# Graphical Correlation of Gaging-Station Records

Manual of Hydrology: Part 1. General Surface-Water Techniques

# Graphical Correlation of Gaging-Station Records

*By* JAMES K. SEARCY

Manual of Hydrology:  Part 1. General Surface-Water
  Techniques

*Methods and practices*
*of the Geological Survey*

UNITED STATES DEPARTMENT OF THE INTERIOR

FRED A. SEATON, *Secretary*


GEOLOGICAL SURVEY

Thomas B. Nolan, *Director*

# CONTENTS

# ILLUSTRATIONS

IV CONTENTS

# TABLES

Page

TABLE 1. Monthly mean discharge, in cubic feet per second, of Leaf River near Collins, Miss., drainage area, 752 square miles_ 82

2. Monthly mean discharge, in cubic feet per second, of Bowie Creek near Hattiesburg, Miss., drainage area, 304 square miles_____ 83

3. Monthly mean discharge, in cubic feet per second, of White River at Greenwater, Wash., drainage area, 216 square miles_____ 86

4. Monthly mean discharge, in cubic feet per second, of Green-water River at Greenwater, Wash., drainage area, 73.9 square miles_____ 86

5. Curve values $(Y_c)$ of monthly discharge, in cubic feet per second, of Bowie Creek near Hattiesburg, Miss., from figure 10_____ 91

6. $Y:Y_c$ ratio for Bowie Creek near Hattiesburg, Miss_____ 92

7. Monthly mean discharge, in cubic feet per second, of Bogue Chitto at Franklinton, La., drainage area, 985 square miles_____ 92

8. Curve values $(Y_c)$ of monthly discharge, in cubic feet per second, of Bogue Chitto at Franklinton, La., from figure 11_ 94

9. $Y:Y_c$ ratio for Bogue Chitto at Franklinton, La_____ 94

10. $Y:Y_c$ ratio for Bowie Creek near Hattiesburg, Miss., from figure 12_____ 95

11. Monthly discharge of Bowie Creek near Hattiesburg, Miss., adjusted for Bogue Chitto_____ 95

12. Monthly mean discharge, in cubic feet per second, of Leaf River at Hattiesburg, Miss., drainage area, 1,760 square miles_____ 99

13. Monthly discharge, in cubic feet per second, Leaf River at Hattiesburg, Miss. minus Leaf River near Collins, Miss___ 99

# MANUAL OF HYDROLOGY: PART 1. GENERAL SURFACE-WATER TECHNIQUES

## GRAPHICAL CORRELATION OF GAGING-STATION RECORDS

By JAMES K. SEARCY

### ABSTRACT

A gaging-station record is a sample of the rate of flow of a stream at a given site. This sample can be used to estimate the magnitude and distribution of future flows if the record is long enough to be representative of the long-term flow of the stream. The reliability of a short-term record for estimating future flow characteristics can be improved through correlation with a long-term record. Correlation can be either numerical or graphical, but graphical correlation of gaging-station records has several advantages.

The graphical correlation method is described in a step-by-step procedure with an illustrative problem of simple correlation, illustrative problems of three examples of multiple correlation—removing seasonal effect—and two examples of correlation of one record with two other records. Except in the problem on removal of seasonal effect, the same group of stations is used in the illustrative problems. The purpose of the problems is to illustrate the method—not to show the improvement that can result from multiple correlation as compared with simple correlation.

Hydrologic factors determine whether a usable relation exists between gaging-station records. Statistics is only a tool for evaluating and using an existing relation, and the investigator must be guided by a knowledge of hydrology.

## INTRODUCTION

Gaging-station records supply the basic data needed to evaluate and design water-resources developments. In addition to being a historical record of the streamflow that occurred at the gaging-station site, the streamflow record is a sample that often can be used to estimate the magnitude and distribution of future flows.

The reliability of the streamflow sample for estimating future flows at the gaging-station site depends upon how well the sample represents the flows that will occur during the life of a development. One important measure of representativeness is the length of the streamflow

67

sample. Frequently, streamflow data are needed for sites at which no gaging stations have been operated or for which the records are much too short to be representative of future flow.

Statistical correlation is a means by which short-term records can be adjusted to represent a long-term record. Correlation is the process of establishing a mutual relation between a variable and one or more related variables. For gaging-station records, the usual variables are the short-term gaging-station record and one or more long-term gaging-station records. The correlation is termed simple correlation if the relation is between two variables, and is termed multiple correlation if the relation is between one variable and two or more related variables.

Either type of correlation can be graphical or numerical (the two methods differ in procedure rather than in principle), but for gaging-station records, graphical correlation has decided advantages. When logarithms of the discharge are used, fitting a curved relation between gaging-station records by graphical methods is less laborious than fitting a curve by numerical methods.

Graphical methods for simple correlation and for three types of multiple correlation are explained in this report. These methods were developed principally within the Geological Survey by W. B. Langbein, C. H. Hardison, and others. The examples were selected to illustrate the methods and not to show the differences in the results of the various methods. The multiple correlation techniques presented here by no means exhaust the possibilities of multiple correlation. For example, the use of precipitation data in place of a third gaging-station record might prove best in some localities.

## USE OF CORRELATION

The relation established between gaging-station records for a concurrent short period is sometimes used to obtain correlative estimates of monthly discharges at the short-term station corresponding to discharges at the long-term station outside the common period. Correlation is also useful for extending short-term gaging-station records to estimate flows of a given frequency and for analyzing storage-frequency. Often, the information on magnitude and frequency of low flows from a long-term station can be transferred directly to the short-term station through the curve of relation without synthesizing a table of discharges for the short-term station. Such estimates represent a discharge that has the same probability of occurrence as at the long-term station. The frequency data that are transferred directly have the same order of accuracy as if they had been computed from a table of correlative estimates for the same period.

Use of estimates derived from correlation are in some ways an unsatisfactory substitute for actual discharge records. For design purposes, however, a streamflow record, of any length, represents only a sample of the flow of the stream, and as such contains an inherent sampling error. This sampling error, which cannot be reduced by measuring the same sample more accurately, is often large in comparison to the errors that would be introduced by correlation with a longer record. Thus correlation with longer records can serve to reduce the sampling error and to provide a more reliable base from which to estimate the characteristics of future flow.

Langbein and Hardison (1955) demonstrated that the mean discharge determined from a short record can generally be improved by correlating it with a record more than 25 percent longer than the short record because the error introduced by the correlation is usually less than the decrease in sampling error accomplished by extension.

## HYDROLOGIC PRINCIPLES

The runoff from a drainage basin depends upon the climate and upon the inherent characteristics of the basin. In areas of low relief, the climate of nearby drainage basins is much the same, but in mountainous areas the climate varies because of the positions of the mountain ranges and the differences in the altitude of the basins.

The principal drainage basin characteristics that affect the amount and distribution of runoff are location, size, shape, physiography, geology, soils, vegetative cover, and cultural developments. Numerous other basin characteristics have a lesser effect on the runoff from the basin. Any attempt to evaluate the extent to which one of these basin characteristics affects runoff and to relate the runoff of one basin to that of another by measuring the differences in their basin characteristics would be a tremendous undertaking.

The complex interrelations of climate and drainage basin characteristics are integrated in the flow of the stream, and their aggregate effect is measured directly at the stream-gaging station. The measured runoff from drainage basins obviously furnishes the best basis for comparing their runoff characteristics.

Relations between the flow at stream-gaging stations can be based on concurrent daily, weekly, monthly, or annual discharge. Generally, the relations are the same regardless of the time unit used, although the variation of the plotted points decreases as the time unit becomes longer. Selection of the time unit is a hydrologically arbitrary decision.

Monthly discharge, which is the unit used in many water-resources development studies, is a convenient time unit for establishing relations. Monthly discharges are published in the Geological Survey

Water-Supply Papers in the annual series of Surface Water Supply of the United States and are compiled from beginning of record through September 1950 in Water-Supply Papers 1301–1318. Monthly discharges are used in the illustrative problems of this chapter.

### FACTORS AFFECTING THE RELATION BETWEEN GAGING-STATION RECORDS

Although some characteristics of two nearby basins in areas of low relief may differ, their climates are generally similar. This bond of climate is the key to the correlation between gaging-station records. If the bond is missing, as, for example in attempting to correlate a mountain stream and a nearby desert stream, a usable correlation is impossible. The climate of the drainage basins need not be identical, but each basin should have the same chance of receiving a given rainfall.

The distance between the gaging stations adversely affects the degree of relation between gaging-station records, so that other factors being equal, the relation becomes less reliable as the distance between the stations increases. An exception to this is that records from two stations a considerable distance apart on the same stream are generally more closely related than are records from nearby streams.

The relative size of the drainage basins affects their relation, particularly in the time distribution of flows. The runoff of a large basin is generally the summation of high flows from some basins and low flows from others and is thus less variable than the runoff of most of the small, separate basins of which it is composed. Heavy rains near the end of a month substantially increase the monthly runoff from most small basins, whereas large ones receive little increase in runoff until the following month. This condition would cause the plotted points for two consecutive months to depart in opposite directions from a line of average relation. Therefore when basins differ greatly in size, the runoff from one is a poor indication of the concurrent runoff from the other basin. Correlation between stations whose drainage areas differ in size by more than tenfold should be avoided.

Differences in the surficial geology of drainage basins have a pronounced effect on low-flow relations. As compared with basins of low infiltration capacity, those of high infiltration capacity have little direct runoff from light rains but have higher sustained flows during dry periods. Direct runoff from rains that exceed the infiltration capacity of both drainage basins is less affected by the surficial geology and thus tends to become proportional to the size of the drainage areas.

A complete discussion of the factors that affect the relation between gaging stations is beyond the scope of this chapter. The purpose of this brief discussion is to emphasize that the hydrology of the two drainage basins is the basis of their relation—not the statistics used. Nevertheless statistics furnishes a powerful tool for organizing streamflow data to express useful relations.

## STATISTICAL PRINCIPLES

Few users of streamflow records are statisticians. Although many good statistics books are available, one is often confused in applying statistics to a streamflow problem because of the variety of symbols used by the different authors and the lack of reference to streamflow applications. This chapter explains an application of statistics to streamflow problems and the use of graphical correlation methods to determine the relation between gaging-station records; the explanation is not intended to be a treatise on graphical correlation. The technique described in this chapter is basically the same as that described by Langbein and Hardison (1955) and Somers (1954).

The statistical formulas given here are based on the premises that the data are normally distributed and that the observations are entirely random in sequence. (See "Concepts and definitions.") Consecutive daily streamflow data fit neither of these criteria, but the data can be normalized by a logarithmic transformation accomplished by either using the logarithm of the discharge or by plotting the discharge data on logarithmic paper. Streamflow is by nature nonrandom. Each daily discharge is affected by flows for previous days and in turn, affects succeeding daily flows. Monthly discharge is also nonrandom but to a lesser degree than daily discharge. The primary cause of nonrandomness in streamflow data is the lag in the rainfall-runoff relation (carryover effect). A secondary cause is the nonrandomness of precipitation and the tendency of wet and dry years to occur in groups rather than in random sequence.

In addition to being random, statistical data should be homogeneous (drawn from the same population). Streamflow data from the same stream would seem homogeneous by definition, but the nonhomogeneity may result from major developments that affect runoff, or occasionally by radical changes in data-collection methods. The double-mass curve discussed in a separate chapter is one means of testing the homogeneity of streamflow data.

## CONCEPTS AND DEFINITIONS

*Probability.*—Probability is the likelihood of the occurrence of any particular event expressed as a decimal fraction or as a percentage.

A value of zero is assigned to an event that certainly will not occur, and a value of one to an event that is certain to occur. A probability of 0.50 means that an event is as likely to occur as not (a 50–50 chance).

*Individual.*—A unit of data is an individual. In this chapter the mean monthly flow is the unit or individual.

*Population (universe).*—Any set of individuals having a common measurable characteristic constitutes a population; for example, a long series of monthly flows of a stream unaffected by regulation, diversion, or major changes in land use.

*Sample.*—A group of individuals taken from a population constitutes a sample. $N$ is used to represent the number of individuals, or, in correlation, the number of pairs of individuals in the sample. For example, a 30-year record of flows at a gaging station is a sample from the life of the stream. A sample is taken for the purpose of making inferences about the population.

*Random data.*—Random data are data in which the individuals in the sample were selected in such a manner that any individual in the population had an equal chance of being selected on the next draw.

*Normal distribution.*—Many populations are distributed about the mean according to the Gaussian or normal distribution. In such a distribution 68.27 percent of the individuals lie within one standard deviation of the mean; 95.45 percent lie within two standard deviations; and almost all (99.73 percent) lie within 3 standard deviations of the mean.

