + \pi/2) + a'_4 \sin 2(\theta'_{max} + \pi/2) \qquad 5.3$$

$$R'_c = C'_1/C'_2 \qquad 5.4$$

where the prime (') denotes the average value of each coefficient and the parameters computed from them. The average azimuth, ϕ', is readily converted from θ'_{max}.

Maps were drawn for the study region with the summarized values placed at the center of each 2^0 by 2^0 area. Some plotting points fall outside the region bounds. If summary areas did not include BCS locations, no values were plotted at those points. The resulting isolines were drawn to depict the regional variation of those coefficients and parameters displayed except for ϕ' and R'_c. The azimuth, ϕ', is shown by a line through the plotting point, and the degree of ellipticity, R'_c, is written at the point and visually displayed by a shaded ellipse whose ratio of minor to major axes approximates the value of R'_c. These maps are found in the following chapter.

The summarized values of a'_3 and a'_4 are given in tables 2 through 5. These values along with values a'_1 and a'_2 taken from the appropriate maps may be utilized to compute the surface for any summary area. The summary area plotting points are listed by rows and columns as shown on figs. 25 and 26.

Figure 19 shows the distributions of the computed regression coefficients a_1, a_2, a_3 and a_4 for the precipitation variable. Figure 20 shows the corresponding distributions for the effective precipitation variable. Figure 21 shows the distributions of the computed rate of decrease parameters C_1 (major axis) and C_2 (minor axis) for both the precipitation variable and the effective precipitation variable.

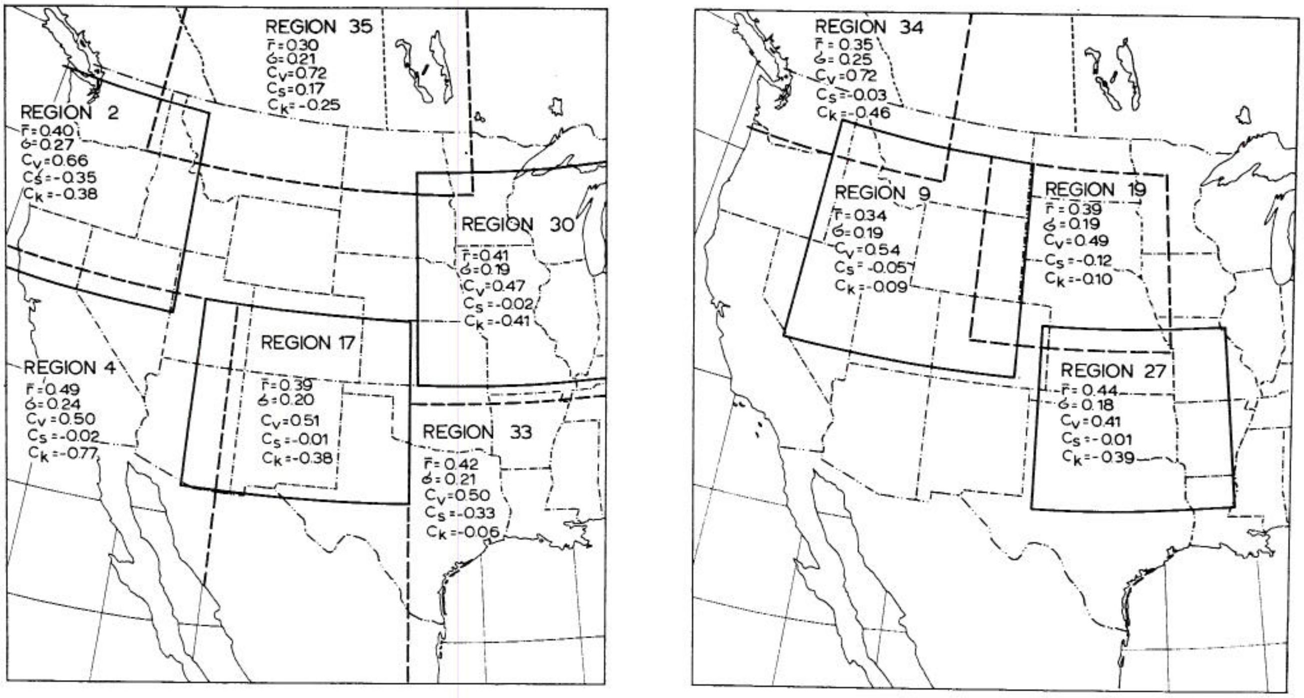

Fig. 13 Delineation of the selected gross subdivisions for precipitation and the statistics for the distributions of the inter-station correlation coefficients between all stations in each of the gross subdivisions.

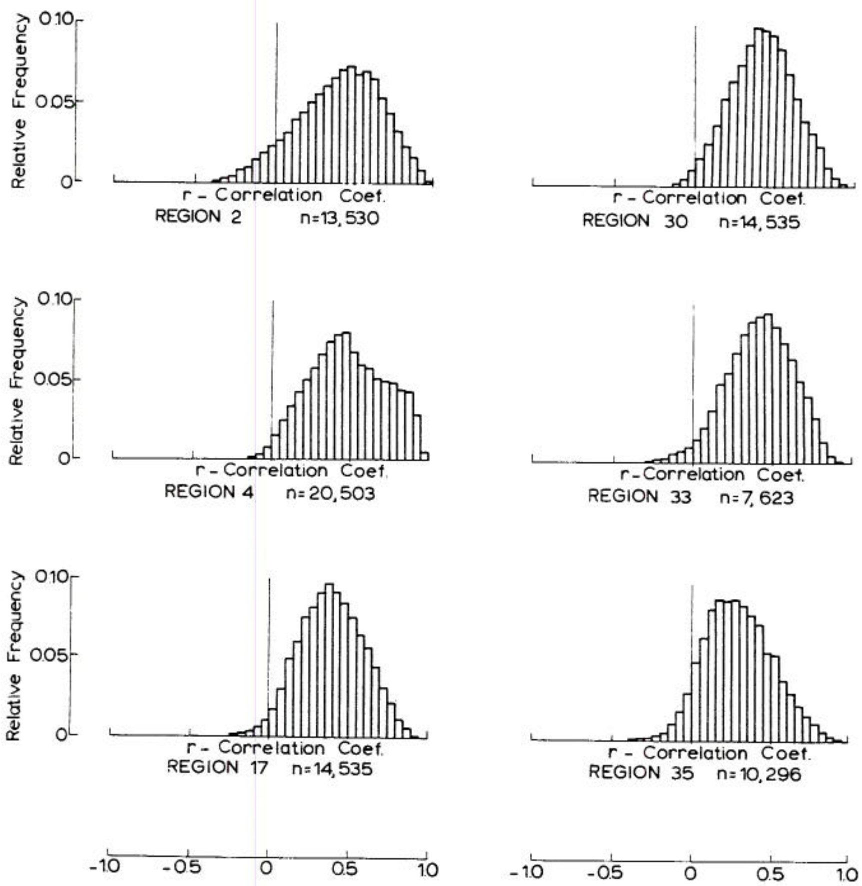

Fig. 14 Distributions of inter-station correlation coefficients for precipitation gross subdivisions, numbers 2, 4, 17, 30, 33 and 35 (n is the number of computed coefficients).

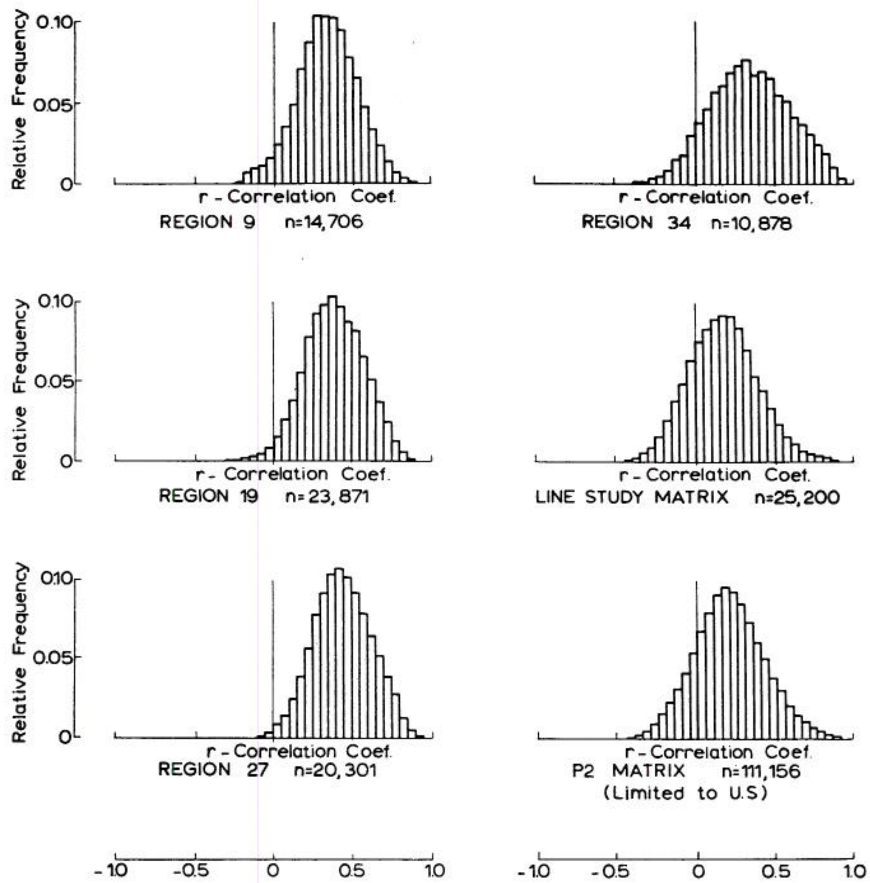

Fig. 15 Distributions of inter-station correlation coefficients for precipitation gross subdivisions, numbers 9, 19, 27, 34, line study subsample, and nonhomogeneous sample (n is the number of computed coefficients).

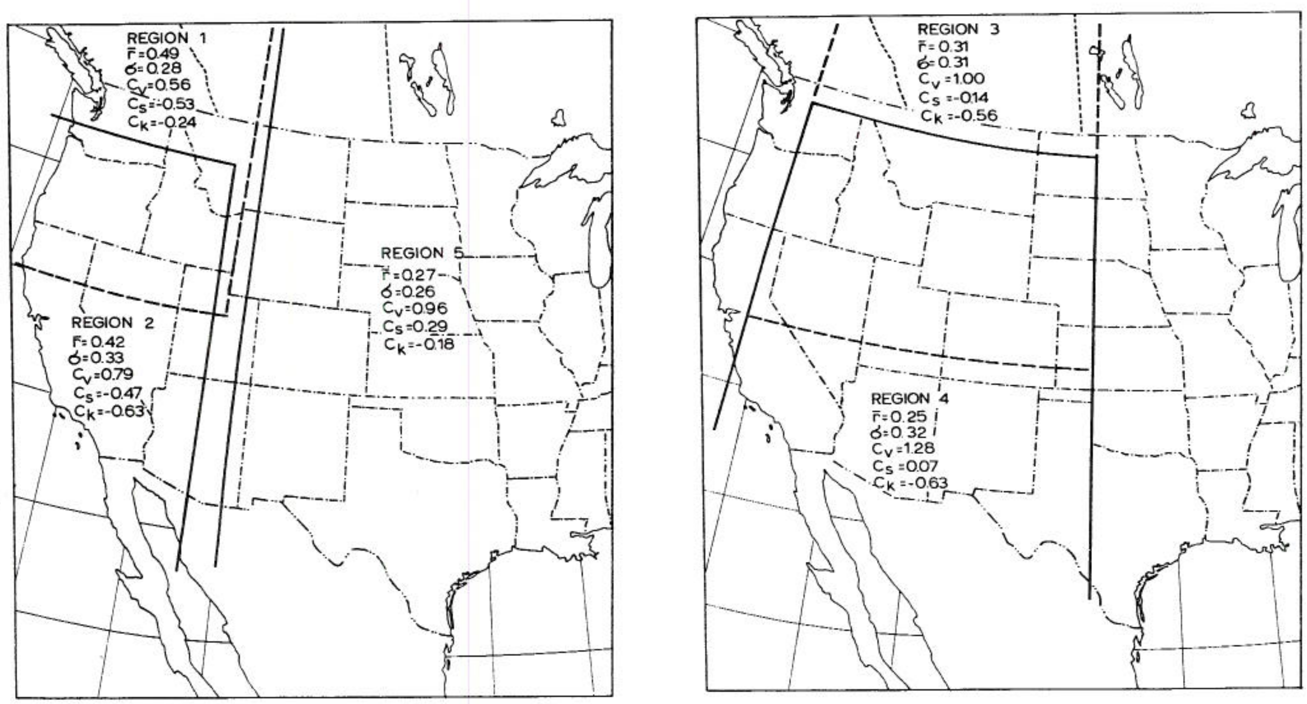

Fig. 16 Delineation of the gross subdivisions for effective precipitation and the statistics for the distributions of the inter-station correlation coefficients between all stations in each of the gross subdivisions.

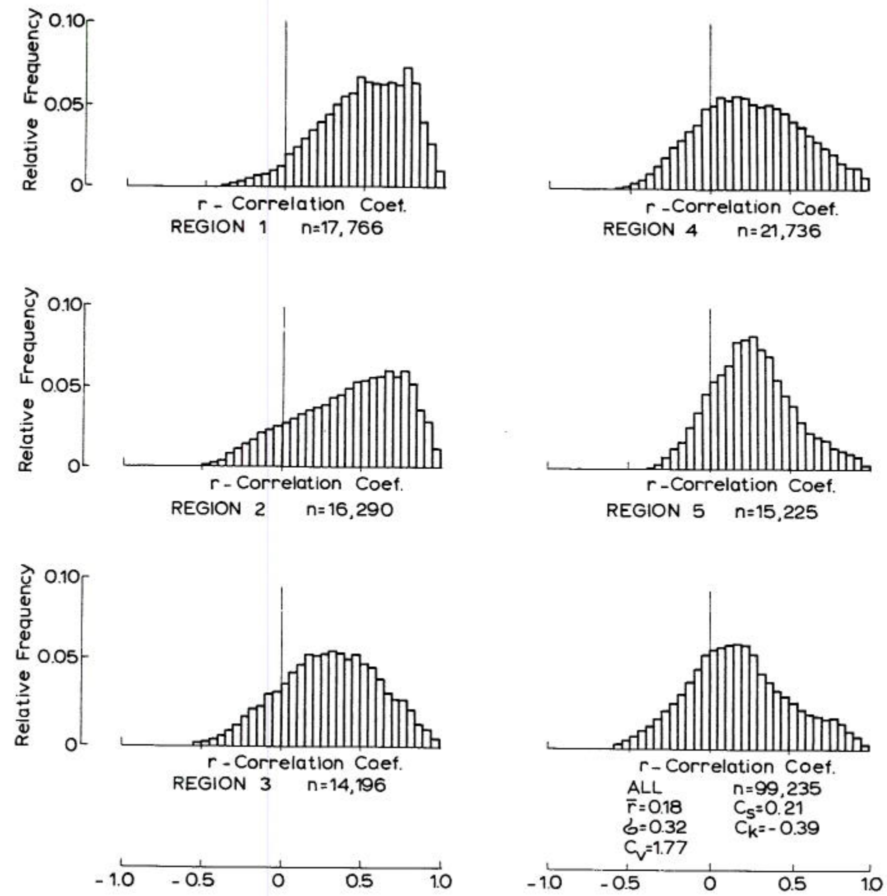

Fig. 17 Distributions of inter-station correlation coefficients for the effective precipitation gross subdivisions and the total area (n is the number of computed coefficients).

