1976

Estimation of low flows in semi-arid basins and related implications in water quality management

David Worden Hubly
Iowa State University

76-28,251

HUBLY, David Worden, 1935-
ESTIMATION OF LOW FLOWS IN SEMI-ARID BASINS AND RELATED IMPLICATIONS IN WATER QUALITY MANAGEMENT.

Iowa State University, Ph.D., 1976
Engineering, sanitary and municipal

Xerox University Microfilms, Ann Arbor, Michigan 48106

Estimation of low flows in semi-arid

basins and related implications in

water quality management

by

David Worden Hubly

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Civil Engineering
Major: Water Resources
Minor: Economics

Approved:

Signature was redacted for privacy.

~~In~~ Charge ~~of Major~~ Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1976

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

$Ac_{k\ell}$ = acres of crop ℓ served by irrigation system k

AW = available water for crops, in percent of total potential available water

B = the estimate of the regression coefficient

BWR = basic water requirements

CON1 = a model variable simulating the consumptive use of crops in areas directly tributary to the study surface stream system

CON2 = a model variable simulating consumptive use of crops in areas not directly tributary to the study surface stream system

d = the Durbin-Watson statistic

d_ℓ = a critical lower limit for the Durbin-Watson statistic

d_u = a critical upper limit for the Durbin-Watson statistic

DE_x = the ditch efficiency of ditch x

EHAT = a predicted residual used in variance equalization

e_i = a sample of the residuals

e_t = the dependent variable in the autocorrelation model

E_i = a population of residuals

E_t = evapotranspiration

E_{tp} = potential evapotranspiration

E_{tpc} = potential evapotranspiration for crop c

E_{tpr} = potential evapotranspiration for reference crop r

ETPH = a model variable simulating the evapotranspiration of phreatophytes

f = consumptive use factor, monthly

f_d = consumptive use factor, daily

F = a ratio or a statistic used in significance evaluations

IA_x = the irrigated acres served by ditch x in a specified study area

k = the number of dependent variables or a monthly consumptive use coefficient or an all purpose counter

k_c = a monthly consumptive use coefficient for crop c

k_p = a consumptive use coefficient for phreatophytes

k_t = a coefficient dependent on air temperature used by SCS to modify k_c values

K = a seasonal consumptive use coefficient or an all purpose summation quantity

K_{co} = a crop consumptive use coefficient

LE = lake evaporation

n_1 = number of positive signs

n_2 = number of negative signs

NTRWA = a model variable representing the daily application of water to cropland in areas not tributary to the study area's surface water system

NTSOST = a model variable synthesized from NTRWA and CON2 simulating soil moisture stress in areas not tributary to the study area's surface water system

P (or p) = percentage of annual sunlight hours occurring in any single month

PE Index = a special index used to estimate evapotranspiration

QDIS = a model variable representing the discharges in each segment

QDIS1 = a time lagged model variable representing the discharges in each segment using an assumed stream flow velocity of 15 miles per day

QDIS2 = same as QDIS1 except velocity is assumed to be 10 miles per day

QDIV = a model variable representing the diversions in a study segment

QDIV1 = a time lagged version of QDIV that assumes the stream velocity is 15 miles per day

QDIV2 = same as QDIV1 except velocity is assumed to be 10 miles per day

QIN = a model variable representing the daily inflows to each study segment

QINHAT = the inflows predicted by the model

QOUT = a model variable representing the daily outflows from each study segment

R^2 = the square of the multiple correlation coefficient

RE = reference evapotranspiration

SAS = Statistical Analysis System, a formal statistical package for regression analysis

SAWE = stream-aquifer water exchanges

SD = standard deviation

SPSS = Statistical Package for the Social Sciences, a second formal statistical package available for regression analysis

SROP = a model variable that simulates the surface runoff from precipitation

$SSinc_i$ = the increase in the explained sum of squares that occurs when variable i is added to the regression model

T = the T statistic and distribution coefficient

TA_x = total acres irrigated by ditch x

T_N = normal temperatures

TRWA = a model variable representing the water applied to areas tributary to the study area's surface water system

TSOST = a model variable, synthesized from TRWA and CON1, simulating the soil moisture stress in areas tributary to the study area's surface stream system

u = a monthly estimate of evapotranspiration

U = a seasonal estimate of evapotranspiration

u' = a monthly estimate of evapotranspiration for phreatophytes divided by 31

$\hat{u}_d$ = a daily estimate of a monthly crop consumptive use estimate

WD_{xt} = water diverted by ditch x on day t

Z = the Z statistic that is used in the analysis of runs, for the evaluation of randomness

α = a measure of significance

Δ (QDIS, QDIV, QIN, or QOUT) = the incremental difference between two consecutive observations of the Q variables

Δs = a set of variables representing or simulating the change in storage activity in the surface stream system

$\Delta s/\Delta t$ = the rate of change in the change in storage variable

δ^2/s^2 = the von-Neumann ratio

ρ = a parameter of the autocorrelation model

σ^2 = the population variance

σ = the population standard deviation

INTRODUCTION

The national concern for the water quality in our streams and rivers has increased the importance of low flow estimation. The Federal Water Pollution Control Act of 1965 [1] required the States to establish Water Quality Standards for surface waters, and the 1972 amendments to the FWPCA [2] (PL92-500) created a new planning regime called Water Quality Management (WQM) that relies on these water quality standards. Both Acts expect the water quality standards to be ultimately realized whenever any stream's flow exceeds some predetermined minimum flow, often called the Critical Low (or Critical Design) Flow (CLF). Furthermore, the water quality management process often employs prediction models which project the quality impacts of point and non-point discharges at the CLF. If the prediction model shows a violation of the water quality standards, one or more dischargers may be required to meet more stringent effluent standards. The resulting alteration and/or addition of processes can involve extremely large resource investments. Because the consequences of a faulty projection are great, the projection models must be accurate. Among the sources of error are the data and the method used to estimate the Critical Low Flow, and these are reasons for increased concern with low flow estimation. The customary method of studying and estimating low flows (1) involves the collection of low-flow discharges at a stream gaging station over a period of time, (2) assumes that these are naturally occurring events essentially unaffected by man's activities, or, if so, will remain relatively unaffected or unaltered in the future, and (3) includes the listing and review of the record minimum low flows, and a statistical analysis using a selected probability method. McKee and Wolf [3] reviewed in 1963 the historical development of low flow estimation, and Wolf [4] made an additional review following passage of the comprehensive 1965 legislation by Congress.

The available low flow estimation methods can produce accurate estimates as long as the assumptions underlying the techniques are essentially fulfilled. Such assumptions, however, are fulfilled far more effectively in the humid and subhumid basins commonly found east of the Missouri-Mississippi Rivers and along the Pacific coast than in the arid and semiarid areas of the western United States where many more constraints are encountered, both natural and manmade. As a result the relatively accurate low flow estimates obtained in the humid and subhumid areas may not be realized in the arid and semiarid areas. This study investigates these low flow estimation problems associated with the arid and semiarid areas.

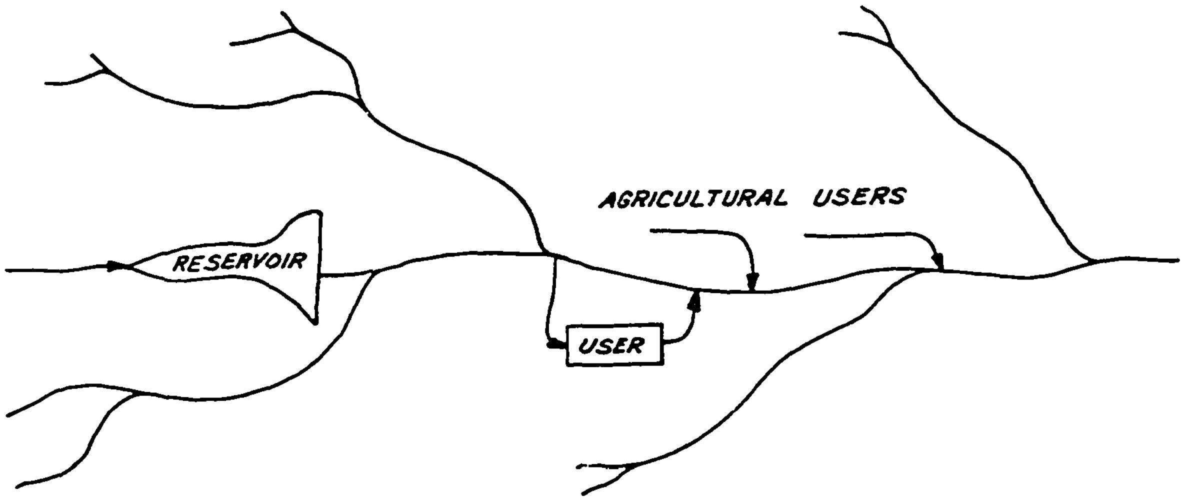
The nature of these low flow estimation problems is best defined by a comparison between the typical humid/subhumid hydrologic structure and the structure found in most arid/semiarid basins. Furthermore, the water scarcity in the arid/semiarid basins has spawned complex economic and institutional structures which create a unique set of low flow estimation problems and solutions. There exists also a basic inseparability of water quality and water quantity. Therefore, the first part of this introductory chapter presents a concise comparison of the typical physical-economic-institutional frameworks found in (1) humid/subhumid basins and (2) arid/semiarid basins. The second part of this chapter then presents and illustrates the specific problems related to low flow estimation.

The General Physical-Economic-Institutional Framework

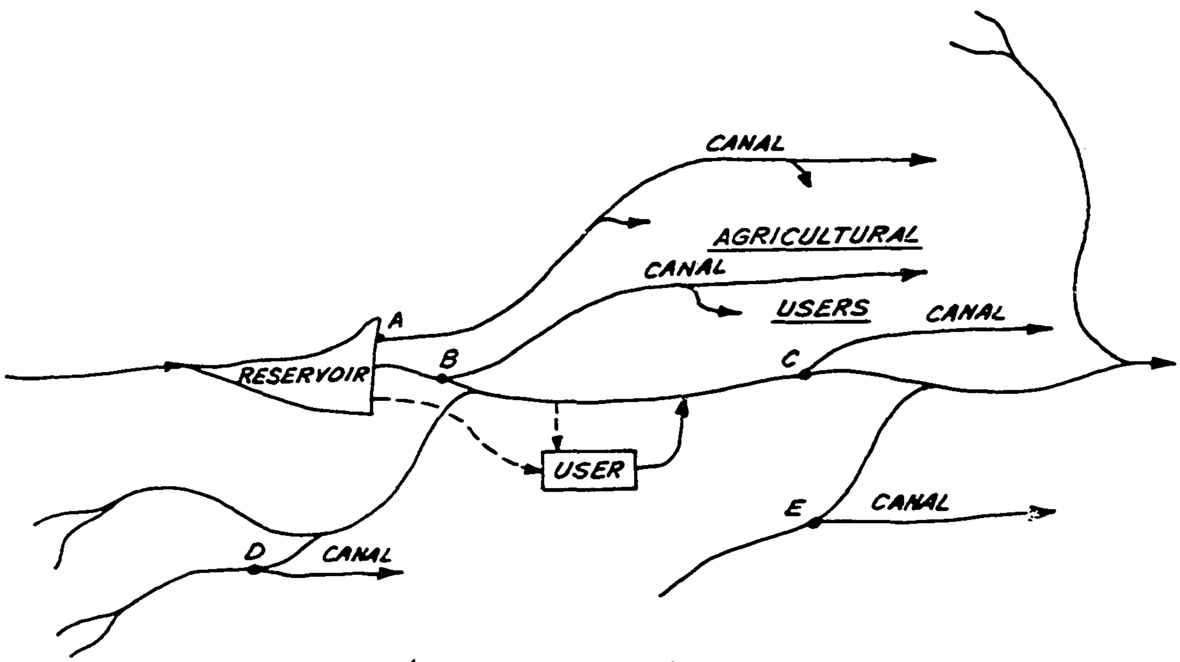
The physical framework

A very basic hydrologic difference between these two basin groups is illustrated with the dendritic patterns shown in Figure 1. In a humid/subhumid basin the water entering the surface system is predominantly surface runoff and groundwater exfiltration, and as the tributary drainage area is increased at successive downstream points the magnitude of the streamflow is also increased. These streams are characterized as continuously "gaining" streams. Alternatively, the arid/semiarid surface water system (downstream of mountain headwater sources) receives little groundwater or surface runoff and acts primarily as a conduit for water entering the system in the more humid headwaters or entering downstream as a result of a local but major precipitation event. Often the streams lose water to the groundwater system, resulting in a decrease in flow magnitude at successive downstream points. These surface systems are characterized as "losing" streams.