The error formulas and most of the tests of significance are based on the assumption that the variable is distributed in the normal form. Unless departure from the normal form is severe the tests of significance and the error formulas are sufficiently accurate.

*Standard deviation (s).*—The standard deviation (sometimes called the root mean square of the deviations) is a measure of the scatter of the items about the arithmetic mean of the sample and is computed by the formula

$$s = \sqrt{\frac{\Sigma(x^2)}{N-1}} \tag{1}$$

where $x=$ deviation of the item from the arithmetic mean $(\overline{X})$

or    $x = X - \overline{X}$

$N=$ number of items

*Degrees of freedom (d.f.).*—A nebulous but important concept in statistics is that of degrees of freedom, but few textbooks attempt a definition. The number of degrees of freedom is, in general, the number of independent observations minus the number of constraints imposed on the system. The constraints imposed on the system are

called lost degrees of freedom (*l.d.f.*).  Thus, we lose one degree of freedom in computing the standard deviation because one constraint (an average) has been placed on the data.  For example, if the average of 10 numbers is 100, 9 of the numbers can have any value, but the 10th number can be only that which makes the average 100; we lose the freedom of choice on one item by specifying the average. The procedure for determining lost degrees of freedom will be shown in each problem.

*Independent variable.*—The independent variable is the one from which another variable is to be estimated or predicted.  It is plotted as the abscissa (*X*-axis).  The independent variable is generally the long-term station when extending streamflow records.

*Dependent variable.*—The dependent variable is the one that is estimated or predicted from given values of the independent variable.  It is plotted as the ordinate (*Y*-axis).  The dependent variable is generally the short-term station when extending streamflow records.

*Scatter diagram.*—Where comparative data (for example, concurrent monthly discharges) at two gaging stations are plotted against each other, the result is a scatter diagram.  The pattern of the scatter indicates the degree of relation between the flow at the two stations. A narrow scatter band indicates a high degree of association between the two sets of data, and greater departures from the relation line indicate lesser degrees of association.

*Plotting paper.*—Logarithmic paper (base 10) is used for plotting the relation between the flows at two gaging stations because for most data the relation on such paper can be approximated by one or more straight lines.

*Line of regression.*—The line or equation expressing a linear relation between two variables is known as the line of regression.  In the numerical method of analysis, the line of regression is generally fitted by the method of least squares.  The line is termed the regression of $Y$ on $X$ if the squares of the $Y$ residuals have been minimized or the regression of $X$ on $Y$ if the squares of the $X$ residuals have been minimized.

*Curve of relation.*—The graphically determined line of regression should preferably be called a curve of relation to disassociate it from a line fitted by the least squares method.

*Equal-yield line.*—The equal-yield line is a line of relation (45° on logarithmic paper) between the flows at two stations, based on the assumption that the runoff at each station is proportional to its respective drainage area.

*Log unit.*—A log unit is one cycle on logarithmic paper.  Log units less than unity indicate the proportional part of a cycle.  For example, one-tenth of the length of a cycle is 0.10 log unit.

*Correlative estimates* $(Y_c)$.—The curve of relation value for $Y$ corresponding to a given value of $X$ is a correlative estimate.

*Residual* $(d)$.—The vertical $(Y)$ difference between the curve of relation and a plotted point is a residual (sometimes called error). It is the actual value of $Y$ minus the calculated (estimated) value of $Y$, or $d = Y - Y_c$. The residual carries either a plus or a minus sign.

*Standard error of estimate* $(S_e)$.—The standard error of estimate is a measure of the variation of the points about the curve of relation. It is similar to the standard deviation except that the curve of relation takes the place of the arithmetic mean. If the distribution of the scatter is normal, lines drawn a vertical distance equal to one standard error of estimate above and below the line or curve of regression and parallel to it, will encompass about two-thirds (68.27 percent) of the points. Parallel lines two standard errors of estimate above and below the line of regression will encompass 95.46 percent of the points, and lines three standard errors of estimate away will encompass 99.73 percent of the points. In correlating gaging-station records, the standard error of estimate is conveniently expressed in log units, rather than in units of $Y$ or in percentage of the mean.

For a straight-line graph the mathematical expression for the standard error of estimate is

$$S_e = \sqrt{\frac{\Sigma (d)^2}{N-2}} \qquad (2)$$

where

$$d = \text{residual } (Y - Y_c)$$
$$N = \text{number of items}$$

For any curve fitted to observed data, $S_e$ may be computed by the formula

$$S_e = \sqrt{\frac{\Sigma (Y - Y_c)^2}{N - l.d.f.}} \qquad (3)$$

where

$l.d.f. = $ lost degrees of freedom or number of constraints placed on the system (2 degrees of freedom are lost by use of a straight line)

$Y \quad = $ observed value of the dependent variable

$Y_c \quad = $ computed or curve value of the dependent variable

*Coefficient of correlation* $(r)$.—The size of the standard error of estimate is a measure of the closeness of the points to the curve of relation, but it does not show the improvement in the estimate of $Y$ that results from using the correlation with $X$ rather than estimating $Y$ as the mean of the $Y$ values. A measure of the usefulness of the curve of relation for estimating $Y$ values may be obtained by relating the standard error of estimate $(S_e)$ to the standard deviation of the $Y$ values $(s_y)$. This relates the variation of the points about the curve of relation to the variation of the points about their arithmetic mean.

For a linear relation, such a measure is the coefficient of correlation ($r$) computed by the formula

$$r=\sqrt{1-\left(\frac{S_e}{s_y}\right)^2} \qquad (4)$$

The value of the coefficient of correlation ranges from 0 for no correlation to $\pm 1.00$ for perfect correlation.  A plus sign is assigned $r$ if the line of regression slopes upward to the right and a minus sign if it slopes downward.  A low coefficient of correlation does not always indicate poor correlation; it may result from fitting a straight line to a good relation that should be described by a curved line.

*Index of correlation* ($\rho$).—For curvilinear relations the measure of correlation is called the index of correlation, or

$$\rho=\sqrt{1-\left(\frac{S_e}{s_y}\right)^2} \qquad (5)$$

The similarity between $r$ and $\rho$ is apparent.  The limits of $\rho$ are 0 for no correlation to 1.00 for perfect correlation.  No plus or minus sign can be assigned to $\rho$ because the sign might vary in different parts of the curve.

*Coefficient or index of correlation of gaging-station records.*—The coefficient, or index, of correlation furnishes a guide to the length of long-term record needed for improving the mean of a short-term record through correlation.  Langbein and Hardison (1955) give criteria for determining the minimum ratio of the long-term to the short-term record for various values of the coefficient or index of correlation.

The usefulness of the correlation, as indicated by $r$ or $\rho$, shows how much improvement in the estimates of $Y$ results from using the correlation, as compared with estimating $Y$ as the mean of all $Y$ values.  This concept is correct for many types of data, but as streamflow data usually have a definite seasonal pattern, for example, the most likely discharge for October would be the mean of the October discharges rather than the mean of all monthly discharges.  Thus, the actual improvement in estimating $Y$ from the curve of relation is not as great as the coefficient of correlation (computed by equation 4) indicates.  A better measure of the improvement resulting from the correlation of gaging-station records is to compare the standard error of estimate (equation 3) with the standard deviations of the monthly discharge about their respective monthly means rather than about the mean of all the months (Langbein and Hardison, 1955, p. 1176–77 of closure).  The computed standard deviation about the respective

monthly means is adjusted for the 12 degrees of freedom lost (multiply

by $\sqrt{\dfrac{N}{N-12)}}$ and substituted for the value of $s_y$ in the formula $r$ or

$\rho = \sqrt{1 - \left(\dfrac{S_e}{s_y}\right)^2}$. In the illustrative correlation of the section, "Remov-

ing seasonal effect," the index of correlation adjusted for monthly means is 0.94 as compared with an index of 0.98 using $S_e$ computed by formula 3. In some correlations the adjustment for deviations of the data about the monthly mean makes a great difference in the value of $r$ or $\rho$.

## THE GRAPHICAL METHOD

The first step in expressing the relation between pairs of associated measurements is to plot the measurement of one variable against the corresponding measurement of the other variable. Unless the variables are precise measurements of two quantities that bear a fixed relation to each other, the plot will present a "shotgun" pattern, technically known as a scatter diagram. The cluster of the points indicates the closeness of the relation or perhaps the accuracy of the measurements, whereas the shape of the pattern indicates the nature of the relation. A straight line is used if it fits the points fairly well—the resulting relation is called linear. If the points lie on a curved line the relation is called curvilinear. The graphical method differs from the numerical method in the way in which the line or curve is fitted to the points. Neither method explains why the points scatter.

In graphical correlation, the shape of the relation line depends to a large extent upon the type of plotting paper used. Logarithmic paper is convenient for graphical correlation of gaging-station records because it tends to normalize the streamflow data and to transform the common curvilinear relation between gaging-station records to one or more straight lines. The Geological Survey uses 5-inch cycle logarithmic paper with either 3 by 2 cycles or 4 by 3 cycles. The 5-inch cycle fits the 20-graduations-to-the-inch face of an engineers' scale so that direct readings can be made to hundredths of a log cycle.

Generally, the relation between two gaging-station records can be satisfactorily expressed on logarithmic paper by 2 straight lines, 1 for the high-flow part of the relation and 1 for the low-flow part. For many streams east of the Mississippi River, the 2 lines intersect at from 1 to 2 times the average discharge. If the low-flow and the high-flow regression lines intersect at sharp angles, they should be connected by a transition curve. Drawing a transition curve is purely arbitrary but it is hard to conceive of a relation between two gaging-station records that changes abruptly at some specified discharge.

If the angle between the lines is very flat, a transition curve is unnecessary.

The relation between 2 records can sometimes be described by a single straight line and occasionally the relation between 2 records requires a third straight line at extremely low flows. In general, straight lines are drawn through the data and little weight is given to the position of points near the extreme upper or lower end of the line. Although the points sometimes seem to define sharply curved or even wavy lines, the line of relation for the succeeding period of equal length will often balance the irregularities of the first period and at the extremities define a hook in the opposite direction from that defined in the first period.

Where the high-flow relation line is poorly defined, the equal-yield (drainage-area ratio) line may be used as a guide in drawing the relation line. The high-flow relation line and the equal-yield line for most gaging-station correlations lie close together and are about parallel. This is not true where the storage capacity (both surface and underground) within the two basins differs greatly.

The slope of the low-flow relation line is governed by the surficial geology of the two drainage basins. If the geology of the two basins were identical, no break would occur in the relation line. The streamflow data generally provide good definition of the low-flow relation line except at the extreme low end. When one of the streams ceases to flow or yields less at extreme low flows than the other streams, the low-flow relation line may break at extreme low flows and cross the axis of the better yielding stream.

Streams affected by snowmelt may have a well-defined relation during part of the year that differs from the relation during other periods of the year. Such a condition may occur because of differences in the aspect of the basins with respect to the sun, differences in forest cover, or differences in altitude that affect the rate of snowmelt. The procedure for correlating records where the rate of snowmelt differs between the basins is described in the section "Multiple correlation" under the subheading "Removing seasonal effect."

## COMPARISON OF THE GRAPHICAL AND NUMERICAL METHODS

A principal difference between curve fitting by the graphical method, described here, and the numerical method of least squares is the effect of the remote points on the line of relation. By squaring the distance from the line of relation to the remote point, the method of least squares gives much more weight to the remote point than to one of the more numerous points near the regression line. This moves the line of relation toward the remote point and lessens the deviation of the point from the regression line while increasing the deviation of a point near to but on the other side of the line. As an

exceptional example, the line of relation might be moved to a position between where most of the points lie and where the few remote points lie. Thus, the use of the numerical method for correlation of gaging-station records sometimes biases most of the estimates. Using the mean of a group of points has the same objection but to a lesser degree.

The graphical method uses the median of groups of points to position the line of relation. Thus, the magnitude of a remote point has no influence on the position of the line of relation, except that it is above or below the midvalue of the group. Generally a line of relation positioned by median points gives a better correlative estimate than a least-square line but where the conditions which produced the remote point recur, the median point line of relation gives a poorer estimate than the least-square line.

The graphical correlation procedure described in this manual gives a single line of relation between 2 gaging stations, whereas the numerical method gives 2 lines of best fit depending on which station is used as the dependent variable. Each of the two lines of best fit, called least-square lines, are regressed toward the mean of the dependent variable and the discharges estimated from the least-square line will have less variability than those of the sample used in the correlation. For most uses estimated discharge should maintain the original variability so that frequency presentations based on estimated data will not show streamflow to be less variable than it is simply because the coefficient of correlation was low. The single line of relation, or engineer line, is used in this manual to maintain the original variability of the sample.