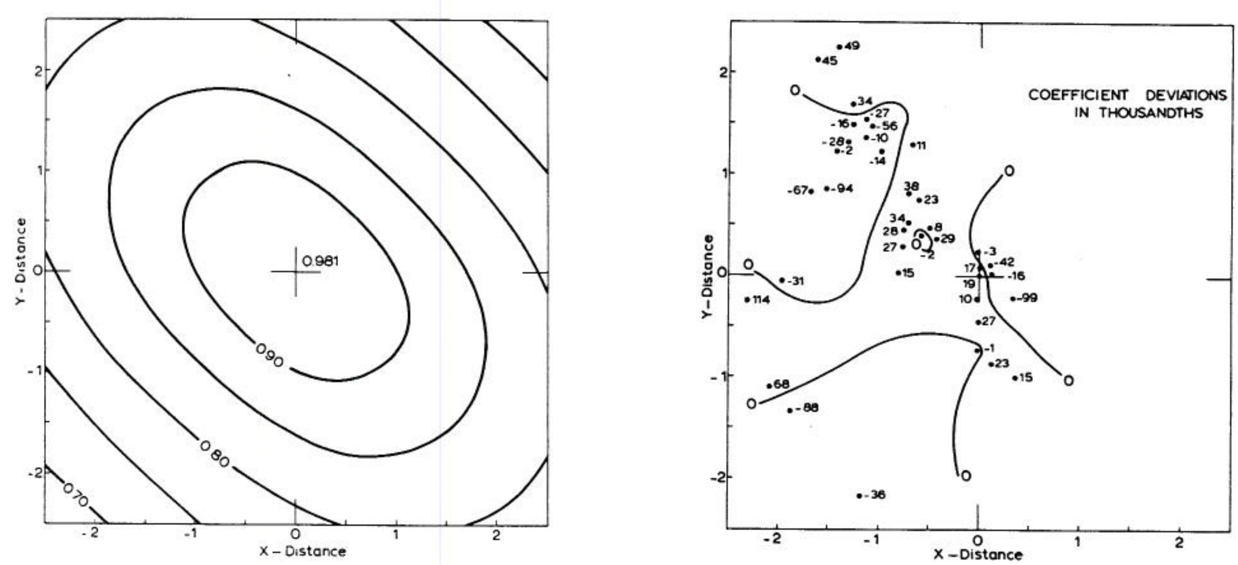

Fig. 18 Isolines of inter-station coefficients for the selected function (EP-series, NS = 153) $Y = 0.98 \operatorname{Exp}[(-0.087 + 0.002 \cos 2\theta - 0.021 \sin 2\theta)d]$, 65.47 percent of the variation explained by regression (Top, left) and the deviations of the actual coefficients (Fig. 1) from the fitted surface (Top, right).

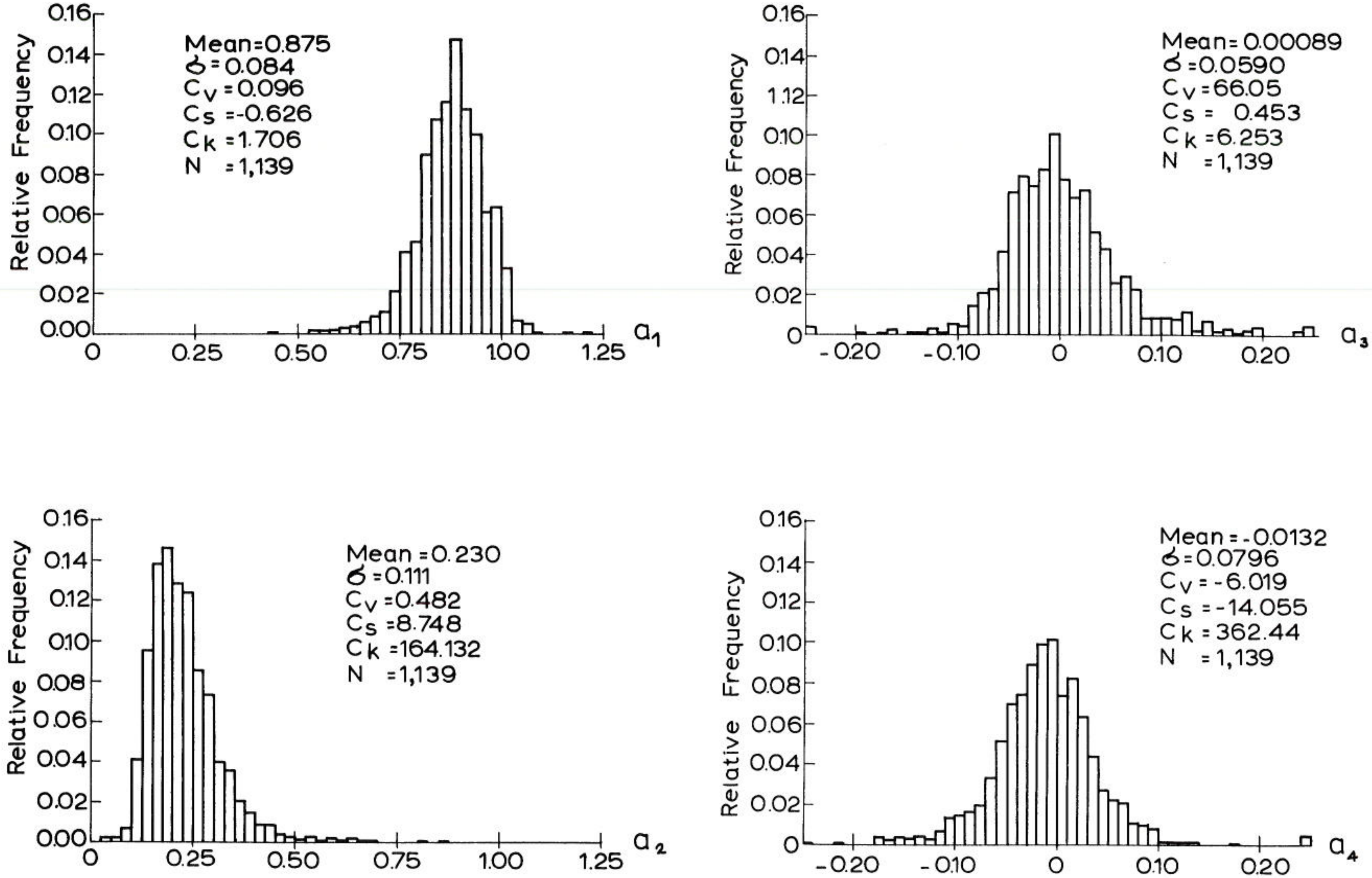

Fig. 19 Distributions of the regression coefficients a_1, a_2, a_3 and a_4 for the precipitation variable.

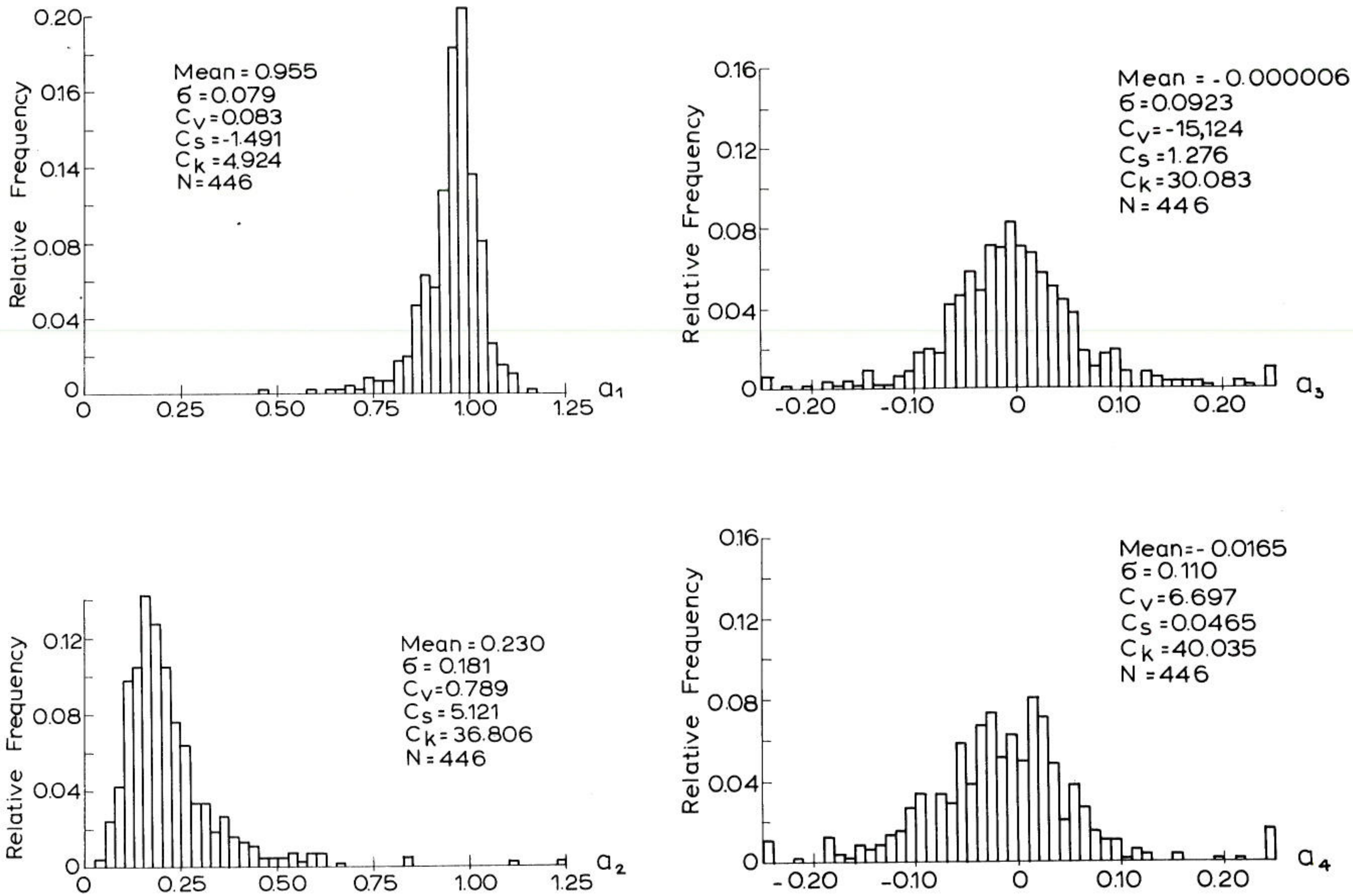

Fig. 20 Distributions of the regression coefficients a_1, a_2, a_3 and a_4 for the effective precipitation variable.

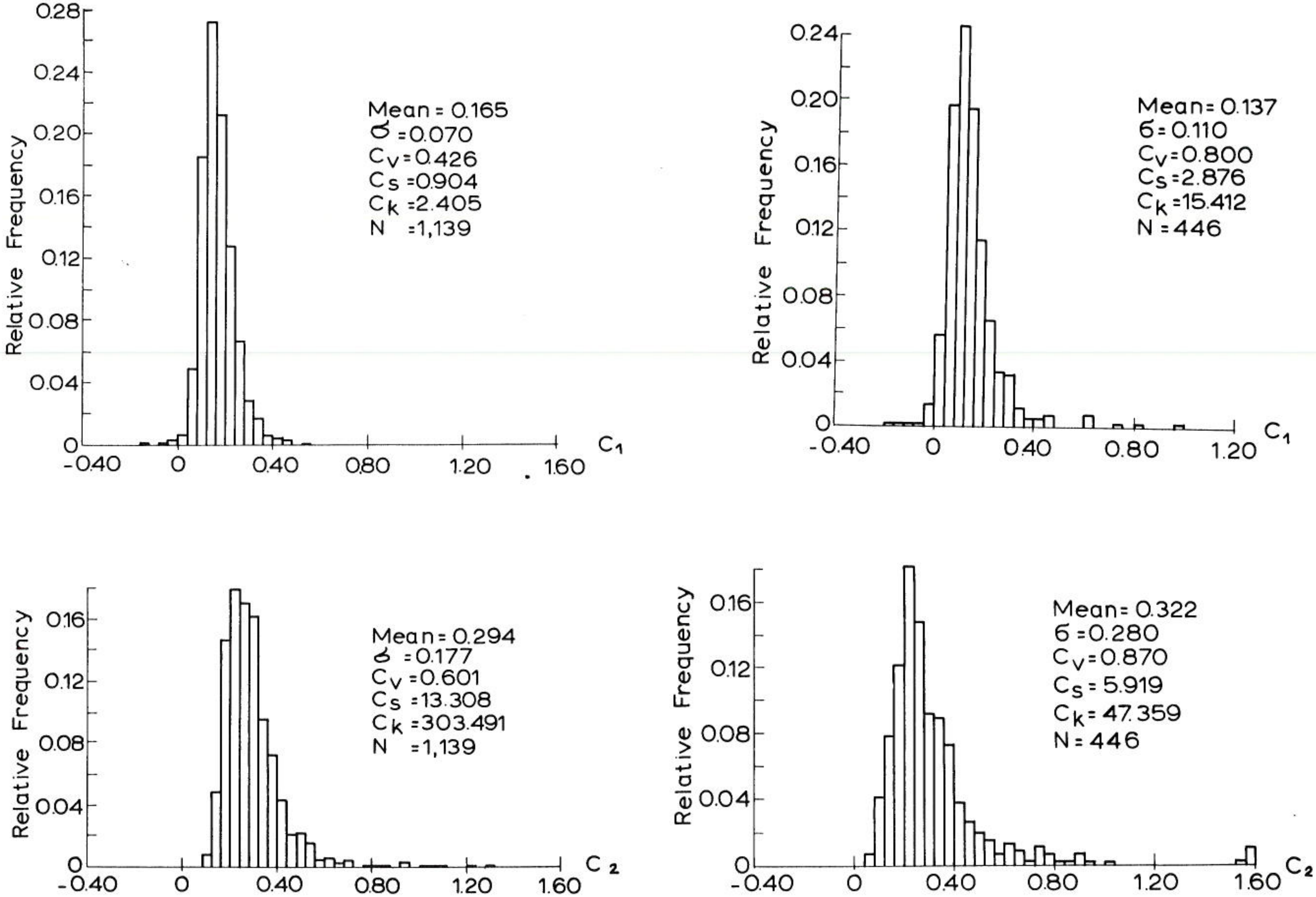

Fig. 21 Distributions of the rate of decrease parameters C_1 (major axis, top) and C_2 (minor axis, bottom) for the precipitation variable (left) and the effective precipitation variable (right).