Reservoirs are common to both groups of basins, and the primary uses of many humid/subhumid reservoirs are often flood control and recreation. In arid/semiarid basins the primary purpose of a reservoir usually is to store water for use in the drier periods of the annual water cycle, and often for consumptive use. As a result the amount of water in storage will vary more dramatically in the arid/semiarid basins.



HUMID/SUBHUMID BASIN



ARID/SEMIARID BASIN

Figure 1. Typical basin dendritics

In addition water users are more likely to divert directly to use from a humid/subhumid stream because the dependable flow magnitude frequently is far greater than the user's need and storage is not required. However, in the arid/semiarid area most users divert from or to a reservoir or store their water in an upstream reservoir prior to use and bring the water to their diversion point as needed.

In the humid/subhumid areas agricultural users generally depend on precipitation for their crop moisture, and in many areas this group of users are primarily dischargers to the surface system because of the need to keep their croplands drained. However, in the arid/semiarid basins the agricultural users must convey their crop water from the surface and/or groundwater systems to their fields. These areas typically contain canal systems that divert large quantities of water at discrete points in the surface water system, e.g. points A, B, and C in Figure 1. The agricultural users also divert water from tributaries, e.g. points D and E in Figure 1, often in large enough quantities to prevent any tributary water from reaching the main channel of the system. This arid region water-use activity effectively breaks the total system into many disconnected subsystems. The agricultural activities in the dry basins usually return only a small portion of the diverted water to the surface system which amplifies the "losing" characteristic.

The conditions described above for the humid/subhumid basins produce a hydrograph of flow versus location along a stream bed that is a smooth continuous curve rising consistently along successive downstream locations. However, in the arid/semiarid basins this hydrograph is usually a broken, sometimes discontinuous, curve that declines along successive downstream points. Figure 2 illustrates the typical differences in these discharge hydrographs.

The economic framework

In humid/subhumid areas water quantity is a relatively minor factor in the production functions of most users because the supply of water usually exceeds the demand as long as the quality of the water exceeds some specifiable minimum level. Acquiring water for use seldom involves a cost to obtain ownership, and changes in use patterns usually produce minimal impact on other water users. However, in the arid/semiarid areas water always is an important production function factor and often limits the nature and/or magnitude of economic activities. Acquiring water usually involves an ownership cost, and the transferability of such ownership creates a water market. These interactions between the water and economic spheres permit significant economic impacts whenever water use patterns change. Any low flow estimation method having techniques that can account for recent changes in water use patterns will allow water quality planners to employ water use pattern changes as a management tool; and the technique will, thereby, create an economic impact.

The institutional framework

Both humid/subhumid and arid/semiarid basins have institutional structures involving the water sphere. The structure in the humid/subhumid basins is oriented towards the orderly allocation of water and is loosely connected to the economic structure. Since the water

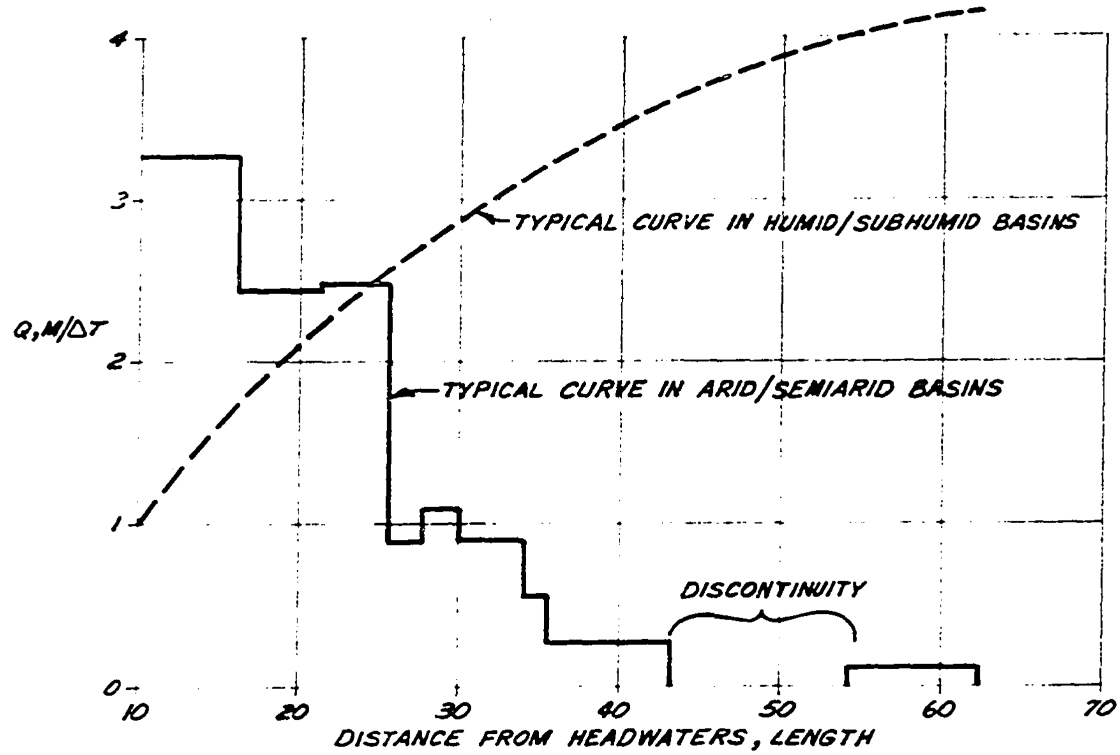


Figure 2. Hydrograph comparisons

supply usually exceeds the demand, the allocation system is not complex and requires minimal administration (the so-called "riparian" doctrine). However, in the arid/semiarid areas the institutional framework is large, complex, and well defined (the "prior-appropriation" doctrine). This framework serves two important purposes: (1) it establishes a transferable property right to use water which then becomes the base unit of the water market, and (2) it regulates the water flows and impoundments to implement the water allocation plan that is established by the water market. These functions require large complex legal and administrative systems which are not found in the humid/subhumid basins.

In summary, the flow magnitudes in humid/subhumid basins are primarily affected by natural events beginning with precipitation, and are minimally affected by the economic and institutional structures. In contrast, the flow magnitudes in arid/semiarid basins are strongly influenced by the economic and institutional structures. In other words, the flow magnitudes in the former basins have a strong stochastic element, and the flow magnitudes in the latter basins contain a strong deterministic, or nonstochastic, element. These basic differences explain why low flow estimation in arid/semiarid basins contains several serious problems not found in the more humid areas. Specific problem areas relating to the arid/semiarid regions are presented in the following discussion.

Low Flow Estimation and Associated Problems in Arid/Semiarid Basins

Each State has defined a Critical Low Flow, CLF, and the Environmental Protection Agency has defined CLF for numerous special cases, e.g. the industrial discharge standards required by PL92-500. The defined CLF's can vary from state to state, e.g. the critical time period in some states is limited to the dry, warm months of summer and fall while other states use a 12-month time period, including therein the winter months. However, most CLF's are defined in terms of duration and probability of occurrence. The most common CLF definition is the lowest average daily flow for a seven consecutive day period that will probably occur on the average once every ten years (the return period or recurrence interval). This flow is called the 7-day, 10-year low flow ($Q_{7/10}$). This CLF definition has evolved from the general acceptance of the following underlying low flow estimation principles.

1. The low flow should be defined over some duration of time because aquatic ecosystems can recover from most short term stress conditions such as those resulting from pollutant "shocks". The 7-day duration is usually expressed in regulations, and apparently is preferred because it will smooth instantaneous, diurnal, and weekly cycles but will respond to monthly, seasonal, and long term trends. Durations of 1, 2, 3, and 30 days have also been suggested as appropriate for low flow estimation.

2. The low flow must contain a risk element because the complete elimination of ecological damage caused by water pollution is economically

infeasible. The risk element is incorporated in the return interval which relates probability to frequency as follows:

$$P = \frac{1}{T_R} , \qquad (1)$$

where P is the probability of nonexceedance and T_R is the average return interval.

3. The length of the return interval is also limited by the nature of the available data "strings". Longer return intervals decrease the accepted risk but require longer data strings for estimation, resulting in fewer stream locations with sufficient record lengths to provide accurate estimates. Return intervals from 2 to 20 years are discussed in the literature, but the 10 year (or 0.1 annual probability) is almost universally accepted as the best compromise between accepted risk and data availability [3, 4].

The estimation of CLF's has been a problem throughout the history of water quality management [3, 4], and currently requires the use of frequency distributions. These methods have been well documented by Riggs [5, 6], the U.S. Corps of Engineers [7], and Yevjevich [8]. The estimation process involves three steps:

1. the selection of the best frequency distribution model,
2. the fitting of the data to the distribution model, and
3. the selection of the CLF from the developed mathematical or graphic model.

Steps 1 and 2 are complex, in a statistical sense, but the final step requires only the calculation or the graphical selection of the flow magnitude for a probability of exceedence equal to 0.9 (conversely, a probability of nonexceedence of 0.1).

A common frequency distribution model is a probability distribution function (PDF) that relates flow magnitude to the probability of exceedence (or nonexceedence) [8]. The PDF is sometimes called a cumulative probability distribution function, and occasionally the term cumulative frequency distribution is used as a synonym for PDF. However, the cumulative frequency distribution term is more appropriate for the description of a frequency analysis of a set of observations so the more precise "probability distribution function" is used herein because a model theoretically expresses the population distribution, not the sample distribution. Numerous PDF's are used in low flow estimation with the following distributions receiving the most attention.

1. Normal [8]
2. Lognormal [9]
3. Gumbel Extreme Value Type I [10]
4. Gumbel Extreme Value Type III [11]
5. Pearson Type III (or Gamma) [7]
6. Maximum Likelihood [7].

Discussions of the distributions can be found in the noted references. Distributions 1, 2 and 6 assume the flow values are normally distributed about a central value which is also the mode, median, and mean. However, stream flow is bounded by a minimum of zero and theoretically has no maximum bound which usually produces a distribution of values where the mode is less than the median, and both statistics are also less than the mean. Such distributions are called right (or positive) skewed, and this skewness characteristic is the reason for the development of distributions 3, 4 and 5 for low flow estimation. Skewed data can sometimes be transformed into a normal distribution by using the Naperian logarithm of each value and a lognormal distribution function.

A PDF can be fitted to a sample of low flow data graphically or mathematically. Both methods are often used in the same analysis to confirm the accuracy of the estimate. The graphical process involves the plotting of observed values followed by the visual fitting of a smooth curve. The ordinate scale includes either actual or logarithm values of the observed flows (low flow data are usually represented as a logarithm), and the abscissa is a probability (or return interval) scale designed for a specific probability function. The abscissa values assigned to each data point are calculated with a "plotting position" equation after the data have been arranged by magnitude. The following two equations are commonly used:

$$P = \frac{2m - 1}{2N} , \tag{2}$$

$$P = \frac{m}{N + 1} , \tag{3}$$

where P = plotting position, m = rank of the observation according to magnitude within data array, and N = the number of observations. Riggs [5] recommends equation 3 while the Corps of Engineers [7] uses both. Yevjevich [8] describes the desired characteristics of a plotting position equation and concludes that equation 3 best satisfies these characteristics.

Graphical techniques cannot be used for distributions having more than two parameters which limits its application to the normal, lognormal, and Gumbel Type 1 distributions. If the empirical data fit the distribution associated with the plotting paper, the resulting curve will be a straight line. Skewed data from a stochastic population will yield concave or convex curves. Irregular curves, e.g. "S" curves, are symptomatic of several different data inconsistencies.

Mathematical fitting techniques involve the estimation of distribution parameters from the empirical data and the subsequent calculation of the frequency curve coordinates using the assumed probability distribution function or factors derived from the function. This technique is appropriate for skewed data that remain skewed after transformation. The Pearson Type III distribution (arithmetic or logarithmic) is the distribution most commonly used with this technique, and low flow data are usually represented as natural logarithms.

The use of probability distribution functions as statistical models implies two basic assumptions about the process being modeled. First, the process is assumed to consist of random phenomena, and the empirical data represent a random sample of this stochastic population. Secondly, the process is assumed to remain unchanged over time (stationary) and over space (homogeneous). While no hydrologic process fulfills these assumptions perfectly, many hydrologic processes are considered to meet these assumptions because of a very slow rate of change and/or an apparent adherence to the laws of chance. The annual low flow events for essentially unregulated streams in humid and subhumid climates are considered to fulfill these basic assumptions because the low flow events are still linked to stochastic events such as rainfall and snowfall and the process is changing very slowly in a lengthy geologic time frame. However, in arid and semiarid climates the low flow events cannot be assumed to fulfill the basic assumptions for two reasons. First, the annual low flow events are usually not associated with a stochastic process such as rainfall but often are associated with a more continuous process of groundwater exfiltration or infiltration. Secondly, man's demand for water in arid and semiarid climates usually exceeds the supply. Subsequently, this has resulted in extensive regulation by man of the surface and the groundwaters. Because of this regulation, the low flow processes are subject to rapid change, and a time series of observations may be a collection of samples from several different populations of events. Therefore, the application of a frequency distribution for low flow estimation in an arid or semiarid basin can lead to a large potential error. This statement and the basic elements of low flow estimation are illustrated in the following example.