Another advantage of the graphical method is that a curved relation can be defined graphically as easily as a linear relation, whereas fitting a curved line to data by the numerical method is complicated. A straight-line relation is seldom applicable throughout the range in discharge, because few gaging-station records have the same relation to another record at both high flows and low flows. The low flow, or dry-weather flow, comes from ground-water storage, whereas the high flow and the medium flow are principally direct runoff from precipitation. In order that the relation between two gaging-station records be linear throughout the range in flow, the effect of surficial geology on low flow must produce the same ratio between the records as the effect of other basin characteristics produces on high flow.

Neither the graphical method nor the numerical method is satisfactory for extending a curvilinear relation beyond the observed data. Extrapolation by either method is based on the dubious assumption that the curvature used to fit the extreme points will fit points that may lie outside the range of present data. Although the numeri-

cal method provides a single extension, the result is as questionable as a curve extended graphically by eye.

## SIMPLE CORRELATION

The procedure for simple correlation of two gaging-station records is described in the following example, correlating two stations in southern Mississippi—station $A$ (Leaf River near Collins, drainage area, 752 square miles) and station $C$ (Bowie Creek near Hattiesburg, drainage area, 304 square miles). In another section these stations, and two others in the vicinity (see fig. 9), are used to illustrate the
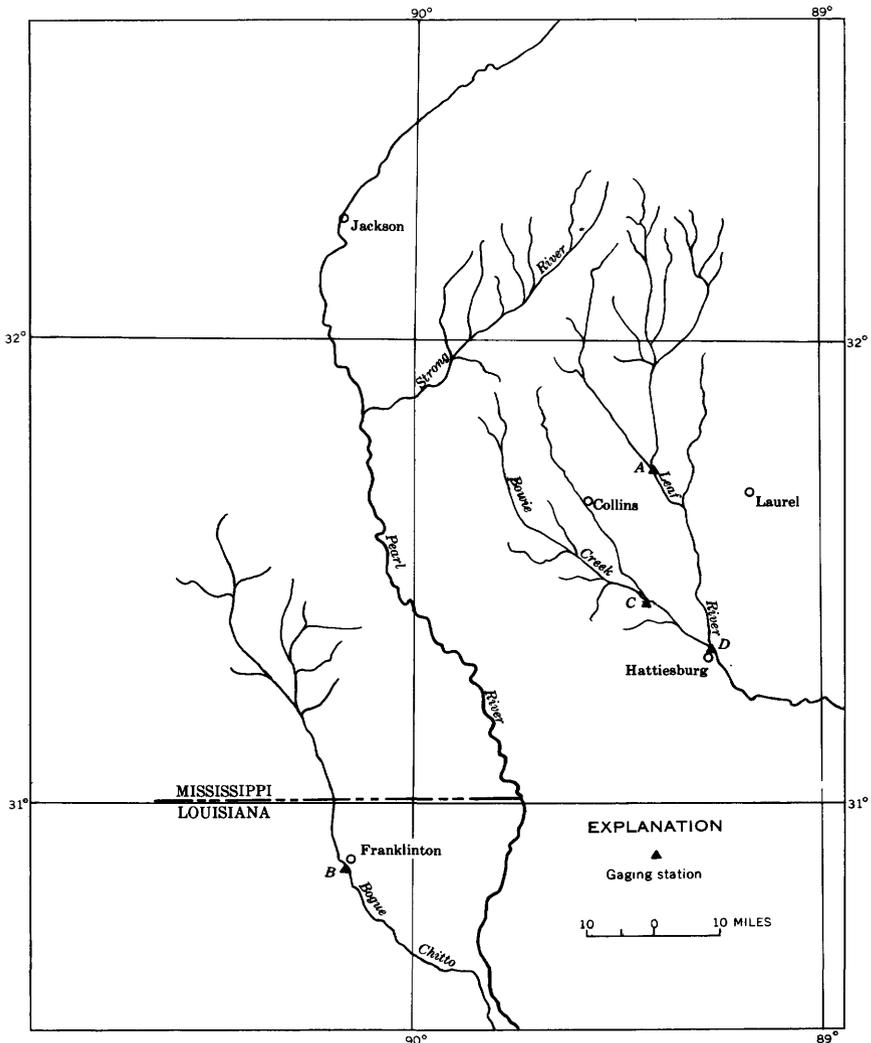


FIGURE 9.—Area in southern Mississippi.

multiple correlation procedures. Southern Mississippi has little variation in relief or in climate, and snowmelt is not a factor in runoff.

## GRAPHICAL CORRELATION WITH TWO VARIABLES

1.  Plot the selected unit (usually monthly discharge in cubic feet per second) of station $A$ (table 1) against the concurrent unit of station $C$ (table 2) on logarithmic paper with the independent variable on the $X$ axis. (See fig. 10). For working with unfamiliar records, distinctive plotting symbols for each calendar month are recommended because this will reveal seasonal patterns in runoff (see "Removing seasonal effect").

2.  Using vertical lines, divide the total distribution of the points in the $X$ direction into 5 to 10 parts, or slices of equal width (8 slices are used in fig. 10).

3.  Graphically determine the median of the points in each slice in in both the $X$ and the $Y$ directions. If the number of points in a slice is even, average (graphically) the middle two points; if odd, average the middle three points, giving the midpoint a weight of two. If a slice contains less than three points, do not determine a median for that slice. A simple method of giving the midpoint a weight of two is to average the midpoint and the point next above, then average the midpoint and the point next below; the average of the two averages is the median.

4.  Repeat steps 2 and 3 for the $Y$ direction. In practice, it is convenient to use different colors for the medians of the $Y$ slices.

5.  Plot the equal-yield line by drawing a dashed 45° line through the point obtained by plotting the drainage areas of stations $A$ and $C$ as used for discharge figures.

6.  Draw straight lines (of relation) through the average of the median points giving reduced weight to the end points. The upper end of the relation is generally about parallel to the equal-yield line. The lower end is generally another straight line which can be connected to the upper line by a short transition curve.

7.  Draw two lines each equidistant (vertically) from the curve of relation throughout its length, one above and one below the curve of relation, so that one-sixth of the points will lie above the upper curve and one-sixth of the points will lie below the lower curve. The upper and lower lines will seldom be at the same distance from the line of relation.
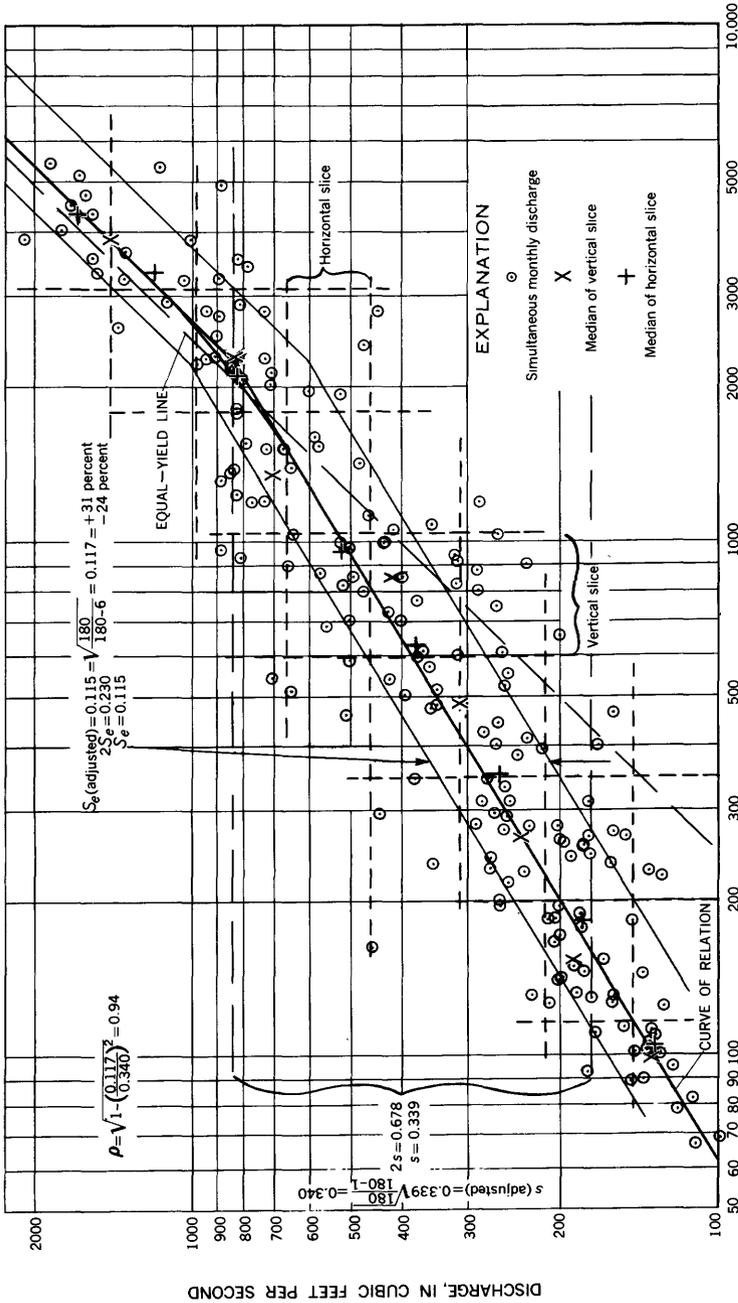
FIGURE 10.—Correlation of monthly discharge, Leaf River and Bowie Creek, water years 1940–54.

TABLE 1.—*Monthly mean discharge, in cubic feet per second, of Leaf River near Collins, Miss., drainage area, 752 square miles*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 93.5 | 90.0 | 263 | 395 | 3,232 | 999 | 1,945 | 2,823 | 429 | 4,373 | 402 | 230 |
| 1941 | 125 | 273 | 2,147 | 860 | 883 | 1,557 | 1,609 | 244 | 146 | 555 | 268 | 101 |
| 1942 | 109 | 127 | 1,366 | 828 | 1,426 | 3,886 | 755 | 461 | 275 | 256 | 514 | 618 |
| 1943 | 237 | 195 | 1,210 | 978 | 1,387 | 3,905 | 1,817 | 293 | 172 | 313 | 111 | 164 |
| 1944 | 89.8 | 346 | 542 | 1,202 | 2,221 | 3,268 | 4,974 | 2,078 | 334 | 167 | 478 | 218 |
| 1945 | 130 | 296 | 878 | 1,404 | 3,436 | 2,416 | 1,793 | 934 | 709 | 603 | 280 | 154 |
| 1946 | 244 | 201 | 546 | 2,269 | 3,556 | 2,284 | 384 | 2,197 | 506 | 590 | 808 | 196 |
| 1947 | 130 | 685 | 703 | 5,465 | 801 | 2,607 | 4,044 | 1,327 | 519 | 179 | 143 | 132 |
| 1948 | 133 | 347 | 976 | 1,043 | 3,272 | 4,500 | 1,520 | 312 | 185 | 185 | 280 | 523 |
| 1949 | 312 | 3,651 | 2,524 | 2,803 | 3,592 | 3,329 | 2,944 | 2,029 | 1,003 | 1,237 | 906 | 858 |
| 1950 | 284 | 228 | 574 | 5,343 | 5,158 | 2,203 | 600 | 1,539 | 446 | 417 | 297 | 231 |
| 1951 | 141 | 189 | 481 | 1,058 | 2,752 | 2,825 | 1,978 | 255 | 267 | 260 | 114 | 150 |
| 1952 | 107 | 127 | 738 | 404 | 1,008 | 1,135 | 771 | 836 | 184 | 102 | 113 | 78.7 |
| 1953 | 67.8 | 103 | 177 | 1,085 | 2,902 | 2,283 | 1,522 | 4,744 | 247 | 660 | 463 | 224 |
| 1954 | 95.8 | 131 | 938 | 943 | 615 | 1,202 | 1,040 | 909 | 145 | 237 | 82.4 | 69.2 |

