Standard Guide for
Displaying Results of Chemical Analyses of Ground Water for Major Ions and Trace Elements—Diagrams Based on Data Analytical Calculations[1]

This standard is issued under the fixed designation D 5877; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide covers methods that graphically display chemical analyses of multiple ground-water samples, discrete values and also those reduced to comprehensive summaries or parameters. Details required by the investigator to fully use the methods are found in the listed references. The methods included in this guide are many of the graphical procedures that were not discussed in two previous guides, Guides D 5738 and D 5754.

NOTE 1—The graphic methods in this guide apply to both raw and transformed data, for example, unaltered medians, maximums, and minimums and transformed means, square-roots, frequency distributions, and so forth. The methods are often computational intensive, requiring the use of a digital computer. Some graphical methods illustrate the results of the statistical analysis of a sample data set. For example, box plots are graphical portrayals of the maximum, minimum, median, 25th percentile, and 75th percentile of one variable, such as the chloride ion from a group of chemical analyses.

Besides chemical components, other variables that may be plotted to show an interdependence with water chemistry include time, distance, and temperature.

1.2 This guide on diagrams based on data analytical calculations is the third of several documents to inform the hydrologists and geochemists about traditional graphical methods for displaying ground-water chemical data.

NOTE 2—The initial guide described the category of water-analysis diagrams that use two-dimensional trilinear graphs to display, on a single diagram, the common chemical components from two or more analyses of natural ground water.

1.2.1 The second guide described the category of water-analysis diagrams that use pattern and pictorial methods as a basis for displaying each of the individual chemical components determined from the analysis of a single sample of natural ground water.

1.3 This guide presents a compilation of diagrams that allows for transformation of numerical data into visual, usable forms. It is not a guide to selection or use. That choice is program or project specific.

1.4 Many graphic techniques have been developed by investigators to illustrate the results of the data analytical computations to assist in summarizing and interpreting related data sets. In this guide, selected graphical methods are illustrated using ground-water chemistry data.

1.5 The basic or original format of each of the graphical techniques given in this guide has been modified in several ways, largely depending upon the data analytical techniques used by the investigators. Several minor modifications are mentioned, some significant revisions are discussed in more detail.

1.6 Notations have been incorporated within many diagrams illustrated in this guide to assist the reader in understanding how the diagrams are constructed. These notations would not be required on a diagram designed for inclusion in a project document.

NOTE 3—Use of trade names in this guide is for identification purposes only and does not constitute endorsement by ASTM.

1.7 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

1.8 *This guide offers an organized collection of information or a series of options and does not recommend a specific course of action. This document cannot replace education or experience and should be used in conjunction with professional judgment. Not all aspects of this guide may be applicable in all circumstances. This ASTM standard is not intended to represent or replace the standard of care by which the adequacy of a given professional service must be judged, nor should this document be applied without consideration of a project's many unique aspects. The word "Standard" in the title of this document means only that the document has been approved through the ASTM consensus process.*

2. Referenced Documents

2.1 *ASTM Standards:*
D 596 Practice for Reporting Results of Analysis of Water[2]

---

[1] This guide is under the jurisdiction of ASTM Committee D18 on Soil and Rock and is the direct responsibility of Subcommittee D18.21 on Ground Water and Vadose Zone Investigations.
Current edition approved Dec. 10, 1995. Published February 1996.

[2] *Annual Book of ASTM Standards*, Vol 11.01.

D 653 Terminology Relating to Soil, Rock, and Contained Fluids[3]

D 1129 Terminology Relating to Water[2]

D 5754 Guide for Displaying the Results of Chemical Analyses of Ground Water for Major Ions and Trace Elements—Trilinear Diagrams for Two or More Analyses[2]

D 5738 Guide for Displaying the Results of Chemical Analyses of Ground Water for Major Ions and Trace Elements—Diagrams for Single Analyses[2]

3. Terminology

3.1 *Definitions*—Except as listed as follows, all definitions are in accordance with Terminology D 653:

3.1.1 *adjacent values (statistics)*—values that fall between the quartile and one step beyond the quartile position, where the interquartile range is from the 25th to 75th percentile of a sample, and a step is equal to 1.5 times the interquartile range (**1**).[4] The same definition applies to hinges (**2**).

3.1.2 *anion*—an ion that moves or would move toward an anode; thus nearly always synonymous with negative ion.

3.1.3 *cation*—an ion that moves or would move toward a cathode; thus nearly always synonymous with positive ion.

3.1.4 *equivalent per million (epm)*—for water chemistry, an equivalent weight unit expressed in English terms, also expressed as milligram-equivalent per kilogram. When the concentration of an ion, expressed in ppm, is multiplied by the equivalent weight (combining weight) factor (see explanation of *equivalent weight factor*) of that ion, the result is expressed in epm.

3.1.4.1 *Discussion*—For a completely determined chemical analysis of a water sample, the total epm value of the cations will equal the total epm value of the anions (chemically balanced). The plotted values on the water-analysis diagrams described in this guide can be expressed in percentages of the total epm (although all illustrations are in milliequivalent per litre) of the cations and anions of each water analysis. Therefore, to use the diagrams, analyses must be converted from ppm to epm by multiplying each ion by its equivalent weight factor and determining the percent of each ion of the total cation or anion.

3.1.5 *equivalent weight factor*—the equivalent weight factor or combining weight factor, also called the reaction coefficient, is used for converting chemical constituents expressed in ppm to epm and mg/L to meq/L (see explanation of *epm* and *meq/L*). To determine the equivalent weight factor, divide the formula weight of the solute component into the valence of the solute component:

$$(equivalent\ weight\ factor) = \frac{(valence\ solute\ component)}{(formula\ weight\ solute\ component)} \quad (1)$$

Then to determine the equivalent weight (meq/L) of the solute component, multiply the mg/L value of the solute component times the equivalent weight factor, as follows;

$$(meq/L\ solute\ component) = (mg/L\ solute\ component) \times (equivalent\ weight\ factor)$$

[3] *Annual Book of ASTM Standards*, Vol 04.08.

[4] The boldface numbers given in parentheses refer to a list of references at the end of the text.

(2)

For example, the formula weight of Ca^{2+} is 40.10 and the ionic charge is 2 (as shown by the 2 +), and for a value of 20 mg/L Ca, the equivalent weight value is computed to be 0.9975 meq/L:

$$(0.9975\ meq/L\ Ca) = (20\ mg/L\ Ca) \times \frac{(2)}{(40.10)} \quad (3)$$

3.1.5.1 *Discussion*—Many general geochemistry publications and water encyclopedias have a complete table of′ equivalent weight factors′ for the ions found in natural ground water (**3, 4**).

3.1.6 *far-out values (statistics)*—values that fall beyond the two-step range (see outside values) (**1, 2**).

3.1.7 *hinge (statistics)*—as used by Tukey (**2**), the upper and lower values of a ranked sample that, along with the median, divide the number of data values into four equal parts. The data at the hinge position includes interpolated values.

3.1.7.1 *Discussion*—Tukey (**2**) used the hinge system for his box and whisker plots and for his hinge plot and related summaries. The hinge method of division is similar to the use of quartiles.

3.1.8 *interquartile or hinge range (statistics)*— the difference between the values at the quartile or hinge extremes (**2**).

3.1.9 *maximum or sample maximum (statistics)*— the value of the variable having the greatest value in a data set (sample).

3.1.10 *milliequivalent per litre (meq/L)*—for water chemistry, an equivalent weight unit expressed in metric terms, also expressed as milligram-equivalent per litre. When the concentration of an ion, expressed in mg/L, is multiplied by the equivalent weight factor (see explanation of *equivalent weight factor*) of that ion, the result is expressed in meq/L.

3.1.10.1 *Discussion*—For a completely determined chemical analysis of a water sample, the total value of the cations will equal the total value of the anions (chemically balanced). The plotted values on the water-analysis diagrams described in this guide are expressed in percentages of the total meq/L of the cations and anions of each water analysis. Therefore, to use the diagrams, analyses must be converted from mg/L to meq/L by multiplying each ion by its equivalent weight factor and determining the percent of each ion of the total cation or anion.

3.1.11 *milligrams per kilogram (mg/kg)*—for water chemistry, a weight-per-weight unit expressed in metric terms. The number of milligrams of solute (for example, Na) per kilogram of solution (water). For example, a 10 000-mg/kg solute is the same as 1% solute in the total 100 % solution. The mg/kg unit is equivalent to ppm according to Matthess (**5**).

3.1.12 *milligrams per litre (mg/L)*—for water chemistry, a weight-per-volume unit expressed in metric terms. The weight in milligrams (10^{-3} g) of the solute within the volume (litre) solution. The weight can be also expressed in micrograms (10^{-6} g). The use of the mg/L unit is the worldwide standard for the analysis and reporting of water chemistry.

3.1.12.1 *Discussion*—The ppm and mg/L values of the constituents in natural ground water are nearly equal (within anticipated analytical errors) until the concentration of the dissolved solids reaches about 7000 mg/L. For highly mineralized waters, a density correction should be used when computing ppm from mg/L (**3**).

3.1.13 *minimum or sample minimum (statistics)*—the value of the variable having the smallest value in a data set (sample).

3.1.14 *natural ground water*—is water positioned under the land's surface, which consists of the basic elements, hydrogen and oxygen (H_2O), and numerous major dissolved chemical constituents, such as calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), carbonate (CO_3), bicarbonate (HCO_3), chloride (Cl), and sulfate (SO_4).

3.1.14.1 *Discussion*—Other major constituents, in special cases, can include aluminum (Al), boron (B), fluoride (F), iron (Fe), nitrate (NO_3), and phosphorus (PO_4). Minor and trace elements that can occur in natural ground water vary widely, but can include arsenic (As), copper (Cu), lead (Pb), mercury (Hg), radium (Ra), and zinc (Zn). In addition, natural ground water may contain dissolved gases, such as hydrogen sulfide (H_2S), carbon dioxide (CO_2), oxygen (O_2), methane (CH_4), ammonia (NH_3), argon (Ar), helium (He), and radon (Rn). Also maybe included are neutrally charged mineral species, such as silicate (SiO_2), naturally occurring organics, such as tannic acids, colloidal materials, and particulates, such as bacteria viruses and naturally charged pollen spores.

3.1.14.2 *Discussion*—Most of the natural ground water is a part of the hydrologic cycle, that is the constant circulation of meteoric water as vapor in the atmosphere as a result of evaporation from the earth's surface (land and ocean), liquid and solid (ice) on and under the land as a result of precipitation from the atmosphere, and as liquid returned to the ocean from the land. A small amount of the ground water may be magmatic water originating from rocks deep within the crust of the earth. Other ground water is connate in that it is trapped in sediments and has not actively moved in the hydrologic cycle for a period measured in geologic time.

3.1.14.3 *Discussion*—While moving through the hydrologic cycle, chemical elements in the water undergo ion exchange, adsorption/desorption, precipitation/dissolution, oxidation/reduction, and other chemical reactions in response to changes in temperature, hydraulic pressure, biological agents, and chemical composition of the water. The chemical composition of natural ground water ranges from that similar to distilled water with a minor amount of dissolved solids to brines with at least 100 000 mg/L dissolved solids (natural occurring brines have been analyzed with more than 300 000 mg/L dissolved chemical solids) **(6)**.

3.1.15 *outlier (statistics)*—observations whose values are quite different from others in the sample (far-out values fall into this category). These may be the most important values in the data set and should be investigated further **(1)**. In the case of a single deletion, the relationship between the change in mean and the deleted observation is linear, whereas, the relationship between the change in standard deviation and the deleted observation is nonlinear or approximately quadratic for the total number of sample observations considerably larger than the standardization variable squared **(7)**. Values as described by Sara **(8)** as unusually high, low, or otherwise unexpected values within the sample.

3.1.15.1 *Discussion*—Outliers **(8)** can be attributed to a number of conditions, including: extreme, but accurately detected, conditions or environmental conditions; sampling errors or field contamination; analytical errors or laboratory contamination; recording or transcription errors; and faulty (water) sample preparation or preservation, or shelf-life exceedance.

3.1.16 *outside values (statistics)*—values that fall between one and two steps beyond the interquartile range (see adjacent values) **(1, 2)**.

3.1.17 *parts per million (ppm)*—for water chemistry, a dimensionless ratio of unit-of-measurement per unit-of-measurement expressed in English terms. One part per million is equivalent to one milligram of solute in one kilogram of solution. For example, if the total weight of the solution (one million ppm) has 99 % solvent and 1 % solute, this is the same as 990 000 ppm solvent and 10 000 ppm solute in the one million parts of solution.

3.1.18 *polar smoothing (statistics)*—this type of smooth, as used on a scatterplot or Piper diagram, improves the visualization of multiple groups of data sets by enclosing a fixed percent (50 or 75 %) of each group with a mathematically determined ellipse **(1, 9, 10, 11, 12, 13)**.

3.1.19 *population (statistics)*—a well-defined set (either finite or infinite) of elements **(14)**.

3.1.19.1 *Discussion*—For ground-water quality data the infinite population is actually the finite sampled population, as it would be impossible, and certainly impractical, to obtain and chemically analysis all of the ground water from an aquifer.

3.1.20 *quantile (statistics)*—the data point corresponding to a given fraction of the data. Similar to percentile, which is the data point corresponding to a given percentage of the data **(15)**.

3.1.21 *quartile (statistics)*—the upper and lower values of a ranked sample that, along with the median, divides the number of data values into quartile percentages or four equal parts (>0 to ≤25,> 25 to ≤50, >50 to ≤75, and >75 to ≤100 %). The data at the quartile position includes interpolated values.

3.1.22 *sample mean (statistics)*—an arithmetic average of a series of values of a data set (sample) **(16)**.

$$\bar{X} \, (sample \; mean) = \sum_{i=1}^{n} \frac{X_i}{n} \qquad (4)$$

3.1.23 *sample median (statistics)*—the value of the middle variable in a data set (sample) arranged in rank order **(16)**. Also, the 50th percentile or the central value of the distribution when the data are ranked in order of magnitude **(1)**.