The minimum average daily flows for seven consecutive days in each calendar year from 1940 through 1970 have been calculated for USGS gaging station 07124000 and is presented in Table 1. The gaging station is located on the Arkansas River near Las Animas, Colorado. The plotting positions and ranks of each value are also shown in Table 1, and the plotting position was calculated with equation 3 above. The selection of the minimum flows and the Log Pearson Type III analysis presented below were accomplished with computer program A969,

Table 1. Annual 7-day low flows, USGS Station 07124000

Year	Flow, cfs	Rank[a]	Plotting position
1940	9.71	19	0.594
1941	13.10	15	0.469
1942	27.40	1	0.031
1943	17.90	10	0.312
1944	21.00	7	0.219
1945	16.40	11	0.344
1946	18.60	9	0.281
1947	21.70	6	0.188
1948	23.60	2	0.062
1949	23.30	3	0.094
1950	19.70	8	0.250
1951	15.90	12	0.375
1952	22.90	5	0.156
1953	9.60	20	0.625
1954	6.07	25	0.781
1955	7.00	23	0.719
1956	2.61	30	0.938
1957	6.20	24	0.750
1958	11.30	17	0.531
1959	4.46	28	0.875
1960	4.61	27	0.844
1961	8.03	22	0.688
1962	9.46	21	0.656
1963	3.34	29	0.906
1964	2.21	31	0.969
1965	6.00	26	0.812
1966	12.30	16	0.500
1967	10.30	18	0.562
1968	14.70	13	0.406
1969	14.30	14	0.434
1970	22.90	4	0.124

[a]This analysis was based on the probability of exceedence, e.g., flows equal to or greater than the specified 7-day low flow magnitude; this ranking is opposite of that used for nonexceedence, where the lowest magnitude is given the rank of 1, etc.

Streamflow Statistics, that is available through the USGS [12]. The following statistics were also calculated:

Mean = 13.12 cfs

Standard Deviation = 7.295 cfs

Skewness = 0.232

Mean of Logs = 1.033

Standard Deviation of Logs = 0.300

Skewness of Logs = - 0.705

The magnitude of the skewness is in the middle of the range of skewness values found among the gaging stations in the Arkansas basin in Colorado that had data strings equal to or greater than 10 years.

The data in Table 1 are plotted on a lognormal grid in Figure 3, and a smooth curve has been visually fitted to the plotted points. The concave downward shape of the curve and the median value being larger than the mean are typical of negatively skewed data. In this example, the log transformation did not eliminate the original data skewness. The changing curvature of the curve may be indicative of an aberration in the data and provides a warning of potential error.

The data in Table 1 are used also to estimate the parameters of a Log Pearson Type III frequency curve, and the computed coordinates are also plotted and connected with a smooth curve in Figure 3. The computed curve does not fit the observed data very well, and the two curves diverge at the high and low flow values. The CLF ($Q_{7/10}$) is estimated from each curve in Figure 3, and the estimates differ by 16% of the larger computed number. If no other potential error is present

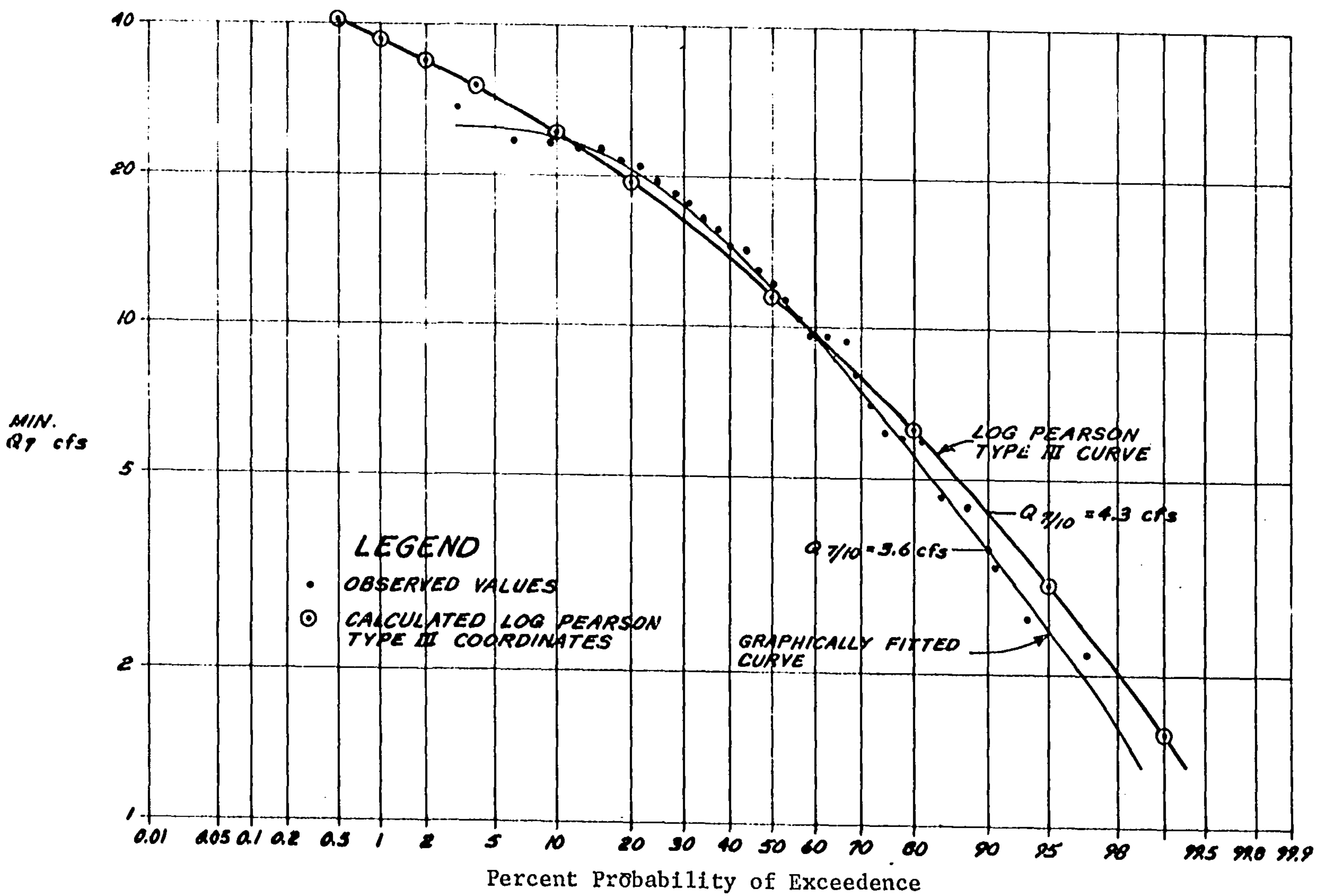


Figure 3. Frequency curves for USGS Station #07124000

this amount of error can be accepted as typical of a hydrologic analysis; however, the additional errors are present in arid/semiarid basins.

The problems noted above may be explained, at least partially, in Figure 4 where the annual 7-day low flows have been plotted in the order of occurrence. The plot suggests that a significant change occurred in the low flow process after 1952, when the mean annual 7-day low flow dropped from 19.3 cfs (1940-1952) to 7.8 cfs (1953-1969). Both time periods contain wet and dry years, so the change probably is related to events in extensive irrigation practices above and below this gage. The last data point, 1970, may be indicating a return to the earlier flow magnitudes, but the number of observations is not sufficient yet to draw any firm conclusions. The change in basin conditions causing the 1952 change could not be identified, but the latter change, if it is confirmed, is probably related to sale of the Las Animas Town Ditch water rights and the subsequent movement of the point of diversion many miles upstream. The data string appears to be time variant which accounts for part of the lack of fit and estimation error demonstrated above, i.e. the process is nonstationary.

Since man's activities in arid and semiarid basins can alter the low flow process, long data strings will probably include samples from several dissimilar populations. This raises serious questions about the usefulness in such areas of an established hydrologic "sacred cow", the long data string. If frequency distribution techniques are to be used for low flow estimation on regulated streams, estimation accuracy can be improved by using shorter but more recent segments of the long data strings even though many degrees of freedom are lost.

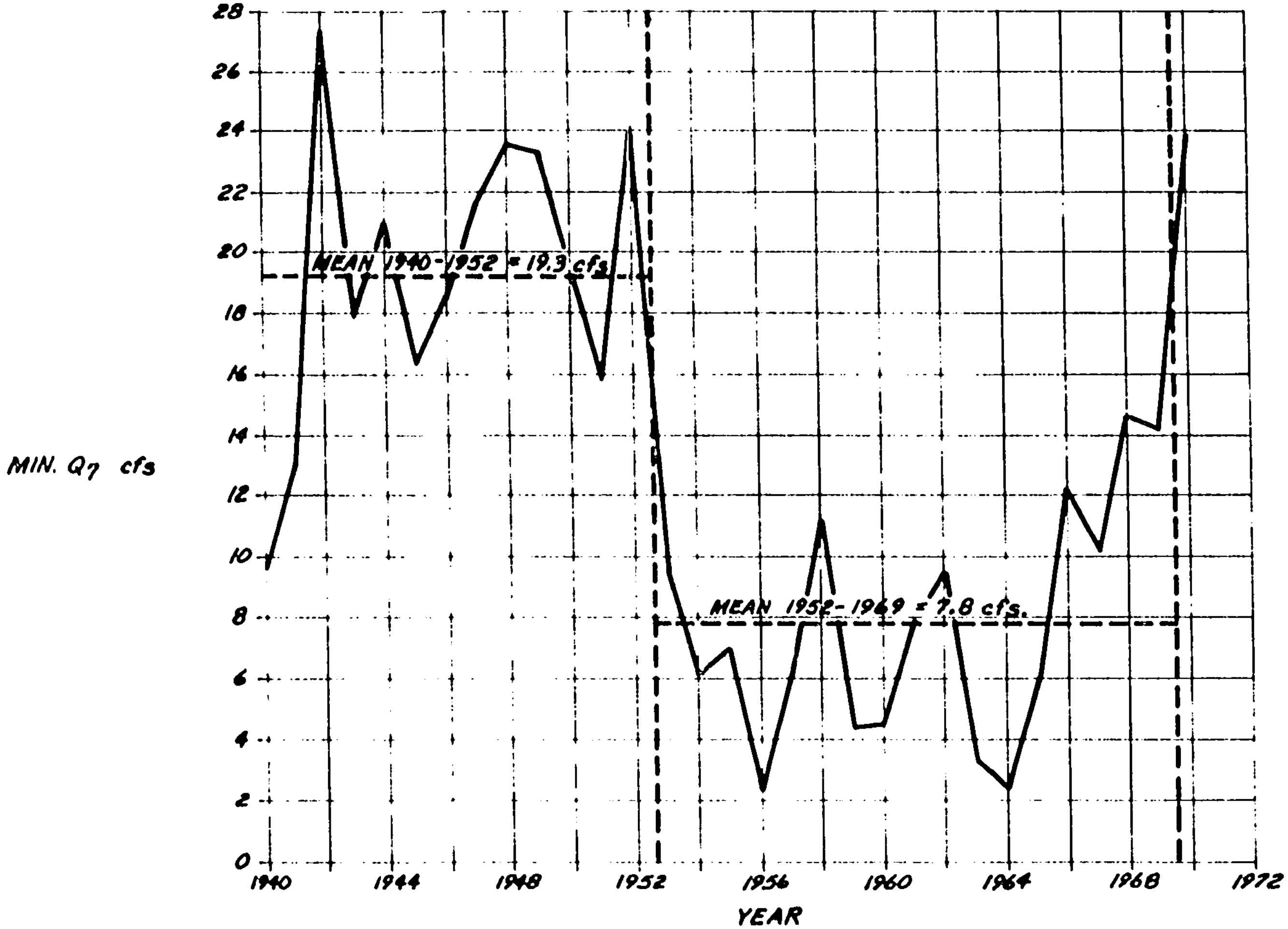


Figure 4. Annual 7-day low flows in order of occurrence — USGS Station #07124000

But the time variant problem is not the only low flow estimation problem encountered in arid and semiarid areas because irrigation and other regulation activities are not uniform along the axis of streams. For example, intensive irrigation usually produces overdrafts on both surface and groundwater supplies during low flow periods; and, therefore, the surface streams will be losing flow to the alluvial aquifer. Nearby stretches[1] or reaches with little irrigation activity can have a reverse surface-groundwater relationship. As a result, the data from a gaging station affected by the first stretch will represent a different low flow process than the data from a gaging station affected by the second stretch or reach. This line variant character of the low flow process makes the use of regional approaches (where a single set of distribution parameters are estimated from several data strings in a region) subject to large potential error. Line variant processes are called nonhomogeneous, as defined by Yevjevich [8].