TABLE 2.—*Monthly mean discharge, in cubic feet per second, of Bowie Creek near Hattiesburg, Miss., drainage area, 304 square miles*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 178 | 138 | 201 | 217 | 1,035 | 435 | 523 | 446 | 279 | 1,553 | 170 | 136 |
| 1941 | 127 | 158 | 711 | 400 | 289 | 792 | 589 | 191 | 181 | 252 | 150 | 130 |
| 1942 | 132 | 210 | 853 | 519 | 483 | 1,014 | 264 | 510 | 256 | 180 | 648 | 368 |
| 1943 | 349 | 258 | 730 | 507 | 839 | 2,086 | 826 | 252 | 201 | 282 | 172 | 457 |
| 1944 | 146 | 275 | 423 | 776 | 979 | 1,353 | 880 | 801 | 254 | 206 | 353 | 251 |
| 1945 | 174 | 441 | 572 | 656 | 787 | 473 | 823 | 317 | 503 | 315 | 203 | 165 |
| 1946 | 273 | 261 | 706 | 941 | 823 | 904 | 241 | 852 | 396 | 504 | 287 | 202 |
| 1947 | 174 | 556 | 402 | 1,867 | 476 | 1,386 | 1,769 | 884 | 344 | 182 | 199 | 227 |
| 1948 | 187 | 379 | 886 | 646 | 893 | 1,699 | 725 | 249 | 211 | 205 | 229 | 254 |
| 1949 | 177 | 1,343 | 900 | 730 | 1,552 | 1,533 | 1,119 | 709 | 431 | 825 | 660 | 495 |
| 1950 | 288 | 235 | 355 | 1,157 | 1,644 | 846 | 376 | 578 | 263 | 231 | 266 | 272 |
| 1951 | 203 | 184 | 344 | 417 | 890 | 941 | 600 | 182 | 177 | 197 | 151 | 189 |
| 1952 | 136 | 159 | 423 | 265 | 522 | 463 | 377 | 317 | 146 | 144 | 134 | 120 |
| 1953 | 111 | 137 | 183 | 350 | 810 | 730 | 678 | 1,596 | 176 | 201 | 158 | 128 |
| 1954 | 123 | 158 | 812 | 319 | 259 | 286 | 264 | 232 | 139 | 161 | 112 | 99.1 |

8. The area between the upper and lower curves of step 7 contains two-thirds of the plotted points. By definition this band has a vertical spread of about two standard errors of estimate. Measure the vertical distance (using the 20 scale on 5-inch log cycle paper) between the upper and lower curves. One-half of this distance is an estimate of the unadjusted standard error of estimate in log units. If the log cycle of the paper used does not fit an engineers' scale, compute the standard error of estimate in log units by looking up the logarithm of the quotient of a discharge on the upper curve divided by the discharge vertically below ʋᴧ ᵢʰe lower curve or by converting the scaled distance to a decimal ratio of the log cycle distance.

9. Adjust the standard error of estimate by multiplying it by $\sqrt{\dfrac{N}{N-l.d.f.}}$. Six degrees of freedom (*l.d.f.*) are lost for the curve described in step 6. The lost degrees of freedom are determined on the assumption that two points are needed to locate the upper line, two points to locate the lower line, and two points to locate the transition curve. A difference of several *l.d.f.* makes little difference in the result unless $N$ (the number of pairs of items) is small.

10. Convert the standard error of estimate first to a ratio and then to a percentage by looking up the antilog and the anticolog. A slide rule is convenient for making this conversion. For example: Assume that one-half of the distance between the curves is 0.10 log units. Set 0.10 on the $L$ (log) scale and read 1.26 on the $D$ scale. The ratio is 1.26 or $+26$ percent. Set the colog 0.90 on the $L$ scale' and read the ratio 0.795 which corresponds to $-20.5$ percent on the $D$ scale. The operation may be performed with a single setting by alining the ends of the scales and reading 0.795 on the $CI$ scale.

11. Draw horizontal lines that exclude one-sixth of the points (in the $Y$ direction) above the top line and one-sixth of the points below the bottom line. One-half of the vertical distance between the lines multiplied by $\sqrt{\dfrac{N}{N-1}}=s_y$, the standard deviation of $Y$.

12. Find the index of correlation ($\rho$) by the formula $\rho=\sqrt{1-\left(\dfrac{S_e}{s_y}\right)^2}$. If the relation is a single straight line, this formula gives the coefficient of correlation ($r$).

### MULTIPLE CORRELATION

Multiple correlation is used to obtain better correlative estimates for the dependent variable than could be obtained by simple correlation alone. Multiple correlation is the use of more than two variables. The additional variables can be records from additional gaging stations, seasonal factors, or precipitation data.

This discussion of multiple correlation is limited to the removal of seasonal effect, to the use of records from two gaging stations for estimating the discharge of a third station, and to the use of records from two gaging stations (located on the same stream) for estimating the discharge of a tributary that enters a stream between the stations.

## REMOVING SEASONAL EFFECT

The relation between the runoff from one basin and that from another basin may vary with the season of the year. This variation is called seasonal effect or monthly effect. The most common example of seasonal effect occurs when two streams receive runoff from snowmelt. A southern exposure of one basin might cause the bulk of the snowpack to melt in April, whereas a northern exposure of an adjacent basin might delay snowmelt until May. Differences in altitude may have a similar effect. During months when snowmelt is not a factor, the relation between the basins might remain fairly constant.

Seasonal effect might also be due to differences in vegetation. During the growing season profuse vegetation in one basin might consume large amounts of water so that the relation for that season differs from that of other seasons. Also, the use of large amounts of water from one stream during the summer could produce a seasonal effect.

Seasonal effect is indicated by the scatter diagram when the points for the same calendar month show à tendency to plot on the same side of and about the same distance from the curve. If the hydrologic differences are sufficiently stable from year to year, their average effect can be determined and removed by the following method, using two stations in western Washington:

1. Plot the selected unit (generally monthly) of discharge of station $E$ (table 3) against the concurrent unit of station $F$ (table 4) on logarithmic paper with the independent variable on the $X$ axis (see pl. 1$A$). The symbols recommended aid in detecting seasonal effect and are essential to the application of the graphical method, but colors can be used instead of the ticks used in plate 1$A$.

2. On any paper, draw about 16 vertical lines at equal intervals (pl. 1$B$) for the seasonal adjustment curve. Each line represents a month, and the year overlaps about two months on each end. Draw a horizontal line to represent zero residual. Graduations on the seasonal adjustment curve are not used for plotting.

3. Draw a preliminary line of relation. The actual position of the line has little effect on the results but in some correlations the size of the adjustments are reduced by allowing for the expected adjusted position of the points.

TABLE 3.—*Monthly mean discharge, in cubic feet per second, of White River at Greenwater, Wash., drainage area, 216 square miles*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1947 | 468 | 710 | 1,462 | 841 | 899 | 735 | 971 | 1,538 | 1,247 | 839 | 576 | 481 |
| 1948 | 1,114 | 1,140 | 800 | 739 | 581 | 499 | 690 | 1,592 | 2,652 | 1,088 | 691 | 467 |
| 1949 | 466 | 643 | 649 | 432 | 581 | 762 | 1,121 | 2,280 | 1,762 | 1,128 | 666 | 473 |
| 1950 | 414 | 960 | 769 | 655 | 694 | 924 | 892 | 1,570 | 2,660 | 1,874 | 862 | 552 |
| 1951 | 818 | 1,448 | 1,536 | 727 | 1,313 | 624 | 1,169 | 1,504 | 1,484 | 1,129 | 638 | 422 |
| 1952 | 556 | 544 | 501 | 330 | 642 | 428 | 970 | 1,407 | 1,249 | 1,004 | 625 | 433 |
| 1953 | 341 | 227 | 243 | 1,043 | 978 | 416 | 665 | 1,149 | 1,302 | 1,404 | 715 | 500 |

TABLE 4.—*Monthly mean discharge, in cubic feet per second, of Greenwater River at Greenwater, Wash., drainage area, 73.9 square miles*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1947 | 77.2 | 163 | 533 | 263 | 245 | 270 | 341 | 408 | 255 | 111 | 58.1 | 47.1 |
| 1948 | 181 | 504 | 320 | 225 | 177 | 135 | 241 | 605 | 747 | 184 | 92.3 | 72.1 |
| 1949 | 86.3 | 156 | 159 | 102 | 171 | 282 | 433 | 833 | 425 | 160 | 67.3 | 45.3 |
| 1950 | 64.7 | 180 | 226 | 203 | 240 | 286 | 291 | 575 | 900 | 371 | 102 | 57.2 |
| 1951 | 94.0 | 328 | 400 | 235 | 327 | 175 | 411 | 528 | 319 | 114 | 57.5 | 43.1 |
| 1952 | 69.5 | 97.2 | 118 | 64.4 | 158 | 118 | 323 | 391 | 218 | 119 | 55.5 | 38.3 |
| 1953 | 32.1 | 32.9 | 35.0 | 276 | 297 | 124 | 206 | 364 | 357 | 191 | 71.9 | 47.7 |

4. With a pair of dividers (or a scale), measure the vertical distance from the preliminary relation line to each October point. Transfer these distances to the October line of the seasonal adjustment curve. Mark the point above or below the zero residual line to match the relation of the plotted point to the preliminary relation line. Repeat this procedure with the points for each month. The advantage of using monthly symbols in step 1 is now evident.

5. Draw horizontal lines on each monthly line so as to exclude one-sixth of the points above the top line and one-sixth of the points below the bottom line. The distance between the lines is, by definition, two standard deviations.

6. Graphically determine the median value for each month by averaging the middle two points, if the number of points is even, or by averaging the middle three points, giving the midpoint a weight of two, if the number of points is odd. Transfer the median and standard deviation lines of the last two months to the opposite end to facilitate drawing the seasonal adjustment curve.

7. Draw a smooth curve to average the median values, but keep the curve within the standard deviation lines. The curve should contain as few bends as possible; for example, the curve from January through April (pl. 1$B$). A mathematical method of drawing this is described in Water-Supply Paper 1541–B, "Double-Mass Curves," by Searcy and Hardison under the section "Fitting curves to cyclic data," by Langbein.

8. The distance from the zero residual line to the curve is the adjustment for the month. Set the adjustment for October on the dividers and in plate 1$A$ apply this adjustment to the points for all Octobers. This adjustment will result in moving the observed points down if the indicated adjustment lies above the zero residual line and moving them up if it lies below. Repeat this procedure for all months. The adjusted value is marked by a solid circle, usually on the same sheet with the preliminary line. To reduce clutter, the preliminary regression line could be erased. For clarity, the adjusted points in the example have been plotted as a separate graph. (See pl. 1.)

9. Repeat the graphical correlation procedure for the two variables given in the section "Simple correlation," using the adjusted position of the points.

In computing $S_e$, 6 degrees of freedom are lost: 2 for the relation line and 4 for the adjustment curve. The degrees of freedom lost in the adjustment curve are determined by spotting the points necessary

to approximately reproduce the curve.   In plate 1$B$ the curve could be redrawn from points on the curve on the October, February, June, and August lines.

Compute the index of correlation ($\rho$) by formula 5 and, for comparison, compute an adjusted index using the standard deviation of the $Y$ values as determined by averaging the standard deviations of the monthly discharges around their respective mean monthly discharges.   The procedure is explained in the section "Concepts and definitions" under "Coefficient or index of correlation for gaging-station records."   In the illustrative problem (pl. 1) the adjusted index of correlation is 0.94 as compared with an index of 0.98 computed from the standard deviation of the $Y$ values around the mean of all the $Y$ values.

For added refinement, the line of regression obtained from the adjusted points (pl. 1) can be used as the preliminary line of regression in step 3 and steps 4–9 repeated.

*Use.*—To use the adjusted relation curve for estimating a value for Greenwater River corresponding to a given discharge of White River—

(a) Enter the curve of step 9 (pl. 1) with the value of White River and obtain the $Y_c$ value of Greenwater River.

(b) Obtain the appropriate adjustment to $Y_c$ from the seasonal adjustment curve (pl. 1$B$).

(c) Add, graphically, the plus or minus adjustment to $Y_c$ to obtain the estimated actual point for Greenwater River.   An alternate procedure is to use the adjustments as ratios and multiply the $Y_c$ discharge by the appropriate monthly ratio.

## GRAPHICAL CORRELATION WITH THREE OR MORE VARIABLES

When the standard error of estimate of simple correlation is too great, using more stations in the correlation will often reduce the magnitude of the standard error of estimate.   The general case will be described by an example, correlating Bowie Creek near Hattiesburg, Miss. (station $C$), as the dependent variable, with Leaf River near Collins, Miss. (station $A$), and Bogue Chitto at Franklinton, La. (station $B$), as the two independent variables.   (See fig. 9.)

On figure 9, station $D$ is shown on the Leaf River at Hattiesburg, Miss., downstream from the confluence of Leaf River and Bowie Creek.   Bowie Creek supplies much of the difference in flow between gaging stations $D$ and $A$.   Thus, a correlation of the Bowie Creek

discharge with the difference in discharges between the Leaf River stations $D$ and $A$ (see section "Special case") would be expected to give better results than a correlation of the Bowie Creek discharge with the discharge from two streams in nearby basins. (See section "General case.") This is demonstrated by comparing the results obtained from the general and the special case.