3.1.23.1 *Discussion*—For an odd number of observations, the sample median is the data point which has an equal number of observations both above and below it. For an even number of observations, it is the average of the two central observations **(1)**.

3.1.24 *sample size (n) (statistics)*—the number of data observations in the sample.

3.1.25 *sample (statistics)*—a subset of elements taken from a population **(14)**. Also, called sampled population, sample data set or data set. The part or subset of a statistical population that if properly chosen may be used to estimate parameters **(16)**.

3.1.25.1 *Discussion*—If the sample is representative of the entire population, important conclusions about the population can often be inferred from analysis of the sample **(15)**. For ground-water quality data, the sample is a finite subset of data

elements from an infinite population.

3.1.26 *smoothing (statistics)*—smoothing techniques are methods of fitting a line through a number of related data values to enhance the perception of understanding the relationship of one variable (*Y*) to another (*X*). By use of mathematical computations, another set of points $(X_i, \hat{Y}_i)$ are determined (several methods are used) and plotted, these are termed the smoothed values. For this guide, the two types of smoothing discussed are line (**1, 9, 10, 11**) and polar (**1, 9, 10, 11, 12, 13**).

3.1.27 *standard deviation or sample standard deviation (statistics)*—the square root of the average of the squares of deviations about the mean of a set of data (**16**).

$$s \ (sample \ standard \ deviation) = \sqrt{s^2}$$
(5)

3.1.28 *statistical analysis*—the art of reducing numerical data and their interrelationships to comprehensible summaries or parameters (**16**).

3.1.29 *variance or sample variance (statistics)*— the square of the standard deviation (**16**). The expected value of the square of deviations of the variable from its expected value (**17**):

$$s^2 \ (sample \ variance) = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{(n-1)}$$
(6)

3.1.30 *water analysis*—a set of data showing the concentration of chemical ions and measure of physical properties determined from a water sample. In this guide, the water analysis normally includes the common constituents and project-dictated parameters as found in natural and human-influenced ground water (see *natural ground water*).

3.1.31 *water-analysis diagram*—for purposes of this guide, a diagram for graphically displaying water-quality analyses and related parameters. These diagrams can be used to assist in the scientific interpretation of occurrence of cations and anions in natural and human-influenced ground water, for example, the interrelationship of a number of water samples within the studied area.

3.1.32 *water sample*—in this guide, a water sample refers to a carefully collected specimen of natural or human-influenced ground water obtained from the aquifer for analyzing the chemical constituents in the water.

3.1.32.1 *Discussion*—In this guide on analytical calculations, a water sample is one element in the entire sample or data set with a sample size of *n* from the entire population. (See *population (statistics), sample (statistics)*, and *sample size (n) (statistics)*.)

4. Summary of Guide

4.1 The significance and use present the relevance of the water-analysis diagrams that pictorially display the results of data analytical computations of chemical constituents and related parameters from natural and human-influenced ground-water sources.

4.2 A summary of the recommended checks for accuracy (quality control of the data) is presented.

NOTE 4—Most of the graphical methods presented in this guide use one or two chemical constituents from each of many analyses. However, several methods require the use of complete analyses. The measure of the quality confidence of the analyses used for these methods must follow the

same level of evaluation as that outlined in Guides D 5738 and D 5754.

4.3 Descriptions and comprehensive illustrations are given for the following water-analysis diagrams.

4.3.1 *Diagrams for a Single Set of Data*:

4.3.1.1 Frequency histogram diagram (**1, 3, 8, 14, 18, 19, 20, 21, 22, 23, 24, 25**) ,

4.3.1.2 Relative frequency histogram diagram (**14, 19, 23, 26**),

4.3.1.3 Rootogram diagram (**27**),

4.3.1.4 Stem and leaf plot (**1**),

4.3.1.5 Dot and line or error plot (**1, 8**),

4.3.1.6 Hinge plot, five-number summaries, and fenced summaries (**2, 27**),

4.3.1.7 Box and whisker or range plots (**1, 2, 8, 18, 20, 21, 27, 28, 29, 30, 31**),

4.3.1.8 Frequency distribution diagram (**17, 19, 20, 32**) ,

4.3.1.9 Cumulative frequency distribution and quantile diagram (**3, 17, 19, 20, 21, 29**),

4.3.1.10 Cumulative percentage diagram (**3**),

4.3.1.11 Probability plot using percent (**34**), and

4.3.1.12 Probability plot using normal quantiles (**1, 34, 40**).

4.3.2 *Diagrams for Two Sets of Related Data*:

4.3.2.1 Simple scatterplot (scattergrams) (**1, 3, 21, 22, 23, 25, 26, 29, 30, 36, 37, 38, 39, 40**),

4.3.2.2 Scatterplot with samples (data sets) from two populations (**1**),

4.3.2.3 Scatterplot with moving medians or means smooth (**1**),

4.3.2.4 Scatterplot with LOWESS smooth (**1, 10, 11, 41, 42**),

4.3.2.5 Scatterplot with polar smooths of samples (data sets) from more than one population (**1, 9**),

4.3.2.6 Scatterplot with absolute differences versus the sample (data set) (**1, 9, 11**),

4.3.2.7 Scatterplot for correlation coefficient (**1, 3, 14, 19, 26, 35, 43, 44**) ,

4.3.2.8 Basic time-series plot (**1, 3, 8, 9, 19, 24, 31, 32, 33, 45, 46, 47, 48, 49**),

4.3.2.9 Time-series plot for multiple data sets (**1**),

4.3.2.10 Elapsed time plot (**1, 8, 9, 50, 51, 52**), and

4.3.2.11 Q-Q plots (**1, 19, 41, 53**).

4.3.3 *Other Diagrams of Interest*:

4.3.3.1 Schoeller nomograph or vertical scale diagram (**3, 5, 38, 54, 56, 57, 58, 59**),

4.3.3.2 Irrigation classification or salinity hazard diagram (**3, 55, 59**),

4.3.3.3 Piper diagram with polar smoothing (**1, 9, 10, 11, 12, 13**),

4.3.3.4 Three-variable pattern plot (**60**),

4.3.3.5 Three-dimension rotational plot (**1, 60, 61**),

4.3.3.6 Ropes three-dimensional diagram (**62, 63, 64**), and

4.3.3.7 Cluster analysis diagrams (**1, 32**).

4.4 Automated procedures (computer-aided graphics) for basic calculations and the construction of the water-analysis diagrams are identified,

4.5 Keywords, and

4.6 A list of referenced documents is given for additional information.

5. Significance and Use

5.1 Each year, many thousands of water samples are collected, and the chemical components are determined from natural and human-influenced ground-water sources.

5.2 An understanding of the relationship between the similarities and differences of these water analyses is simplified by use of data analytical methods and the display of the results of these methods as pictorial diagrams.

5.3 This guide presents a compilation of the diagrams used for illustrating the results of these methods.

5.4 This type of diagram summarizes data from a number of analyses to allow for an objective comparison between the chemical and related parameters.

5.5 The diagrams based on data analytical calculations described in this guide display the following; time and areal trends; maximums, minimums, and means; relationships between chemical and associated parameters; significant outliers; distributions; and a summary of a number of data parameters.

5.6 The objective interpretations of the origin, composition, and interrelationships of ground water are common uses of the diagrams based on data analytical calculations.

5.6.1 The origin of the water may be postulated by the amount and the relationship of the chemical constituents in a sample of water analyses summarized on the diagrams.

5.6.2 The chemical composition of the water can be scrutinized for distinct characteristics and anomalies by use of the diagrams.

5.6.3 A graphical comparison of distinct data sets of chemical analyses allows the investigator to evaluate the interrelationships of the ground water from separate locations.

5.7 This is not a guide for the selection of a diagram for a distinct purpose. That choice is program or project specific.

NOTE 5—For many hydrochemical research problems involving the scientific interpretation of ground water, the 'analytical water-analysis diagram' is only one segment of several methods needed to interpret the data.

6. Selection and Preparation of Data for Plotting on the Analytical Diagrams

6.1 For the data analytical graphical methods described in this guide, transformation of the raw data is often required before analysis. However, several methods, for example, some of the scattergrams, the transformation is accomplished by analytical smoothing of the curve after the data are plotted on the diagram.

NOTE 6—Helsel and Hirsch (1) on pages 253 through 255 discusses the subject of whether to transform or not to transform the response variable (y). The response variable may require transformation because the variance of the residuals is a function of x for much of the ground-water quality data and for hydrology in general. Helsel and Hirsch states that the decision of whether to transform y should generally be based on graphs.

6.2 *Minimum Data Requirements*:

6.2.1 The basic requirements for the analytical methods described in this guide are that the samples are randomly selected and of sufficient size to represent the sampled population and therefore, allow for a meaningful analysis (1).

NOTE 7—A truly random sample is impractical, as ground water samples are from a subsurface population that only can be obtained from

sources that intersect the water table, for example, wells, springs, and tunnels or caves. These sources are not likely to be distributed randomly in three dimensions throughout an aquifer. However, a more refined picture of the entire population is possible as the size of the random sample is increased (1).

6.3 *Recommended Checks for Accuracy of Data Parameters*:

6.3.1 For those methods described in this guide that use a sample (data set) from chemical analyses that are not complete, the individual data values must be carefully reviewed to avoid errors in the results.

NOTE 8—Some of these methods can use a data set consisting of a single constituent, for example, the evaluation of chloride by a histogram. Other methods use two parameters, for example, the evaluation of the relationship of nitrate and dissolved solids by a scattergram. Other methods, such as a ratio evaluation, use data sets consisting of more than two parameters, but less than complete analyses.

6.3.1.1 Erroneous values in a data set (sample) flagged as outliers, become more apparent when using graphical methods, as these values do not plot with the prevalent group of the data points.

6.3.1.2 Erroneous values that fall in the same numerical range as a typical value in the data set are difficult to detect, but are most likely found by a complete validation of the data set (sample) against the original data source.

NOTE 9—To reduce the chance of incorporating erroneous numbers into the data analytical evaluation, the original chemical analyses and related data must be carefully previewed as to proper collection and analytical procedures. In addition, take care to ensure that none of the numbers have been transposed during preparation of the data for the analytical evaluation.

6.3.2 For those methods described in this guide that use a sample (data set) that consists of complete chemical analyses (where all of the major chemical ions in the ground water are determined), a check of the chemical balance should be made to help in the detection of data errors.

6.3.2.1 The chemical balance or chemical equilibrium of a complete analysis is calculated by converting the ions from mg/L to meq/L values and adding the cations together and the anions together. The computation for percent balance is as follows, with 0 (zero) as the optimum percentage value (percentage is determined by multiplying the computed value times 100);

$% \ chemical \ balance \ (+/-)$

$$= \frac{total \ cations - total \ anions \ (meq/L)}{total \ cations + total \ anions \ (meq/L)} \times 100 \qquad (7)$$

NOTE 10—Minor amounts of ions such as fluoride (F), nitrate (NO_3), iron (Fe), and barium (Ba), may occur in natural or human-influenced ground water, but normally do not significantly influence the chemical balance. If any of these ions (for example, NO_3) occur in amounts that alter the chemical balance, they should be included in the computations. Other constituents may occur in minor amounts in a colloidal or suspended state, such as silica (SiO_2), iron hydroxide (Fe), and aluminum compounds (Al), and are not considered in the chemical balance because they are not dissolved constituents.

6.4 *Required Calculations for Diagram Construction*:

6.4.1 The data analytical methods described in this guide use a wide range of computations to analyze the data sets

(samples) and to prepare the data for illustration on the various diagrams.

6.4.2 Because of the many types of equations, they are presented or referred to with the first diagram that discusses the computational method and then cited when used in later diagrams.

7. Water-Analysis Diagrams

7.1 *Introduction*—This guide provides methods that furnish helpful graphical summaries of the results of data analysis of water samples. These methods include procedures that graphically display a single data set (sample), two sets of directly related data, and multiple sets of data.

7.1.1 Helsel and Hirsch **(1)**, describe many of the data analytical methods for use in the study of water resources. The book explains many graphical procedures to illustrate chemical analyses and related ground-water data.

7.1.2 In the description by Helsel and Hirsch **(1)** they state one of the most frequent tasks when analyzing data is to describe and summarize those data in forms which convey their important characteristics.

7.1.3 Helsel and Hirsch also said that "Graphs are essential for two purposes: (*1*) to provide insight for the analyst into the data under scrutiny, and (*2*) to illustrate important concepts when presenting the results to others" **(1)**.

NOTE 11—Many other excellent publications are available for the statistical study of natural ground water; most of those are referred to in the text and listed in the bibliography.

NOTE 12—The criteria for the selection, error check of data values, and preparation of the data sets (samples) used for plotting on many of the data analytical diagrams is described in Section 6.

7.2 *Diagrams for a Single Set of Data*— These diagrams graphically illustrate the results of an analytical examination of a single set of data (sample) selected from a number of chemical analyses of natural and human-influenced ground water.

7.2.1 *Histograms*—This type of diagram is a vertical bar-graph for showing the distribution of a variable. The length of the individual bars represents the frequency of the data values within each subdivision of the total parameter range.

7.2.1.1 *Frequency Histogram Diagram*—One type of diagram has the bars representing the occurrence numbers on the *Y*-axis plotted against the subdivided parameter values on the *X*-axis (see Fig. 1 and Table 1). The parameter values are subdivided so that there are no open or zero intervals **(1, 3, 8, 14, 18, 19, 21, 22, 23, 24, 25)**.

7.2.1.2 *Relative Frequency Histogram Diagram*—The relative frequency histogram has the occurrence number (as shown on the frequency histogram) replaced by the percentage value to show the distribution of a ranked sample data set (see Fig. 2 and Table 1) **(14, 19, 23, 26)**.