TABLE 1

COMPARISON OF FUNCTIONS

QUANTITY	FUNCTION 4.6			FUNCTION 4.14			FUNCTION 4.15		
	1	2	Both Sets	1	2	Both Sets	1	2	Both Sets
A. PRECIPITATION									
Mean of the percentage of the total variation explained	48.41	51.58	50.00	54.08	56.41	55.25	53.12	55.60	54.36
Mean maximum value	0.763	0.770	0.766	0.912	0.902	0.907	0.875	0.859	0.867
Number of Maximum Values > 1.0	0	0	0	2	1	3	0	1	1
Greatest variation explained (%)	71.08	82.05	82.05	76.47	84.74	84.74	76.21	85.02	85.02
Least variation explained (%)	24.23	3.03	3.03	24.00	10.16	10.16	24.25	8.53	8.53
B. EFFECTIVE PRECIPITATION									
Mean of the percentage of the total variation explained	48.39	45.98	47.18	51.53	48.22	49.87	52.04	49.65	50.84
Mean maximum value	0.881	0.891	0.886	0.961	0.996	0.977	0.946	0.972	0.959
Number of maximum values > 1.0	0	0	0	3	5	8	2	2	4
Greatest variation explained (%)	74.19	82.12	82.12	73.99	75.90	75.90	75.21	79.67	79.67
Least variation explained (%)	14.13	7.85	7.85	18.78	11.69	11.69	18.02	11.03	11.03

TABLE 2. TABULATION OF THE VALUES OF a'_3 FOR THE PRECIPITATION VARIABLE

Row	Column 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-0.330	-0.004	------	------	------	------	------	-0.266	------	------	------	------	------	------	------	------	------	------	------	------
2	0.015	------	0.248	0.061	------	-0.067	------	-0.038	-0.326	-0.025	0.077	0.073	-0.016	------	------	------	------	------	------	------
3	-0.096	-0.008	------	0.086	-0.050	0.054	0.100	0.024	0.029	-0.063	-0.032	-0.070	-0.094	------	------	------	------	------	------	------
4	-0.008	-0.004	0.138	0.066	0.123	0.076	0.048	0.022	-0.089	-0.049	-0.048	0.001	-0.014	-0.034	-0.058	-0.026	------	------	------	------
5	------	------	0.115	0.072	0.040	0.010	0.018	-0.032	-0.064	-0.066	-0.096	-0.018	-0.031	-0.028	-0.030	-0.050	------	------	------	------
6	------	------	0.068	-0.010	0.017	-0.006	0.048	0.026	-0.013	-0.008	-0.046	-0.039	-0.033	-0.060	-0.046	-0.073	------	------	------	------
7	------	------	-0.016	-0.004	-0.012	0.017	-0.016	-0.005	0.022	0.061	-0.034	-0.029	0.016	0.014	-0.025	-0.024	-0.058	-0.083	-0.170	------
8	------	-0.001	0.047	0.026	------	-0.035	0.022	0.006	0.048	0.018	-0.021	0.029	0.021	0.013	0.005	-0.002	-0.043	-0.062	-0.070	------
9	------	------	0.063	0.054	0.118	------	------	0.015	0.002	0.105	0.173	0.045	0.012	0.017	-0.031	-0.033	-0.056	-0.064	-0.044	------
10	------	------	0.041	0.020	-0.020	0.068	0.068	-0.029	------	0.036	-0.023	0.012	0.021	0.031	-0.021	-0.003	-0.030	-0.051	-0.041	-0.027
11	------	------	------	0.006	-0.010	0.004	0.069	-0.007	0.001	0.001	-0.007	-0.012	0.009	------	0.002	-0.021	-0.045	-0.049	-0.010	------
12	------	------	------	------	0.044	0.040	0.036	-0.025	-0.005	-0.020	-0.007	0.018	0.027	-0.012	-0.028	-0.001	-0.028	-0.038	-0.030	------
13	------	------	------	------	------	------	------	------	0.031	-0.000	0.029	------	-0.021	0.019	-0.022	-0.000	-0.017	-0.013	-0.051	-0.034
14	------	------	------	------	------	------	------	------	------	------	------	0.035	------	-0.007	-0.016	-0.011	-0.009	-0.079	0.020	------
15	------	------	------	------	------	------	------	------	------	------	------	------	------	------	0.002	0.073	------	------	------	------

TABLE 3. TABULATION OF THE VALUES OF a'_4 FOR THE PRECIPITATION VARIABLE

Row	Column 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.279	0.086	------	------	------	------	------	-2.029	------	------	------	------	------	------	------	------	------	------	------	------
2	0.089	------	0.311	0.349	------	0.019	------	0.104	0.074	0.029	0.069	-0.018	-0.125	------	------	------	------	------	------	------
3	-0.004	-0.037	------	-0.038	0.006	0.007	0.071	0.006	-0.171	-0.037	0.009	-0.009	-0.073	------	------	------	------	------	------	------
4	0.005	-0.017	-0.050	-0.058	0.096	-0.005	-0.008	-0.025	-0.030	-0.016	-0.005	-0.033	-0.041	-0.051	-0.060	-0.095	------	------	------	------
5	------	------	-0.059	-0.042	-0.063	-0.035	-0.028	-0.019	-0.002	-0.005	0.024	0.017	-0.002	-0.023	-0.024	-0.031	------	------	------	------
6	------	------	-0.087	-0.123	-0.071	0.002	0.029	-0.049	-0.019	0.024	0.035	0.006	0.030	-0.011	-0.025	-0.007	------	------	------	------
7	------	------	-0.048	-0.039	0.007	0.053	0.025	0.010	-0.047	-0.062	0.014	0.056	0.020	0.024	-0.026	-0.031	-0.036	-0.036	-0.083	------
8	------	0.015	-0.023	-0.046	------	-0.017	-0.034	-0.013	0.001	-0.041	-0.030	0.033	0.016	0.008	-0.013	-0.033	-0.053	-0.046	-0.028	------
9	------	------	0.001	0.006	-0.069	------	------	-0.064	-0.014	-0.022	-0.001	-0.042	-0.004	-0.014	-0.004	-0.005	-0.006	-0.048	-0.090	------
10	------	------	-0.009	0.050	0.060	-0.123	0.056	-0.001	------	0.058	0.014	0.034	-0.035	-0.014	-0.012	-0.005	-0.016	-0.023	-0.074	-0.085
11	------	------	------	0.027	0.042	0.015	0.006	0.051	0.030	-0.024	-0.003	0.016	-0.013	------	-0.039	-0.016	-0.040	-0.084	-0.070	------
12	------	------	------	------	0.006	0.048	-0.045	0.009	0.003	-0.006	-0.015	-0.014	-0.020	-0.002	0.000	-0.013	-0.024	-0.034	-0.028	------
13	------	------	------	------	------	------	------	------	-0.013	-0.010	-0.053	------	-0.013	-0.027	-0.014	-0.015	-0.003	-0.003	-0.003	-0.053
14	------	------	------	------	------	------	------	------	------	------	------	-0.009	------	-0.024	-0.044	0.003	0.020	0.093	0.003	------
15	------	------	------	------	------	------	------	------	------	------	------	------	------	------	0.025	0.053	------	------	------	------

TABLE 4. TABULATION OF THE VALUES OF a'_3 FOR THE EFFECTIVE PRECIPITATION VARIABLE

Row	Column 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------
2	------	------	------	------	------	------	-0,019	------	------	------	------	------	------	------	------	------	------	------	------	------
3	------	------	-0,091	0,075	0,040	0,107	0,081	------	------	------	------	------	------	------	------	------	------	------	------	------
4	------	------	------	0,038	0,009	0,014	0,021	0,014	------	0,175	------	------	------	------	------	------	------	------	------	------
5	------	------	0,015	0,021	------	-0,016	0,021	-0,018	-0,013	------	------	-0,047	-0,088	------	------	------	------	------	------	------
6	------	------	0,033	0,049	-0,061	0,001	-0,052	-0,064	0,074	0,218	0,153	-0,110	-0,115	-0,013	-0,029	0,113	------	------	------	------
7	------	0,020	-0,017	-0,029	-0,022	------	-0,065	------	0,022	0,070	0,255	0,068	------	-0,096	0,074	0,061	------	------	------	------
8	------	------	-0,026	-0,041	------	-0,016	-0,038	------	0,010	-0,042	0,090	0,061	-0,177	-0,115	0,036	0,053	0,023	-0,068	------	------
9	------	------	0,006	0,014	0,022	------	------	0,012	0,011	0,161	0,042	-0,229	-0,004	-0,106	-0,117	0,015	-0,014	-0,041	-0,029	------
10	------	------	------	-0,031	0,004	------	------	0,294	0,124	-0,042	-0,032	-0,088	------	------	------	-0,006	-0,030	-0,017	-0,036	-0,022
11	------	------	------	-0,022	-0,049	-0,065	------	0,098	-0,185	0,014	------	0,015	------	------	0,005	------	-0,040	0,016	------	-0,028
12	------	------	------	------	------	------	------	------	-0,060	-0,134	0,003	0,089	0,132	-0,151	------	0,016	0,019	------	------	------
13	------	------	------	------	------	------	------	------	0,885	------	------	------	------	0,019	0,029	0,014	0,002	0,002	------	------
14	------	------	------	------	------	------	------	------	------	------	------	------	------	-0,067	0,012	0,164	------	------	------	------
15	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

TABLE 5. TABULATION OF THE VALUES OF a'_4 FOR THE EFFECTIVE PRECIPITATION VARIABLE

Row	Column 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------
2	------	------	------	------	------	------	0,074	------	------	------	------	------	------	------	------	------	------	------	------	------
3	------	------	0,146	0,391	-0,007	0,061	0,110	------	------	------	------	------	------	------	------	------	------	------	------	------
4	------	------	------	-0,111	0,008	-0,006	-0,000	0,034	------	0,247	------	------	------	------	------	------	------	------	------	------
5	------	------	-0,058	-0,103	------	-0,000	0,006	-0,023	-0,023	------	------	-0,012	0,103	------	------	------	------	------	------	------
6	------	------	-0,050	-0,048	-0,050	0,010	0,002	-0,033	0,016	-0,182	0,209	0,011	0,080	-0,038	0,043	0,033	------	------	------	------
7	------	0,005	-0,017	-0,046	-0,082	------	0,002	------	-0,047	0,008	-0,190	-0,019	------	-0,124	0,030	0,072	------	------	------	------
8	------	------	0,026	0,011	------	-0,005	0,011	------	-0,018	-0,067	-0,048	-0,054	0,048	-0,088	-0,006	0,010	-0,015	-0,043	------	------
9	------	------	0,029	0,055	0,033	------	------	-0,184	-0,014	0,006	-0,002	-0,177	0,042	-0,058	-0,034	-0,023	0,015	-0,026	-0,112	------
10	------	------	------	0,034	0,020	------	------	0,120	1,014	0,064	0,075	-0,061	------	------	------	-0,044	-0,010	-0,061	-0,062	0,007
11	------	------	------	0,026	0,022	0,051	------	0,362	0,128	-0,024	------	-0,092	------	------	-0,129	------	-0,155	-0,122	------	-0,024
12	------	------	------	------	------	------	------	------	-0,099	-0,124	0,029	0,025	0,052	-0,030	------	0,018	0,011	------	------	------
13	------	------	------	------	------	------	------	------	-1,106	------	------	------	------	-0,045	-0,016	-0,033	-0,020	0,036	------	------
14	------	------	------	------	------	------	------	------	------	------	------	------	------	0,012	-0,121	-0,077	------	------	------	------
15	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

CHAPTER VI

DISCUSSION AND SUMMATION OF RESULTS

1. Variation Due to Regression. The selected function which was used to explain the variation of inter-station correlation coefficients within the area about each block control station has the form (Eq. 4.15)

$$r_e = a_1 e^{-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d}$$

where θ and d are expressions of direction and distance, respectively. This function was utilized for both study variables, annual precipitation (P) and annual effective precipitation (EP). The average value of the variation explained in 1139 fittings of the function for precipitation was 57.40 percent. The average value of the variation explained in 446 fittings of the function for effective precipitation was 59.83 percent.

The fit seems to be better around the edge of the region. This is caused by the lack of points to one side of the block control station. This increase, however, is not of a magnitude to overshadow the definite trend to better fittings east of the Rocky Mountains.

The regional distributions of the percentages of variation explained by regression analysis (E_{var}) are presented in fig. 22 for precipitation and in fig. 23 for effective precipitation. The percentage of the variation explained in inter-station correlation coefficients appears to be more randomly distributed over the study region for effective precipitation than for precipitation. A continuous expanse of the mountainous western United States from central California across Idaho and southward along the Rocky Mountains has less than 50 percent of the variation in the coefficients explained by the fitted function. The areas for effective precipitation where the variation explained is less than 50 percent also lie in very mountainous terrain but fall in separated areas. The selected function provides a reasonable explanation of the regional pattern in inter-station correlation coefficients except for the mountainous areas outlined above.