Other combinations of records from the same stream system could be used; for example, discharge for Leaf River at Hattiesburg could be correlated with the sum of the Collins station and Bowie Creek discharges, or the sum of the Collins station and Bowie Creek discharges could be averaged with those of a downstream station, Leaf River at McClain, and the average used to correlate with the Hattiesburg discharge. Addition of other stations does not necessarily improve the correlation. On this particular stream, averaging the records for the downstream station at McClain with the sum of records of the two upstream stations would probably give poorer results than using the sum of the two upstream stations alone. Actual correlation by both combinations is the only way to tell which combination gives the best result.

### GENERAL CASE

Records of two long-term stations, $A$ and $B$, are correlated with the record of a short-term station, $C$, as follows:

1. Correlate station $A$, (Leaf River near Collins, Miss., table 1) as abscissa, with station $C$ (Bowie Creek near Hattiesburg, Miss., table 2). This is the correlation of figure 10.

2. Obtain the curve value $(Y_c)$ for station $C$ (Bowie Creek) corresponding to each value of station $A$ (Leaf River near Collins, table 1). (See table 5.)

3. Compute the station $C$ (Bowie Creek) ratio, $Y:Y_c$, from tables 2 and 5. (See table 6.)

4. Correlate station $A$ (Leaf River near Collins, table 1) as abscissa with station $B$ (Bogue Chitto at Franklinton, La., table 7). (See fig. 11.)

5. Obtain the curve value $(Y_c)$ for station $B$ (Bogue Chitto) corresponding to each value of station $A$ (Leaf River near Collins, table 1). (See table 8.)
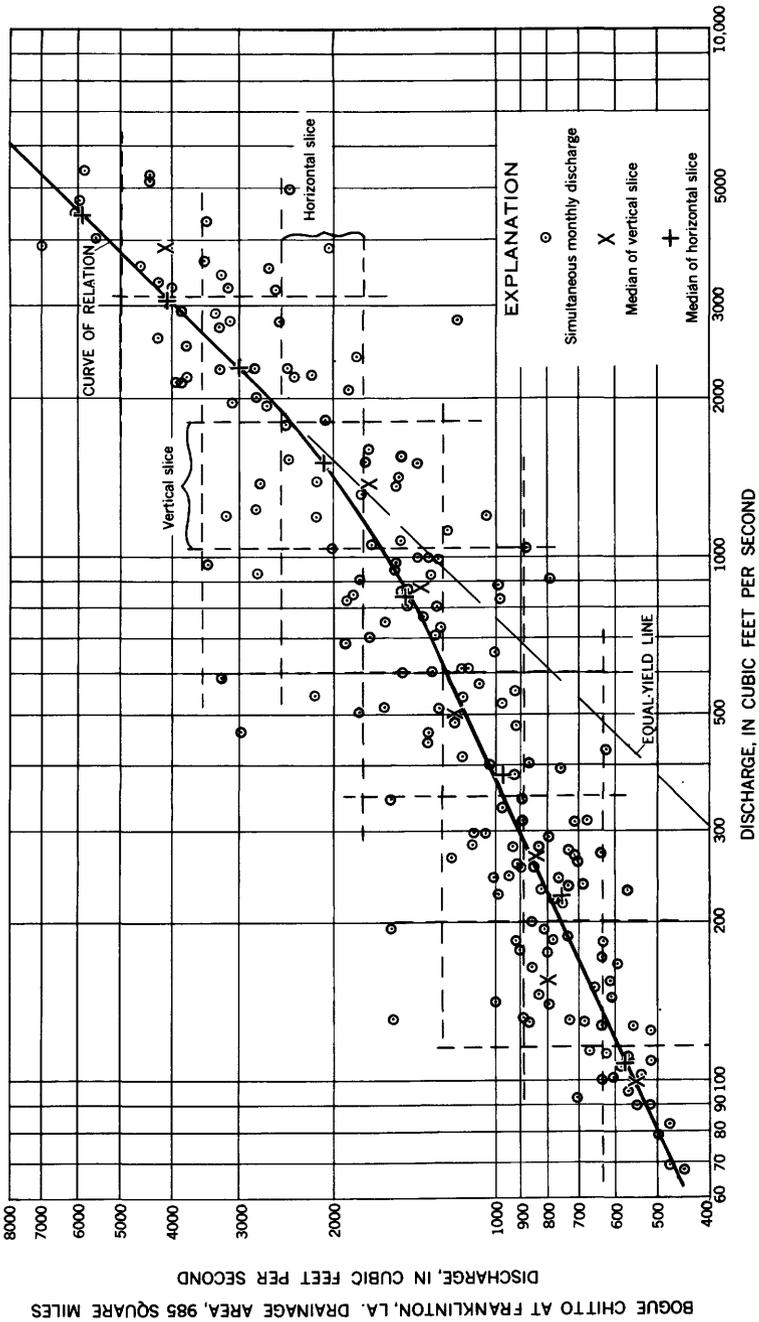
FIGURE 11.—Correlation of monthly discharge, Leaf River and Bogue Chitto, water years 1940–54.

TABLE 5.—*Curve values* ($Y_c$) *of monthly discharge, in cubic feet per second, of Bowie Creek near Hattiesburg, Miss., drainage area, 304 square miles, from figure 10*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 127 | 125 | 235 | 299 | 1,210 | 512 | 792 | 1,070 | 313 | 1,630 | 301 | 217 |
| 1941 | 152 | 240 | 856 | 473 | 481 | 680 | 698 | 225 | 166 | 366 | 238 | 134 |
| 1942 | 140 | 153 | 622 | 462 | 640 | 1,440 | 439 | 329 | 242 | 232 | 350 | 389 |
| 1943 | 221 | 198 | 578 | 508 | 630 | 1,460 | 758 | 250 | 183 | 261 | 141 | 178 |
| 1944 | 125 | 276 | 360 | 575 | 878 | 1,220 | 1,850 | 840 | 271 | 180 | 335 | 211 |
| 1945 | 155 | 252 | 479 | 632 | 1,280 | 938 | 750 | 497 | 422 | 385 | 244 | 172 |
| 1946 | 225 | 201 | 362 | 896 | 1,330 | 898 | 295 | 870 | 347 | 379 | 457 | 197 |
| 1947 | 155 | 414 | 421 | 2,030 | 454 | 998 | 1,510 | 610 | 350 | 188 | 164 | 156 |
| 1948 | 158 | 276 | 508 | 530 | 1,220 | 1,670 | 670 | 260 | 192 | 192 | 244 | 353 |
| 1949 | 260 | 1,360 | 968 | 1,060 | 1,340 | 1,240 | 1,110 | 820 | 518 | 588 | 489 | 470 |
| 1950 | 246 | 216 | 372 | 1,990 | 1,920 | 870 | 382 | 678 | 321 | 309 | 252 | 218 |
| 1951 | 163 | 194 | 338 | 534 | 1,040 | 1,070 | 800 | 232 | 237 | 234 | 144 | 169 |
| 1952 | 137 | 153 | 430 | 302 | 518 | 555 | 443 | 464 | 191 | 134 | 143 | 116 |
| 1953 | 106 | 136 | 186 | 542 | 1,090 | 898 | 668 | 1,770 | 227 | 403 | 329 | 214 |
| 1954 | 129 | 156 | 498 | 500 | 389 | 575 | 530 | 489 | 166 | 222 | 118 | 107 |

TABLE 6.—Y:Y_e ratio for Bowie Creek near Hattiesburg, Miss.

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 1.402 | 1.104 | 0.855 | 0.726 | 0.855 | 0.850 | 0.660 | 0.417 | 0.891 | 0.953 | 0.565 | 0.627 |
| 1941 | .836 | .658 | .831 | .846 | .601 | 1.165 | .844 | .849 | 1.090 | .689 | .630 | .970 |
| 1942 | .943 | 1.373 | 1.371 | 1.123 | .755 | .704 | .601 | 1.550 | 1.058 | .776 | 1.851 | .946 |
| 1943 | 1.579 | 1.303 | 1.263 | .998 | 1.332 | 1.429 | 1.090 | 1.008 | 1.098 | 1.080 | 1.220 | 2.567 |
| 1944 | 1.168 | .996 | 1.175 | 1.350 | 1.115 | 1.109 | .476 | .954 | .937 | 1.144 | 1.054 | 1.190 |
| 1945 | 1.123 | .750 | 1.194 | 1.038 | .615 | .504 | 1.097 | .638 | 1.192 | .818 | .832 | .959 |
| 1946 | 1.213 | 1.299 | 1.950 | 1.050 | .619 | 1.007 | .817 | .979 | 1.141 | 1.330 | .628 | 1.025 |
| 1947 | 1.123 | 1.343 | .955 | .920 | 1.048 | 1.389 | 1.172 | 1.449 | .983 | .968 | 1.213 | 1.455 |
| 1948 | 1.184 | 1.373 | .744 | 1.219 | .732 | 1.017 | 1.082 | .958 | 1.099 | 1.068 | .939 | .720 |
| 1949 | .681 | .988 | .930 | .689 | 1.158 | 1.236 | 1.008 | .865 | .832 | 1.403 | 1.350 | 1.053 |
| 1950 | 1.171 | 1.088 | .954 | .581 | .856 | .972 | .984 | .853 | .819 | .748 | 1.056 | 1.248 |
| 1951 | 1.245 | .948 | 1.018 | .781 | .856 | .879 | .750 | .784 | .747 | .842 | 1.049 | 1.118 |
| 1952 | .993 | 1.039 | .984 | .877 | 1.008 | .834 | .851 | .683 | .764 | 1.075 | .937 | 1.034 |
| 1953 | 1.047 | 1.007 | .984 | .646 | .743 | .813 | 1.015 | .902 | .775 | .499 | .480 | .598 |
| 1954 | .953 | 1.013 | 1.631 | .638 | .666 | .497 | .498 | .474 | .837 | .725 | .949 | .926 |

TABLE 7.—Monthly mean discharge, in cubic feet per second, of Bogue Chitto at Franklinton, La., drainage area, 985 square miles

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 700 | 518 | 701 | 762 | 2,554 | 1,277 | 2,655 | 1,178 | 624 | 3,433 | 1,033 | 569 |
| 1941 | 514 | 639 | 3,817 | 1,497 | 987 | 1,498 | 1,728 | 761 | 830 | 914 | 716 | 637 |
| 1942 | 514 | 554 | 1,532 | 1,880 | 1,515 | 2,044 | 1,609 | 1,329 | 731 | 853 | 1,271 | 1,131 |
| 1943 | 731 | 816 | 3,179 | 1,533 | 2,743 | 6,999 | 2,066 | 796 | 634 | 675 | 566 | 851 |
| 1944 | 547 | 896 | 1,148 | 2,144 | 2,197 | 4,002 | 2,452 | 1,871 | 973 | 591 | 919 | 750 |
| 1945 | 681 | 1,043 | 1,451 | 3,264 | 3,227 | 1,813 | 923 | 1,317 | 1,290 | 1,306 | 922 | 608 |
| 1946 | 1,005 | 857 | 2,164 | 5,829 | 2,628 | 2,802 | 5,562 | 2,354 | 1,790 | 3,213 | 1,278 | 1,561 |
| 1947 | 865 | 1,901 | 1,700 | 1,999 | 1,463 | 4,219 | 1,400 | 1,768 | 1,608 | 905 | 1,003 | 1,554 |
| 1948 | 883 | 1,560 | 3,409 | 2,505 | 3,138 | 6,015 | 3,822 | 896 | 779 | 916 | 830 | 971 |
| 1949 | 715 | 3,479 | 3,735 | 4,402 | 4,574 | 4,206 | 3,495 | 2,783 | 1,330 | 2,788 | 1,787 | 1,843 |
| 1950 | 1,100 | 986 | 1,074 | 1,696 | 4,399 | 3,730 | 3,090 | 2,419 | 1,335 | 1,148 | 1,094 | 821 |
| 1951 | 794 | 727 | 1,186 | 866 | 3,281 | 3,107 | 1,357 | 897 | 1,210 | 915 | 664 | 654 |
| 1952 | 582 | 638 | 1,258 | 1,509 | 1,401 | 1,232 | 1,738 | 973 | 629 | 603 | 624 | 499 |
| 1953 | 448 | 538 | 799 | 1,539 | 3,314 | 2,429 | 875 | 5,933 | 941 | 999 | 2,971 | 778 |
| 1954 | 567 | 729 | 2,769 | | 1,152 | 1,038 | | 791 | 607 | 685 | 473 | 474 |

6. Compute the station $B$ (Bogue Chitto) ratio, $Y:Y_c$, from tables 7 and 8.    (See table 9.)

7. On logarithmic paper, plot the ratios (from step 6) representing station $B$ deviations as the abscissa against ratios (from step 3) representing station $C$ deviations as the ordinate.    (See fig. 12.)    Draw a curve of relation, remembering that the sum of the residuals should equal zero; that the residuals in any section of the line should equal zero; and that the residuals should be minimized.    In figure 12 the position of a few low points suggests a hook in the line.    However, the evidence is insufficient to warrant departing from a straight line.