NOTE 13—The relative frequency percentages are computed as follows:

$$relative\ frequency = number\ of\ occurrences\ in\ interval/total\ sample\ size \tag{8}$$

For example: on Fig. 2, from a sample data set of 54, six values are >50 and ≤60 and the relative frequency is 11.1 % (6/54).

7.2.1.3 *Rootogram Diagram*—The rootogram is a histo-
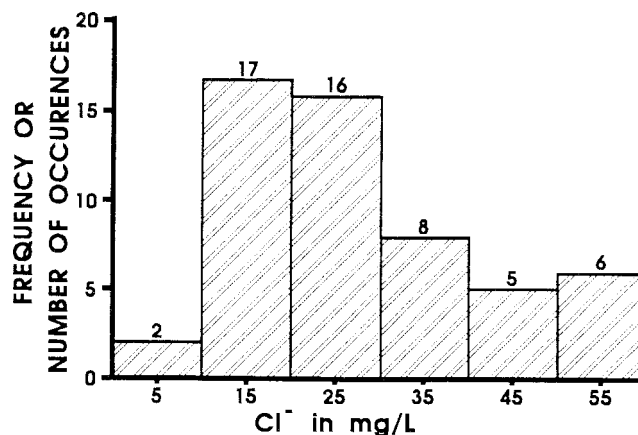
NOTE 1—Analyses selected from Ref **(67)**. See Table 1.
FIG. 1 Histogram

TABLE 1 Chloride Values (mg/L) in Ascending Order (67)^A

7	18	19	23	26	31	40	53
9	18	20	23	27	33	43	55
11	19	20	24	27	35	48	56
11	19	20	25	28	36	48	57
15	19	20	25	29	36	49	59
16	19	21	25	29	38	50	...
17	19	22	26	30	39	51	...

^{A}*Basic Statistics*—$n = 54$, minimum = 7, mean = 29.3, median = 25.5, maximum = 59, sample standard deviation = 13.5, and sample standard error = 1.35.
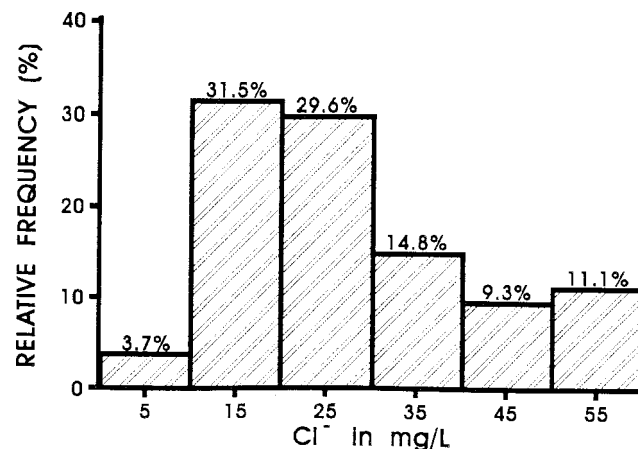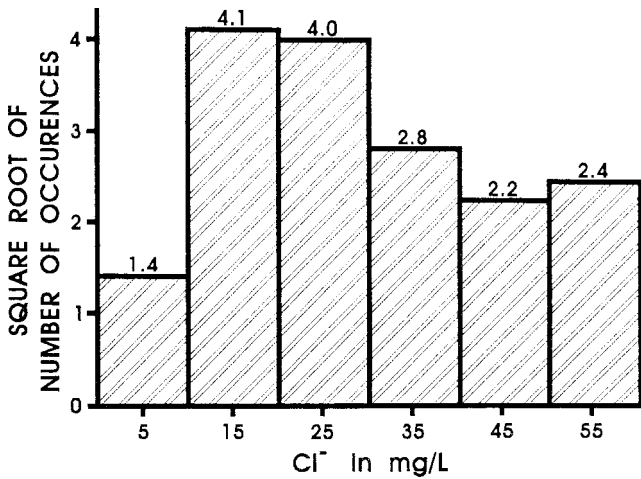

NOTE 1—Analyses from Ref **(67)**. See Table 1.
FIG. 2 Relative Frequency Histogram

gram where the frequency for each interval is plotted as square root **(27)**. For example, the frequency or number of occurrences on Fig. 1 are shown as square root values on Fig. 3. This tends to make the interval differences look smaller than on the traditional histogram and is useful where some intervals have a lopsided large number of occurrences (see Fig. 3 and Table 1).
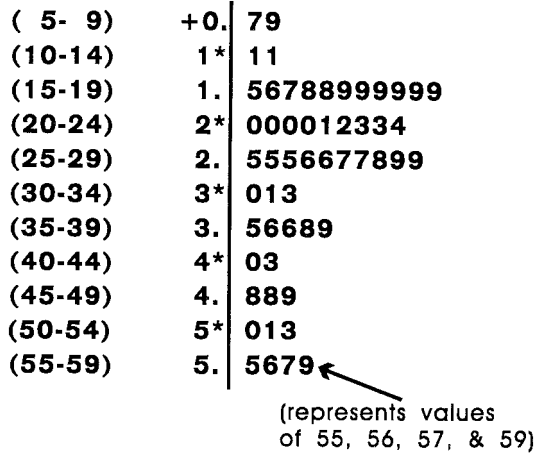
7.2.2 *Stem and Leaf Plot*—The stem and leaf plot is similar to a histogram rotated 90° **(1)**. An advantage of this plot is that the real values are placed on the diagram (see Fig. 4 and Table 1).

7.2.3 *Dot and Line or Error Plots*—The dot and line plot is used to represent the mean and standard deviation (or standard error) of a sample (see Fig. 5 and Table 1) **(1, 8)**.
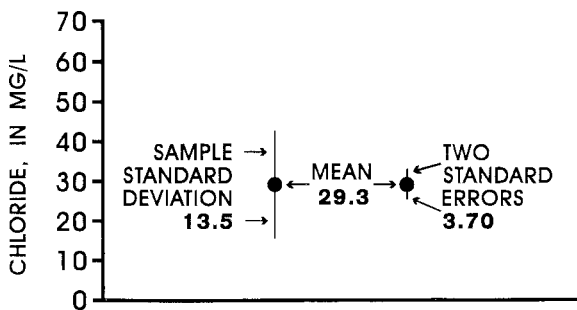
NOTE 1—See Ref (2). Analyses from Ref (67), see Table 1.
FIG. 3 Rootogram

(range in mg/L)

```
( 5- 9)    +0. │ 79
(10-14)     1* │ 11
(15-19)     1. │ 56788999999
(20-24)     2* │ 000012334
(25-29)     2. │ 5556677899
(30-34)     3* │ 013
(35-39)     3. │ 56689
(40-44)     4* │ 03
(45-49)     4. │ 889
(50-54)     5* │ 013
(55-59)     5. │ 5679
```

(represents values of 55, 56, 57, & 59)

NOTE 1—See Ref (1) and Table 1 for analyses.
FIG. 4 Stem and Leaf Plot of Chloride Concentration

NOTE 1—See Ref (1) and Table 1 for analyses.
FIG. 5 Dot and Line Plot of Chloride Concentration

7.2.3.1 The dot represents the mean of the sample and the line represents plus or minus one standard deviation (see definitions) or plus or minus one or more standard errors beyond the mean, computed as follows:

$$s.e. = \frac{s}{\sqrt{n}} \qquad (9)$$

where:
s.e. = standard error,
s = standard deviation, and
n = sample size.

7.2.3.2 The dot and line plots are useful only when the data are actually symmetric and as a simple method to display differences in several samples.

7.2.4 *Hinge Plot, Five-Number Summaries, and Fenced Summaries*—The hinge plot and related summaries were discussed by Tukey (**2, 27**). This method uses the median of the sample data set because the mean value could be influenced by exotic values (outliers and far-out values) that may be included in the data set.

7.2.4.1 The hinge plot requires that the sample data set be ranked from lowest to highest value. From this ranked data set, five numbers are determined, one of which is the median value. Two other values are the low and high extremes. Half-way between (actual or interpolated) the median and the extreme values are the two 'hinges'. From this ranked sample data set and the five related numbers, the hinge plot can be constructed (see Fig. 6 and Table 2).

7.2.4.2 Clearly the hinge plot would be cumbersome for even moderately sized sample data sets. For these data sets the five-number summary can convey the results of the statistical analysis (see Fig. 6).

7.2.4.3 When the sample data set contains outside and far-out values, a fenced summary can be used to emphasize

NOTE 1—Adapted from Refs (**2**) and (**27**). See Table 2 for analyses.
FIG. 6 Hinge Plot, Five-Number and Fenced Summaries

7

TABLE 2 Chloride Values (mg/L) in Ascending Order (65)[A]

|							
5	10	15	17	21	23	42	143
7	11	15	18	21	26	43	...
9	11	15	18	21	26	44	...
9	13	16	20	23	27	127	...

[A]*Basic Statistics*—n = 29, minimum = 5, mean = 27.5, median = 18, maximum = 143, sample standard deviation = 31.0, sample standard error = 5.8, 10th percentile = 9, 25th percentile = 12, 75th percentile = 26, 90th percentile = 44, L hinge = 13, and U hinge = 26.

these extreme values (see Fig. 6) **(2, 27)**.

7.2.5 *Box and Whisker or Range Plots*— A box and whisker plot or boxplot presents a concise graphical display for summarizing the distribution of a sample and is useful when comparing the attributes of several samples **(1, 2, 8, 18, 20, 21, 27, 28, 29, 30, 31)**. Three variations of the boxplot are shown in Fig. 7, these are the simple hinged, standard quartiled, and truncated quartiled. These boxplots use the data given in Table 2 **(66)**.

NOTE 14—Besides the three boxplots discussed, there are many variations from Tukey's **(2)** original hinged box and whisker diagram. Tukey **(2)** discusses a schematic plot that is similar to the standard quartile boxplot as shown in Fig. 7 **(2, 27)**. This variation limits the length of the whiskers and uses identified symbols (for example, symbol with name of data source) for the outside and outliers or far outside values. Several variations discussed by Helsel and Hirsch **(1)** are for displaying the median confidence intervals within the box and boxplots of censored data where only those values above or below a threshold are displayed. A method of visibly emphasizing the confidence interval is to construct notches on the sides of boxplot at the position of the confidence interval and having the box narrower within the interval. Other methods include the placing of parentheses or a shaded area within the boxplot to represent the confidence interval.

NOTE 15—In general, the boxplots provide the following as a visual summary:
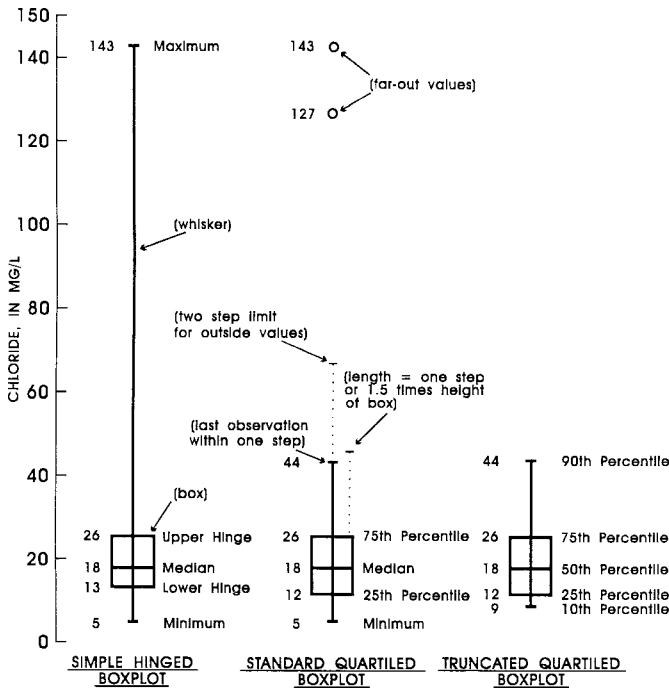
NOTE 1—See Ref **(2)** and quartile system **(1)**. Analyses from Ref **(65)**, see Table 2.

FIG. 7 Boxplots, Using Hinged System

(*1*) The center (median) of the sample is shown by a line within the box. One modification of the boxplot gives the mean line in combination with the median **(18)**.

(*2*) The variation or spread of the sample is shown by the height of the 75th quartile box above and 25th quartile box below the median (or upper hinge and lower hinge boxes).

(*3*) The skewness of the sample is apparent by comparing the relative size of the 75th quartile box with the 25th quartile box. The simple hinge boxplot gives similar results. The quartile skewness (qs) can be computed as follows, where P is the data value at the quartile or hinge position, for example, 26 mg/L for the 75th percentile or upper hinge on Fig. 7 (*1*);

$$qs \; (skewness) = \frac{(P_{0.75} - P_{0.50}) - (P_{0.50} - P_{0.25})}{P_{0.75} - P_{0.25}} \tag{10}$$

The skewness of the standard quartiled boxplot in Fig. 7 is computed as follows:

$$+ \, 0.1428 = \frac{(26 - 18) - (18 - 12)}{26 - 12} \tag{11}$$

The value of *qs* is positive (+ 0.1428), that shows a slight right-skewed distribution for the sample. A negative value would demonstrate a left-skewed sample. However, a [mdit]qs[med] value approaching zero (0) indicates a normally distributed sample.

Besides the above equation, several other computations are commonly used to measure skewness, those are discussed by Helsel and Hirsch **(1)**

(*4*) The presence or absence of unusual values is conspicuous by the length of the whiskers of the simple hinged boxplot or by symbols used on the standard quartiled boxplot. For the standard quartile boxplot, the outside values plot between the one and two steps above the 75th quartile or below the 25th quartile (a step is 1.5 times the height of the box). The outside values are identified by a symbol, such as an "*" (asterisk), at the proper scaled locations and in line with the whiskers. In addition, are the far outside or outlier values that fall above or below the two-step range. These outliers are identified by plotting a symbol different from the outside symbol, such as an "o" (open circle).

NOTE 16—A useful plot, which can be used with the boxplot, is the dot plot **(2)**. This plot consists of plotting dots for the data values parallel to and at the same scale as the boxplot. This type of plot will emphasize subgroups of the sample data set, whereas, the boxplot (or schematic) will cover over any groupings.