The frequency distributions of the percentages of the variation explained by each fitted surface are shown in fig. 24. The distribution for effective precipitation has a larger standard deviation. Greater proportions of the percentages for effective precipitation are less than 40 percent and greater than 80 percent than those for precipitation. Apparently the several factors influencing inter-station correlations of effective precipitation have greatly diversity of effects over the region. The factors operating to reduce correlation coefficients are varying continuously and excentricities favor reductions in correlation rather than increases. Significant changes in evapotranspiration and river basin characteristics affecting effective precipitation can be greatly localized. Thus, the tendency for more smaller percentages of explained variations in effective precipitation may be explainable by localized extreme deviations from the fitted surfaces.

The ensemble frequency distributions for the deviations from the fitted surfaces for both variables are given in fig. 24. Approximately 75 percent of the deviations for P and EP fall within one standard deviation of the mean, but the distribution of the deviations for EP has a higher negative skew. The greater proportion of coefficients falling much below the fitted surface supports the likelihood of more excentricities in the inter-station correlation coefficients of effective precipitation and supports the point concerning the smaller percentages of the variation explained.

2. Orientation of the Inter-station Correlation Surface and Its Elongation. The orientation and configuration of the three-dimensional surfaces fitted to values of inter-station correlation coefficients are best presented by drawing isolines of equal coefficients as shown in fig. 18. However, this could not be done for every station. Instead, the axis of maximal correlation was determined, and the ellipticity parameter ($R_c = C_1/C_2$) was used to indicate the approximate ellipticity of the isolines. The orientation of the axis of maximal correlation and the approximate ellipticity of the isolines are shown in summarized form in figs. 25 and 26. The area averaging of the parameters (summarized form) is presented in Chapter V, Section 3.

A value of R_c approaching 1.0 indicates that the factors affecting the regional variation of the inter-station correlation coefficients are nearly isotropic in character. As R_c approaches zero, the isolines close and elongate; that is, the difference in the rates of decrease of the correlation coefficients with distance along the major and minor axes become greater.

Precipitation. -- The orientation of the major axes vary over the study region, but there are some patterns (Fig. 25). The orientations of the major axes do not coincide generally with the flux of moisture. On the Pacific Coast where the flux of moisture is from west to east off the Pacific Ocean, the general orientation is orthogonal to the flux and parallel to the alignment of the topographic barriers. The over-riding effects of orographic uplift seems to rotate the expected major axes 90^0. Where the mountain chains are broken and over high unbroken plateaus, there is a tendency for the axes to align with the west to east upper wind pattern.

Topographic barriers tend to control the orientations of the major axes inland until the eastern slopes of the Rocky Mountains are reached. Then, over the plains east of the Rocky Mountains the major axes assume east or northeast orientations. Three factors likely contribute to the east or northeast orientations in the southeastern part of the area: the general west to east wind pattern at higher levels, the general southwest to northeast orientation of fronts, and the northern to northeastern flux of moisture from the Gulf of Mexico. The fronts tend

Fig. 22 Regional distribution of the percentage of the variation in the inter-station correlation coefficients explained by regression for the annual precipitation variable.

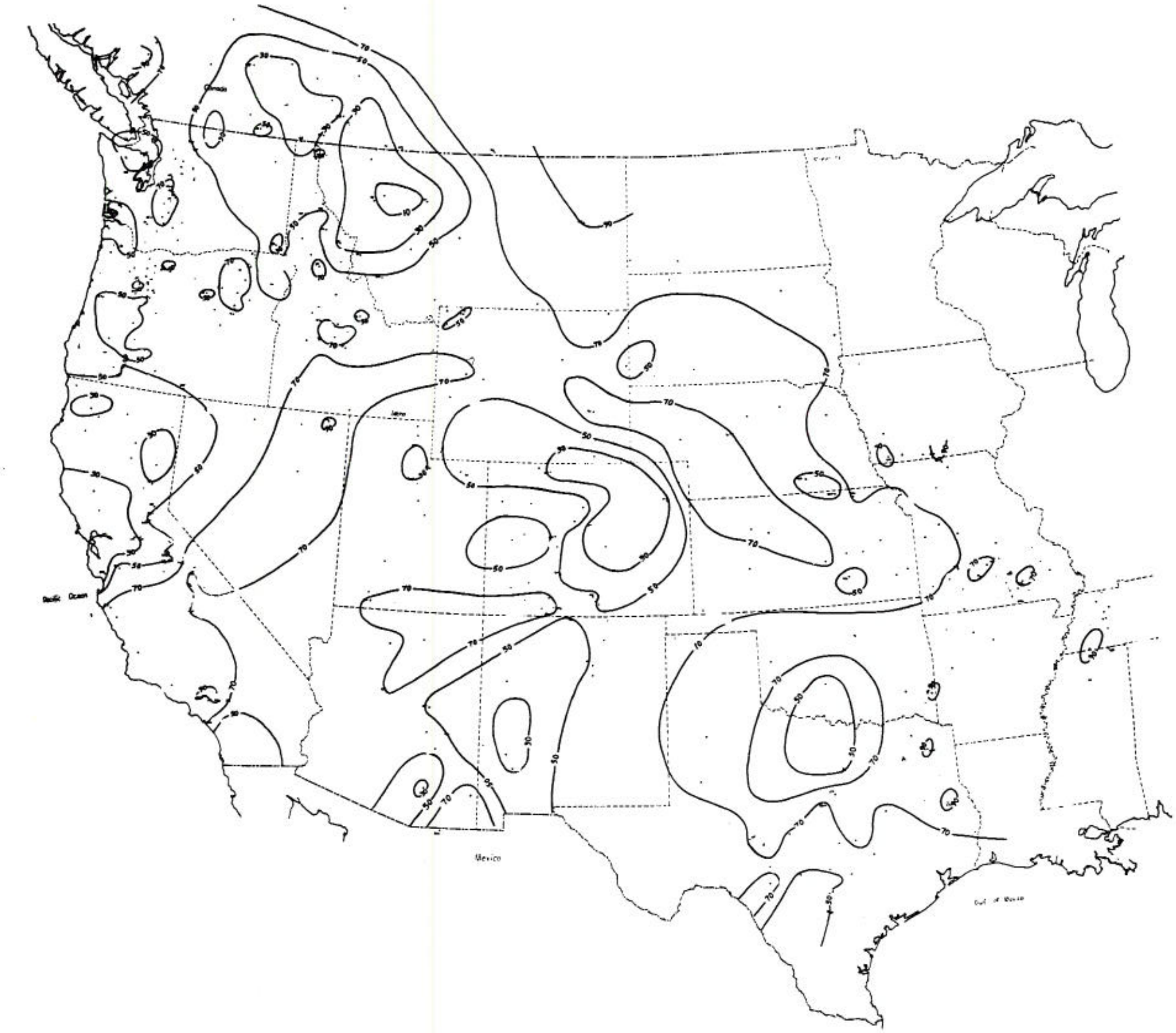

Fig. 23 Regional distribution of the percentage of the variation in the inter-station correlation coefficients explained by regression for the annual effective precipitation variable.

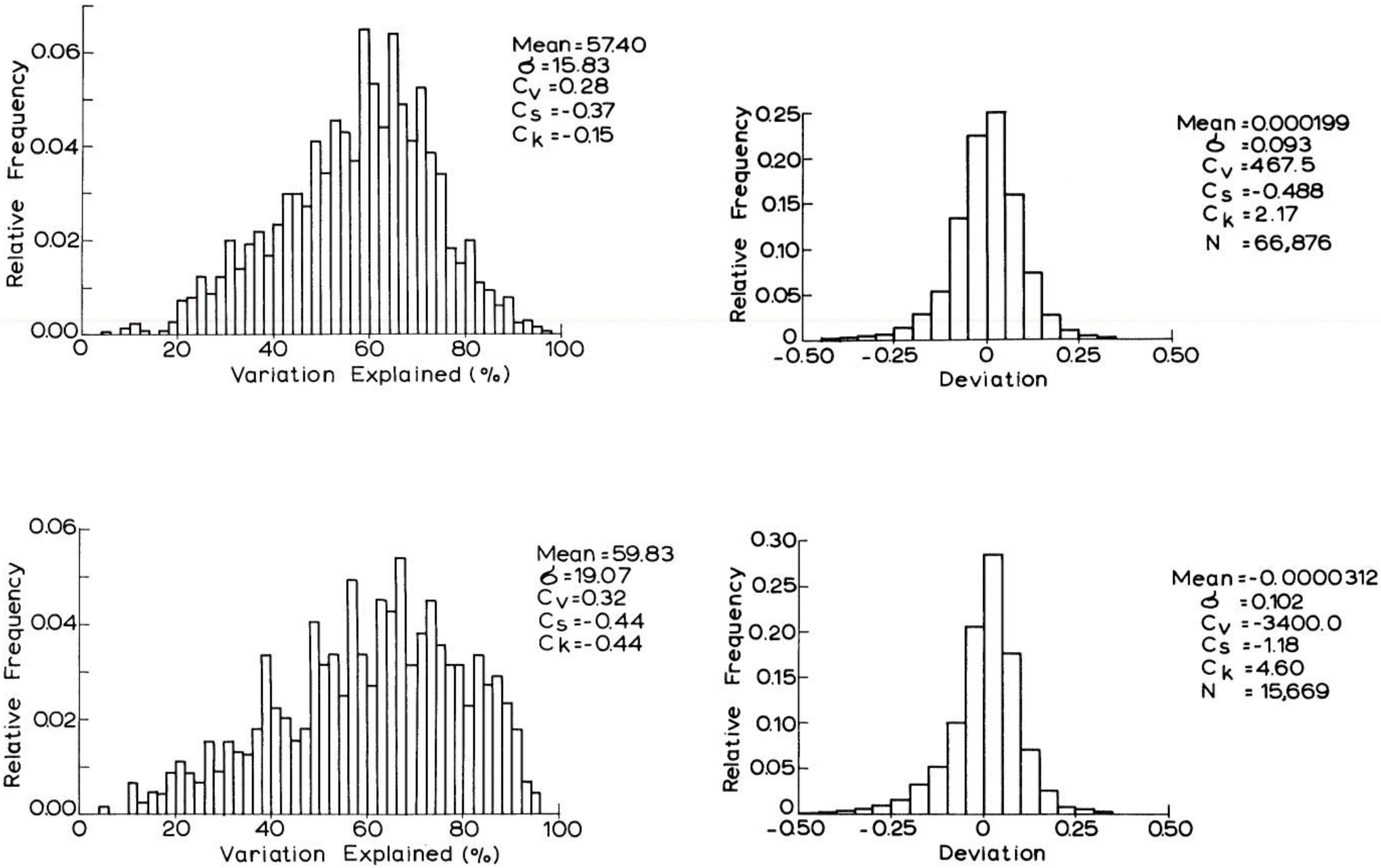

Fig. 24 Distributions of the percent of the variation due to regression for precipitation (Top, left) and effective precipitation (Bottom, left) and the distributions of the deviations of the observed inter-station correlation coefficients from the fitted surfaces for precipitation (Top, right) and effective precipitation (Bottom, right).

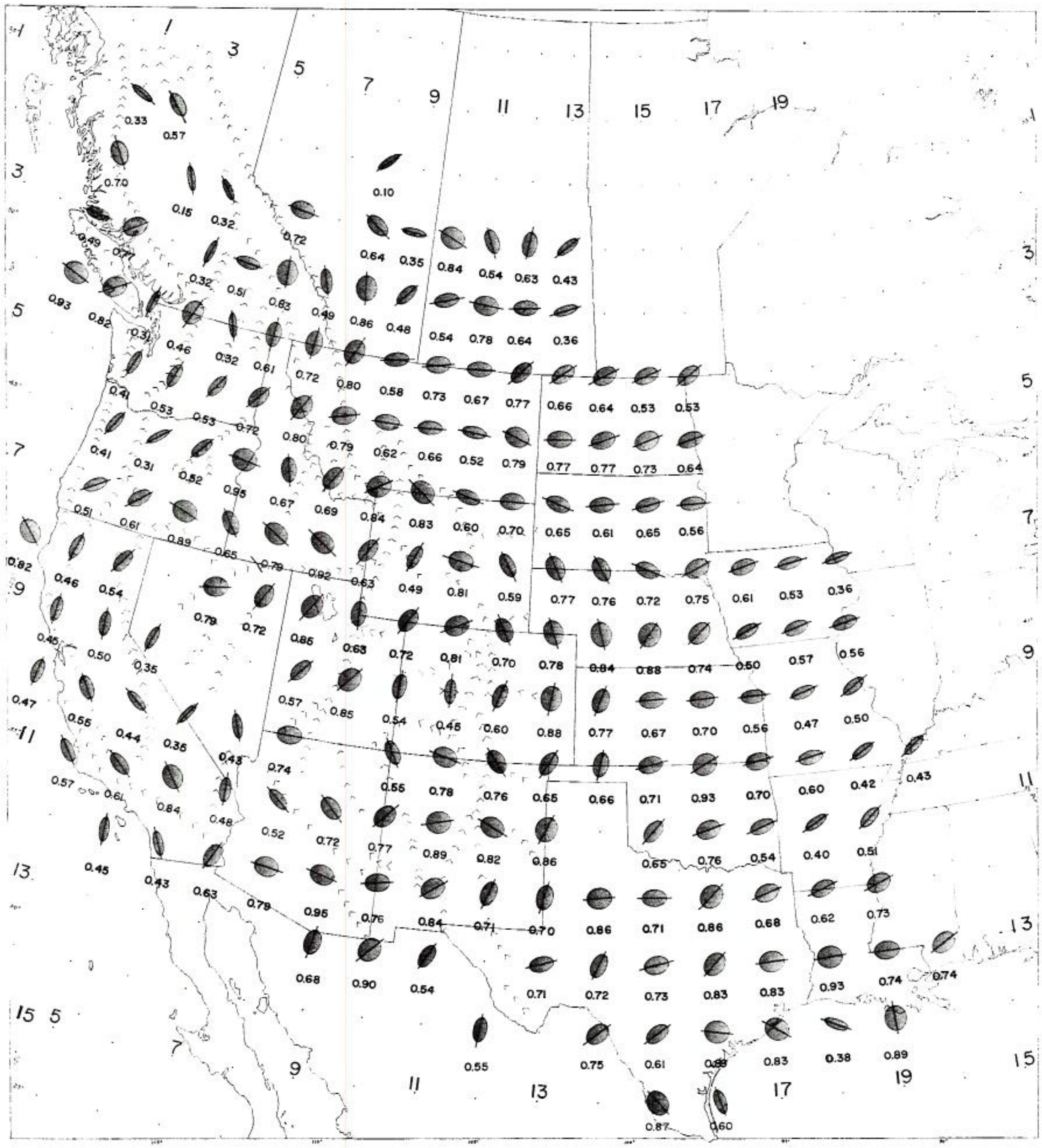

Fig. 25 Regional distribution of the azimuths of the axes of the maximal correlation coefficients and the values of the ratio (C'_1 / C'_2) for the annual precipitation variable.