8. Select $Y:Y_c$ ratios, from the curve of step 7 (fig. 12), for station $C$ (Bowie Creek) corresponding to the $Y:Y_c$ ratios for station $B$ (table 8).    These data are tabulated in table 10.

9. Adjust station $C$ discharges (table 2) by dividing by the ratios (table 10) selected from the curve.    The adjusted discharges are tabulated in table 11.
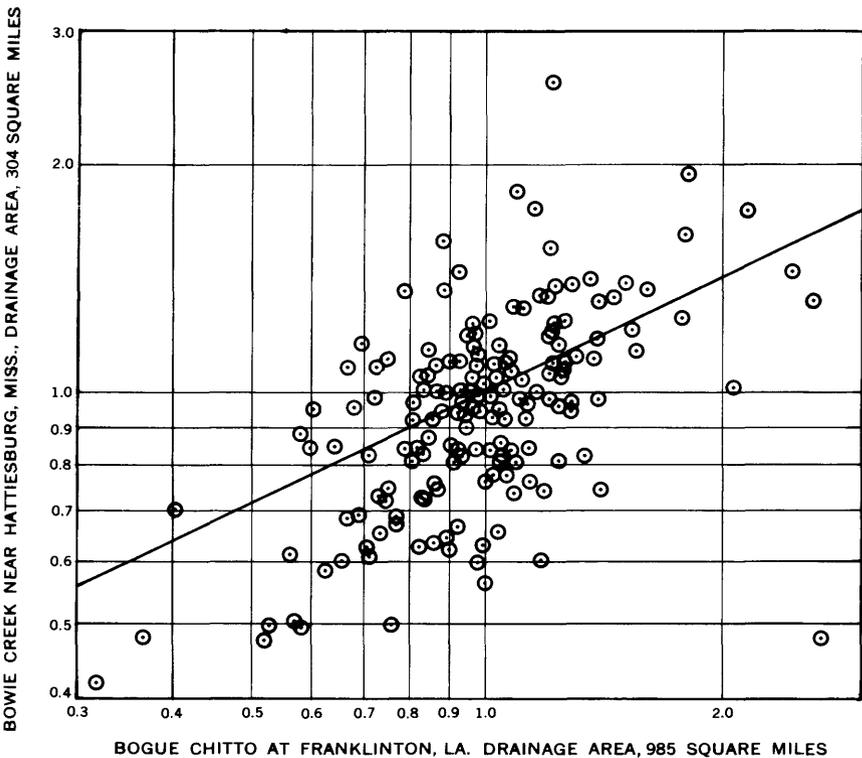


FIGURE 12.—Ratio $Y:Y_c$ Bogue Chitto and Bowie Creek, water years 1940–54.

TABLE 8.—Curve values (Y_c) of monthly discharge, in cubic feet per second, of Bogue Chitto at Franklinton, La., drainage area, 985 square miles, from figure 11

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 538 | 528 | 856 | 1,020 | 4,220 | 1,580 | 2,530 | 3,680 | 1,060 | 5,700 | 1,030 | 802 |
| 1941 | 612 | 865 | 2,820 | 1,460 | 1,490 | 2,140 | 2,180 | 824 | 659 | 1,190 | 860 | 557 |
| 1942 | 578 | 618 | 1,940 | 1,430 | 2,000 | 5,080 | 1,360 | 1,090 | 870 | 842 | 1,150 | 1,250 |
| 1943 | 816 | 748 | 1,780 | 1,570 | 1,960 | 5,120 | 2,400 | 896 | 702 | 920 | 580 | 690 |
| 1944 | 527 | 960 | 1,170 | 1,780 | 2,500 | 4,280 | 6,500 | 2,720 | 948 | 698 | 1,110 | 785 |
| 1945 | 626 | 899 | 1,480 | 980 | 4,500 | 3,160 | 2,370 | 1,530 | 1,320 | 1,240 | 878 | 676 |
| 1946 | 825 | 758 | 1,180 | 2,980 | 4,640 | 3,000 | 1,010 | 2,880 | 1,140 | 1,220 | 1,420 | 747 |
| 1947 | 626 | 1,300 | 1,320 | 7,160 | 1,410 | 3,410 | 5,300 | 1,900 | 1,150 | 720 | 650 | 628 |
| 1948 | 630 | 960 | 1,570 | 1,640 | 4,270 | 5,880 | 2,090 | 920 | 728 | 728 | 878 | 1,160 |
| 1949 | 920 | 4,780 | 3,290 | 3,670 | 4,700 | 4,360 | 3,820 | 2,640 | 1,580 | 1,820 | 1,510 | 1,460 |
| 1950 | 880 | 800 | 1,210 | 7,000 | 6,780 | 2,880 | 1,230 | 2,120 | 1,070 | 1,050 | 900 | 802 |
| 1951 | 645 | 738 | 1,120 | 1,650 | 3,600 | 3,680 | 2,580 | 840 | 860 | 850 | 590 | 665 |
| 1952 | 570 | 618 | 1,350 | 1,030 | 1,600 | 1,720 | 1,380 | 1,440 | 728 | 559 | 587 | 497 |
| 1953 | 462 | 561 | 718 | 1,680 | 3,800 | 3,000 | 2,090 | 6,200 | 828 | 1,280 | 1,100 | 792 |
| 1954 | 541 | 625 | 1,530 | 1,540 | 1,250 | 1,780 | 1,640 | 1,510 | 658 | 818 | 509 | 469 |

TABLE 9.—Y : Y_c ratio for Bogue Chitto at Franklinton, La.

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 1.301 | 0.981 | 0.819 | 0.747 | 0.605 | 0.808 | 1.049 | 0.320 | 0.589 | 0.602 | 1.003 | 0.709 |
| 1941 | .840 | .739 | 1.354 | 1.025 | .662 | .700 | .793 | .924 | 1.259 | .768 | .833 | 1.144 |
| 1942 | .889 | .896 | .790 | 1.315 | .758 | .402 | 1.183 | 1.219 | .840 | 1.013 | 1.105 | .905 |
| 1943 | .896 | 1.091 | 1.786 | .976 | 1.399 | 1.367 | .861 | .888 | .903 | .734 | .976 | 1.233 |
| 1944 | 1.038 | .933 | .981 | 1.207 | .758 | .935 | .370 | .688 | 1.026 | .847 | .828 | .955 |
| 1945 | 1.088 | 1.160 | .980 | 1.083 | .717 | .574 | 1.035 | .861 | .977 | 1.053 | 1.050 | .899 |
| 1946 | 1.218 | 1.131 | 1.834 | 1.095 | .566 | .934 | .914 | .817 | 1.570 | 2.634 | .900 | 2.090 |
| 1947 | 1.382 | 1.462 | 1.288 | .814 | 1.038 | 1.237 | 1.049 | .931 | 1.398 | 1.257 | 1.543 | 2.475 |
| 1948 | 1.402 | 1.625 | 2.171 | 1.219 | .735 | 1.023 | .670 | .974 | 1.070 | 1.258 | .945 | .837 |
| 1949 | .777 | .728 | 1.135 | .683 | .973 | .965 | 1.001 | 1.054 | .842 | 1.532 | 1.183 | 1.262 |
| 1950 | 1.250 | 1.232 | .888 | .629 | .649 | 1.295 | 1.215 | 1.141 | 1.248 | 1.093 | 1.216 | 1.024 |
| 1951 | 1.231 | .985 | 1.059 | 1.028 | .911 | .844 | 1.198 | 1.068 | 1.407 | 1.076 | 1.125 | .983 |
| 1952 | 1.021 | 1.032 | .932 | .841 | .876 | .716 | .983 | .676 | .864 | 1.079 | 1.063 | 1.004 |
| 1953 | .970 | .959 | 1.113 | .898 | .872 | .810 | .832 | .957 | 1.136 | .780 | 2.701 | .982 |
| 1954 | 1.048 | 1.166 | 1.810 | .999 | .922 | .583 | .534 | .524 | .922 | .837 | .929 | 1.011 |

TABLE 10.—$Y:Y_c$ ratio for Bowie Creek near Hattiesburg, Miss., from figure 12

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 1.14 | 0.998 | 0.910 | 0.870 | 0.782 | 0.900 | 1.03 | 0.570 | 0.770 | 0.748 | 1.00 | 0.850 |
| 1941 | .92 | .862 | 1.17 | 1.02 | .819 | .840 | .897 | .962 | 1.13 | .88 | .920 | 1.07 |
| 1942 | .948 | .958 | .897 | 1.15 | .872 | .640 | 1.09 | 1.10 | .92 | 1.00 | 1.06 | .958 |
| 1943 | .958 | 1.11 | 1.34 | .996 | 1.19 | 1.17 | .937 | .948 | .958 | .86 | .99 | 1.11 |
| 1944 | 1.02 | .966 | .998 | 1.04 | .872 | .965 | .614 | .835 | 1.02 | .922 | .917 | .98 |
| 1945 | 1.06 | 1.08 | .998 | 1.04 | .850 | .760 | 1.02 | .930 | .995 | 1.03 | 1.03 | .956 |
| 1946 | 1.11 | 1.07 | 1.35 | 1.05 | .759 | .965 | .958 | .905 | 1.26 | 1.62 | .96 | 1.45 |
| 1947 | 1.18 | 1.21 | 1.14 | .902 | 1.03 | 1.12 | 1.03 | .962 | 1.19 | 1.13 | 1.25 | 1.57 |
| 1948 | 1.19 | 1.28 | 1.47 | 1.11 | .86 | 1.02 | .838 | .996 | 1.04 | 1.13 | .97 | .918 |
| 1949 | .882 | .859 | 1.07 | .834 | .992 | .99 | 1.00 | 1.03 | .922 | 1.24 | 1.09 | 1.12 |
| 1950 | 1.12 | 1.12 | .945 | .798 | .808 | 1.14 | 1.10 | 1.07 | 1.12 | 1.11 | 1.10 | 1.02 |
| 1951 | 1.11 | .998 | 1.03 | 1.02 | .960 | .92 | 1.10 | 1.04 | 1.19 | 1.10 | 1.07 | .992 |
| 1952 | 1.01 | 1.03 | .968 | .920 | .940 | .848 | .998 | .822 | .932 | 1.10 | 1.04 | 1.00 |
| 1953 | .995 | .985 | 1.06 | .956 | .938 | .900 | .918 | .980 | 1.07 | .888 | 1.64 | .99 |
| 1954 | 1.03 | 1.085 | 1.35 | 1.00 | .962 | .766 | .738 | .728 | .96 | .920 | .96 | 1.00 |

TABLE 11.—Monthly discharge of Bowie Creek near Hattiesburg, Miss., adjusted for Bogue Chitto

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 156 | 138 | 221 | 249 | 1,324 | 483 | 508 | 782 | 362 | 2,076 | 170 | 160 |
| 1941 | 138 | 183 | 608 | 392 | 353 | 943 | 657 | 199 | 160 | 286 | 163 | 121 |
| 1942 | 139 | 219 | 951 | 451 | 554 | 1,584 | 242 | 464 | 278 | 180 | 611 | 384 |
| 1943 | 364 | 232 | 545 | 509 | 705 | 1,783 | 882 | 266 | 210 | 328 | 174 | 412 |
| 1944 | 143 | 285 | 424 | 705 | 1,123 | 1,402 | 1,433 | 959 | 249 | 223 | 385 | 256 |
| 1945 | 164 | 408 | 573 | 631 | 926 | 622 | 807 | 341 | 506 | 306 | 197 | 173 |
| 1946 | 246 | 244 | 523 | 896 | 1,084 | 937 | 252 | 941 | 314 | 311 | 299 | 139 |
| 1947 | 147 | 460 | 353 | 2,070 | 462 | 1,666 | 1,717 | 919 | 289 | 161 | 159 | 145 |
| 1948 | 157 | 296 | 603 | 582 | 1,038 | 1,548 | 865 | 250 | 203 | 181 | 236 | 277 |
| 1949 | 201 | 1,563 | 841 | 875 | 1,565 | 742 | 1,119 | 688 | 467 | 665 | 606 | 442 |
| 1950 | 257 | 210 | 376 | 1,450 | 2,035 | 1,023 | 342 | 540 | 235 | 208 | 242 | 267 |
| 1951 | 183 | 184 | 334 | 409 | 927 | 546 | 545 | 175 | 149 | 179 | 141 | 191 |
| 1952 | 135 | 154 | 437 | 288 | 555 | 811 | 378 | 386 | 157 | 131 | 129 | 120 |
| 1953 | 112 | 139 | 173 | 366 | 864 | 373 | 739 | 1,629 | 164 | 226 | 96.3 | 129 |
| 1954 | 119 | 146 | 601 | 319 | 269 | | 358 | 319 | 145 | 175 | 117 | 99.1 |

10. Correlate adjusted station $C$ discharges (table 11) with station $A$ discharges (table 1). (See fig. 13.) The lost degrees of freedom are eight—six for the correlation plus two for the $Y:Y_c$ curve.