The line variant element introduces potential error, but the nature of stream regulation and irrigation diversions can introduce even greater potential errors. Estimates of CLF are usually required for effluent discharge points located between two gaging stations so interpolation between the two gage estimates is required to develop the estimate. Increases and decreases in stream flow usually occur as discrete steps along the stream's flow line, but a straight line

---

[1]A stream "reach" is herein defined as the stream segment between two adjacent points where a discrete hydraulic or hydrologic characteristic changes, and a "stretch" is a series of reaches between two specified boundaries. The stretches are usually bounded by confluences or gages. In this dissertation the boundaries are gages.

interpolation between stations assumes a uniform continuous change. Therefore, in the vicinity of large discrete flow changes, the low flow estimates will have large errors unless specialized interpolation methods are used.

Low flows in arid and semiarid basins are often very small numbers. Zero flows are not uncommon. The smallness of these flows has two implications in low flow analysis. First, zero flows preclude the direct use of logarithmic distributions which are the preferred distributions. The USGS developed a variation of the Log Pearson Type III analysis that estimates the parameters with the zero flows excluded and then adjusts the distribution to account for the missing zero flows. However, a description of the procedure could not be acquired by the author so the statistical validity of the procedure has not been evaluated. The second problem caused by the small magnitude of the low flows is an increase in the potential error in the gage data. Many gages in arid and semiarid areas must record a broad range of flows, and the streambeds are often wide and flat so that control sections for low flows are difficult or impossible to construct. As a result, the stage-discharge relationship at the gage can change dramatically over short time periods, introducing additional errors in the data and yielding less reliable information for low flow analysis.

This introductory discussion has defined a serious problem area concerning low flow estimation in arid and semiarid climates. The problems arise from the nonstationarity and nonhomogeneity of the low flow processes, discrete flow changes of large magnitudes, and low flow measurement errors. The research documented in this dissertation presents an alternate approach to low flow estimation in arid and semi-arid areas and develops part of the new tools required to implement the alternative approach. Any new approach must satisfy several criteria that are considered desirable in low flow estimation techniques. These criteria are:

1. the duration element must be included,

2. the probability of occurrence element must also be retained,

3. the technique must not be so complex and expensive that it will not be used by water quality management agencies and their personnel,

4. the method in so far as possible must use available data, and

5. the method must be responsive to changing conditions.

DEFINITION OF THE RESEARCH PROBLEM

The goal at the start of the study was to investigate all elements of the problem and its solution (physical, economics, and institutional) with the end product being a complete and implementable low flow estimation method for arid/semiarid regions. As the study progressed it became apparent that it would not be possible to achieve this goal within the constraints of available time and funds. Therefore, the immediate goal of this study became a technical one, and efforts were focused on that part of the problem. The remaining portions of the problem were placed in a future research context. The following discussion develops three research areas within the main subject area, and a general scope of study is developed for the two future research projects. Then a detailed definition is developed for the research reported in this first phase of the overall problem area.

The Total Research Problem

An accepted approach to research involves the proposal of a hypothesis concerning a problem followed by the testing of this hypothesis. The basic hypothesis for this low flow estimation problem has been developed from a set of conditions often found in arid and semiarid basins but seldom found in humid and subhumid basins.

In arid and semiarid basins the river flows are usually regulated according to a "water law" and a set of "water rights" that have been legally established according to the water law. As a result of this regulated condition, the flow (and especially the low flow) passing any particular point is often determined by the nature of the water rights downstream. Thus, in the estimation of low flows in regulated basins it may be more appropriate to utilize downstream variables instead of the upstream variables. This concept is distinctly different from the humid/subhumid basin approach where flow magnitudes are related to numerous upstream hydromorphic and geomorphic factors. Several studies have evaluated some of these relationships [13, 14], and Orsborn [15] has presented a concise review of these methodologies.

Three classes of water law are found in the United States which allocate water according to (1) the Riparian Doctrine, (2) the Prior Appropriations Doctrine, or (3) administration of legislative mandates. The second Doctrine is the most common approach found in arid and semi-arid areas in the U.S. and will be the only method considered in this study. The many variations among State water laws is documented by Hutchins [16].

Under the Prior Appropriation Doctrine, the water law recognizes the right to divert water for beneficial use with priorities established among rights on the basis of initial time of diversion. A diversion right (the senior right) has priority over all rights (the junior rights) that were initiated later in chronological time as established in a court of law. These priorities are the basis for the regulation of river flows by controlling which rights are permitted to divert water. A regulatory authority establishes a "call date" which is the initial appropriation date of the controlling right, i.e. the right that will divert all of the streamflow available at its point of diversion. The controlling right is often only partially fulfilled.

The call dates vary with the available flow and can change daily. Call dates are often changed intuitively by the authority, such decisions being based on years of experience. Furthermore, the date can also be changed when a diverter is not receiving the magnitude permitted by his full right and places a "call on the river". Calls by senior rights force junior rights to stop diverting. So the streamflow passing a particular point during low flow conditions usually consists of the flow elements necessary to satisfy senior downstream rights. The basic hypothesis underlying this research is that low flows in regulated basins can be estimated from a conjunctive analysis of the set of water rights used in the regulation of the basin and the hydrologic and hydraulic characteristics of the surface water system. However, this hypothesis is constrained to include only those regulated basins where at least one diversion or a minimum flow requirement establishes a controlling flow sink. This condition is commonly experienced in over-appropriated basins during low flow conditions.

A low flow estimation procedure incorporating water rights should include two basic analytical steps.

Step 1. Predict the controlling water right(s) and identify all senior rights that would also be diverting water.

Step 2. Estimate the low flow at the desired point using a water balance model for the stream segment between the estimation point and the nearest downstream controlling water right. This procedure was chosen for further study because it fulfills the five criteria for a new estimation method described in the previous chapter. The duration and probability elements will be included in Step 1, and the response to changing conditions will be included in Step 2. Furthermore, the method utilizes available data and available analytical techniques. However, tools are not currently developed for either step; and, therefore, each step raises several research topics.

Future Research Projects

Separate research efforts can be pursued in (1) the development of a Step 1 methodology and (2) the examination of the economic implications of implementing this new estimation procedure. General scopes of study for both areas are discussed below.

Step 1 study

Two separate methods for accomplishing Step 1 have been identified. These are:

A. Apply a frequency analysis to the call dates of record, and

B. Correlate a key gage flow with the call date record and use a frequency analysis on that gage's data string.

Applying a frequency analysis to the call dates will require the evaluation of fit between available distributions and the data. Such research would also require the evaluation of methods for including the duration element. Furthermore, a 0.1 nonexceedence probability of a flow may not correspond to a 0.1 probability of occurrence for a specific call date so this question would also need evaluation. Also, call date data may not be as readily available or as accurate as gage data in some basins. Because of these potential problems the second method may be the preferred approach.

A question that arises in considering method B is how to select the best gage for correlation to the call date data. One criterion should be to select a gage that has a historical data string showing minimum effect from regulation practices; however, this criterion may force selection of a gage far upstream from the stretch being analyzed which might adversely affect the accuracy of the analysis. At the beginning of this study a correlation between USGS gage No. 07099400, near Pueblo, Colorado, and the call date record for the Arkansas River basin was attempted, and the results showed little correlation. The cause of the lack of correlation appeared to be the large quantities of imported water[1] being transported past this gage in the river. Imported water is stored in the reservoirs in the headwaters of the basin and is released on demand for river transport to the diversion point of the water owner. Each release is called a "river run", and several river runs are usually in progress at any time during the common critical low flow periods. Reservoir release records are available which could be used to predict river run hydrographs at gage stations. Subtracting these predicted hydrographs from the gage data should produce an estimate of the natural flow passing the gage. The accuracy of the estimate will depend on the accuracy of the hydrograph model. Luckey and Livingston [17] have developed a routing model for the Arkansas River in Colorado that estimates the reservoir release hydrographs, and additional development of this model is currently in

---

[1] "Imported water" is water that has been diverted across a basin divide, often called transmountain diversions. Imported water is common in basins with high water demand and a scarce supply.

progress. Since the river run hydrograph estimation is a key element in developing a usable procedure for Step 1, further work on the Step 1 problems is postponed until the Luckey and Livingston model is fully developed.

Economic Study

The connection which exists between the water quality management activity of low flow estimation and the water quantity allocation systems developed under the Prior Appropriation Doctrine creates a research need in the economics area with support from the legal and engineering fields. The research should evaluate:

1. the impact of the introduction of water quality management into the water allocation system on the socio-economic structure of the arid/semiarid regions, and

2. the water quality management activities in which water allocation modifications will be a feasible alternative.

If the water quality management role in water allocation always appears to result in a regional or national social benefit/social cost ratio less than one then the implementation of a more responsive low flow estimation technique will be inadvisable. On the opposite end of the spectrum changes in the water allocation system may never be economically feasible alternatives so the adoption of the new technique will depend only on the benefit/cost relationships within the water quality area. In between these two extremes lie many combinations of water quality management policy and socio-economic impacts through the water allocation system, and clarification in this middle area is the general purpose of this second phase, economic-institutional study.

Water quality and water allocation can interface in several ways if the low flow estimation technique can respond immediately to changes in the allocation system. One mechanism involves relocation of senior water right diversions to a point below the effluent discharge point, thereby increasing the amount of flow receiving the discharge. Changes in beneficial use and consumptive use would not be required to implement this mechanism so water quality management would not have to become a legally defined beneficial use. This mechanism would undoubtedly require a payment by the discharger to the water rights owners to accomplish the relocation or the outright purchase of the necessary rights by the discharger. A second mechanism can develop if water quality becomes a legally defined beneficial use. Both mechanisms would probably include the development of a new "minimum flow right" concept which can be generally defined as a right to have a specific flow magnitude met or exceeded at a discharge point whenever the minimum flow right is in priority. Owners of junior minimum flow rights might also affect the allocation system by blocking the relocation of senior rights that would adversely affect their right. This mechanism will require only the Step 1 analysis of the two step low flow estimation procedure, but its development is improbable because several constitutional, judicial, and legislative actions will be required. Since the first mechanism requires little change in the existing allocation systems the following discussion of the economic study assumes the first mechanism is the implementation interface.

The economic study should include two distinct study phases. Initially the study should develop predictive tools which can then be used in a normative approach to water quality management. This second phase will probably be limited to a demonstration of how to use, in a normative manner, the predictive tools developed in the first phase.

The predictive phase of the study will first require the development of two cost models for 1) estimating treatment and/or process costs as a function of treatment effectiveness and 2) estimating the cost of acquiring water rights. The first cost model will be relatively easy to develop because the Environmental Protection Agency (EPA) has published extensive reference material in this area.

However, the second cost model will require considerable collection of data, and the definition of the model structure will be more complex. The past research in this area is limited and difficult to locate so this cost model will require the evaluation of past water right transfers. The value of a water right that is being purchased for readjudication is primarily a function of three factors, 1) the right's historical yield record, 2) the amount of water the right has historically used consumptively, i.e. the amount of water not returned to the surface water system, and 3) the temporal occurrence of that consumptive use. The historical yield record can be stated in terms of frequency of occurrence for various quantities, e.g. average annual yield in ac. ft./year; and this analysis can be developed for recorded right transfers from readily available data. The consumptive use can be calculated using an engineering analysis based on known past uses of the water right or can be drawn from the court records involving the historical transfers if the testimony adequately covered this area. These records will also determine or establish the temporal use period. The amount of consideration involved in each transfer can probably be obtained from the court records or, more likely, extracted from the county revenue tax records. Evaluation of this element of the cost model may be difficult because many water right transfers are complex, e.g. several parties may be involved or nonmonetary consideration may be appraised at a level above or below true market value. Furthermore, water right transfers can include large legal and engineering costs that must be estimated. The mathematical model that best represents the functional relationship between these three cost elements must also be determined in the study. The development of this cost model for the valuation of water rights is considered to be a sizable research effort by itself.

Once these two cost models are developed, the predictive phase of the economic study must relate the two activities, water quality management and water rights transfer, to regional resource allocation. This critical second step will require the development of an allocation model that can predict the distribution of capital, labor, and resources with various levels of competition for water rights from the water quality management sector. The structure of this allocation model is not intuitively obvious, so an initial task in this research element will be the selection of the best analytical procedure. Three approaches will deserve serious consideration, and they are: 1) optimization, 2) simultaneous equations, and 3) simulation models. At this early stage of problem definition the linear programming optimization model appears to be the preferred analytical tool because the allocation process is usually defined in terms of optimizing some personal benefit such as profit, but the application of the econometric simultaneous equation approach on a regional basis may prove to be easier to apply or more effective. These three approaches have large data requirements so the data bank limits may force the use of a less data intensive regional approach.