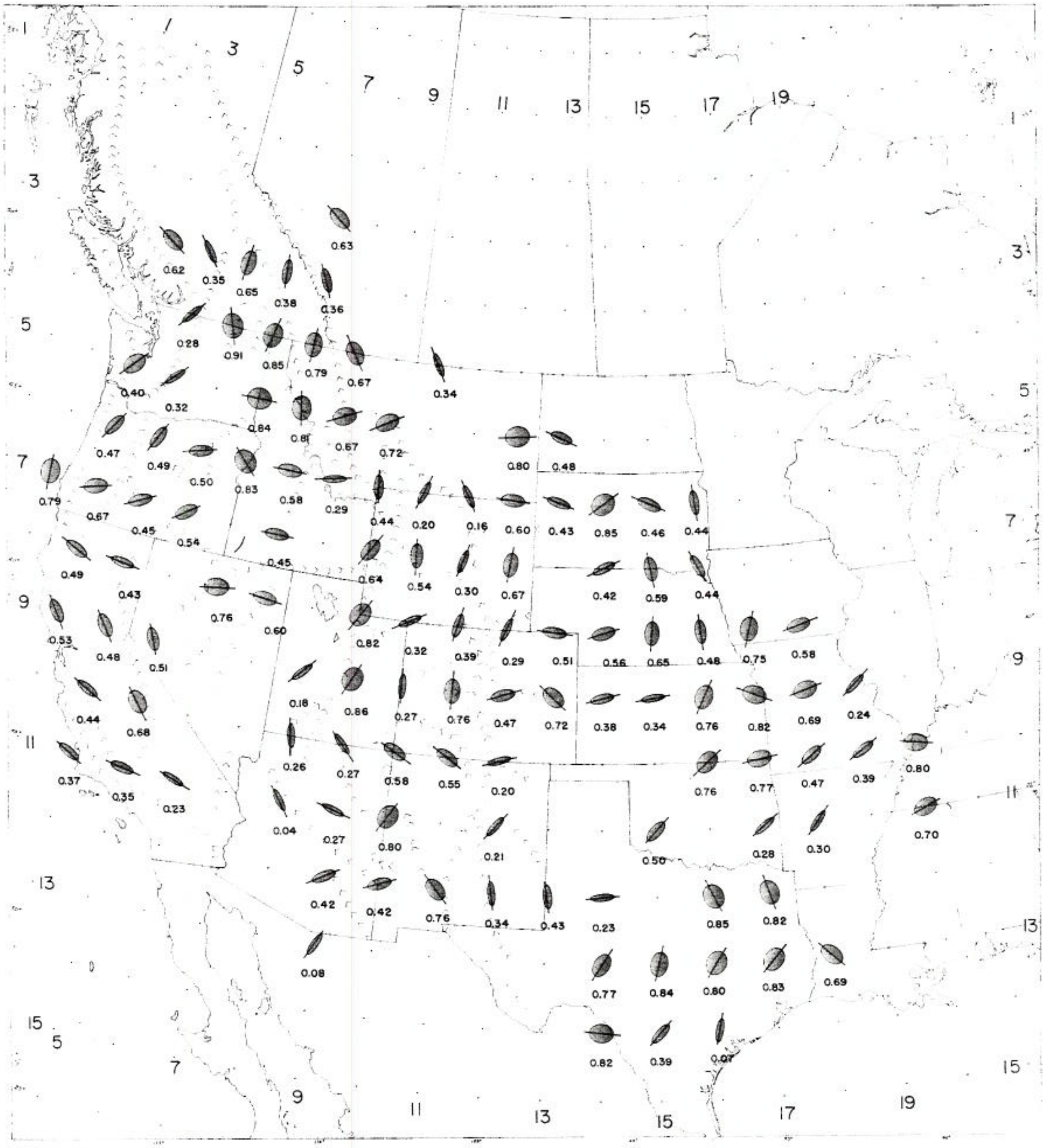

Fig. 26 Regional distributions of the azimuths of the axes of the maximal correlation coefficients and the values of the ratio (C_1'/C_2') for the annual effective precipitation variable.

to act as orographic barriers which move from north to south lifting the Gulf moisture. The upper circulation carriers convective cells eastward along the front, adding to them any Pacific moisture spilling over the Rocky Mountains.

From South Dakota northward, except near the edge of the investigated area, there is a consistent west to east orientation. This area lies along the favored tracks of low cells which move from west to east. In this region, the dominant source of moisture is the North Pacific, and the moisture is diffused in passing over the mountains.

The degree of ellipticity portrayed by the shapes of the ellipses drawn at each summary point shows some irregularity over the study region. The greatest elongation of the ellipses appears to be associated with areas of greatest orographic influence, particularly along the Pacific Coast. However, the ratio, R_c, becomes smaller again as the distance east of the Rocky Mountains increases. This may be caused by increased frontal activity and the availability of undiffused moisture from the Gulf of Mexico.

Effective Precipitation. -- The orientations of the major axes of the annual effective precipitation inter-station correlation coefficients are much more diverse than are those of annual precipitation. Where there are pronounced orographic effects upon precipitation, one finds that effective precipitation axes of maximal correlation tend to parallel those of precipitation. However, over much of the map there is considerable dissimilarity between results for the two variables. Much of this dissimilarity may be explainable by the distribution of the effective precipitation station locations which is quite sparse and irregular, especially in the northeast and southwest sections of the study region.

However, a survey of the total map indicates a more general orientation of the major axes perpendicular to the slope of the topography and consequently perpendicular to the channels of the stream network. This indicates that the meteorologic factors are likely masked by the effects of evapotranspiration and river basin characteristics. The greatest rates of decrease of the inter-station correlation coefficients are parallel to the slope of the land. Both evapotranspiration and river basin characteristics vary most in the same direction. The result is that evapotranspiration effects are more closely related across the slope than down it. As pointed out in Chapter III, Section 3, dissimilar smoothing effects contribute to lower inter-station correlations. In areas where the surface slopes have multiple orientations and exposures, the contrast in the rates of decrease along the major and minor axes are diminished. This does not necessarily mean that the coefficients are of larger magnitude but only that there is less absolute difference between the highest and lowest coefficients in the surface control block.

3. Significance of the Regression Coefficient, a_1. Theoretically the inter-station correlation coefficient at the block control station is 1.0. However, with a_1 fixed at 1.0, the explained variation was very low. Allowing a_1 to vary gave an additional degree of freedom in the surface fitting. The explained variation increased as a result. The regional distributions of a_1' are shown in figs. 27 and 28 for precipitation and effective precipitation, respectively. The values of a_1 represent the maximum value for each fitted surface and is a multiplying factor before the exponential term (Eq. 4.15). The mean values of a_1 are 0.875 for precipitation and 0.955 for effective precipitation.

The larger the values of a_1 the greater the initial inter-station correlation coefficients. The rate of decrease also depends on the magnitudes of the coefficients of a_2, a_3, and a_4 (Eq. 4.15). The frequency distribution of a_1 are given in figs. 19 and 20 for the P and EP variables, respectively. The standard deviations of the two distributions are essentially equal, but the EP distribution of a_1 has a much greater number of values larger than 1.0. For the EP variable, 25.3 percent of the values of a_1 exceeded 1.00 but only 5.6 percent exceeded 1.05. For the precipitation variables, 4.1 percent of the values of a_1 exceeded 1.00 and only 1.5 percent exceeded 1.025. The mean value of a_1 for effective precipitation, 0.955, is not an unreasonable value.

The negative exponential function is a continuously decreasing function of constantly decreasing slope. It is, therefore, incapable of changes of slope to accommodate reduced rates of change of slope at short distances and larger rates of decrease at greater distances. Thus, for effective precipitation when correlations are very high at short distances, the computed intercept value, a_1, often is greater than 1.0. Precipitation coefficients seem to drop off rapidly at the shorter distances and then persist at the greater distances. This would force the intercept to fall below 1.0. The higher values of inter-station correlation coefficients at the shorter distances come about through the smoothing effects of the evapotranspiration process which may give way to changes in topographic factors at intermediate distances (75 miles or so).

The rate of decrease in the inter-station correlation coefficients for precipitation is so great at very short distances as to almost appear as a discontinuity though from physical considerations none is expected. Marwitz [1965] points out in his discussion of the accuracy and representativeness of standard rain gages that they exhibit catch deficiencies that increase with concurrent wind speed and are greater for snow than for rain. Shielding of gage sites vary the concurrent wind speeds for different proximities of shielding and wind directions. McGuiness [1963] concluded that the average error of determining the mean rainfall on small watersheds was related to the amount of rainfall and to the density of the rain gage network. This points to the integrating and smoothing factors at work to produce similarity in effective precipitation measurements at different points, and to the unreliability of the point rainfall measurements to represent the experience of the area about it or at other points.

These and other factors tend to support the idea of a discontinuity in the inter-station correlation coefficients of precipitation around the BCS. Since precipitation measurements are not integrated measurements in the sense of effective precipitation, but are point measurements subject to error and greater variability, this inconsistency may be assigned to sampling techniques and measurement errors. This could possibly produce a built in reduction factor for the inter-station correlation coefficients.

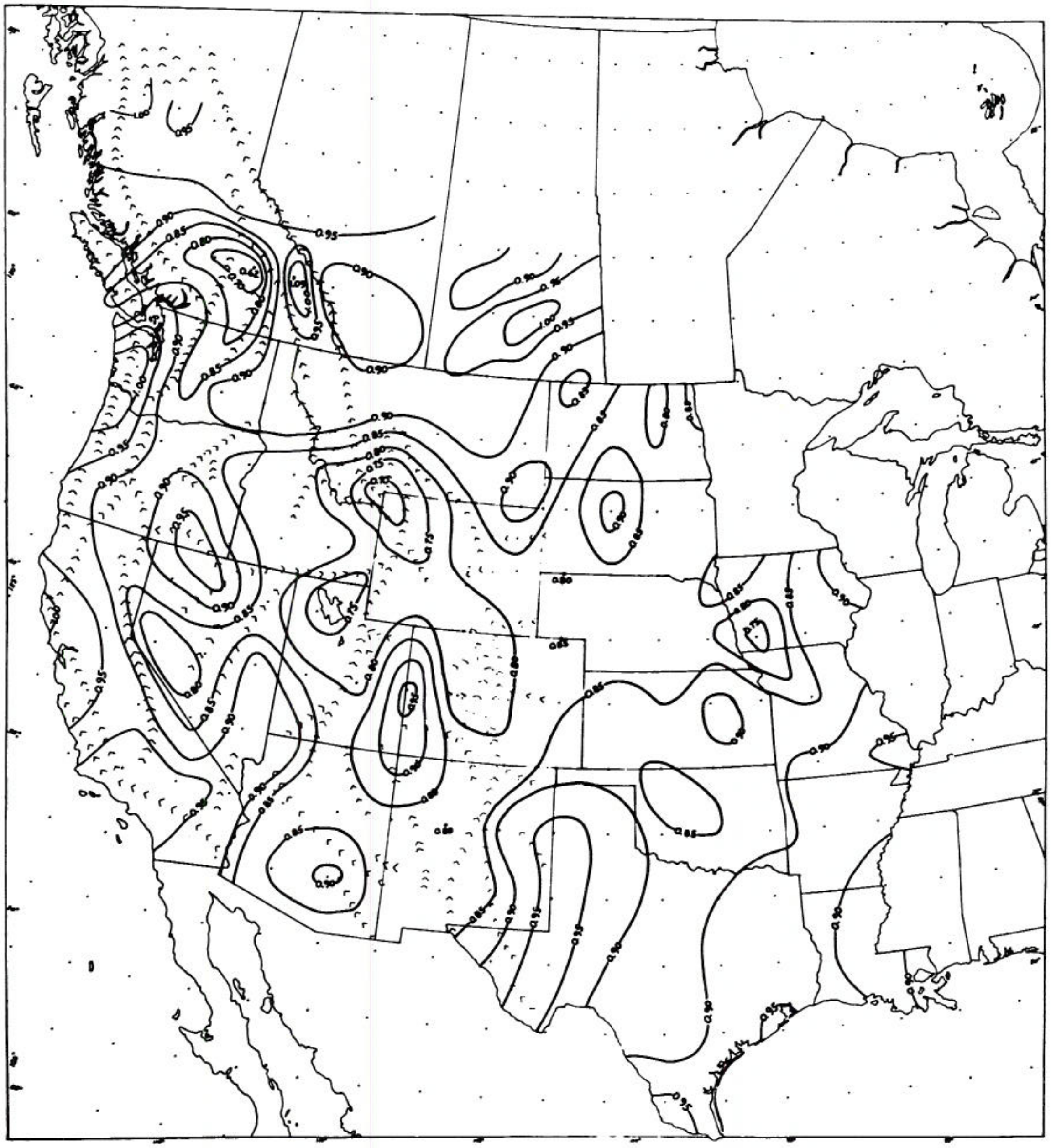

Fig. 27 Regional distribution of the regression coefficient a_1' for the annual precipitation variable.