*Use.*—To use the relation curve for estimating station $C$, given the discharge for station $A$ and station $B$:

(a) Enter the curve of step 4 (fig. 11) with the value of station $A$ (Leaf River near Collins) and obtain the $Y_c$ value for station $B$ (Bogue Chitto).

(b) Compute $Y:Y_c$ for station $B$.

(c) Enter the ratio curve of step 7 (fig. 12) with the ratio of step (b) and read the $Y:Y_c$ value for station $C$ (Bowie Creek).

(d) Read the adjusted discharge of station $C$ (Bowie Creek) from the curve of step 10 (fig. 13).

(e) Multiply the adjusted discharge of station $C$ (step (d)) by the $Y:Y_c$ ratio from step (c) to obtain the discharge of station $C$.

### SPECIAL CASE

One of a variety of special cases is illustrated by the following example: Gaging stations are located on the main stem of the Leaf River above and below the tributary, for which correlative estimates are desired. (See fig. 9.) In such a situation, the difference between the discharge of the upstream station and of the downstream station would furnish the best independent variable for correlating with the dependent station record.

The monthly discharge of upstream station $A$, Leaf River near Collins (table 1), is subtracted from the records of downstream station $D$ at Hattiesburg (table 12), with the results tabulated in table 13.

The differences in monthly discharge (table 13) are correlated with the monthly discharges of Bowie Creek (table 2). The methods of simple correlation are used for this purpose. The results are shown in figure 14.

### SUMMARY

Gaging-station records are the best basis for estimating the magnitude and the distribution of future flows. Often, only a short-term record is available at the site of a proposed water resources development. Such a record is less likely to represent the future flow of a stream than a long-term record, but through correlation it can be adjusted to represent the long-term period of a nearby long-term station.
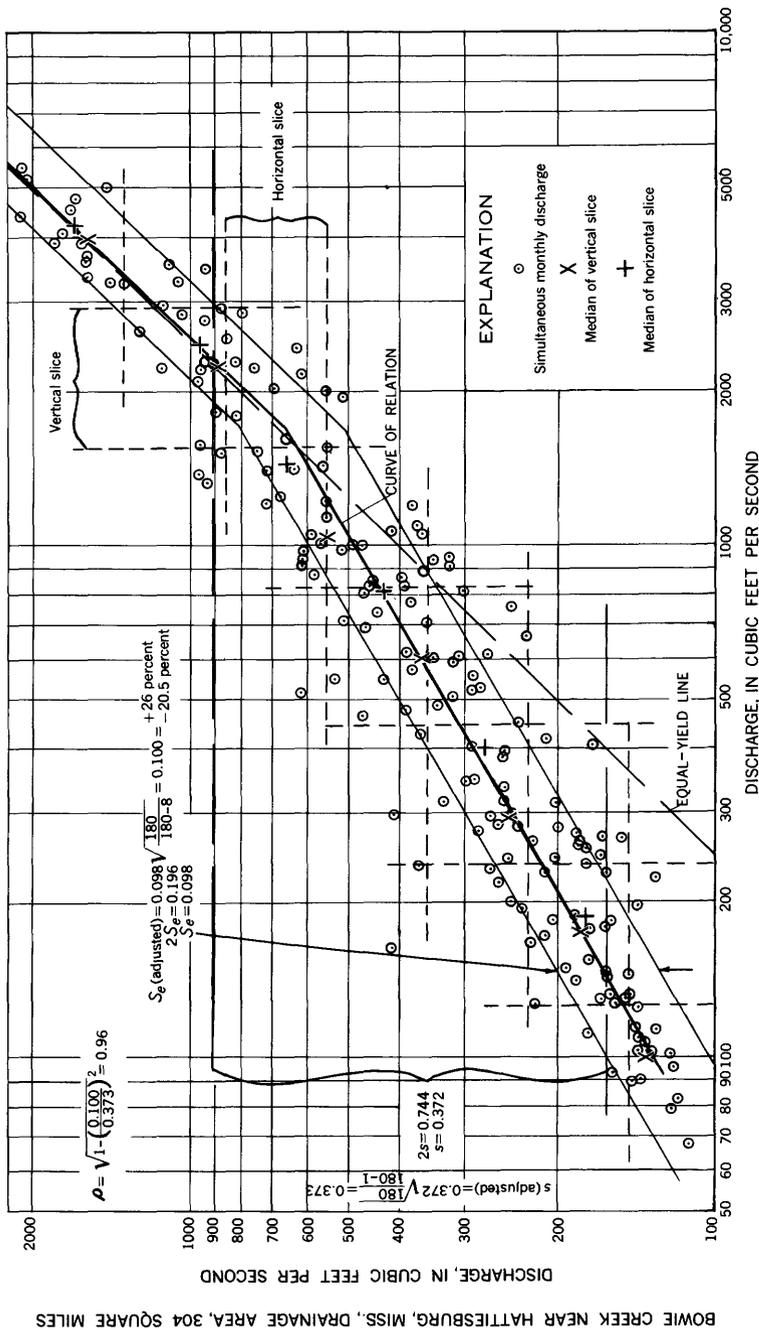
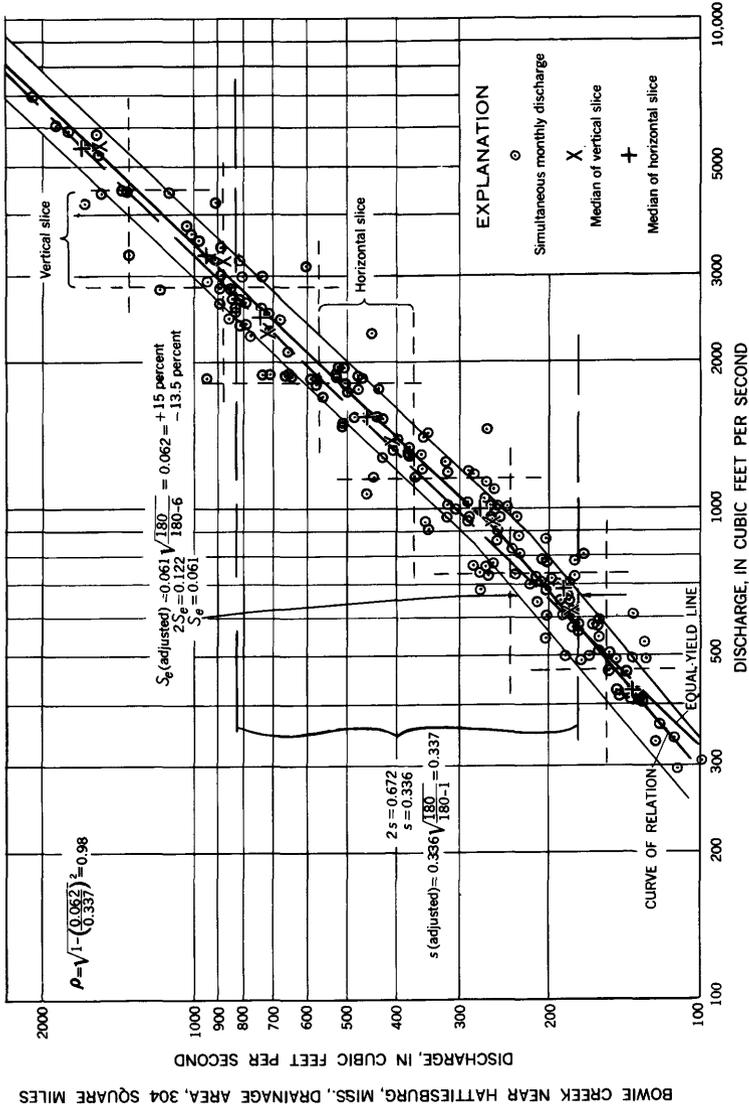FIGURE 13.—Correlation of monthly discharge, Leaf River and Bowie Creek (adjusted), water years 1940-54

FIGURE 14.—Correlation of monthly discharge, Bowie Creek and Leaf River basin between Hattiesburg and Collins, water years 1940–54.

TABLE 12.— *Monthly mean discharge, in cubic feet per second, of Leaf River at Hattiesburg, Miss., drainage area, 1,760 square miles*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 670 | 511 | 871 | 1,097 | 7,067 | 2,555 | 3,820 | 5,124 | 1,605 | 10,260 | 1,213 | 843 |
| 1941 | 622 | 870 | 4,484 | 2,198 | 1,831 | 4,222 | 3,467 | 948 | 801 | 1,422 | 777 | 509 |
| 1942 | 519 | 776 | 3,824 | 2,790 | 2,976 | 7,571 | 7,894 | 1,972 | 1,055 | 879 | 2,409 | 1,788 |
| 1943 | 1,182 | 1,153 | 3,110 | 2,934 | 4,098 | 10,910 | 4,404 | 1,195 | 855 | 1,087 | 604 | 1,246 |
| 1944 | 515 | 1,334 | 1,827 | 3,481 | 5,789 | 7,830 | 8,425 | 5,083 | 1,219 | 879 | 1,778 | 1,238 |
| 1945 | 691 | 1,471 | 2,741 | 3,520 | 5,834 | 4,188 | 4,339 | 2,138 | 2,200 | 1,634 | 1,062 | 655 |
| 1946 | 987 | 931 | 2,454 | 5,222 | 6,229 | 5,534 | 4,400 | 5,141 | 1,910 | 2,406 | 2,006 | 916 |
| 1947 | 713 | 2,396 | 2,033 | 11,570 | 6,695 | 7,121 | 10,040 | 4,548 | 1,959 | 900 | 842 | 942 |
| 1948 | 742 | 1,668 | 4,017 | 2,922 | 6,120 | 10,040 | 4,041 | 1,275 | 918 | 977 | 1,166 | 625 |
| 1949 | 1,046 | 6,985 | 6,740 | 5,401 | 8,900 | 7,792 | 7,428 | 4,565 | 2,788 | 3,727 | 2,796 | 2,611 |
| 1950 | 1,329 | 1,062 | 1,796 | 8,163 | 9,401 | 5,064 | 1,947 | 3,346 | 1,420 | 1,157 | 1,065 | 916 |
| 1951 | 689 | 690 | 1,396 | 2,392 | 5,416 | 4,700 | 5,129 | 967 | 1,054 | 982 | 581 | 808 |
| 1952 | 604 | 706 | 2,275 | 1,466 | 2,934 | 3,013 | 2,072 | 1,800 | 676 | 521 | 536 | 446 |
| 1953 | 366 | 514 | 846 | 2,500 | 5,304 | 5,293 | 3,990 | 10,230 | 983 | 1,535 | 1,012 | 761 |
| 1954 | 434 | 648 | 4,192 | 2,206 | 1,632 | 2,172 | 2,507 | 1,884 | 610 | 820 | 423 | 377 |

TABLE 13.— *Monthly discharge, in cubic feet per second, Leaf River at Hattiesburg, Miss., minus Leaf River near Collins, Miss.*

| Water year | October | November | December | January | February | March | April | May | June | July | August | September |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1940 | 576 | 421 | 608 | 702 | 3,835 | 1,556 | 1,875 | 2,301 | 1,176 | 5,887 | 811 | 613 |
| 1941 | 497 | 597 | 2,337 | 1,338 | 948 | 2,665 | 1,858 | 704 | 655 | 867 | 509 | 408 |
| 1942 | 410 | 649 | 2,458 | 1,962 | 1,550 | 3,685 | 1,139 | 1,511 | 780 | 623 | 1,895 | 1,170 |
| 1943 | 945 | 958 | 1,900 | 1,956 | 2,711 | 7,005 | 2,587 | 902 | 683 | 774 | 493 | 1,082 |
| 1944 | 425 | 988 | 1,285 | 2,279 | 3,568 | 4,562 | 3,451 | 3,005 | 885 | 712 | 1,300 | 1,020 |
| 1945 | 561 | 1,175 | 1,863 | 2,116 | 2,398 | 1,772 | 1,546 | 1,204 | 1,491 | 1,031 | 782 | 501 |
| 1946 | 743 | 730 | 1,908 | 2,953 | 2,673 | 3,250 | 2,546 | 2,944 | 1,404 | 1,816 | 1,198 | 720 |
| 1947 | 583 | 1,711 | 1,330 | 6,105 | 1,894 | 4,514 | 5,016 | 3,221 | 1,440 | 721 | 699 | 810 |
| 1948 | 609 | 1,321 | 3,041 | 1,879 | 2,848 | 4,540 | 5,996 | 963 | 733 | 792 | 886 | 1,102 |
| 1949 | 734 | 3,334 | 4,216 | 2,598 | 5,308 | 4,463 | 2,521 | 2,536 | 1,785 | 2,490 | 1,890 | 1,753 |
| 1950 | 1,045 | 334 | 1,222 | 2,820 | 4,243 | 2,861 | 4,484 | 1,807 | 974 | 740 | 768 | 685 |
| 1951 | 548 | 501 | 915 | 1,334 | 2,664 | 1,875 | 3,347 | 712 | 787 | 722 | 467 | 658 |
| 1952 | 497 | 579 | 1,537 | 1,062 | 1,926 | 1,878 | 3,151 | 964 | 492 | 419 | 423 | 367 |
| 1953 | 298 | 411 | 669 | 1,415 | 2,402 | 3,010 | 1,301 | 5,486 | 736 | 875 | 549 | 537 |
| 1954 | 338 | 517 | 3,254 | 1,263 | 1,017 | 970 | 2,468 | 975 | 465 | 583 | 341 | 308 |

Graphical correlation is a practical method of establishing a relation between a long-term record and a short-term record. The relation established for a concurrent short-term period is used to adjust the short-term record to that of the long-term period. When the relation between stations is good, the short-term record is improved by correlating it with records that are more than 25 percent longer than the short-term record; as the degree of relation decreases, increased length of the long-term record is required to improve the short-term record; and if the relation becomes poor no amount of long-term record can improve the short-term record. The degree of correlation is measured by the index or coefficient of correlation. (See p. 75.)