7.2.5.1 The following equations are for determining the lower and upper hinge values for the simple hinged boxplot **(1)**.

$$h_L(lower \; hinge) = X_L, \text{ where } L = \frac{integer \; ((n+3)/2)}{2} \tag{12}$$

$$h_U (upper \; hinge) = X_U, \text{ where } U = (n+1) - L \tag{13}$$

7.2.5.2 The following equations are for determining the 25th, 50th, and 75th percentile values (*p*) for the standard and truncated quartiled boxplots.

$$P_{0.25} \; (lower \; quartile) = X_{(n+1)} \times 0.25 \tag{14}$$

$$P_{0.50} \; (median) = X_{(n+1)} \times 0.50 \tag{15}$$

$$P_{0.75} \; (upper \; quartile) = X_{(n+1)} \times 0.75 \tag{16}$$

Non-integer values of $X_{(n+1)} \times$ percentile factor imply linear interpolation between adjacent values of X, for example, a value of 7.5 implies that the value of the quartile is ($X_7 + (0.50 \times (X_8 - X_7))$).

7.2.6 *Frequency Distribution Diagram*— The frequency diagram shows the distribution of a ranked sample data set by plotting the relative frequency, percentage, or actual count of each interval on the related data value axis (Fig. 8 and Table 1) **(3, 15, 16, 17, 19, 20, 32)**. These plot points are connected by lines to create a frequency distribution curve. A diagram
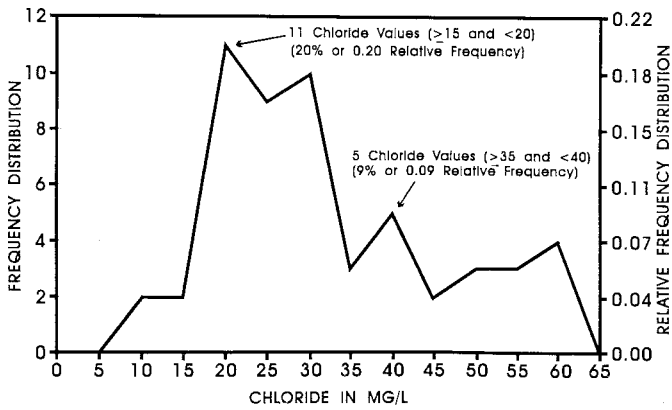
NOTE 1—Analyses from Ref (67), see Table 1.
**FIG. 8 Frequency Distribution Diagram**

showing a similar type of information is the relative frequency histogram.

NOTE 17—The interval values consist of a range of data values, for example on Fig. 8, the plot position at the chloride scale number of 20 represents chloride values that are ≥15 and <20 mg/L.

7.2.7 *Cumulative Frequency Diagram*—The purpose of the cumulative frequency diagram is to show a cumulative curve of the frequency distribution of a sample data set by adding the frequency number of each group to that of the preceding group until the total sum is included on the curve. The total is normally 100 % of the observations (16). The cumulative frequency scale can be presented as relative cumulative frequency (0 to 1.0) or percentages (0 to 100 %) or actual cumulative count (0 to *n*).

7.2.7.1 *Cumulative Frequency Distribution and Quantile Diagram*—This type of diagram gives the cumulative frequency of one variable or one ion for a sample data set, for example, the distribution of the concentration of chloride in mg/L from a number of water analyses (see Fig. 9, Table 1) (**1, 3, 15, 17, 21, 34, 36**).

NOTE 18—Either axis can represent the cumulative frequency distribution or the actual values of the sample data set. The scale for the distribution can be given as the cumulative frequency (count), relative
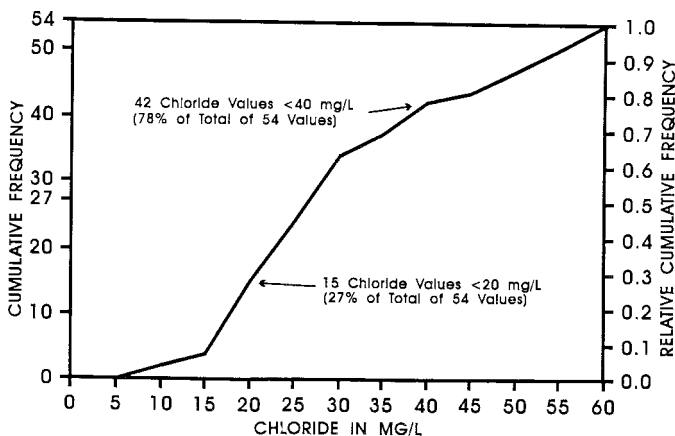
cumulative frequency (quantile), or cumulative percentage (percentiles—quantiles times 100). On the quantile diagram, as described by Helsel and Hirsch (**1**), the scale for quantiles is the same as the relative cumulative frequency or cumulative percentage. The scale for the sample data set is the actual value of the data variable, for example, Cl in mg/L. The plot positions for constructing the frequency curve are the intersections of the *Y*-axis with the relative variable values. For example, the position at the 20-mg/L chloride line represents 27 % (0.27 or 15 values) of the sample data set that has a value of <20 mg/L chloride.
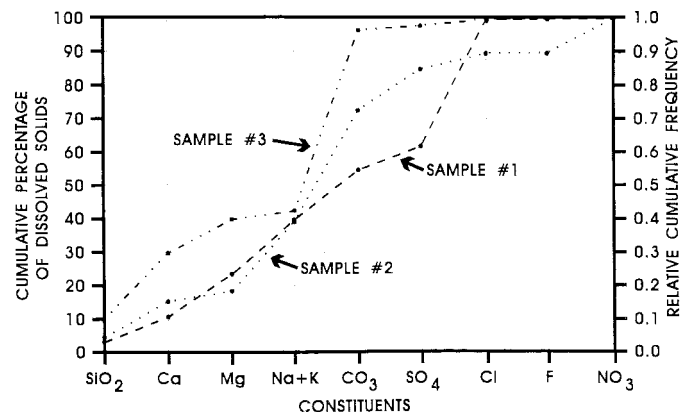
7.2.7.2 *Cumulative Percentage Diagram*— A variation in the use of the diagram is to show the cumulative percentage or relative cumulative frequency of the ions of separate water analyses (**3**). This diagram shows each analysis as a cumulative distribution curve and can be used to compare and differentiate the types of water from the individual ground-water analyses (see Fig. 10).

NOTE 19—As an example on Fig. 10, the left vertical *Y*-scale represents the cumulative percentage and the right scale the relative cumulative frequency of the dissolved solids in a water analysis. The total is always 100 % or 1.0. The related horizontal section of the diagram has no numerical scale and represents the individual ions spaced at even intervals.

7.2.8 *Probability Plot*—Probability is the likelihood of occurrence of an event, where zero is impossibility and one is certainty (**16**). The probability or the relative frequency of a particular occurrence, can be shown on two-dimensional graphs. This type of plot expresses the theoretical distribution of the sample data as a straight line, so that departures from the distribution can be easily observed. A plotted point on the diagram is a representation of the actual data value against the probability of occurrence of that data value (**1, 8, 17, 35**).

NOTE 20—For computing the plot positions on the probability diagram, the data first must be ordered from the smallest (*i* = 1) to the largest (*i* = n) numerical value. Computations required for plotting a single plot point of the actual value (*Y*-axis) with the paired probability of occurrence value (*X*-axis), depend upon the horizontal scale used on the diagram (percentages or normal quantiles), however, all convey the same results.

NOTE 21—The straight line for showing the theoretical distribution of the sample data set is constructed by connecting the intercept of the mean value of *Y* and the 0 value of the normal quantile of *p* (or 50 % line) with the intercept of the mean value of *Y* + one standard deviation and the 1 quantile line (or 84.14 % line) (see Fig. 11 and Fig. 12, and Table 3). This line can be extended to the outer boundaries of the diagram.
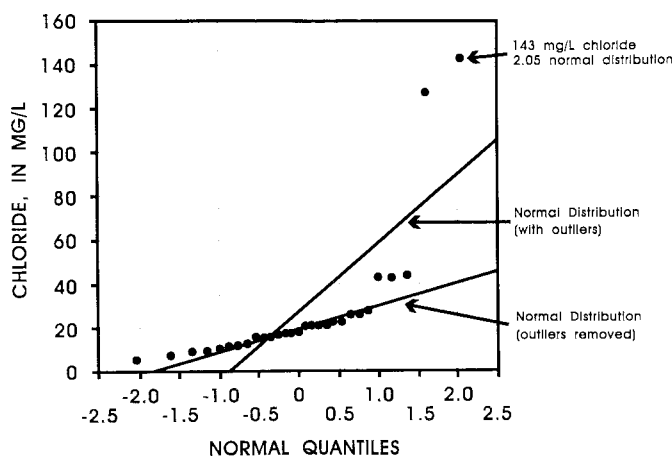

NOTE 1—Analyses from Ref (67), see Table 1.
FIG. 9 Cumulative Frequency Diagram

NOTE 1—Adapted from Ref (3). Analyses from Ref (3).
FIG. 10 Cumulative Percentage Diagram

% PERCENT CHANCE OF EXCEEDANCE
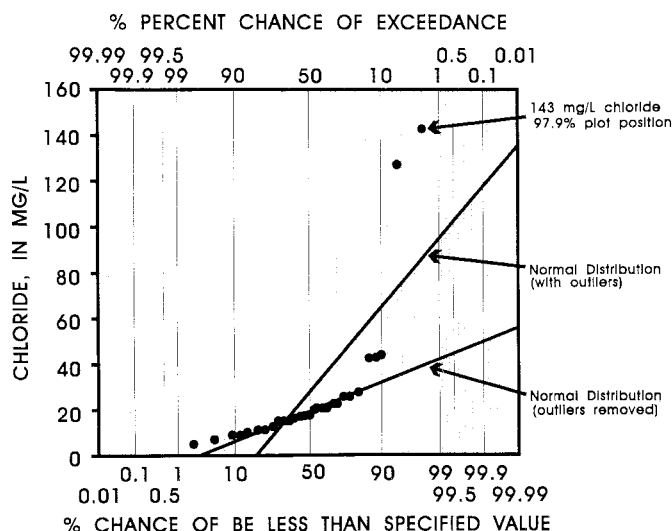
NOTE 1—Analyses from Ref (65), see Table 3.
FIG. 11 Probability Plot Using Percent

NORMAL QUANTILES

NOTE 1—Analyses from Ref (65), see Table 3.
FIG. 12 Probability Plot Using Quantiles

NOTE 22—Probability diagrams for other types of data, such as the residuals from two related variables, for example, chloride and sulfate, can be shown by this category of plot (1).

7.2.8.1 *Probability Plot Using Percent*— The percent probability of occurrence method is nonlinear and is scaled in percentages on the *X*-axis (see Fig. 11 and Table 3). The plot position for the *X*-axis is computed using the following equation, which is the standard in Canada and Europe (34);

$$p_i = \frac{(i - 0.4)}{((n + 0.2) * 100)} \qquad (17)$$

where:
p_i is the percent probability of occurrence. The *Y*-axis is the actual data value or a transformed value, such as the log base ten. Special probability paper can be used for these plots.

7.2.8.2 *Probability Plot Using Normal Quantiles*—The normal quantile method is linear with the *X*-axis scaled in even units (see Fig. 12 and Table 3). The quantile plot values for the *X*-axis are obtained from the standard normal distribution tables (a computerized approximation is also available) by use

of the p_i value from the above equation (1, 34). These tables are available in most basic statistics textbooks (35). The *Y*-axis for the actual data values can be linear or in log base-ten.

NOTE 23—On the probability plot that uses percentages (Fig. 11), one sample standard deviation (s) on either side of the mean includes 68.27 % of the sample data (34.14 % below the mean and 34.14 % above the mean). On the probability plot using quantiles (see Fig. 12), one sample standard deviation on either side of the mean is shown by the − 1.0 and + 1.0 scale on the *X*-axis. Also, two sample standard deviations are 95.45 % on Fig. 11 or − 2.0 and + 2.0 on Fig. 12.

7.3 *Diagrams for Two Sets of Directly Related Data*—These are diagrams that graphically illustrate the results of an analytical examination of two sets of data (samples) selected from a number of chemical analyses of natural and human-influenced ground water and directly related parameters.

7.3.1 *Scatterplots (Scattergrams)*—These plots consist of rectangular two-dimensional diagrams with *X*- and *Y*-axes designed to show the correlation between two related variables. The scatterplot is widely used and is one of the more powerful tools for data analysis (9). Many variations exist for this type of plot (1, 29).

NOTE 24—The *X*- and *Y*-scale factors of the scatterplots can be linear-linear, linear-log, log-linear, or log-log depending on the nature and spread of the data to be displayed. Actual data, the absolute residual, or transformed values can be plotted. For example, data can be transformed to the logarithms of the values or to power functions by use of the ladder of powers (1) (in the form of $y = x^\theta$, where x is the untransformed data, y is the transformed data, and θ is the power exponent). The techniques of overlying a best fit line (termed smoothing) to estimate the center of the data on a scatterplot, include the graphical nonmathematical method of drawing a line through the plotted points, or the mathematical methods of moving median or mean, simple linear regression, and the LOWESS method that uses weighted least squares equations. A variation in LOWESS smoothing is′ polar smoothing′ where polygons are computed to envelope individual groups of data (such as source aquifers) on the plot (1).

7.3.1.1 *Simple Scatterplot*—This type of scatterplot has the first variable on the *X*-axis plotted against the related second variable on the *Y*-axis (1, 3, 21, 22, 23, 25, 26, 29, 30, 36, 37, 38, 39, 40). The paired plot point is represented by a symbol on the diagram, for example, a large dot. The *X*- and *Y*-scales can be linear, or one or both logarithmic. Fig. 13 is a linear plot of Cl on the *Y*-axis against the related SO$_4$ on the *X*-axis (see Table 4). A variation of the scatterplot would be to use several symbols, for example, to show the source of the variables as from several individual aquifers. A 1:1 ratio line of the *X*-variable versus the *Y*-variable can be drawn on the diagram to show the simplest relationship between the two variables.