When the allocation model is completed the normative phase of the study can then provide some water quality management guidelines. The predictive tools can be used to evaluate the following water quality management impacts in terms of rational and/or regional norms [18]:

1. the effect on the value of goods and services,
2. the effect on economic efficiency,
3. the effect on income distribution and the nature of the labor force,
4. the effect on the stability of the regional economic base, and
5. the effect on current demographic trends.

The qualitative and/or quantitative definition of these effects is the desired end product of this second phase, the economic research plan, and the first phase results can then become the input to decision processes in both the water quality management and the water allocation areas.

Current Research Problem

The basic hypothesis underlying the research into Step 2 is the acceptance of an hydrologic budget model as a suitable water balance model. Hydrologic budget models are based on mass balance analysis of all or part of an hydrologic unit such as basins, aquifers, or flowing and standing surface water systems. The hydrologic budget model developed in this study uses a stream segment as the basic hydrologic unit.

A physical model of a stream segment is depicted in Figure 5. The inflow to the segment from the adjoining upstream segment is designated QIN[1], and this inflow, the dependent variable, is estimated from a function including all of the other inflows and outflows, the independent variables, shown in Figure 5. The segment outflow designated QOUT is the ultimate sink for the stream segment flow and will be either a known diversion or a known minimum flow requirement. The flow QDIS represents the known or accurately estimated discrete discharges serving as additional inflow to the stream segment. These additional inflows would include inflows from gaged tributaries, industrial and publicly-owned wastewater discharges, and any other discharge which has been measured or can be accurately estimated. Conversely, the flow QDIV represents the known or accurately estimated diversions which include irrigation, municipal, and industrial diversions. The Distributed Inflows shown along the upper surface edge of the stream segment represent the unknown or difficult to estimate inflows which include

---

[1]Whenever possible descriptive acronyms are employed that are compatible with the variable designations used in the computer analysis.

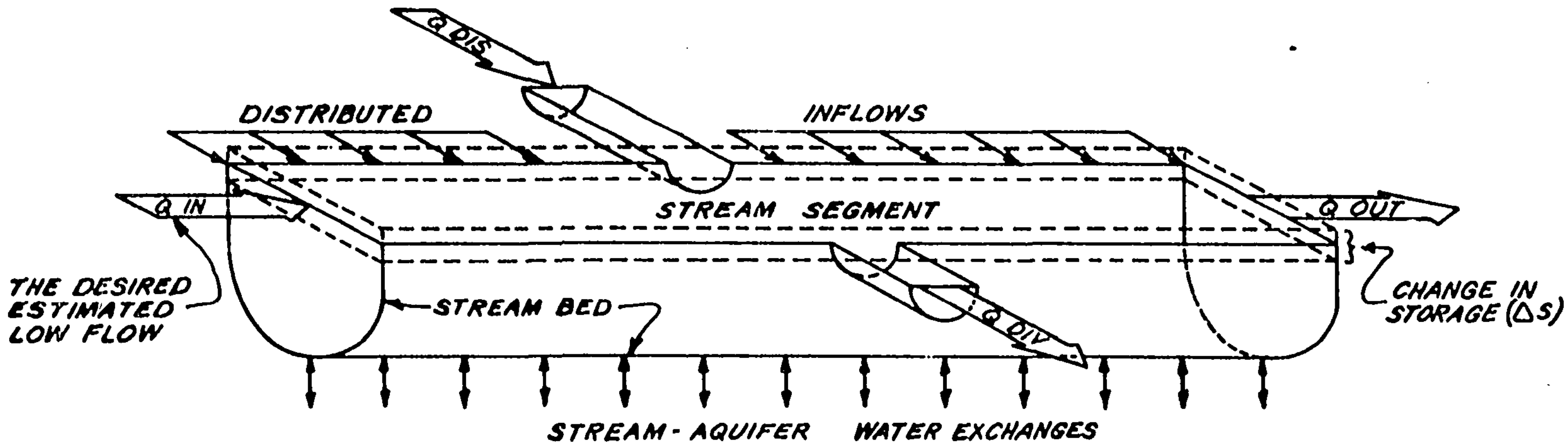


Figure 5. Physical structure of water balance model

surface runoff from precipitation or irrigation tail water[1] that enters either the stream segment or the ungaged tributaries, and waste irrigation water discharged by canals. The Stream-Aquifer Water Exchanges (SAWE) represent either groundwater infiltration or groundwater exfiltration occurring at the wetted stream bed surface and are also distributed along the stream bed length. Because the stream segment surface area is small during low flow events (excepting segments containing active reservoirs), the inflow from precipitation on the stream surface and the outflow from stream surface evaporation are assumed to be negligible and are not included as separate variables in the model. This assumption is reasonable because any impact from these two variables will be assimilated into the Distributed Inflow variable(s).

The stream bed and the planes denoting each end of the stream bed are fixed surfaces, but the stream surface can rise or fall. As a result of stream surface changes, the amount of water stored in the stream segment changes, and this potential change in storage, Δs, is also represented in Figure 5. Low flow conditions are typically near an inflow-outflow equilibrium so the Δs variable should assume a value of approximately zero at low flow; however, the estimation of model parameters must utilize data drawn from a broader range of flows which will include a Δs effect, so the basic model must include this variable.

---

[1] Irrigation tail water is water applied to a field that runs off before it can evaporate or percolate into the soil.

The mass balance equation used in hydrologic budget models is:

$$\Sigma \text{ Inflow} = \Sigma \text{ Outflows} + \Delta \text{ Storage} \tag{4}$$

where increases in storage are positive (+) values and decreases are negative (-) values. The variable groups shown in Figure 5 are substituted into equation 4, and the equation is rearranged to produce:

$$\begin{aligned} \text{QIN} = {} & \text{QOUT} + \text{QDIV} - \text{QDIS} - \text{DISTRIBUTED INFLOWS} \\ & + \text{STREAM-AQUIFER EXCHANGES} + \Delta\text{S} \end{aligned} \tag{5}$$

Exchanges from the stream to the aquifer are assumed to be positive (+) by the model. Equation 5 is used in this study to test the basic hypothesis stated on page 33.

If equation (5) proves to be an acceptable water balance model, the model will fulfill the "response to changing conditions" criteria (#5) in the preceding chapter. For example, assume a low flow QIN has been estimated using the active diversions selected in Step 1 as the independent variables QOUT and QDIV and a calibrated water balance model from Step 2. Subsequent to the estimation of the low flow QIN one of the diversions is moved a considerable distance upstream or downstream from the location of QIN, thereby altering the true low flow QIN. A reasonably good estimate of the effect of the diversion relocation can be obtained immediately by removing that diversion from the independent variables, and this revised estimate can be improved by recalibrating the water balance model as soon as a year or two of new data can be collected. Changes in discrete discharges, land use, and irrigation practices can also be incorporated in a similar manner.

Three criteria were selected to test the basic Step 2 hypothesis stated above. These criteria are:

1. the expected error must fall within the typical range of hydrologic predictions, usually 20-25% of the mean, and should approach the expected error range of low flow estimation methods in current use,

2. the signs (+ or -) of the estimated parameters must conform with the relationships shown in equation 5, and

3. the model must be stable in the low flow ranges, as evidenced by reduced variance between observed and predicted values.

The utilization of these criteria is discussed more completely in the next two chapters.

PROCEDURES

The basic procedure followed in this study is to select the preferred analytical method and then estimate the parameters of a water balance model for a portion of an existing regulated river basin. The adequacy of this sample model is then evaluated in terms of the criteria presented in the preceding chapter, and this evaluation is presented in the next chapter. This chapter discusses the selection of the analytical method, the selected basin segment, the development of the independent variables and the analytical techniques utilized in developing and evaluating the water balance model.

Selection of Analytical Method

Three analytical methods were considered for this study,

(1) Simulation,

(2) Mathematical Programming, and

(3) Multiple Regression.

Simulation involves the development of a mathematical model based on an intimate knowledge of the simulated system(s). Simulation models often contain several smaller models of well-defined subsystems, and this aggregation feature makes simulation a versatile technique that can analyze a macro system composed of dissimilar micro systems. For example, simulation can be used to analyze a system containing both linear and nonlinear submodels. In addition simulation models can include feedback or feedforward relationships and optimization submodels. Simulation usually involves a three-step process. First, the composite model is developed from a knowledge of the basic system. Then the model is calibrated using existing data so that the model produces a dependent variable string similar to the observed values when the model is stimulated with the observed independent variable strings. Finally, the calibrated model is stimulated with lengthy data strings generated from probability distributions and random numbers. The probability distributions are developed from the existing data strings of the independent variables. Hillier and Lieberman [19] present an overview of simulation methods.

The versatility of simulation is usually the reason for its selection as an analytical method; however, the method also contains several features that make it unattractive for this study. First of all, the method requires intimate knowledge of the subsystems, and many of these subsystems in a river segment cannot be defined with enough accuracy to justify the use of simulation without incurring large data acquisition costs. Furthermore, the calibration of a simulation model involves the discrete changes of individual parameters within a sequence of simulation runs; and this process continues until the simulated dependent variable string best fit the observed values. While statistical methods can be used to define "best fit", the method does not guarantee an optimum "best fit" in a multiparameter model unless all combinations are investigated. As a result, simulation is often a very expensive method. The calibration process also does not provide an estimate of parameter sensitivity so an additional sensitivity analysis must be performed. The calibration process also does not provide an estimate of the adequacy of the basic model structures. As a result, a model that adequately simulates an observed data string may produce large errors when extended beyond the conditions that produced the observed data string. Because of these analytical and cost problems, simulation was not selected as the basic analytical method for this study.

Multiple regression and mathematical programming are both optimization methods (the latter is actually a group of optimization methods) that can provide "best estimates" of the water balance model parameters using some preselected optimization criterion. Multiple regression uses the optimization criterion that the best parameter estimates are provided by the model that minimizes the sum of the squares of the differences between the predicted and observed values of the dependent variable corresponding to each set of independent variable data. This criterion is called the "least squares" criterion.

Mathematical programming contains several analytical methods that can utilize other optimization criteria. For example, linear programming could be used with the optimization criterion of minimizing the sum of the absolute value of the difference between the observed and predicted values of the dependent variable. Mathematical programming also contains a method, quadratic programming, that can use the same least squares criterion used in multiple regression. However, this additional flexibility in choosing an optimization criterion requires compromises in the nature and amount of information that can be extracted from the application of most mathematical programming methods. For example, in applying a mathematical programming technique the parameters of the water balance model would be the variables in the mathematical programming model, and the variables of the first model would become the parameters of the second model. Since the parameters of the mathematical programming model would be random variables, the analysis should estimate the effect of the probability element, but such an analysis is very awkward in even the best developed mathematical programming methods. Similarly, the optimal solution of the mathematical programming model would provide point estimates of the desired parameters but would not provide significance or confidence interval estimates without a considerable amount of extra analysis. As a result of these compromises, the use of mathematical programming would also be expensive, and the examination of several water balance model configurations would be very time consuming.

Multiple regression, on the other hand, does provide several statistics about the parameter estimates in an efficient manner including the two mentioned in the above example. These statistics are discussed later in the presentation of the techniques used in this study. Furthermore, the computer software required to perform a regression analysis is also better developed and more readily available than the software required for a mathematical programming analysis. Finally, no strong argument exists for the adoption of an optimization criteria other than least squares; and, given certain assumptions about the variables, the least squares criterion adds three desirable characteristics to the analysis, which are:

1) the least squares parameter estimates will be the minimum variance unbiased estimates among the class of linear estimates (this is the Gauss-Markov property),

2) the least squares estimates may also be the maximum likelihood estimates, and

3) the method using least squares can be extended to include nonlinear models.

Some of the assumptions required to realize these characteristics are not fulfilled by stream gage data so these advantages cannot serve as the basis for choosing the least squares criterion for this study, but the possibility of incorporating even a portion of these features does encourage such a choice.

As a result of all of the above considerations and after reviewing the nature of the research problem, the linear multiple regression method is selected as the basic analytical method for this study. The regression procedure is used in the following model development process.

1) Select a sample basin that contains most of the problems usually associated with regulated basins as discussed earlier and construct a water balance model for the sample basin.

2) Select a time period(s) from which observations can be drawn and a time interval between observations.

3) Construct one or more variables from available data in the sample basin and selected time period that will represent each of the model elements defined by equation 2 in the previous chapter.

4) Estimate the parameters in equation 2 using multiple linear regression with each set of variables for a unit time period representing a single observation.

5) Evaluate the estimated parameters in terms of the criteria presented in the previous chapter using the tools described below.