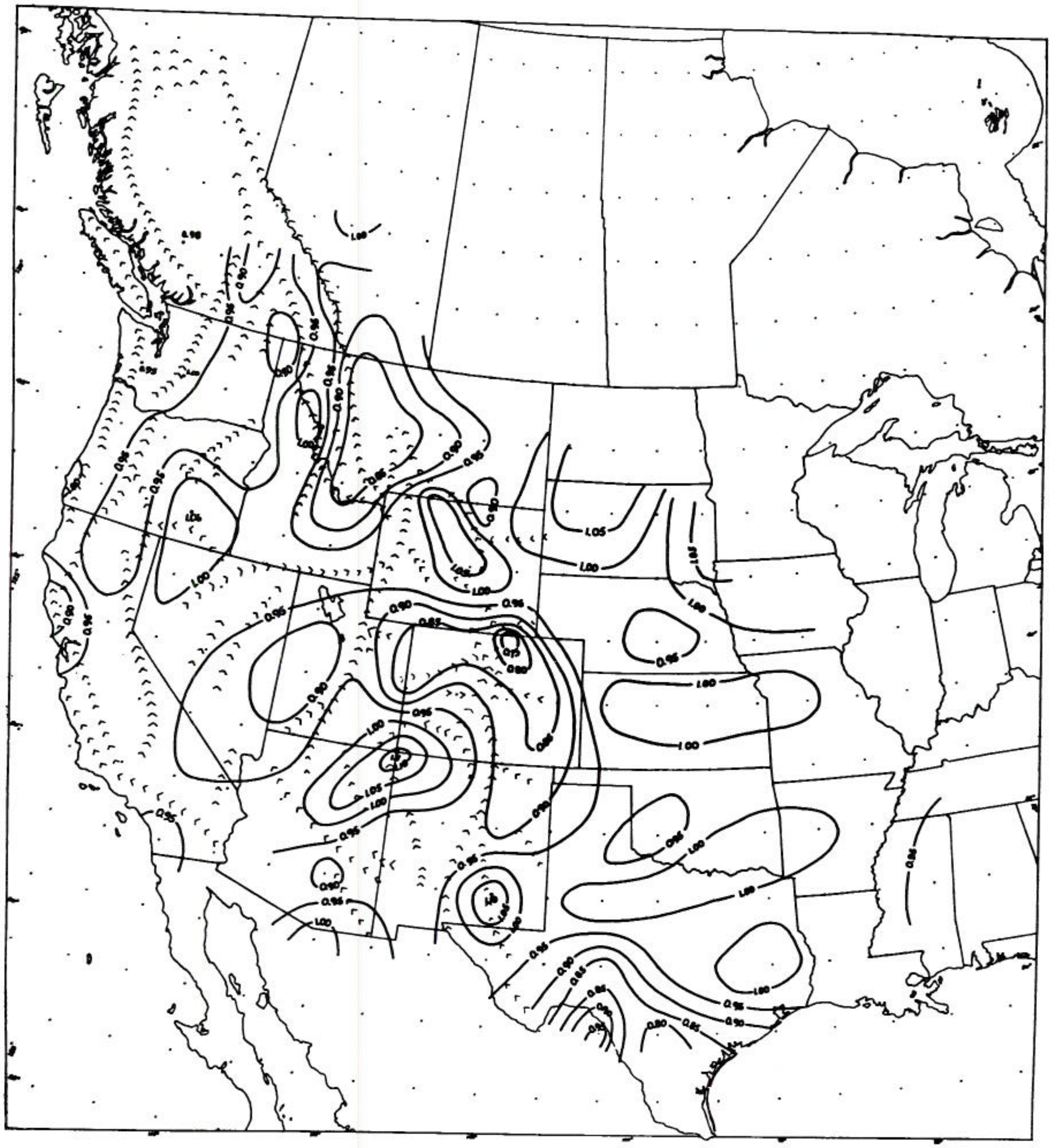

Fig. 28 Regional distribution of the regression coefficient a_1' for the annual effective precipitation variable.

4. Significance of the Regression Coefficients a_2, a_3, and a_4. The regression coefficient a_2 is the basic coefficient that indicates the rate of decrease of correlation with distance. The coefficients a_3 and a_4 coupled with the sine and cosine terms alter the basic rate, a_2, according to the influence of direction.

If a_2 for precipitation is equal to or slightly less than a_2 for effective precipitation, then one may conclude that generally the regional inter-station correlation coefficients of effective precipitation are higher and more persistant with distance than the coefficients for precipitation.

The regional distributions of the coefficient a_2' are shown in figs. 29 and 30 for precipitation and effective precipitation, respectively. The regression coefficients are much more uniformly distributed for precipitation than for effective precipitation. The values of a_2' for the P variable are smallest in the southern and southeastern portions of the area generally increasing northward and northwestward. This contrasts sharply to the patterns in a_2' for the EP variable. The lowest values are found on the western third of the area with a zone of very large and highly variable values running from Arizona northward into the northeast quarter of the study region. This zone encompasses some very mountainous terrain, and the area also has a very sparse and irregular distribution of runoff station locations.

The regional distributions seem to indicate a large area for EP that has coefficients of a_2 higher than those for P. This is opposite to the conclusion one would draw from the comparison of the frequency distributions in figs. 19 and 20. The distributions are the ensembles of all computed coefficients a_2. The disparity in station densities accounts for the seeming contradiction. If interpolations between points on the map are reasonable, then one would conclude that a_2 is generally larger for effective precipitation than for precipitation over the eastern two-thirds of the study region.

The coefficients a_3 and a_4 not only alter the basic rate of decrease as shown by a_2 but also determine the orientation of the axis of maximal correlation. The angle θ_{max} is given by eq. 4.19, and eqs. 4.21 and 4.22 evaluated for θ_{max} give the rate of decrease parameters, C_1 and C_2.

5. Significance of the Parameters, C_1 and C_2. The parameters C_1 and C_2 determine the rate of decrease of inter-station correlation coefficients with distance along the major and minor axes, respectively. The larger the values of C_1 and C_2 the greater the rate of decrease of the inter-station correlation coefficients. Hence, the variability of the coefficients within areas of higher parameters is also larger. Greater differences between C_1 and C_2 also indicate a larger degree of variability in the regional correlation coefficients. The degree of ellipticity parameter ($R_c = C_1/C_2$) indicates the contrast in the rates of decrease with distance along the major and minor axes. The smaller the value R_c the greater the contrast in the rates of decrease along the orthogonal axes.

The regional distributions of C_1' and C_2' are shown on figs. 31 and 32 for precipitation and effective precipitation, respectively. There is considerably more variability in both these parameters for effective precipitation than for precipitation. The greatest variability occurs in a narrow zone extending from Arizona across the Rocky Mountains into the northeast section of the study region. This narrow zone contains several isolated extremely high values of C_1' and C_2'. The parameters are more uniformly distributed for precipitation with an area of high values occurring in the extreme northwest section of the study region. There is little coincidence of high values observed for the two variables, and a tendency for low values for precipitation to occur with high values for effective precipitation and vice versa. River basin characteristics and evapotranspiration processes apparently contribute greatly to the increased variability in the parameter C_1' and C_2' for effective precipitation.

The frequency distributions of C_1 and C_2 are given in fig. 21 for precipitation and effective precipitation, respectively. The respective mean values are 0.165 and 0.137 for C_1 and 0.294 and 0.322 for C_2. The mean values indicate that generally the correlations along the maximal axes persist more for runoff than for precipitation, but that the opposite is true concerning the minimal axes. The mean value of C_2 for effective precipitation is somewhat increased by a few exceptionally high values of C_2, but a comparison of figs. 25 and 26 confirms that there is a much greater contrast between the rates of decrease for the major and minor axes of effective precipitation than for precipitation.

6. Comparison of Fitted Surfaces. A comparison of the fitted surfaces at different points cannot readily be made from the values of the coefficients and parameters alone. The rates of decrease of inter-station correlation coefficients at a given distance from the block control station along the major and minor axes are given, respectively, by the following equations:

$$\delta_1 = \frac{\partial r_e}{\partial d} = -a_1 C_1 e^{-C_1 d} \qquad 6.1$$

$$\delta_2 = \frac{\partial r_e}{\partial d} = -a_1 C_2 e^{-C_2 d} . \qquad 6.2$$

The product of three terms is involved in each of the above equations making direct comparison difficult without evaluating the equations. The coefficient a_1 is the value of the zero intercept and is a multiplying factor that may alter the effects of the rate of decrease parameters C_1 and C_2.

However, comparisons of any two fitted surfaces may be made by the use of the coefficient a_1, the azimuth, and the following two computed quantities for each surface:

$$r_1 = a_1 e^{-2.5\, C_1} \qquad 6.3$$

$$r_2 = a_1 e^{-2.5\, C_2} \qquad 6.4$$

where r_1 and r_2 are the values of the inter-station correlation coefficients at a distance 2.5 units (150 nautical miles) from the BCS along the major and minor axes, respectively. As an example, the values of ϕ', a_1', C_1' and C_2' were taken from figs. 25, 27, and 31 for two locations and the results are presented in Table 6. Rows 5 and 6 indicate that the inter-station correlation coefficients of the surface fitted at 123°W are higher than those at 103°W. The maximum value (at the BCS) and the values r_1' and r_2' were all higher at 123°W than at 103°W. Values of

TABLE 6

Detailed Comparison of Surface Fits

Row	Quantity	POINT LOCATIONS	
		123° W, 41° N	103° W, 41° N
1	ϕ'	13.13°	160.96°
2	a_1'	0.91	0.85
3	C_1'	0.09	0.19
4	C_2'	0.20	0.24
5	r_1' (Eq. 6.1)	0.73	0.53
6	r_2' (Eq. 6.2)	0.55	0.47
7	$(a_1' - r_1')$	0.18	0.32
8	$(a_2' - r_2')$	0.36	0.38

C_1' and C_2' in rows 3 and 4 indicate greater rates of decrease at 103°W than at 123°W. The absolute changes along the major and minor axes are given in rows 7 and 8. If C_2' for the point at 123°W had been only slightly larger, than a portion of the surface at 123° would have had values lower than a portion of the surface at 103°W.

Generally, the inter-station correlation coefficients for effective precipitation are higher near the BCS than those of precipitation and about 40 percent of the surfaces are everywhere greater than the corresponding fitted surface for precipitation. Approximately 55 percent of the fitted surfaces for effective precipitation have minimum values greater than the corresponding surfaces for precipitation. An even greater percentage of r_1 values are higher for effective precipitation pointing out again the greater ellipticity of the fitted surfaces for effective precipitation.

7. Wave Effect in Precipitation Occurrences. Precipitation events often are stated to occur in the manner of waves along the storm path. The theory is that precipitation will occur at a point with a given intensity and then diminish in intensity as the storm moves along only to build again to a somewhat lower intensity. The decaying and rebuilding of the precipitation producing storms could occur more than once. If this occurs with constant wave lengths for similar storms, stations in the storm path should produce inter-station correlation coefficients having a damped cyclic oscillation similar to that of a damped cosine function. There does not appear to be any evidence supporting the hypothesis of constant waves in this study.

The distributions of the precipitation correlation coefficients (Figs. 14 and 15) are all very similar in form except for those of gross subdivisions 2, 4, and 35. Those of the gross subdivisions 2 and 4 have significantly greater portions of the higher values and the gross subdivision 35 has a greater proportion of the lower values compared to the other distributions. Examination of station locations in gross subdivisions 2 and 4 show high densities of stations in one section of the subdivision and larger sections of very low densities. The density of the stations is least for the gross subdivision 35, but reasonably well distributed. Thus, an absence of smaller inter-station distances (higher correlation values) and a relatively greater proportion of longer inter-station distances result. The distribution of coefficients for the gross subdivision 35 has a skewness opposite to those of the gross subdivisions 2 and 4.

The stations of the line study matrix subsample (Fig. 15) were selected along four parallels of latitude and six meridians with no stations within the areas between the lines. The 472-station sample with inconsistent precipitation records (P2 Matrix on Fig. 15) is more uniformly distributed. Both these samples (Chapt. 5, Sect. 2) are from the total study region except the portion in Canada. Inter-station distances for the line study sample should be generally greater than those of the 472-station sample. Hence, one would expect the line study distributions to exhibit lower inter-station correlation coefficients. Figure 15 shows this to be the case.

The frequency distributions from the matrices for effective precipitation exhibit essentially the same characteristics as those of precipitation (Fig. 17). The frequency distribution of subdivisions 1 and 2 graphically portray the effects of higher station densities when compared to the other distributions from areas of lower station densities. The distribution for the whole region shows the influence of the great preponderance of very large inter-station distances.

The number of negative correlation coefficients observed in both the line studies and the matrix computations is small in proportion to the total number of coefficients, and the line studies show that the negative coefficients occur only at extreme distances. If the damped cosine oscillation were approximated, a greater number of negative values with magnitudes significant at the 95 percent level would be expected. All the frequency distributions show very low percentages of negative correlation coefficients significantly different from zero at the 95 percent level.

The frequency distributions and their relation to inter-station distances and the lack of a high percentage of negative correlation coefficients do not admit the hypothesis of a damped oscillatory pattern. These points do, however, support the theory that the variation of inter-station correlation coefficients with distance is continuously decreasing asymptotically to zero.

8. Miscellaneous Comments. Gatewood and others [1964] describe a use of inter-station correlation coefficients to determine the extent of regional homogeneity in runoff. They were testing to see if the entire southwest area of the study region was homogeneous. A correlation coefficient of $r = 0.7$ was adopted by them as the minimum required to indicate homogeneity between two stations. The whole of the southwest was found not to be homogeneous, but subareas were delineated that were considered homogeneous.

This study certainly shows that areas about a given point of interest, based on the above criteria, are not homogeneous for extensive distances and not uniformly about the point. Using the mean values of a_1 and C_1, values greater than $r = 0.7$ can only be expected up to distances of approximately 41 miles and 67 miles for precipitation and effective precipitation, respectively. The distances quoted are along the axes of maximal correlations and transverse distances are usually much less.

One of the basic problems in defining a homogeneous area is to preserve only the essential statistical features of the area. The approach often used is to classify a homogeneous region as one in which the probability of occurrence of an event (in time) is constant for all stations within the region [Thom, 1940; Alexander, 1963]. In hydrologic variables, a spatial trend is generally recognized. Hence, there is a need to establish boundaries for homogeneous regions for a given hydrologic variable. Alexander [1963] has pointed out that "the only feasible way to define such boundaries is to assume a (cumulative) probability of the event occurring within the region." In order to prevent the overlapping of homogeneous regions, the probability for division of the regions will have to be set at 50 percent [Alexander, 1963].

Gatewood and others [1964] based their division upon the similarity of variation of the runoff variable as measured by the linear inter-station correlation coefficients. Correlation coefficients of $r = 0.7$ may be found for each of two stations in common with a station between them. However, the same two stations may not have a coefficient of that magnitude when they are correlated with each other. All stations in a region do not correlate with each other at or above $r = 0.7$. Furthermore, homogeneous areas based upon $r = 0.7$ are not mutually exclusive and depend upon the choice of stations to be correlated. It is further questionable that a region which does not exhibit isotropic characteristics (has elliptical isolines of inter-station correlation coefficients) should be considered homogeneous.

Double-mass plots and other statistical tests are often used to test the consistency of observed data. In Chapter II, Section 2, some comments were made about inaccuracies and inconsistencies in research data. It appears that the described surface fitting technique may point out inconsistent station records. Two approaches could be utilized. The regression equations could be fitted as in this study. Any great decrease in the percentage of variation explained by regression analysis from the regional mean might indicate inconsistent data, and could warrant further detailed study if there is no special physical reason for the discrepancy. A second approach could be to scrutinize the deviations from the fitted regression surface. Any excessive deviations might point to inconsistent data at that station.