7.3.1.2 *Scatterplot with Samples (data sets) from Two Populations*—When the data being analyzed comes from two populations, these data can be flagged on the scatterplot by use of individual plot symbols (1). In addition, smooths, by any of the previously described methods, can be applied separately to each population of data values. These X_iY_i and $X_i\hat{Y}_i$ values from the two populations can be plotted on the same scatterplot for comparison.

7.3.1.3 *Scatterplot with Moving Medians or Means Smooths*—The simplest method of mathematical smoothing data on a scatterplot is by use of either a ′moving′ median or mean (1). A window size or range of plot points on the *X*-axis

10

TABLE 3 Values for Probability Plots (65)^{AB}

5/1/2.0 (−2.05)	10/5/15.7 (−1.00)	15/9/29.5 (−0.54)	17/13/43.2 (−0.17)	21/17/56.8 (0.17)	23/21/70.5 (0.54)	42/25/84.2 (1.00)	143/29/97.9 (2.05)
7/2/5.5 (−1.60)	11/6/18.8 (−0.87)	15/10/32.9 (−0.44)	18/14/46.6 (−0.09)	21/18/60.3 (0.26)	26/22/74.0 (0.64)	43/26/87.7 (1.16)	...
9/3/8.9 (−1.35)	11/7/22.6 (−0.75)	15/11/36.3 (−0.35)	18/15/50.0 (0.00)	21/19/63.7 (0.35)	26/23/77.4 (0.75)	44/27/91.1 (1.35)	...
9/4/12.3 (−1.16)	13/8/26.0 (−0.64)	16/12/39.7 (−0.26)	20/16/53.4 (0.09)	23/20/67.1 (0.44)	27/24/80.8 (0.87)	127/28/94.5 (1.60)	...

^{A}Explanation of Numbers in Table:

9/4/12.3 (−1.16) = chloride value/ascending order/percentage plot position (normal quantile plot position)

where:

Chloride value is in mg/L units,

Ascending order from smallest chloride value to largest value, and

Percentage plot position = (order number− 0.4)/(total number + 0.2) (normal quantile plot position) = standard normal distribution.

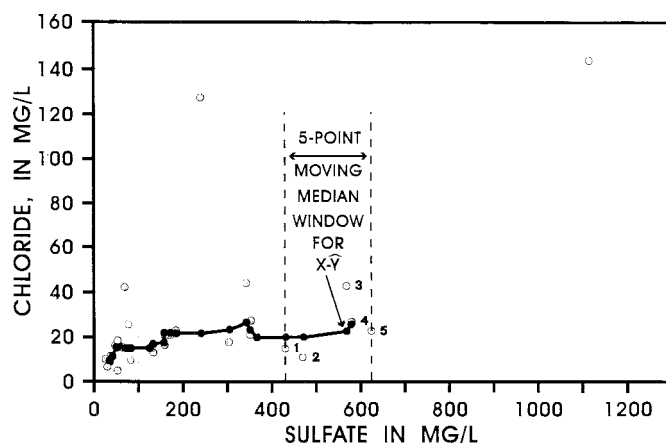^{B} Basic Statistics:

With two outliers (Cl = 127 and 143):

n = 29, mean = 27.4, sample standard deviation = 31.0

With two outliers removed:

n = 27, mean = 19.5, sample standard deviation = 10.2
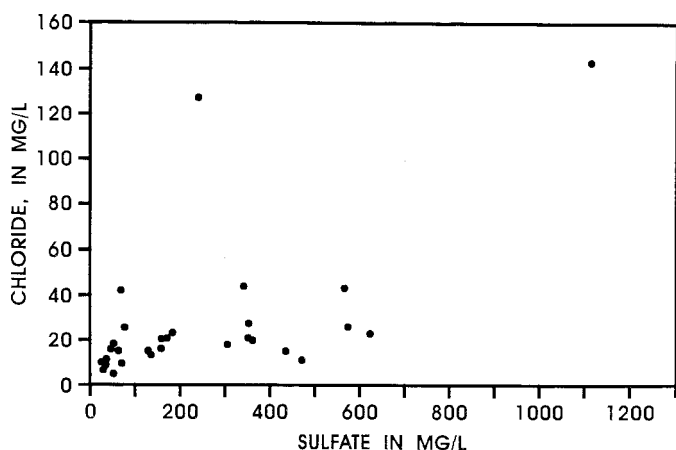
NOTE 1—Adapted from Ref (**1**). Analyses from Ref (**65**), see Table 4.
FIG. 13 Simple Scatterplot

NOTE 1—Adapted from Ref (**1**). Analyses from Ref (**65**), see Table 4.
FIG. 14 Scatterplot Smoothing by Moving Median

TABLE 4 Chloride/Ordered Sulfate Values (mg/L) (65)

7/29	16/40	42/69	13/130	23/183	21/349	11/473	143/1114
10/29	18/48	9/70	17/158	127/242	27/350	43/566	...
11/33	5/55	26/78	21/160	18/305	20/365	26/578	...
9/38	15/68	15/126	21/172	44/341	15/431	23/619	...

is chosen for computing the median or mean value, for example five points. In Fig. 14, the first computed value would be at X_3 position and would be the median or mean of Y_1 through Y_5 (plot position $X_3 − \hat{Y}_3$, where $\hat{Y}$ is the median or mean value of Y_1 through Y_5 at the X_3 plot position). The last value would be at the X_{27} position and the median or mean of Y_{25} through Y_{29} (plot position $X_{27} − \hat{Y}_{27}$).

7.3.1.4 *Scatterplot with LOWESS Smooth*—The LOWESS method of smoothing (locally weighted scatterplot smoother or robust locally weighted regression) uses an averaging algorithm. The procedure was developed by Cleveland (**10**). This is an important method that is computational intensive and requires the use of readily available computer routines for fitting at least 2^n weighted least squares equations[5] (**1, 10, 11, 41, 42**).

---

[5] Computer program available from Wadsworth Advanced Book Program, Monterey, CA.

NOTE 25—An abbreviated summary of the basic LOWESS method of smoothing is as follows (see p. 176 (**11**));

(*1*) For each pair of X_i, Y_i values (central pair), the LOWESS method produces a fitted value, $\hat{Y}_i$, using weighted least squares equations.

(*2*) The weight function (*f*) used in the computation of the weighted least squares value $\hat{Y}_i$ is defined by the investigator and is a real number between 0 and 1 that specifies what segment of X_i values (the values nearest to and including the central X_i, Y_i pair) will be used in the weighted least squares computation of $\hat{Y}_i$. For example, in a data set n, where n = 30, an *f* function value of 0.50 gives an integer value for 0.50 times 30 of 15. A window (envelope) surrounding the central X_i, Y_i pair includes 15 X_i, Y_i pairs, with the 15th X_i, Y_i pair at the boundary of the window. Those 15 values within the window are used to compute the $\hat{Y}_i$ value, using the weighted least squares equation.

(*3*) The least squares equation is weighted according to the distance of each X value within the window from the X_i value at the center of the window. The X_i value positioned at center (where $\hat{Y}_i$ is computed) has the greatest weight, while the value at the edge has the least.

(*4*) Smoothness of the fitted LOWESS line is controlled by increasing or decreasing the value of the weight function *f*. As the value of *f* is increased, more $X_i Y_i$ values are included into the weighted least squares computation that determines each value of $\hat{Y}_i$, consequently causing the resultant fitted line to become smoother. However, a larger value of *f* may not be the best fit.

(*5*) The best value of *f* can be determined by computing and plotting the residual values of $Y_i − \hat{Y}_i$. When the line formed by the plotted residual values is nearly straight and horizontal, the value of *f* is the best fit.

(*6*) The technique of determining the optimum *f* value could require many additional computations of the $\hat{Y}_i$ values and the related residual

values of $Y_i - \hat{Y}_i$, each time plotting and evaluating the resultant LOWESS curves. To be practical, the technique requires the availability of the LOWESS computer software.

(*7*) When relatively large data sets are involved in a scatterplot analysis, the LOWESS smooth is a logical analytical method of presenting the information needed to interpret the relationship between the two variables, *X* and *Y*. Individual LOWESS smooths of large multiple data sets (such as several data sets of arsenic and pH that originates from different regions **(13)**, can be conveniently compared to help in determining regional similarities and differences.

(*8*) The clarity of the scatterplot can be improved when large amounts of data are involved by not plotting the original X_iY_i plot points with the LOWESS smooths $(X_i\hat{Y}_i)$.

7.3.1.5 *Scatterplot With Polar Smooths of Samples (Data Sets) from More Than One Population*—This type of smooth, as used on a scatterplot, improves the visualization of multiple groups of data sets by enclosing a fixed percent (for example, 50 or 75 %) of each group with a mathematically determined ellipse **(1, 9)**. The polar smooth is a variation of the LOWESS smooth technique and requires the use of readily available computer routines for fitting at least 2^n weighted least squares equations.

NOTE 26—When a large number of values are included in each data set, although the plot symbols are different for each group, there can be great difficulty in visually separating the associated data sets to allow for comparison and interpretation. This separation can be improved by use of polar smooths of the data groups, that are described in the following summary **(1)**;

(*1*) The center of each data set is determined by computing the median values for both the *X* and *Y* parameters.

(*2*) All the other data points of the group are described by determining the angle and radius in relationship to the median values of *X* and *Y*, thereby placing the data into polar coordinates.

(*3*) A LOWESS smooth is performed on the data while in polar coordinates, and then the data are transformed back into the original units.

(*4*) While in polar coordinates, a specified percentage of the data are below the smooth line, for example 50 % (the 50 % of the data that was closest to the median *X* and *Y*).

(*5*) Those data values that were below the 50 % smooth line, after being transformed back to the original units, are then enveloped by an ellipse, the shape of which is dictated by the distribution of the plotted data values surrounding the median. This method of representing the data sets is termed the polar smooth.

(*6*) The same method can be used to determine a line enclosing the closest 75 % of the data, termed an outer polar smooth.

7.3.1.6 *Scatterplot With Absolute Differences Versus the Sample (Data Set)*—A scatterplot of absolute differences (absolute residuals) is a measure of the spread or deviation of the data set, and also a method of evaluating the smoothness factor, *f* **(1, 9, 11)**. This is a method of examining the variation of data throughout the range of values and assists in determining the repeatability of the laboratory analysis or field data collection. Of important use of the scatterplot is for the measure of the precision of the smoothness parameter *f* for the LOWESS smooths (see 7.3.1.4).

NOTE 27—This method of plotting the residual values can be used together with any of the techniques of smoothing. The preceding steps are followed;

(*1*) The smoothness values, $X_i\hat{Y}_i$, are determined for the previously determined data set, the absolute difference values.

(*2*) The R_i (residual) value is computed $(Y_i - \hat{Y}_i)$ for each Y_i.

(*3*) The R_i value is plotted against the related X_i creating the basic scatterplot.

(*4*) The X_iR_i values are smoothed using one of the previously discussed methods of mathematical smoothing.

(*5*) The smoothed curve is examined to determine whether the spread of the absolute residuals changes significantly over the range of the data.

(*6*) For the LOWESS smooth, if the smoothed curve formed by the plotted residual values is nearly straight and horizontal, the value of the weight function, *f*, is determined to be as the best fit.

7.3.1.7 *Scatterplot for Correlation Coefficient*—This type of scatterplot is a simple *X* and *Y* plot of two variables used as a visual tool to better understand the correlation coefficients. The correlation coefficient is a measure of the strength of association or goodness of fit between two continuous variables and is independent from the scatterplot. Three measures of correlation are in common use for water resources data, these are the ranked-based Kendall's tau and Spearmen's rho and linear-based Pearson's *r* **(1, 3, 14, 19, 26, 35, 43, 44)**.

NOTE 28—The three correlation coefficients are statistical information that is independently computed on ranked or linear data pairs that also can be plotted on a scattergram. For example, Kendall's tau equation, is shown as follows:

$$Kendall's\ (\tau)\ correlation\ coefficient = \frac{P - M}{n(n-1)/2} \qquad (18)$$

where:
P = number of times the *Y*'s increase as the *X*'s increase,
M = number of times the *Y*'s decrease as the *X*'s increase, and
n = total number of pairs (*X, Y*) of data samples.

The computations required to determine the correlation coefficient for small (≤ 10) and large (>10) sample data sets are given in pp. 209 through 220 of Ref **(1)**
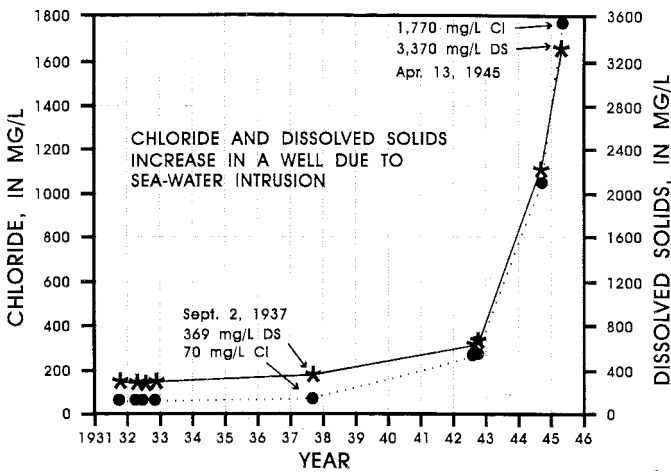
7.3.2 *Time- or Trend-Series Plot*—The time- or trend-series plot consists of a rectangular two-dimensional diagram with the independent time variable normally on the *X* axis and the dependent variable, such as Cl, on the *Y* axis. The plot position is represented by a symbol, such as a dot. The plot symbols can be connected with a line or unconnected. This type of scatterplot is used to reveal the ′trend′ or rate of increase or decrease in magnitude of a periodically determined variable over an interval of time **(1, 3, 8, 9, 19, 24, 31, 32, 45, 48, 49)**.