The Study Sample Area

A segment of the Arkansas River in Colorado was chosen as the sample basin, and the location of the segment is shown in Figure 6. The study segment is bounded by USGS Gage 07109500 located near Avondale, Colorado, and by USGS Gage 07130500 below John Martin Dam near Hasty, Colorado. This river segment is a desirable study subject for several reasons.

1) The basin is intensively irrigated near the river from the Royal Gorge above Canon City to the Colorado-Kansas boundary. This irrigation has created a very large demand for water resulting in decreed water rights for more than seven times the basin's average annual yield. As a result water scarcity is a common event. The river is often reduced to zero flow by senior water rights.

2) The study segment is located near Pueblo, Colorado, which is a very large growing industrial center. The expanding water needs for industry and the associated population has created a very competitive market for water rights which is a desirable condition for economic studies.

3) This segment lies within the area influenced by two large water resource developments, the John Martin Reservoir and the Fryingpan Arkansas project. Since water resource development is

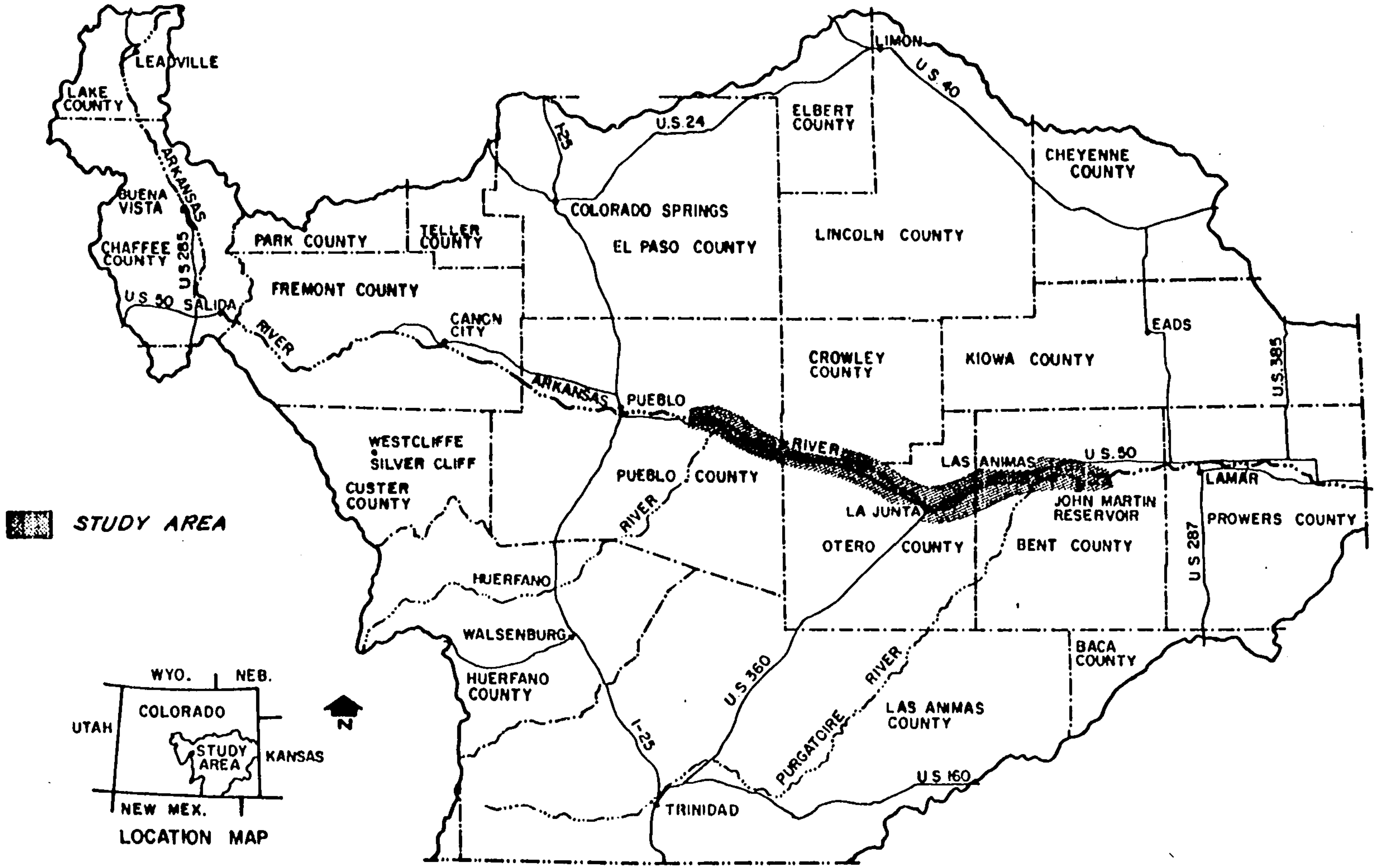


Figure 6. Location of sample study area

common in scarce water areas, the study segment shares this common element with the general population of regulated basins.

4) The intense competition for water and the water resource developments have spawned several hydrologic studies of this basin which provide a good source of data for this study. Some of these studies are discussed below.

The study segment is presented in more detail in Figure 7 with the river flowing from left to right. The study segment was divided into five subsegments that are designated as Stretches 1 through 5. The stretch boundaries are defined by USGS gaging stations, and throughout the study the upstream gage data in each stretch is used as the stretch inflow, QIN, and the downstream gage data is used as the controlling sink, QOUT. In the four downstream stretches, the QIN data is also the QOUT data for the adjacent upstream stretch. Table 2 lists the USGS data. In the balance of this dissertation gage identification numbers are shortened to the four middle digits shown in parentheses in Table 2. The use of five stretches instead of a single long segment permits the estimation of paramters for each individual stretch or for the whole segment.

The Arkansas River and its tributaries in the study area are incised into sedimentary rocks of the Upper Cretaceous period. The river has carved a wide, gently sloping valley and has filled the valley with alluvial deposits which form a large aquifer that is hydraulically well connected with the surface waters. The irrigated areas are generally confined to the area overlying the aquifer so the

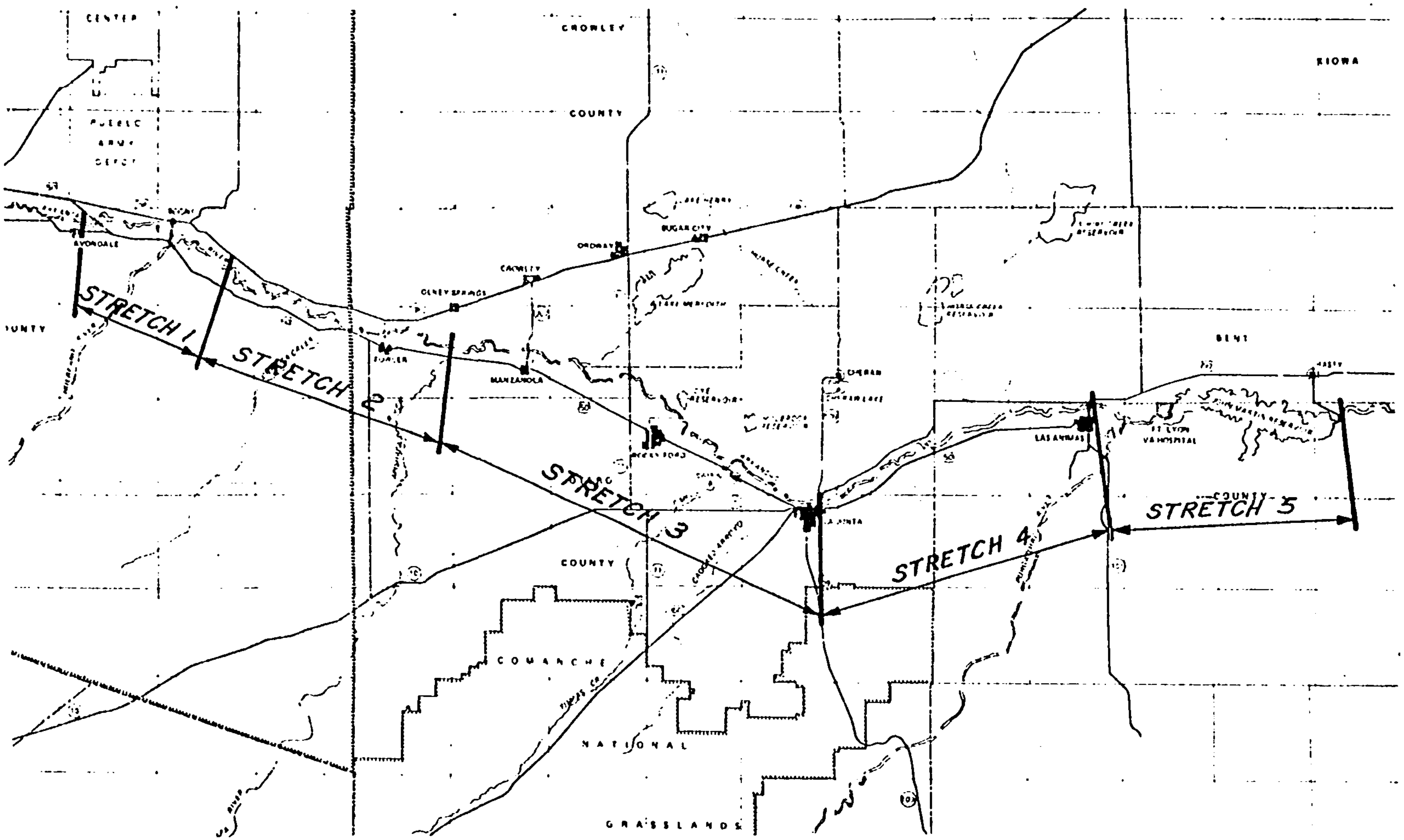


Figure 7. The study segment

Table 2. Definition of stretches

Stretch number	USGS ID# upstream boundary gage	USGS ID# downstream boundary gage	Stretch length in miles
1	07(1095)00	07(1170)00	17.9
2	07(1170)00	07(1197)00	21.2
3	07(1197)00	07(1230)00	30.1
4	07(1230)00	07(1240)00	24.6
5	07(1240)00	07(1305)00	23.1

water lost to deep percolation from the fields will eventually return to the river.

The climate in the segment area is semiarid with annual rainfall being less than 14 inches. Rocky Ford, which is located near the center of the segment, receives an average annual precipitation of 12.31 inches which includes 23.1 inches of snow. Precipitation in the area usually occurs as intense storms of short duration, and flash floods are a common hazard. Temperatures are generally high, and humidity is low.

A schematic diagram of the surface water system near the Arkansas River has been prepared for the river segment from Pueblo Reservoir to the outflow gage in stretch 5 below John Martin Dam and is presented in Figure 8. The schematic shows all surface waterways including the Arkansas River, the tributaries, standing bodies of water, and all irrigation ditches including waste lines. The acres of land covered