Simple correlation is the establishment of a relation between two variables (in this example gaging-station records); multiple correlation is the establishment of a relation between a dependent variable (the record to be estimated) and two or more independent variables (generally long-term records).

Graphical correlation of streamflow data is accomplished on logarithmic paper because the range in data can be plotted without undesirable condensation of the scale, because the relations between gaging-station records tend to plot as one or more straight lines, and because the residuals from logarithmic relations tend to be distributed according to the normal curve.

A relation exists between gaging-station records because of a common climate, and the shape of the relation is fixed by the hydrology of the drainage basins. Statistics furnishes only a useful method for defining an existing relation, it cannot provide the relation.

### LITERATURE CITED

Langbein, W. B., and Hardison, C. H., 1955, Extending streamflow data: Am. Soc. Civil Engineers Proc., Paper 826.
Somers, W. P., 1954, Statistical analysis of streamflow relationships, 22d Western Snow Conf., 1954: Salt Lake City, Proc., p. 26–32.

○

EXPLANATION

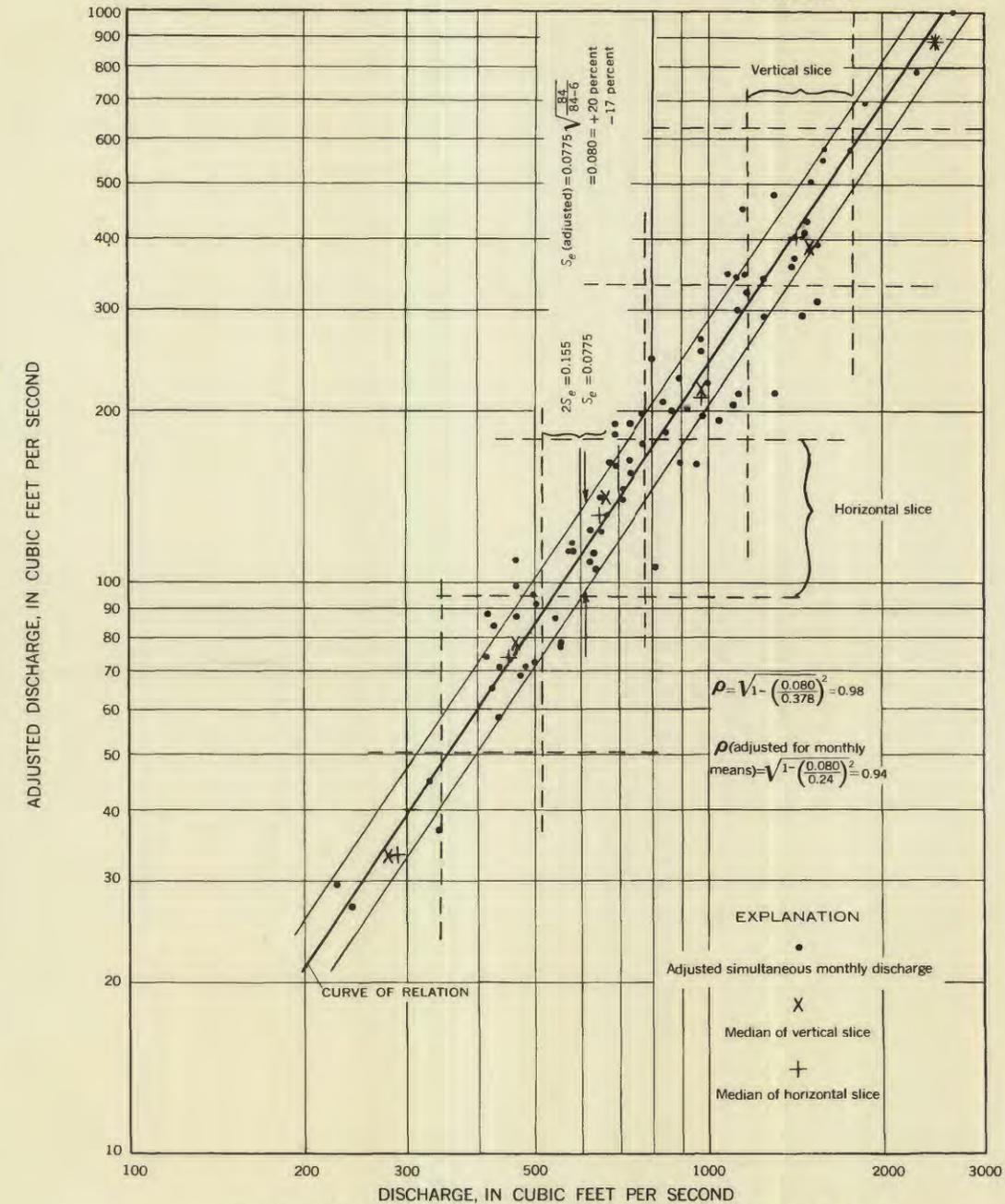| | | |
|---|---|---|
| Oct. | Nov. | Dec. | Jan. |
| Feb. | Mar. | Apr. | May |
| June | July | Aug. | Sept. |

+ Median

$2s = 0.753$
$s = 0.376$

$s(\text{adjusted}) = 0.376\sqrt{\frac{80}{84-1}} = 0.378$

0 Residual

2s

EQUAL-YIELD LINE

PRELIMINARY CURVE OF RELATION

$\frac{0.080}{X.24}$

2s

A. Observed Greenwater River discharge

B. Seasonal adjustment curve

GREENWATER RIVER AT GREENWATER, WASH., DRAINAGE AREA, 739 SQUARE MILES

DISCHARGE, IN CUBIC FEET PER SECOND

Aug. Sept. Oct. Nov. Dec. Jan. Feb. Mar. Apr. May June July Aug. Sept. Oct. Nov.

DISCHARGE, IN CUBIC FEET PER SECOND
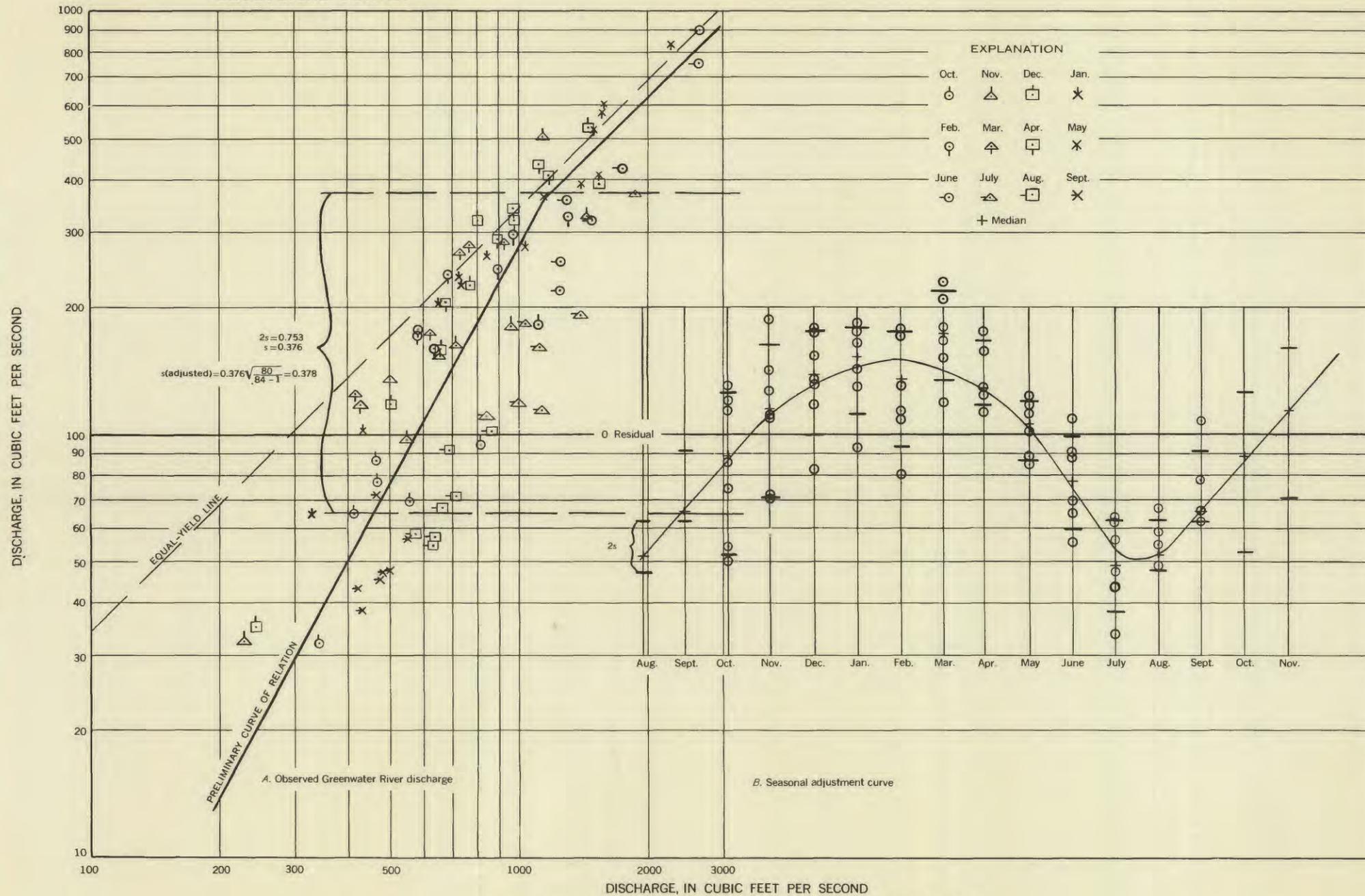
WHITE RIVER AT GREENWATER, WASH., DRAINAGE AREA, 216 SQUARE MILES

DISCHARGE SHOWING SEASONAL EFFECT

Vertical slice

$S_e(\text{adjusted}) = 0.0775\sqrt{\frac{84}{84-5}}$
$= 0.080 \pm 20 \text{ percent}$
$-17 \text{ percent}$

$2S_e = 0.155$
$S_e = 0.0775$

Horizontal slice

$\rho = \sqrt{1 - \left(\frac{0.080}{0.378}\right)^2} = 0.98$

$\rho(\text{adjusted for monthly means}) = \sqrt{1 - \left(\frac{0.080}{0.24}\right)^2} = 0.94$

GREENWATER RIVER AT GREENWATER, WASH., DRAINAGE AREA, 739 SQUARE MILES

ADJUSTED DISCHARGE, IN CUBIC FEET PER SECOND

CURVE OF RELATION

EXPLANATION

• Adjusted simultaneous monthly discharge

X Median of vertical slice

+ Median of horizontal slice

DISCHARGE, IN CUBIC FEET PER SECOND

WHITE RIVER AT GREENWATER, WASH., DRAINAGE AREA, 216 SQUARE MILES

DISCHARGE AFTER ADJUSTING FOR SEASONAL EFFECT

CORRELATION OF MONTHLY DISCHARGE, WHITE RIVER AND GREENWATER RIVER, WASHINGTON, WATER YEARS 1947–53