NOTE 29—Many variations exist for this type of plot, for example, the scale for the dependent variable can be in log units. A sliding scale can be used for multiple data sets of dependent variables having a wide range of values. Vertical lines projecting from the base of the plot to the variable position on the *Y*-axis can be substituted for the plot points. The independent time scale can be replaced by a distance variable, such as horizontal distance in the aquifer or vertical depth in a test hole **(49)**.

7.3.2.1 *Basic Time-Series Plot*—The basic time-series scatterplot has the independent variable, time, plotted against one dependent data set variable, such as Cl. Normally, the dependent variable will be determined at one ground-water site, for example, a well, over time (see Fig. 15).

7.3.2.2 *Times-Series Plot for Multiple Data Sets*—Time-series plots illustrating more than one data set are useful for comparing the similarities and differences of dependent variables of the same type over time, such as, the Cl and dissolved solid values from a ground-water site (see Fig. 15) **(1)**. The dependent variables can be parameters of different types that show fluctuations over time, for example, Cl values and water levels from one well. Each dependent variable can be identified by a separate symbol on the plot.
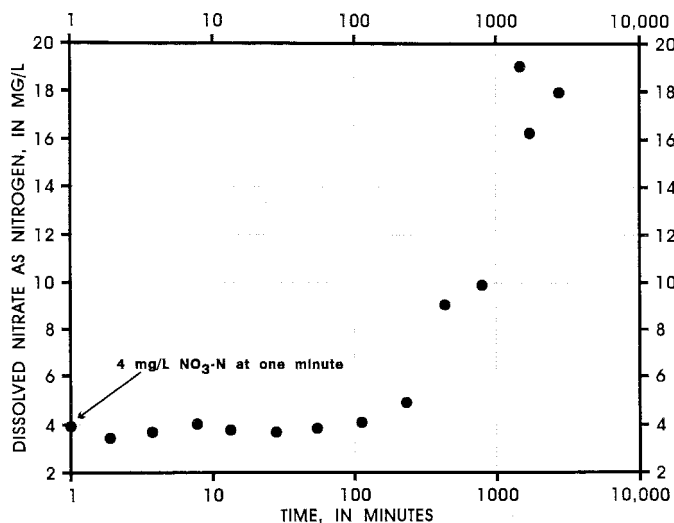
NOTE 1—Analyses from Ref (68).
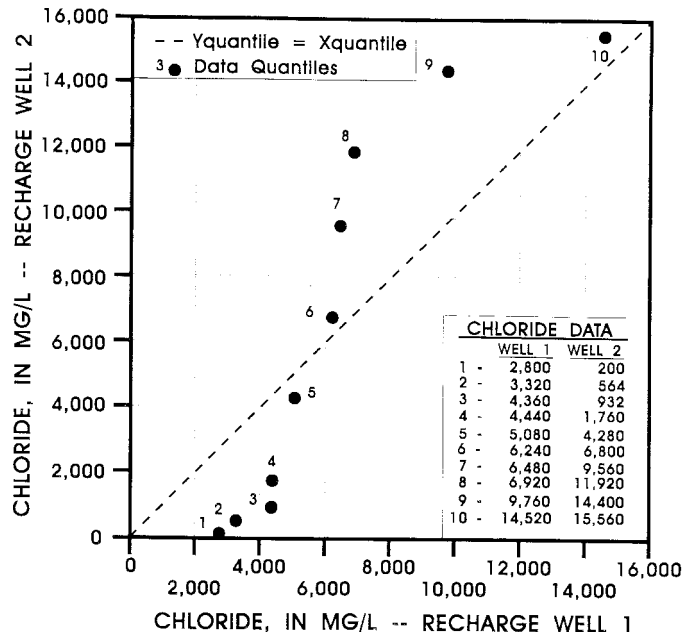**FIG. 15 Time-Series Plot**

7.3.4 *Q-Q Plots*—The Q-Q (quantile-quantile) diagram is a plot that displays the relationship between two sample data sets. Directly related data sets will plot as a straight line. Q-Q plots are similar to probability plots except that the X-axis and Y-axis are the paired values for two individually ranked sample data sets (see Fig. 17) (**1, 12, 19, 53**).

NOTE 31—If the sizes of the two data sets are equal, the data are ranked separately and the ordered values are plotted on the Q-Q diagram (X_1, Y_1 to X_n, Y_n). This was the method used on Fig. 18 and is simply a scatterplot of the paired values from the two data sets. The quantiles are involved when the sizes of the two data sets are not equal. When the sizes are not equal (n = size of smaller set, m = size of larger set), the quantiles for the n data set are computed ($p_i = (i - 0.4)/(n + 0.2)$) and are used to interpolate the values for larger data set by dividing the larger set into n equally spaced parts (**1**). These paired n and interpolated m data values are plotted on the Q-Q diagram.
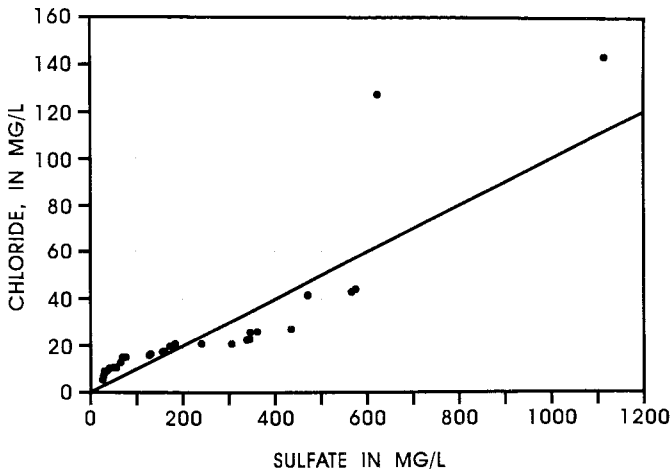
7.4 *Other Diagrams of Interest*—These include diagrams that graphically illustrate the results of an analytical examination of multiple sets of data (samples) selected from a number of chemical analyses of natural and human-influenced ground water and directly related parameters.

7.4.1 *Schoeller Nomograph or Vertical Scale Diagram*—This procedure is a semilogarithmic diagram where the concentrations of the ions are plotted as logarithmic values on the vertical scale. The related horizontal axis of the diagram has no numerical scale and represents the individual ions spaced at even intervals. These ions can be arranged along the horizontal axis in a user-preferred order. This type of diagram allows for the cation and anion constituents of many ground-water analyses to be shown on one diagram so that similarities and differences can be visually evaluated. Two versions of this diagram are in common usage.

7.4.1.1 The original version as presented by Schoeller has the vertical ion scale in meq/L logarithmic units (**38, 54, 55, 56,**

7.3.3 *Elapsed Time Plot*—The elapsed-time plots consist of rectangular two-dimensional diagrams with the independent elapsed-time dimension normally on the X-axis and the dependent variable, such as Cl, on the Y-axis (**1, 8, 9, 50, 51, 52**). This type of scatterplot is similar to the time-series plot for revealing the trend of a variable, however, the scale is divided into time intervals, such as minutes, hours, days, and so forth. The elapsed-time scale can be in logarithmic units (see Fig. 16) or in linear units with the zero (0) representing the beginning of a monitored event positioned on the left side of the plot. These plots are usually for a smaller real-time period than the time-series.

NOTE 30—Many variations exist for this type of plot, for example, the scale for either or both the independent and dependent variables can be in logarithmic units. Multiple scales can be used for the time interval, for example, the elapsed time and the actual date (as on the time-series plot). For comparison, multiple dependent variables can be plotted on a single diagram, for example, Cl values from several ground-water sites or Cl and water-level values from one well.

NOTE 1—Adapted from Ref (**52**).
FIG. 16 Elapsed Time Plot

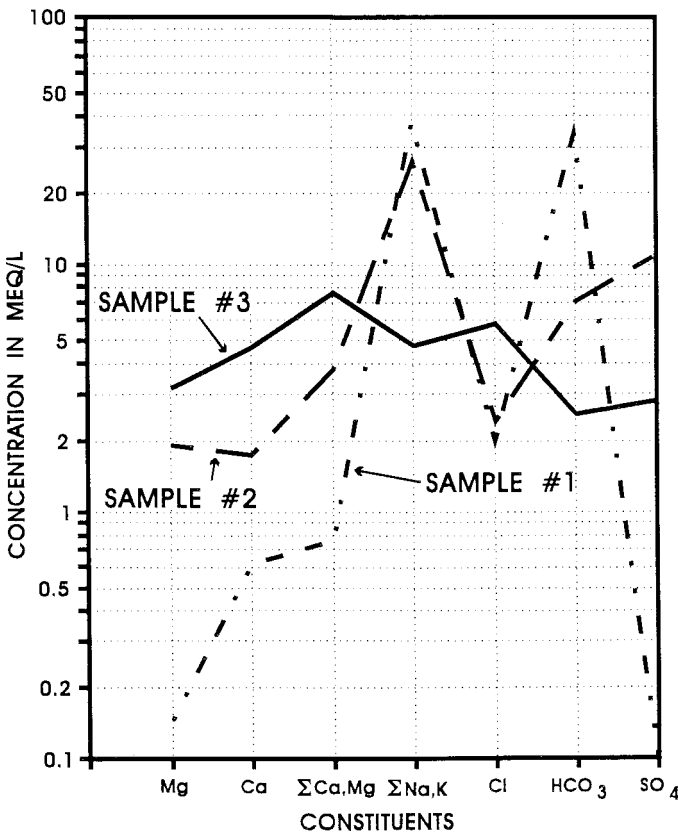NOTE 1—Analyses from Ref (**53**).
FIG. 17 Q-Q Plot

13

NOTE 1—Analyses from Ref (**65**), see Table 4.
FIG. 18 Q-Q Plot

57, 58). This method is preferable for additional calculations (**5**) (see Fig. 19).

7.4.1.2 The modified version, as given by Hem (adapted by Vorhis) (**3**), has two vertical scale units. One scale unit has the ion values in meq/L logarithmic units. The second scale unit has the constituents in mg/L logarithmic units with the scale for each ion shifted so that the mg/L units are horizontally parallel to the equivalent meq/L units (see Fig. 20).

7.4.2 *Irrigation Classification or Salinity Hazard Diagram*—This is a two-dimensional plot with the specific

NOTE 1—Adapted from Ref (**58**). Analyses from Ref (**3**).
FIG. 19 Semilogarithmic Diagram

CONSTITUENTS IN MG/L
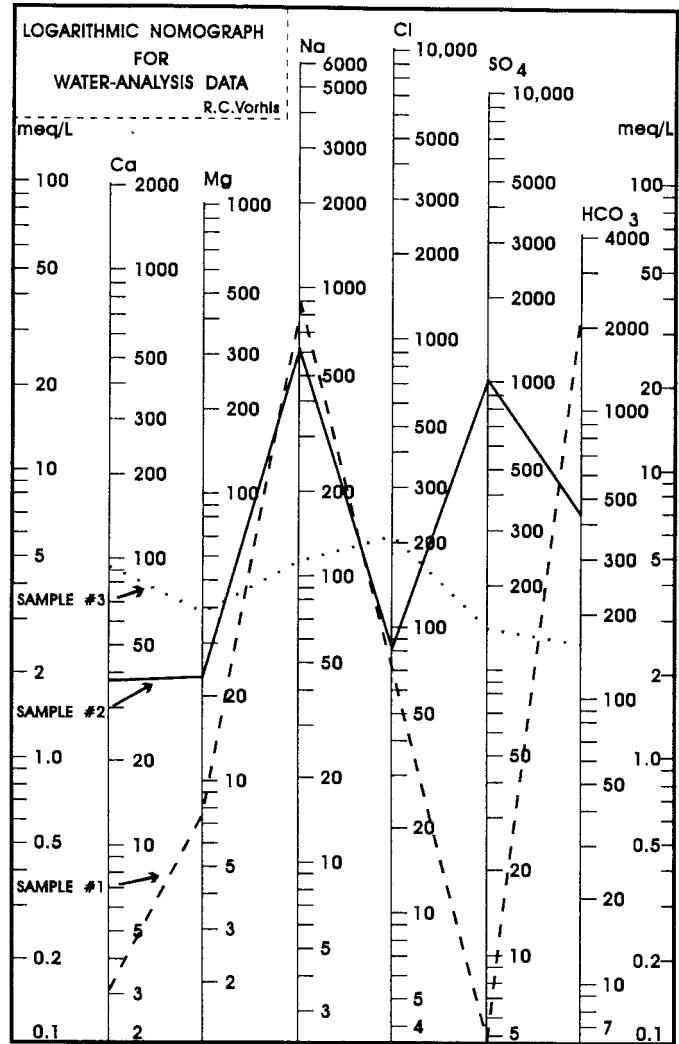
NOTE 1—Adapted from Ref (**3**). Analyses from Ref (**3**).
FIG. 20 Semilogarithmic Diagram

conductivity of the water sample on the *X*-axis and the' sodium-adsorption ratio' on the *Y*-axis. The diagram is divided into 16 areas used to categorize a water as to its suitability for irrigation. These areas are separated by the amount of salinity in the water and the effects of ion-exchange on the soil (see Fig. 21) (**3, 55, 59**).

NOTE 32—The sodium-adsorption ratio (SAR) was developed by the U.S. Salinity Laboratory (**59**) to predict the degree to which irrigation water tends to enter into cation-exchange reactions in soil. A high value for SAR implies a hazard of sodium replacing calcium and magnesium, a situation that ultimately damages the soil structure (**3**). The SAR is computed as follows;

$$SAR = \frac{(Na^+)}{\sqrt{\frac{(Ca^{2+}) + (Mg^{2+})}{2}}}, where\ ions\ are\ meq/L \qquad (19)$$

On the irrigation-classification diagram (see Fig. 21) analyses plotting in the Area S1/C1 are the most suitable for irrigation, while those in the Area S4/C4 the least suitable.