Foster [1944] described what he considered to be a climatic discontinuity in the vicinity of Omaha, Nebraska, based on finding of an east-west axis of maximal inter-station correlation coefficients for precipitation. This study has shown that maximal east-west axes are found at other stations in all directions from Omaha (Fig. 25). If there had been such a "climatic discontinuity," it would seem that such a consistency in the patterns of inter-station correlation coefficients would not exist over such a broad area. Apparently, other more general factors are operating to create maximal and minimal axes of correlation.

Correlation coefficients computed for this study are based on the total record of each station. Therefore, each coefficient is computed with a maximum number of pairs. None are computed with less than thirty pairs. The use of different lengths should have no serious effect on the comparability of the coefficients [Foster, 1944; McDonald and Green, 1960].

On the average, approximately 40 percent of the variation remains unexplained. A certain amount of this can be attributed to random fluctuations and inconsistent data. The higher amounts of unexplained variation are found generally in the mountainous regions. Many studies have indicated definite improvements in the results of regression analysis from the inclusion of an elevation variable. Further study might be done incorporating into the regression model a variable based upon elevation.

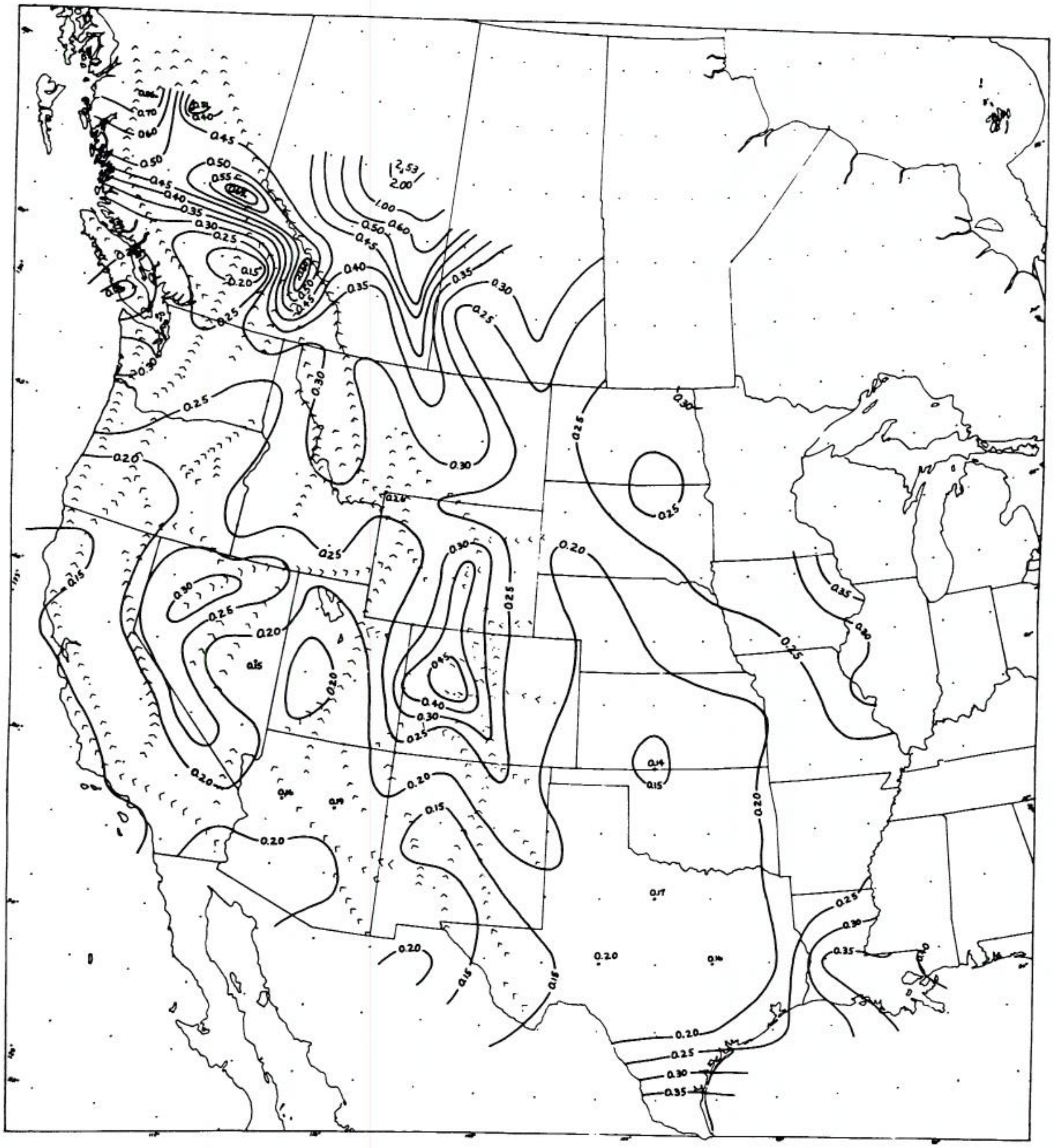

Fig. 29 Regional distribution of the regression coefficient a_2' for the annual precipitation variable.

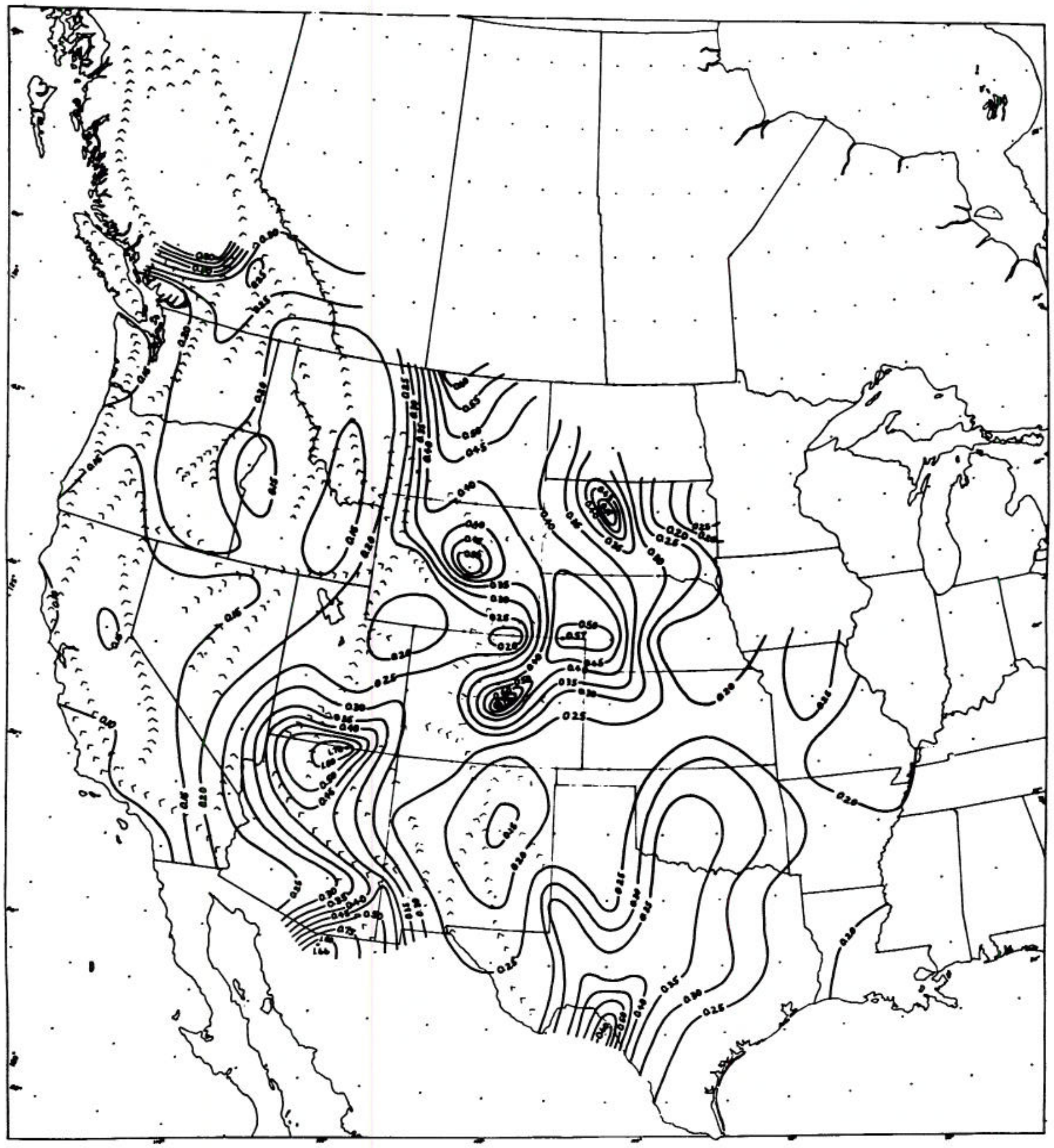

Fig. 30 Regional distribution of the regression coefficient a_2' for the annual effective precipitation variable.

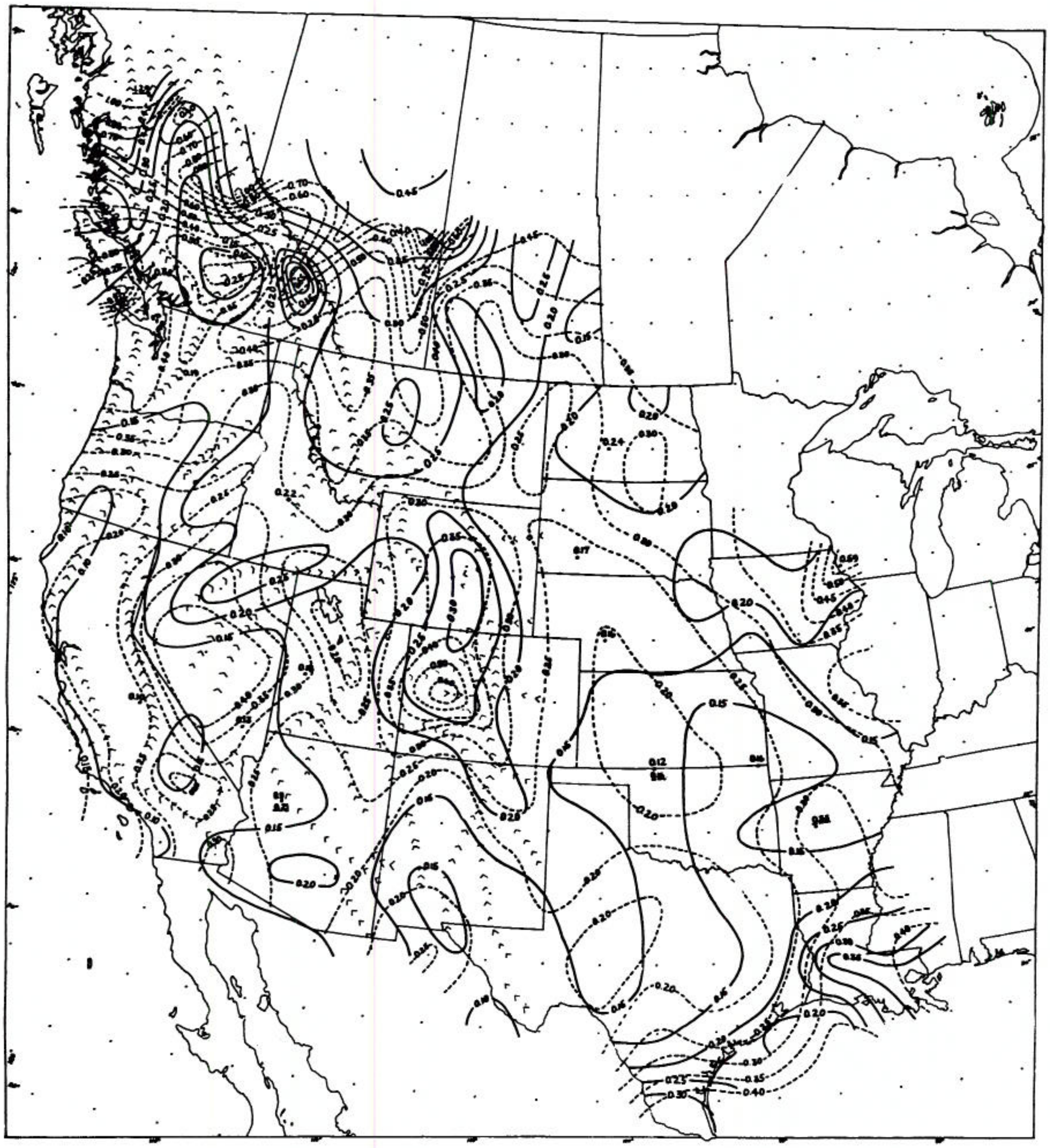

Fig. 31 Regional distributions of the parameters C_1' (solid line) and C_2' (dashed line) for the annual precipitation variable.