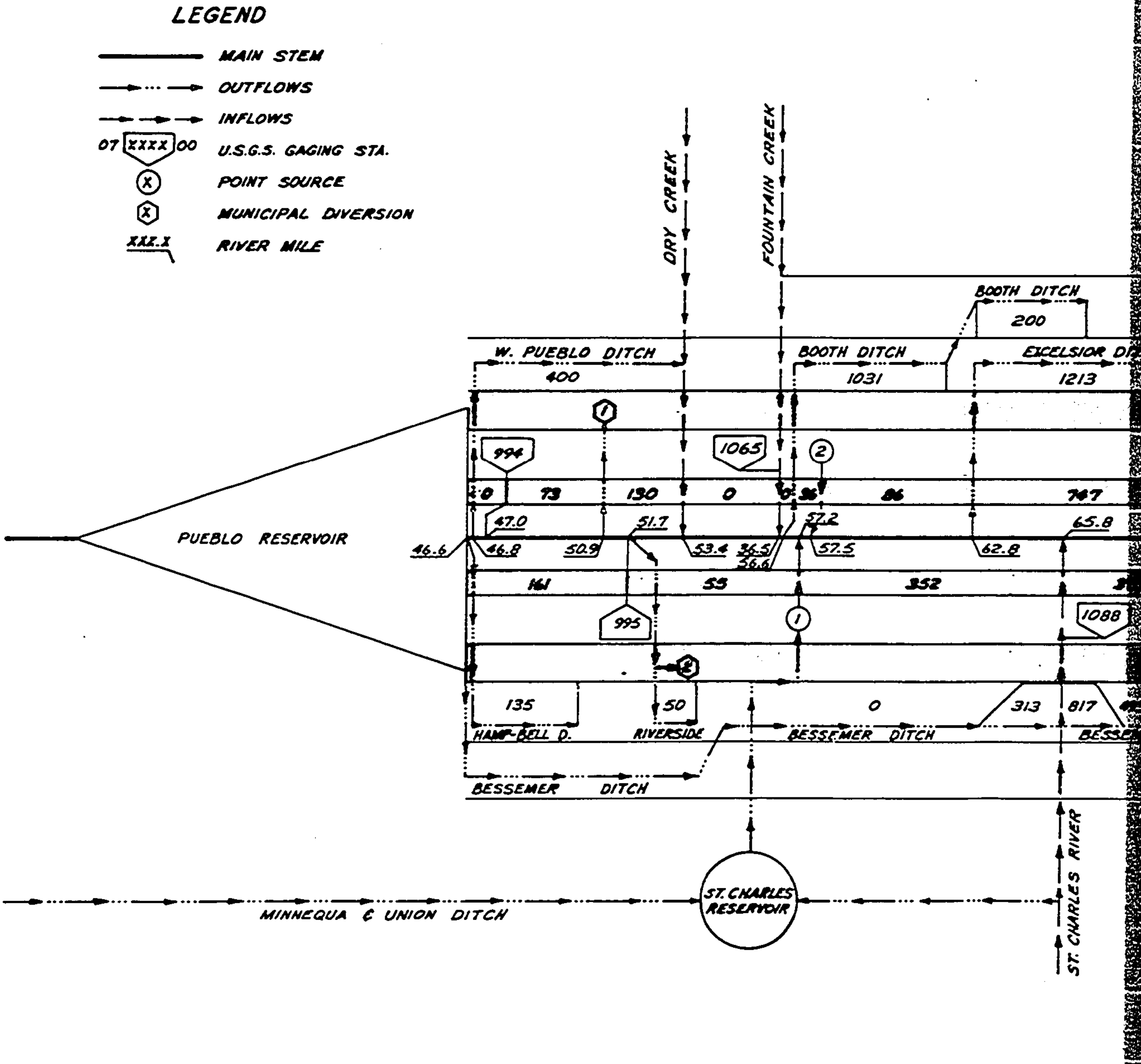
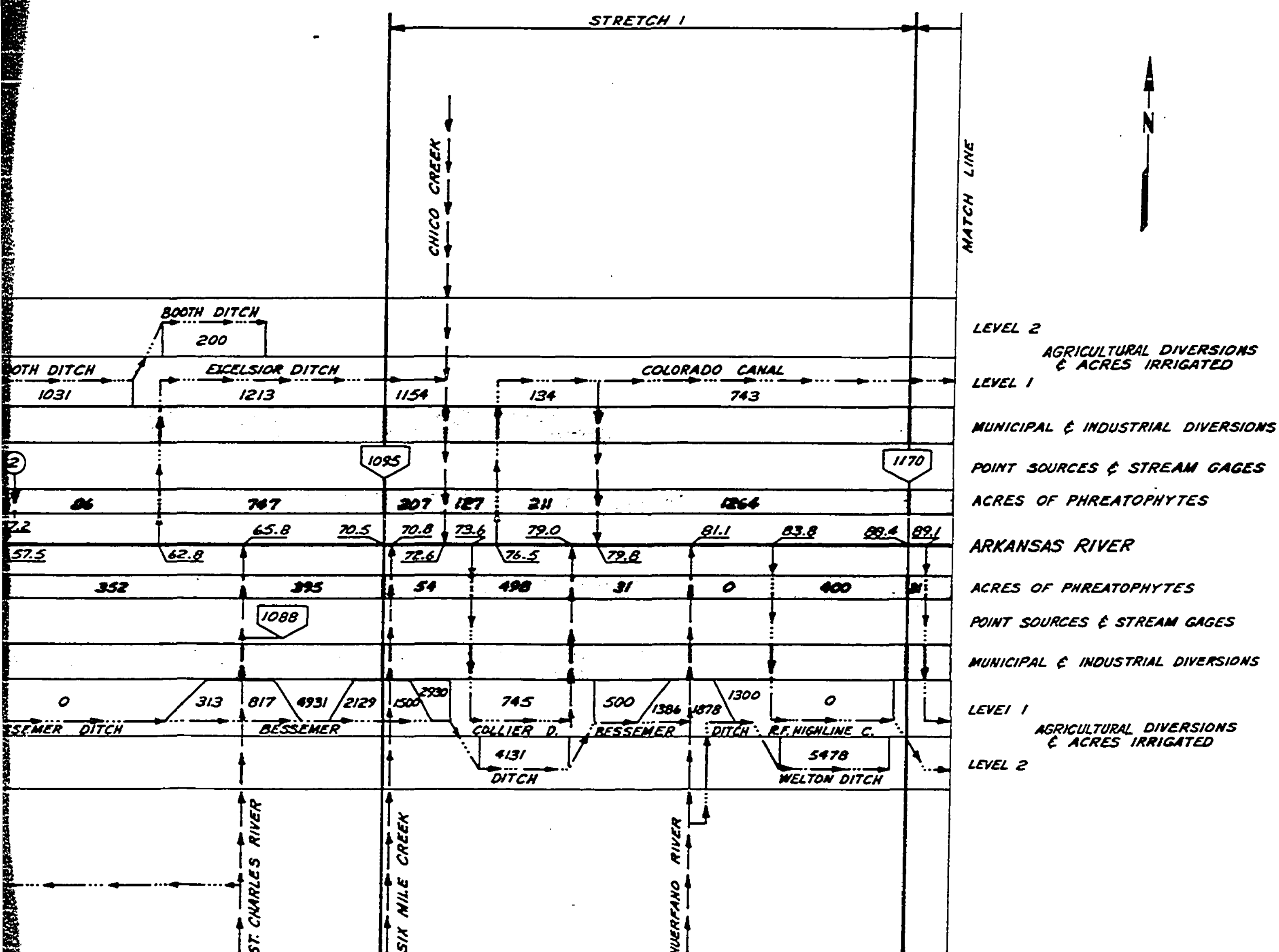


Figure 8a. Schematic diagram of the surface water system from Pueblo Reservoir to below John Martin Dam, upstream reach



STRETCH I
MATCH LINE
N
CHICO CREEK
BOOTH DITCH
200
BOOTH DITCH
1031
EXCELSIOR DITCH
1213
1154
COLORADO CANAL
134
743
1095
1170
86
747
307
127
211
1264
65.8
70.5
70.8
73.6
79.0
81.1
83.8
88.4
89.1
57.5
62.8
72.6
76.5
79.8
352
395
54
498
31
0
400
1088
0
313
817
4931
2129
1500
2930
745
500
1384
1878
1300
0
BESSEMER DITCH
BESSEMER
COLLIER D.
BESSEMER
DITCH
R.F. HIGHLINE C.
4131
DITCH
5478
WELTON DITCH
ST. CHARLES RIVER
SIX MILE CREEK
HUERFANO RIVER
LEVEL 2
AGRICULTURAL DIVERSIONS & ACRES IRRIGATED
LEVEL 1
MUNICIPAL & INDUSTRIAL DIVERSIONS
POINT SOURCES & STREAM GAGES
ACRES OF PHREATOPHYTES
ARKANSAS RIVER
ACRES OF PHREATOPHYTES
POINT SOURCES & STREAM GAGES
MUNICIPAL & INDUSTRIAL DIVERSIONS
LEVEl 1
AGRICULTURAL DIVERSIONS & ACRES IRRIGATED
LEVEL 2

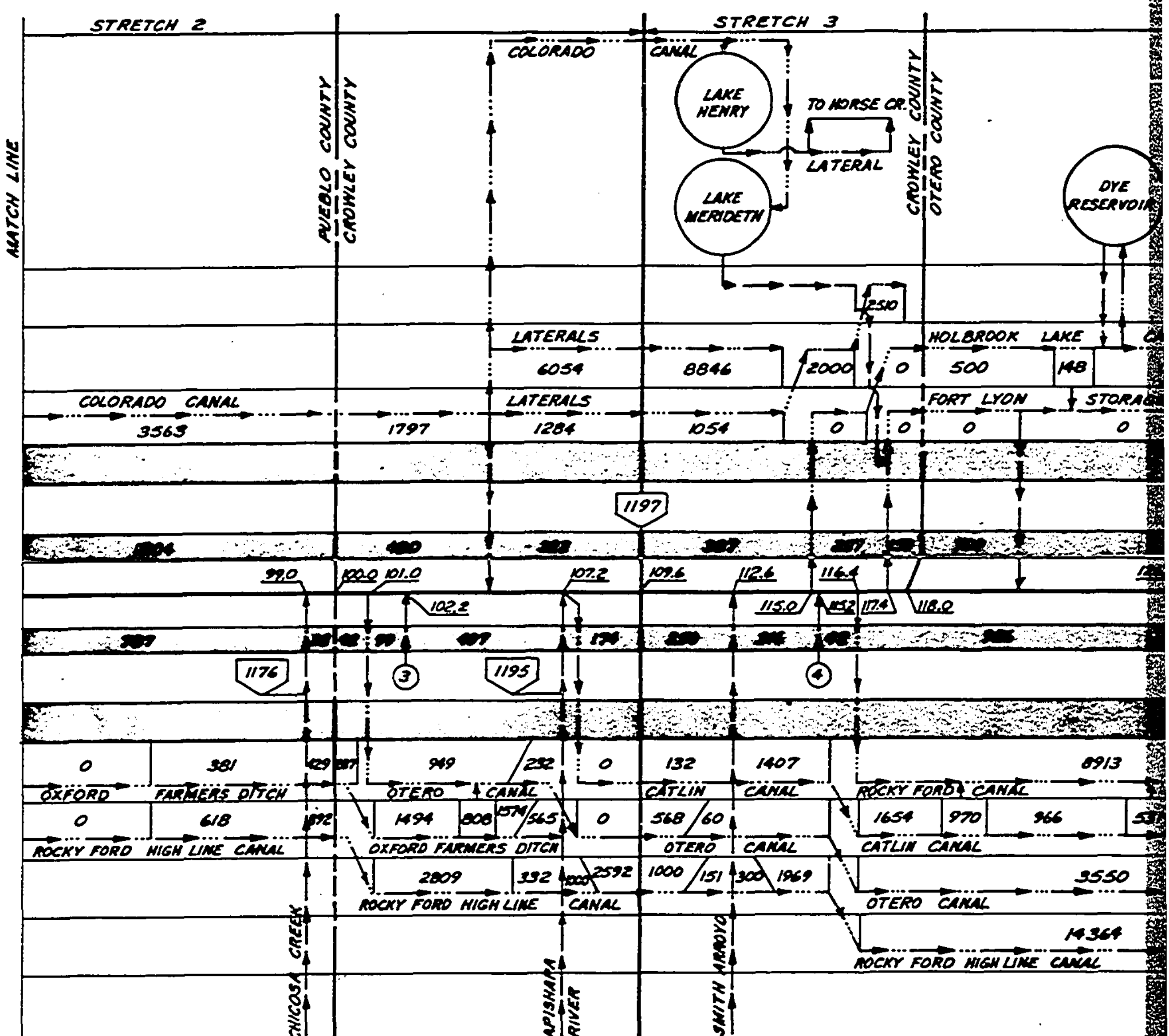
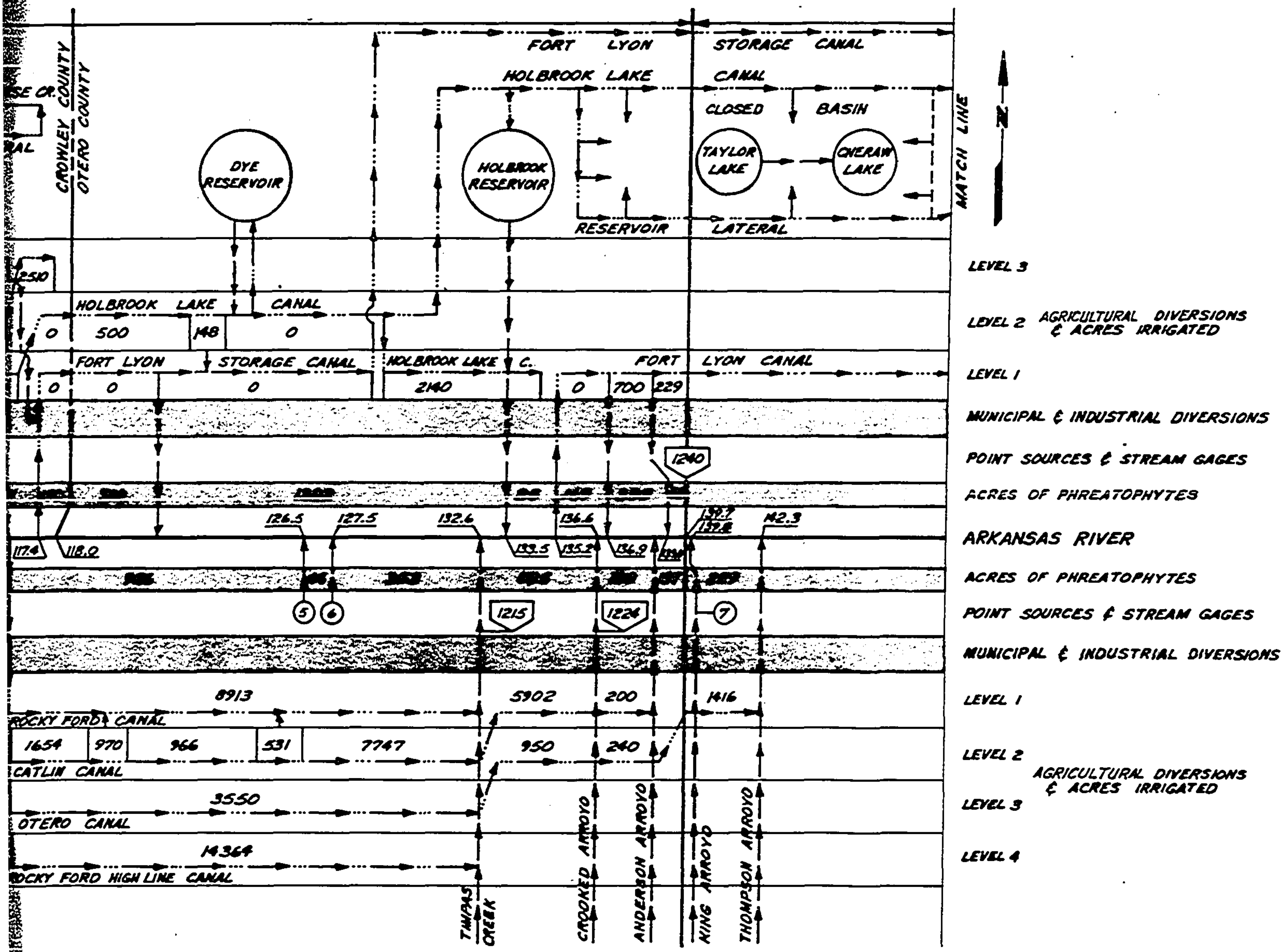