7.4.3 *Piper Diagram With Polar Smoothing*— This type of smooth, as used on a Piper diagram, improves the visualization
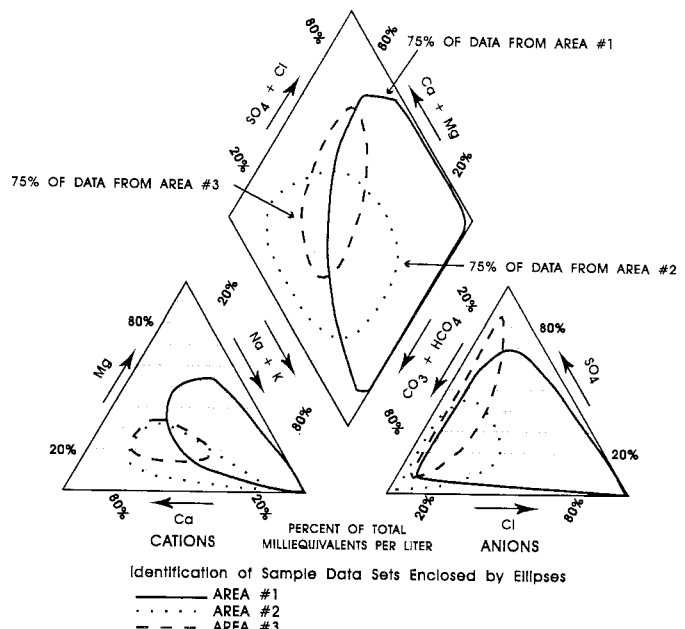
SPECIFIC CONDUCTANCE, IN MICROMHOS PER CM
AT 25 DEGREES CENTIGRADE

WELL NO.	SAR	SP.COND.
#1	12.1	660
#2	3.2	875
#3	8.5	1510
#4	2.4	2500
#5	0.5	997
#6	16.1	2430
#7	7.1	3940
#8	19.2	2880
#9	0.6	164
#10	0.8	511
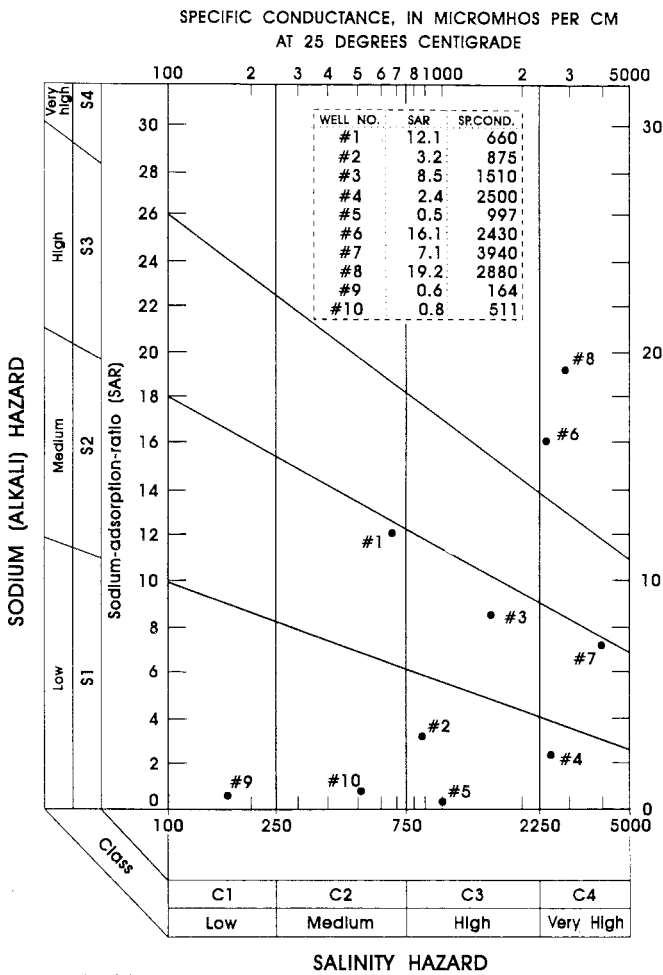
Note 1—Adapted from Ref (**3**). Analyses from Ref (**3**).
FIG. 21 Irrigation Classification Diagram

Note 1—Adapted from Ref (**12**).
FIG. 22 Piper Diagram With Polar Smoothing
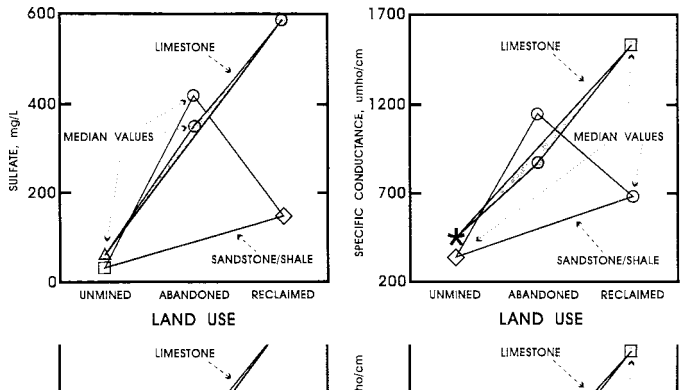
Note 1—Adapted from Ref (**60**).
FIG. 23 Three-Variable Pattern Plot

of multiple groups of data sets by enclosing a fixed percent (for example, 50 or 75 %) of each group with a mathematically determined ellipse (see Fig. 22). The polar smooth is a variation of the LOWESS smooth technique that requires the use of readily available computer routines for fitting at least 2^n weighted least squares equations (**1, 9, 10, 11, 12, 13**).

Note 33—The technique of polar smoothing is described in 7.3.1.5. Welch (**13**) used this method on Piper diagrams to examine ground-water analyses from three physiographic regions, each regional data set consisting of many analyses. The three regional ellipses computed by the polar smooth method were placed on one diagram allowing for easy comparison of the data sets.

7.4.4 *Three-Variable Pattern Plot*—This is a two-dimensional plot with a water quality variable on the *Y*-axis and a land use variable on the *X*-axis (see Fig. 23). The third variable is the rock type of the aquifer. The water-quality variable is scaled in standard units, the land use variable is evenly spaced (unscaled), and the rock type is shown by a pattern (**60**).

Note 34—The diagram, as given by Helsel (**60**), uses the median value of a number of samples for the selected water quality constituent from stream base flow (ground-water) plotted against three land use components (unmined, abandoned, and reclaimed basins). The third variable

consists of two rock types, limestone and sandstone/shale. In addition, the plot symbol shows those values that are significantly different from each other at the 5 % probability level, when compared to others in the same rock type or land use category (**60**).

7.4.5 *Three-Dimension Rotational Plot*— A pseudo-three-dimensional plotting procedure is available for visual observation on the screens of desktop computers. Three water quality variables are plotted in the three dimensions (*X, Y,* and *Z*). The interrelationship of these variables can be observed by rotating the plot on the computer screen (see Fig. 24) (**1, 61**).

7.4.6 *Ropes Three-Dimensional Diagram*— The Ropes diagram was developed to show the change in the quality of water over a period of time. The method can be thought of as a combination of the Schoeller-type diagram (*X*- and *Y*-axes) with a time-series plot in the third dimension (*Z*-axis) (see Fig. 25). This technique may be valuable for analyzing changes over time for a series of monitoring wells at a waste disposal site (**62, 63, 64**).
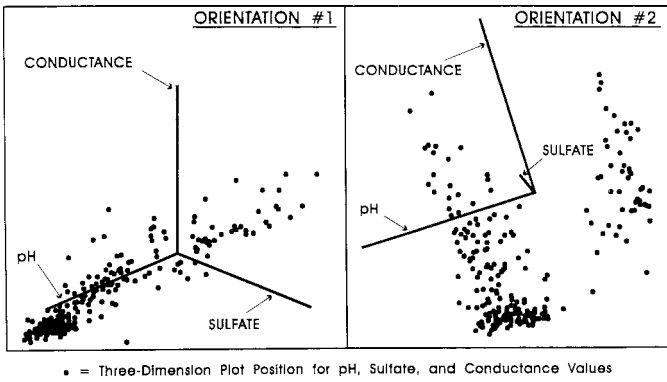
NOTE 1—Adapted from Ref (1). Data from Ref (60).
FIG. 24 Three-Dimensional Rotational Plots

NOTE 1—Adapted from Ref (62).
FIG. 25 Ropes Three-Dimensional Diagram

NOTE 35—The diagram has the logarithmic constituent value scale on the X-axis, an evenly spaced unscaled constituent identification on the Y-axis, and a scaled or unscaled time (or distance) on the Z-axis. The chemical constituents can be plotted in mg/L or meq/L units on the logarithmic scale. This is a time-specific series of diagrams, each with a number of site-specific analyses. These diagrams are interconnected by lines drawn between constituents from the same site (along the Z-axis). For example, on Fig. 25, for monitoring site Number 1, the plot position for the Na value on the earliest date diagram is connected to the plot position for the Na value of site Number 1 on the next dated diagram.

Fig. 25 is a condensed version of the diagram, as normally the logarithmic scale would have two additional log units (for example, 0.1, 1.0, 10.0, 100.0, and 1000.0) and would have eight or more chemical constituents. For sample data sets with many analyses, statistically determined values (for example, the mean, median, quartiles, sample standard deviation, LOWESS smooth, and so forth) can be plotted instead of the actual constituent values.

7.4.7 *Cluster Analysis Diagrams*—The cluster analysis is a procedure for arranging a number of objects in homogeneous subgroups based on their mutual similarities and hierarchical

relationships (16). This procedure was developed by taxonomists to deduce the lineage of living creatures from their characteristics and similarities (14). Geologists have used the method for the correlation of rock units and fossils (14). The use of the procedure to help in the interpretation of water-quality data is uncommon. Eriksson describes the method of computation of the correlation coefficients and construction of the diagrams (32).

NOTE 36—Eriksson shows two diagrams to represent the matrix of correlation coefficients of water-quality analyses. The first is a pseudo three-dimensional figure that he characterizes as a cluster description (32). The second is a dendrogram where the correlation coefficients are used for ranking and construction of the diagram (14, 32). This diagram shows the association of the ions by building up the clusters as a dendrogram. Of the two, the dendrogram seems more useful for the interpretation of ground-water quality data. An example taken from Eriksson is shown by Fig. 26. Davis includes listings of the FORTRAN IV procedures for computing the matrix of the correlation coefficients (14).
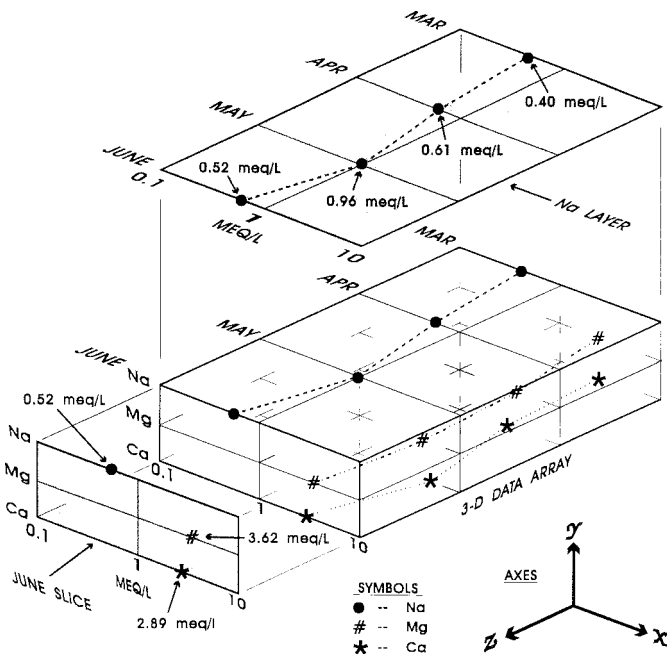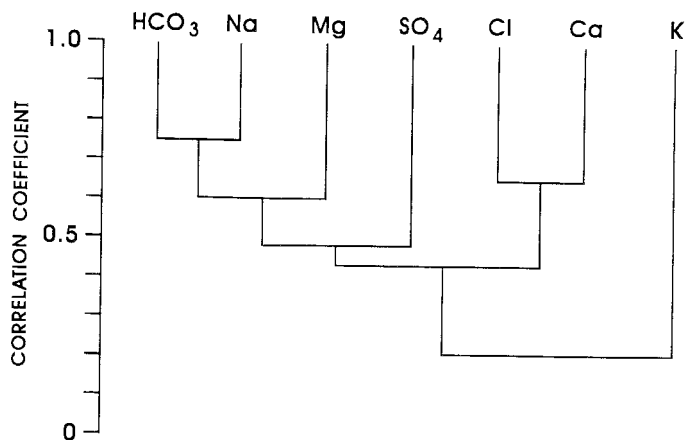
7.5 *Automated Procedures for Diagrams Based on Data Analytical Calculations*:

NOTE 37—Literature searches and verbal inquires found computerized procedures for the histogram, relative frequency histogram, scattergram, box and whisker, trend-series, time-series, polar smoothing for analyses plotted on the Piper trilinear, Schoeller semilogarithmic, and salinity hazard diagrams. Although procedures for the other water-analysis diagrams discussed in this guide are probably available, none could be found. Additional sources of computerized procedures can be added in the subsequent publication of this guide.

7.5.1 A relational ground-water database with an extensive collection of ground water, geology, and areal map procedures (GEOBASE 6.0) for desktop computers is available from Earthware of California. Included in this program package are routines for the display of the Schoeller semilogarithmic and salinity hazard diagrams (55).[6]

7.5.2 A package of software (ROCKSTAT) from Rockware Scientific Software[7] contains routines for the plotting of

---

[6] For further information contact Earthware of California, 30100 Town Center Drive # 196, Laguna Niguel, CA 92677.

[7] For further information contact Rockware Scientific Software, 4251 Kipling Street, Suite 595, Wheat Ridge, CO 80033.