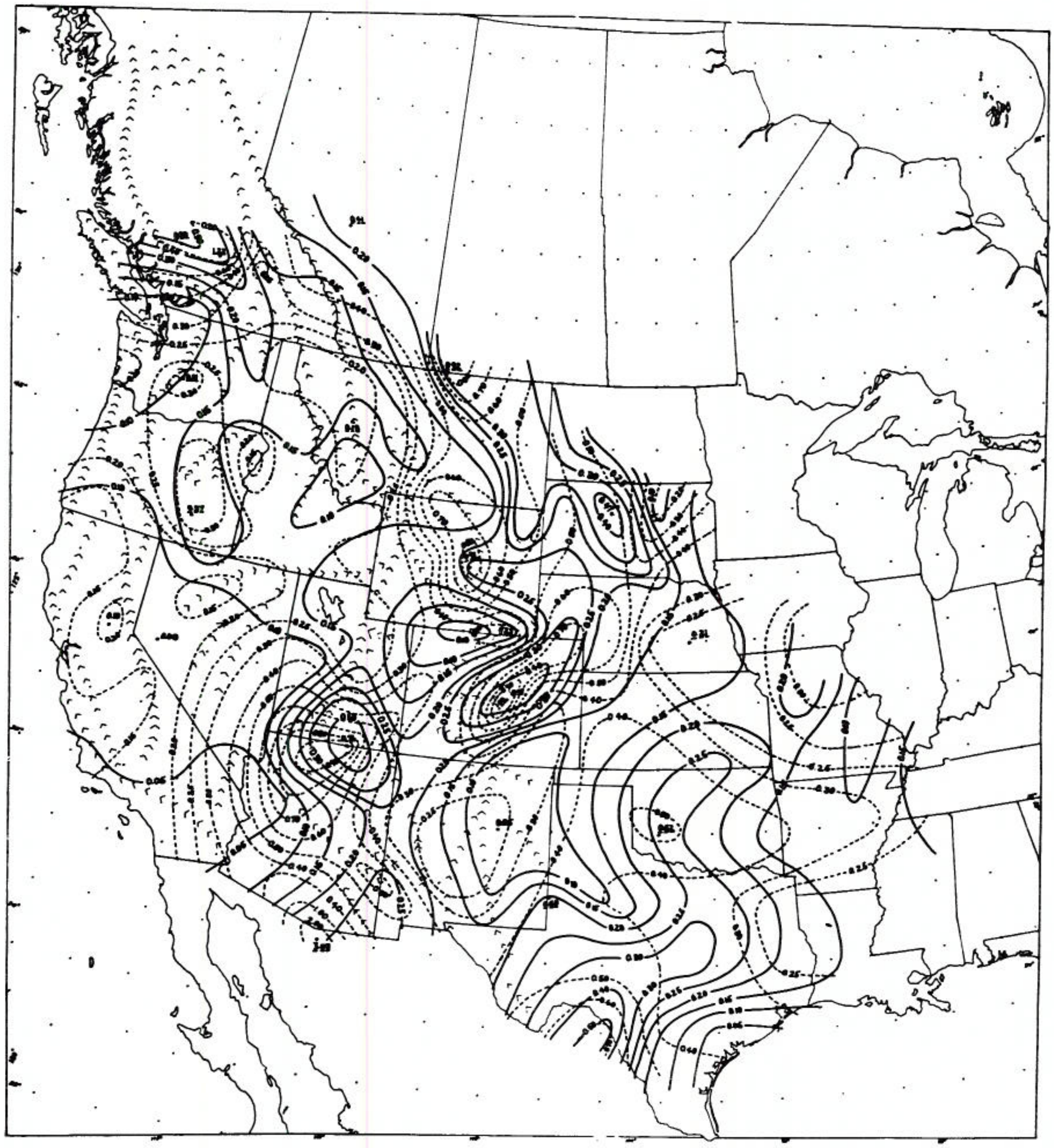

Fig. 32 Regional distributions of the parameters C_1' (solid line) and C_2' (dashed line) for the annual effective precipitation variable.

CHAPTER VII

CONCLUSIONS

AND

RECOMMENDATIONS FOR FURTHER STUDY

This has been an exploratory study and in some aspects more qualitative than quantitative. However, certain conclusions may be drawn.

1. The selected regression model describes adequately the regional patterns of inter-station correlation coefficients in both variables: annual precipitation and annual effective precipitation.

2. Meteorologic and hydrologic factors causing inter-station correlations in the variables studied are not isotropic, and, therefore, isolines of equal inter-station correlation coefficients about a given point may be generally approximated by ellipses.

3. The inter-station correlation coefficients in annual effective precipitation are generally greater than those for precipitation.

4. Although the degree of association within annual effective precipitation is generally greater than in annual precipitation, there is more variability in the correlation coefficients of annual effective precipitation than in those of precipitation.

5. The orientations of the axes of maximal correlations do not afford a ready means of tracing the flux of moisture or a means of identifying its source.

6. Patterns of inter-station correlation coefficients in annual effective precipitation do not generally follow those of precipitation except in areas with unbroken orographic barriers transverse to the flux of moisture.

7. No damped relationship of constant cycles was found to exist between inter-station correlation coefficients and distance for variables on an annual basis. Thus, the hypothesis of a constant wave pattern of storm regeneration is not supported by the results of this study.

8. There is strong evidence that inter-station correlation coefficients of the study variables decrease constantly with distance to approach asymptotically the value of zero.

Further investigations warrant being made in the following areas:

1. Study the distribution of the deviations from the regression surfaces to determine the zone of best fit for the function utilized;

2. Investigate further the variation of values of the coefficients about the correct value of 1.0 at the block control station and at very short distances therefrom;

3. Increase the size of the control blocks to test the applicability of the regression model over larger areas;

4. Make more detailed studies of the regression model applied to selected blocks having more uniformly spaced and densely distributed observation stations in order to relate the model more specifically to hydrologic factors;

5. Reduce the length of the observation interval (seasonal or source moisture division) and investigate the changes in the coefficients and the orientation of the axes of the regression model;

6. Further explore the applicability of the technique to the testing of the consistency of observed data and to the delineation of homogeneous subregions;

7. Incorporate an elevation variable in the regression model and explore its implications; and,

8. Perform theoretical studies of the physical processes producing precipitation and runoff to formulate stochastic models for the regional variations of inter-station correlation coefficients.

BIBLIOGRAPHY

1. Alexander, G. N., 1963, Using probability of storm transposition for estimating frequency of rare floods: Journal of Hydrology, v. 1, no. 1, p. 46-57.

2. Bannon, J. K., Matthewman, A. G., and Murray, R., 1961, The flux of water vapor due to the mean winds, and the convergence of this flux over the Northern Hemisphere in January and July: Royal Meteorological Society Quarterly Journal, v. 87, no. 374, p. 502-512.

3. Benson, M. A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: United States Geological Survey Water-Supply Paper 1580-B, 64 p.

4. Benton, G. S., Blackburn, R. T., and Snead, V. O., 1950, The role of the atmosphere in the hydrologic cycle: American Geophysical Union Transactions, v. 31, no. 1, p. 61-73.

5. Benton, G. S., and Estoque, M. A., 1954, Water-vapor transfer over the North American Continent: Journal of Meteorology, v. 11, no. 6, p. 462-477.

6. Box, G. E. P., 1954, The exploration and exploitation of response surfaces: some general considerations and examples: Biometrics, v. 10, no. 1, p. 16-60.

7. Caffey, J. E., 1965, Regional correlations in annual river flow and in annual precipitation: Ph. D. Dissertation, Colorado State University, Fort Collins, Colorado.

8. Cornish. E. A.. Hill, G. W., and Evans, M. J., 1961, Inter-station correlations of rainfall in Southern Australia: Melbourne, Australia, Commonwealth Scientific and Industrial Research Organization, Division of Mathematical Statistics, Technical Paper no. 10, 16 p., 1 plate.

9. Dorroh, J. H., Jr., 1946, Certain hydrologic and climatic characteristics of the Southwest: Albuquerque, New Mexico, University of New Mexico Publications in Engineering no. 1, 64 p.

10. Foster, E. E., 1944, A climatic discontinuity in the areal correlation of annual precipitation in the Middle West: American Meteorological Society Bulletin, v. 25, no. 9, p. 299-306.

11. Gatewood, J. S., et. al., 1964, General effects of drought on water resources of the Southwest: United States Geological Survey Professional Paper 372-B, 55 p.

12. Horton, R. E., 1943, Hydrologic interrelations between lands and oceans: American Geophysical Union Transactions, v. 24, Part II, p. 753-764.

13. Julian, P. R., 1961, A study of statistical predictability of stream runoff on the upper Colorado River Basin: Research Paper of the High Altitude Observatory, University of Colorado, Boulder, Colorado.

14. McDonald, J. E., 1957, A critical evaluation of correlation methods in climatology and hydrology: Atmospheric Physics Institute Science Report, University of Arizona, no. 4, 36 p.

15. McDonald, J. E., and Green, C. R., 1960, Comparison of rank-difference and product-moment correlation of precipitation data: Journal of Geophysical Research, v. 65, no. 1, p. 333-336.

16. McGuinness, J. L., 1963, Accuracy of estimating watershed mean rainfall: Journal of Geophysical Research, v. 68, no. 16, p. 4763-4767.

17. Markovic, R. D., 1964, Theoretical frequency functions of best fit to distributions of annual precipitations and mean annual river flows: Unpublished Master's Thesis, Colorado State University, Fort Collins, Colorado, 92 p.

18. Marlatt, W., and Riehl, H., 1963, Precipitation regimes over the upper Colorado River: Journal of Geophysical Research, v. 68, no. 24, p. 6447-6458.

19. Marwitz, J. D., 1965, Autocorrelation and regional correlation of summer rainfall: Unpublished Master's Thesis, Colorado State University, Fort Collins, Colorado, 56 p.

20. Merriman, M., ed., 1920, American civil engineers handbook: New York, John Wiley and Sons, Inc., 1955 p.

21. Miller, K. S., 1963, Engineering mathematics: New York, Dover Publications, Inc., 417 p.

22. Parzen, E., 1962, Stochastic processes: San Francisco, Holden-Day, Inc., 324 p.

23. Schleusener, R. E., and Crow, L. W., 1961, Analysis of precipitation data in the Upper Colorado River Basin: Civil Engineering Section, Colorado State University, Fort Collins, Colorado, no. CER61RAS52, 34 p.

24. Stenhouse, N. S., and Cornish, E. A., 1958, Inter-station correlations of monthly rainfall in South Australia: Melbourne, Australia, Commonwealth Scientific and Industrial Research Organization, Division of Mathematical Statistics, Technical Paper no. 5, 22 p.

25. Thom, H. C. S., 1940, On statistical analysis of rainfall data: American Geophysical Union Transactions Part II, p. 490-499.

26. Thomas, H. E., 1962, Meteorologic phenomenon of drought in the Southwest: United States Geological Survey Professional Paper 372-A, 43 p.

27. Visher, S. S., 1943, Some influences upon American climate of the ocean and the gulf: An analysis of climatic changes in "Climate and Man," 1941: American Meteorological Society Bulletin, v. 24, no. 3, p. 79-84.

28. Yevdjevich, V. M., 1963a, Patterns in sequence of annual river flow and annual precipitation: Berkeley, California, International Association of Scientific Hydrology, Publication no. 63, p. [illegible].

29. Yevdjevich, V. M., 1963b, Fluctuations of wet and dry years, Part I. Research data assembly and mathematical models: Colorado State University, Fort Collins, Colorado, Hydrology Paper no. 1, 55 p.

30. Yevdjevich, V. M., 1964, Fluctuations of wet and dry years, Part II. Analysis by serial correlation: Colorado State University, Fort Collins, Colorado, Hydrology Paper no. 4, 50 p.

Key Words: Hydrology, Inter-station Correlation, Correlation, Response Surface, Multiple Regression, Least Squares, Annual River Flows, Annual Effective Precipitation, Annual Precipitation.

Abstract: Two large samples of annual river flow and annual precipitation are investigated for patterns in inter-station correlation coefficients about five station locations. The regional variation in inter-station correlation coefficients is reasonably explained for both variables by the regression equation $r_e = a_1 \exp[-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d]$ where r_e is the estimate of the inter-station correlation coefficient, d is the inter-station distance, and θ is the expression of direction. On the average, approximately 60 percent of the regional variation is explained by the selected function. Isolines of equal coefficients for the fitted surfaces approximate ellipses indicating that meteorologic and hydrologic factors affecting the regional correlations are not isotropic. Factors such as evapotranspiration and river basin characteristics tend to cause the regional patterns for the annual river flow to differ from those observed for annual precipitation.

Reference: Caffey, James E., Colorado State University, Hydrology Papers No. 6 (June 1965) "Regional Correlations in Annual River Flow and in Annual Precipitation."

Key Words: Hydrology, Inter-station Correlation, Correlation, Response Surface, Multiple Regression, Least Squares, Annual River Flows, Annual Effective Precipitation, Annual Precipitation.

Abstract: Two large samples of annual river flow and annual precipitation are investigated for patterns in inter-station correlation coefficients about five station locations. The regional variation in inter-station correlation coefficients is reasonably explained for both variables by the regression equation $r_e = a_1 \exp[-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d]$ where r_e is the estimate of the inter-station correlation coefficient, d is the inter-station distance, and θ is the expression of direction. On the average, approximately 60 percent of the regional variation is explained by the selected function. Isolines of equal coefficients for the fitted surfaces approximate ellipses indicating that meteorologic and hydrologic factors affecting the regional correlations are not isotropic. Factors such as evapotranspiration and river basin characteristics tend to cause the regional patterns for the annual river flow to differ from those observed for annual precipitation.

Reference: Caffey, James E., Colorado State University, Hydrology Papers No. 6 (June 1965) "Regional Correlations in Annual River Flow and in Annual Precipitation."

Key Words: Hydrology, Inter-station Correlation, Correlation, Response Surface, Multiple Regression, Least Squares, Annual River Flows, Annual Effective Precipitation, Annual Precipitation.

Abstract: Two large samples of annual river flow and annual precipitation are investigated for patterns in inter-station correlation coefficients about five station locations. The regional variation in inter-station correlation coefficients is reasonably explained for both variables by the regression equation $r_e = a_1 \exp[-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d]$ where r_e is the estimate of the inter-station correlation coefficient, d is the inter-station distance, and θ is the expression of direction. On the average, approximately 60 percent of the regional variation is explained by the selected function. Isolines of equal coefficients for the fitted surfaces approximate ellipses indicating that meteorologic and hydrologic factors affecting the regional correlations are not isotropic. Factors such as evapotranspiration and river basin characteristics tend to cause the regional patterns for the annual river flow to differ from those observed for annual precipitation.

Reference: Caffey, James E., Colorado State University, Hydrology Papers No. 6 (June 1965) "Regional Correlations in Annual River Flow and in Annual Precipitation."

Key Words: Hydrology, Inter-station Correlation, Correlation, Response Surface, Multiple Regression, Least Squares, Annual River Flows, Annual Effective Precipitation, Annual Precipitation.

Abstract: Two large samples of annual river flow and annual precipitation are investigated for patterns in inter-station correlation coefficients about five station locations. The regional variation in inter-station correlation coefficients is reasonably explained for both variables by the regression equation $r_e = a_1 \exp[-(a_2 + a_3 \cos 2\theta + a_4 \sin 2\theta)d]$ where r_e is the estimate of the inter-station correlation coefficient, d is the inter-station distance, and θ is the expression of direction. On the average, approximately 60 percent of the regional variation is explained by the selected function. Isolines of equal coefficients for the fitted surfaces approximate ellipses indicating that meteorologic and hydrologic factors affecting the regional correlations are not isotropic. Factors such as evapotranspiration and river basin characteristics tend to cause the regional patterns for the annual river flow to differ from those observed for annual precipitation.

Reference: Caffey, James E., Colorado State University, Hydrology Papers No. 6 (June 1965) "Regional Correlations in Annual River Flow and in Annual Precipitation."