NOTE 1—Adapted from Ref (32).
FIG. 26 Cluster Analysis Dendrogram

water-analysis diagrams on a desktop computer. Included are versions of the histogram, relative frequency histogram, and scattergram (23).

7.5.3 A package of water-quality diagrams is available through documentation from the USGS (29). Included in the package are versions of the box and whisker plots and of scattergrams.[8]

7.5.4 A comprehensive environmental data base management system (GIS/KEY) includes a version of the time-series diagram (48).[9]

7.5.5 Loftis, Phillips, and Ward, in 1989, developed a water-quality statistical software package, WQSTAT II, for microcomputers (31). These procedures use data in ASCII or Lotus format to produce box-and-whisker, histograms, trend-series, and time-series plots, along with summary statistics.[10]

8. Keywords

8.1 chemical analysis; chemical ions; graphical display; ground water; ground water quality; statistical analysis; water-analysis diagram

---

[8] For further information contact USGS Books and Report Sales Federal Center, Box 25425, Denver, CO 80225.

[9] For further information contact GIS Solutions, 1800 Sutter Street, Suite 830, Concord, CA 94520-2500.

[10] For further information contact Dept. of Agric. and Chem. Eng., Colorado State University, Fort Collins, CO 80523.

REFERENCES

(1) Helsel, D. R., and Hirsch, R. M., *Statistical Methods in Water Resources, Studies in Environmental Science 49*, Elsevier, Amsterdam, 1992.

(2) Tukey, J. W., *Exploratory Data Analysis*, Addison-Wesley, Reading, 1977.

(3) Hem, J. D., "Study and Interpretation of the Chemical Characteristics of Natural Water," *U.S. Geological Survey*, Water-Supply Paper 2254, 1989.

(4) van der Leedan, F., Troise, F. L., and Todd, D. K., *The Water Encyclopedia, Geraghty & Miller Ground-Water Series, 2nd Edition*, Lewis Publishers, Inc., Chelsea, 1990.

(5) Matthess, G., *The Properties of Groundwater*, John Wiley & Sons, New York, 1982, pp. 299–335.

(6) Davis, S. N., and DeWiest, Roger, J. M., *Hydrogeology*, John Wiley & Sons, New York, 1966, pp. 77–83 and 96–128.

(7) Rajagopal, R.," Influence of Outlying Observations on Selected Estimates of Parameters of Distribution," *Ground Water*, Vol 26, No. 3, 1988, pp. 325–332.

(8) Sara, M. N., *Standard Handbook for Solid and Hazardous Waste Facility Assessments*, Lewis Publishers, Boca Raton, 1993, pp. 11-1 to 11-72.

(9) Cleveland, W. S., and McGill, R., "The Many Faces of a Scatterplot," *Journal of the American Statistical Association*, Vol 79, 1984, pp. 807–822.

(10) Cleveland, W. S., "Robust Locally Weighted Regression and Smoothing Scattergrams," *Journal of the American Statistical Association*, Vol 74, 1979, pp. 829–836.

(11) Cleveland, W. S., *The Elements of Graphing Data*, Wadsworth Advanced Books and Software, Monterey, CA, 1985.

(12) Helsel, D. R., and Slack, J. R., "Diamond in the Rough: Enhancements to Piper Diagrams," submitted to *Journal of Ground Water*, will be available as *U.S. Geological Survey*, Open-File Report, 1994.

(13) Welch, A. H., "Geochemistry of Ground Water in Carson and Eagle Valleys, Western Nevada and Eastern California," *U.S. Geological Survey*, Open-File Report 93-33, 1993.

(14) Davis, J. C., *Statistics and Data Analysis in Geology*, John Wiley & Sons, New York, 1973.

(15) Spiegel, M. R., *Schaum's Outline of Theory and Problems of Statistics*, McGraw-Hill Book Co., New York, 1961.

(16) Bates, R. L., and Jackson, J. A., *Glossary of Geology, Third Edition*, American Geological Institute, Alexandria, 1987.

(17) Yevjevich, V. M., *Probability and Statistics in Hydrology*, Water Resources Publications, Fort Collins, 1972.

(18) Komor, S. C., and Anderson, H. P., Jr., "Nitrogen Isotopes as Indicators of Nitrate Sources in Minnesota Sand-Plain Aquifers,"
Ground Water, Vol 31, No. 5, 1993.

(19) Haan, C. T., *Statistical Methods in Hydrology*, The Iowa State University Press, Ames, 1977.

(20) du Toit, S. H. C., Styn, A. G. W., and Stumpf, R. H., *Graphical Exploratory Data Analysis*, Springer Verlag, New York, 1986.

(21) Tufte, E. R., *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, 1983.

(22) Sen, Z., and Al-Dakheel, A., "Hydrochemical Facies Evaluation in Umm Er Radhuma Limestone, Eastern Saudi Arabia," *Ground Water*, Vol 24, No. 5, 1986, pp. 626–635.

(23) Rockware, *ROCKSTAT Statistical Software*, RockWare Scientific Software, Wheat Ridge, CO, 1993.

(24) Nielsen, D. M. (editor), *Practical Handbook of Ground-Water Monitoring*, Lewis Publishers, Chelsa, 1991, pp. 555–558.

(25) Bachman, L. J., "Field and Laboratory Analyses of Water from the Columbia Aquifer in Eastern Maryland," *Ground Water*, Vol 22, No. 4, 1984, pp. 460–467.

(26) Robbins, H., and Van Ryzin, J., *Introduction to Statistics*, Science Research Associates, Inc., Chicago, 1975.

(27) Tukey, J. W. (Merriam, D. F., editor), "Some Further Inputs, in Geostatistics, a Colloquium," Plenum Press, New York, 1970, pp. 163–174.

(28) Rogers, R. J., "Geochemical Comparison of Ground Water in Areas of New England, New York, and Pennsylvania," *Ground Water*, Vol 27, No. 5, 1989, pp. 690–712.

(29) Briel, L. I., "Documentation of a Multiple-Technique Computer Program for Plotting Major-Ion Composition of Natural Waters," *U.S. Geological Survey*, Open-File Report 93-74, 1993.

(30) Barcelona, M. J., Wehrmann, H. A., and Varljen, M. D.," Reproducible Well-Purging Procedures and VOC Stabilization Criteria for Ground-Water Sampling," *Ground Water*, Vol 32, No. 1, 1994, pp. 12–22.

(31) Loftis, J. C., Phillips, R. D., Ward, R. C., and Taylor, C. H., "WQSTST II: A Water Quality Statistics Package," *Ground Water*, Vol 27, No. 6, 1989, pp. 866–873.

(32) Eriksson, E., *Principles and Applications of Hydrochemistry*, Chapman and Hall, London, 1985, pp. 124–163.

(33) Ritzi, R. W., Jr., Wright, S. L., Mann, B., and Chen, M., "Analysis of Temporal Variability in Hydrogeochemical Data Used for Multivariate Analysis," *Ground Water*, Vol 31, No. 2, 1993, pp. 221–229.

(34) Cunnane, C., "Unbiased Plotting Positions—A Review," *Journal of Hydrology*, Vol 37, 1978, pp. 205–222.

(35) Ullman, N. R., *Statistics, An Applied Approach*, Xerox College Publishing, Lexington, 1972.

(36) Waite, T. D., and Payne, T. E., "Uranium Transport in the Subsurface

Environment Koongarra—A Case Study, in Metals in Groundwater," Allen, H. E., Perdue, E. M., and Brown, D. S. (eds), Lewis Publishers, Boca Raton, 1993, pp. 363–393.

(37) Osmond, J. K., and Cowart, J. B., "Ground Water, in Ivanovich, M., and Harmon, R. S., (eds), Uranium Series Disequilibrium, Applications to Environmental Problems," Clarendon Press, Oxford, 1982.

(38) Collins, A. G., *Geochemistry of Oilfield Waters*, Elsevier, Amsterdam, 1975, p. 496.

(39) Dubrovsky, N. M., Neil, J. M., Welker, M. C., and Evenson, K. D., "Geochemical Relations and Distribution of Selected Trace Elements in Ground Water of the San Joaquin Valley, California," *U.S. Geological Survey*, Water-Supply Paper 2380, 1991.

(40) Fetter, C. W., *Applied Hydrogeology, Second Edition*, Merrill Publishing Co., Columbus, 1988, pp. 325–366.

(41) Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A., *Graphical Methods for Data Analysis*, PWS-Kent Publishing Co., Boston, 1983.

(42) Cleveland, W. S., and McGill, M. E., (editors), *Dynamic Graphics for Statistics*, Wadsworth, Inc., Belmont, CA, 1988.

(43) Kendall, M. G., "A New Measure of Rank Correlation," *Biometrika 30*, 1938, pp. 81–93.

(44) Kendall, M. G., *Rank Correlation Methods*, 4th edition, Charles Griffin, London, 1975.

(45) Kaiou, T., and Bing, J., "The Karst Thermal Mineral Water Resources and Their Medicinal Effects," *International Association of Hydrogeologists, Proceedings of IAH 21st Congress*, Vol XXI, Part 2, 1988, pp. 975–982.

(46) Gottman, J. M., *Time-Series Analysis: A Comprehensive Introduction for Social Scientists*, Cambridge University Press, New York, 1981.

(47) Drever, J. I., *The Geochemistry of Natural Waters*, 2nd edition, Prentice-Hall, Englewood Cliffs, 1988.

(48) GIS/Key, *GIS Solutions Environmental Database Management System*, GIS/Solutions, Concord, CA, 1993.

(49) Pucci, A. A., Jr., and Owens, J. P., "Geochemical Variations in a Core of Hydrogeologic Units Near Freehold, New Jersey," *Ground Water*, Vol 27, No. 6, 1989, pp. 802–812.

(50) Starr, R. C., and Gillham, R. W., "Denitrification and Organic Carbon Availability in Two Aquifers," *Ground Water*, Vol 31, No. 6, 1993, pp. 934–947.

(51) Gupta, S. K., Lau, L. S., and Moravcik, P. S.," Ground-Water Tracing with Injected Helium," *Ground Water*, Vol 32, No. 1, 1994, pp. 96–102.

(52) Eccles, L. A., Klein, J. M., and Hardt, W. F.," USGS Scientists Bring California Water Supply into Compliance with Federal Regulations," *Water Well Journal*, Vol XXXI, No. 2, 1977, pp. 42–45.

(53) Green, C., "Sea Water Intrusion in California, Appendix B, Laboratory Studies and Research Relative to Investigational Work for Prevention and Control of Sea Water Intrusion," *State of California, Department of Water Resources, Bulletin No. 63, Appendix B*, Report by Los Angeles County Flood Control District, Dec. 20, 1954, 1957, pp. 89–112.

(54) Forti, P., and Francavilla, F., "Hydrodynamics and Hydrochemical Evolution of Gypsum Karst Aquifers: Data from the Emilia Romagna Region," *International Association of Hydrogeologists, Proceedings of IAH 21st Congress*, Vol XXI, Part 1, 1988, pp. 219–224.

(55) GEOBASE, *The Relational Groundwater Database With Graphics Output*, Earthware of California, Laguna Niguel, 1992.

(56) Freeze, R. A., and Cherry, J. A., *Groundwater*, Prentice-Hall, Inc., Englewood Cliffs, 1979, pp. 237–302.

(57) Schoeller, H., "Utilité de la Notion des Exchanges de Bases pour la Comparison des Eaux Souterraines," *Société Géologie Comptes Rendus Sommaire et Bulletin, France*, Série 5, Vol 5, 1935, pp. 651–657.

(58) Schoeller, H., "Notions Sur la Corrosion Interne des Canalisations D'eau," *Ann. Ponts Chaussées Mem.*, Paris, No. 8, 1938, pp. 199–282.

(59) U.S. Salinity Laboratory Staff, *Diagnosis and Improvement of Saline and Alkali Soils*, U.S. Department of Agriculture, *Handbook 60*, 1954.

(60) Helsel, D. R., "Mine Drainage and Rock Type Influences on Eastern Ohio Streamwater Quality," *Water Resources Bulletin*, No. 19, 1983, pp. 881–887.

(61) Donoho, A., Donoho, D. L., and Gasko, M., *MACSPIN Graphical Data Analysis Software*, D^2 Software Inc., Austin, 1985.

(62) Ropes, L. H., Morgan, C. O., and McNellis, J. M.," FORTRAN IV Program for Synthesis and Plotting of Water-Quality Data," *Kansas Geol. Survey*, Spec. Distrib. Pub. 39, 1969.

(63) Lowell, B. H., Morgan, C. O., and McNellis, J. M.," Brief Descriptions and Examples of Output from Computer Programs Developed for Use with Water Data in Kansas," *Kansas Geol. Survey*, Spec. Distrib. Pub. 48, 1970.

(64) McNellis, J. M., Morgan, C. O., and Lowell, B. H.," Computer-Produced Tables, Maps, and Diagrams as Tools in the Interpretation of Brine Data from Southeastern Kansas," *Journal of Chemical Geology, Special Issue Geochemistry of Subsurface Brines*, Elsevier Publishing Co., Amsterdam, Vol 4, No. 1/2, 1969, pp. 303–324.

(65) Krothe, N. C., and Bergeron, M. P., "Hydrochemical Facies in a Tertiary Basin in the Milligan Canyon Area, Southwest Montana," *Ground Water*, Vol 19, No. 4, 1981, pp. 392–399.

(66) Velleman, P. F., and Hoaglin, D. C., *Applications, Basics, and Computing of Exploratory Data Analysis*, Duxbury Press, Boston, 1981.

(67) Hall, W. D., and Turk, L. J., "Aquifer Evaluation using Depositional Systems: An example in North-Central Texas," *Ground Water*, Vol 13, No. 6, pp 269–274.

(68) Poland, J. F., Garrett, A. A., and Sinnott, A.," Geology, Hydrology, and Chemical Character of Ground Waters In the Torrance-Santa Monica Area, California," *U.S. Geological Survey*, Water-Supply Paper 1461, 1959, p. 425.