Figure 8b. Schematic diagram of the surface water system from Pueblo Reservoir to below John Martin Dam, middle reach



FORT LYON STORAGE CANAL
HOLBROOK LAKE CANAL
CLOSED BASIN
CROWLEY COUNTY
OTERO COUNTY
DYE RESERVOIR
HOLBROOK RESERVOIR
TAYLOR LAKE
CHERAW LAKE
RESERVOIR LATERAL
MATCH LINE
N
LEVEL 3
LEVEL 2 AGRICULTURAL DIVERSIONS & ACRES IRRIGATED
LEVEL 1
MUNICIPAL & INDUSTRIAL DIVERSIONS
POINT SOURCES & STREAM GAGES
ACRES OF PHREATOPHYTES
ARKANSAS RIVER
ACRES OF PHREATOPHYTES
POINT SOURCES & STREAM GAGES
MUNICIPAL & INDUSTRIAL DIVERSIONS
LEVEL 1
LEVEL 2 AGRICULTURAL DIVERSIONS & ACRES IRRIGATED
LEVEL 3
LEVEL 4
2510
HOLBROOK LAKE CANAL
0 500 148 0
FORT LYON STORAGE CANAL
HOLBROOK LAKE C.
FORT LYON CANAL
0 0 0 2140 0 700 229
1240
117.4 118.0 126.5 127.5 132.6 133.5 135.2 136.6 136.9 142.3
1215 1224
5 6 7
8913 5902 200 1416
ROCKY FORD CANAL
1654 970 966 531 7747 950 240
CATLIN CANAL
3550
OTERO CANAL
14364
ROCKY FORD HIGH LINE CANAL
TIMPAS CREEK
CROOKED ARROYO
ANDERSON ARROYO
KING ARROYO
THOMPSON ARROYO

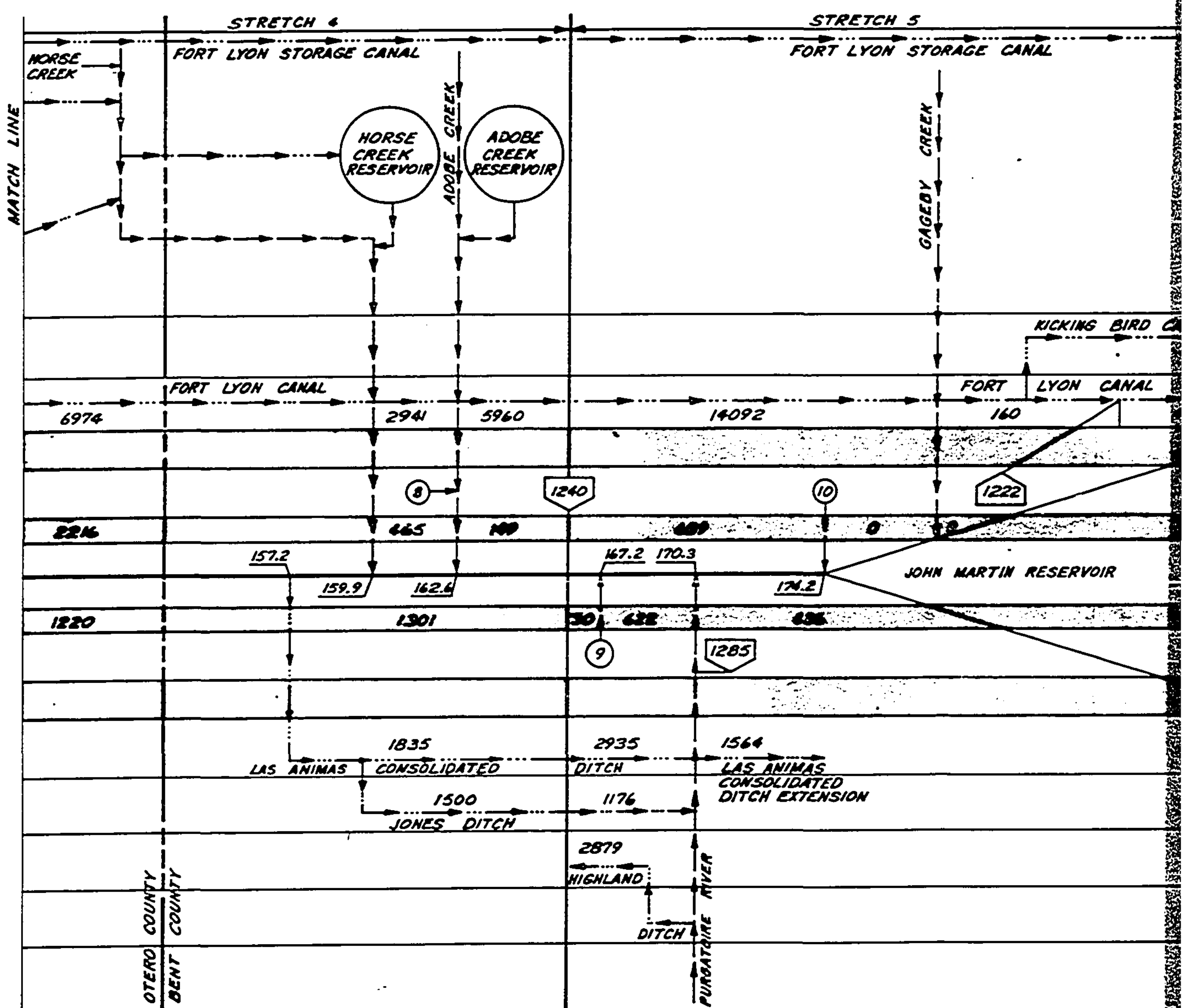
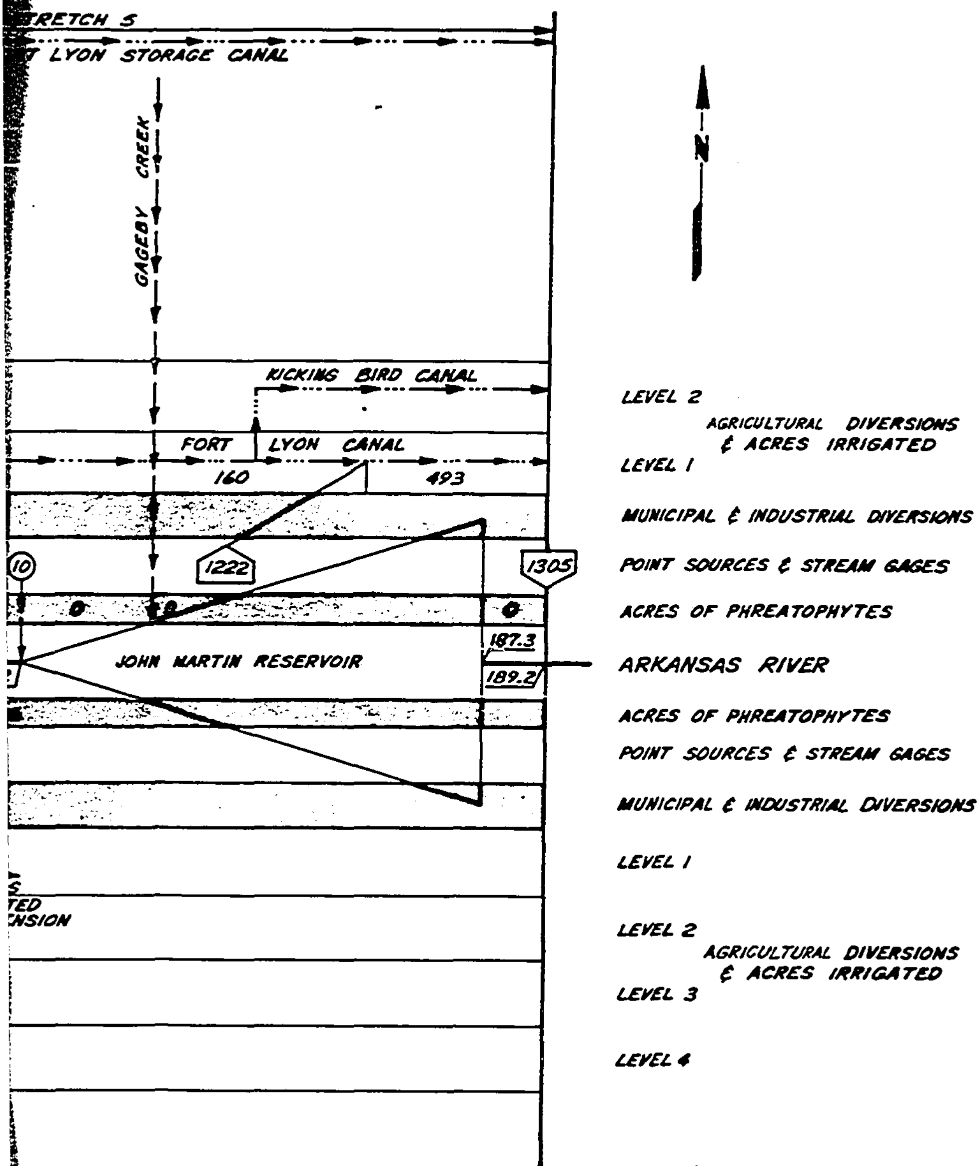


Figure 8c. Schematic diagram of the surface water system from Pueblo Reservoir to below John Martin Dam, downstream reach



MUNICIPAL & INDUSTRIAL DIVERSIONS

1. PUEBLO WTP - NO. SIDE
2. PUEBLO WTP - SO. SIDE

DISCHARGERS

1. CF&I STEEL
2. PUEBLO STP
3. FOWLER STP
4. MANZANOLA STP
5. AMERICAN CRYSTAL SUGAR & FARMLAND FOODS
6. ROCKY FORD STP
7. LA JUNTA STP
8. LAS ANIMAS FISH HATCHERY
9. LAS ANIMAS STP
10. FORT LYON STP

by high water consumption plants, i.e. the phreatophytes, and the irrigated croplands are also shown in Figure 8.

USGS maps (7-1/2 minute quadrangle, scale 1:24000) were used to define drainage areas, and the irrigated and phreatophyte lands are grouped according to drainage. Most of the phreatophyte areas are adjacent to the main channel of the Arkansas River with small areas also existing along the St. Charles River, Sixmile Creek, and the Purgatoire River.

The irrigated lands are arranged in a hierarchy according to their remoteness from the Arkansas River. The areas directly tributary to a stream are shown in the schematic as nearest to the stream and the lands tributary to a canal are shown as second, third, or fourth levels depending on how many canals separate the land from the stream. The tributary acres (first level) can also be divided into those areas with discrete return flow structures and those areas contributing along a reach of the stream but this study did not require such a detailed definition of the irrigated land. Some second level land is directly tributary to the streams through discrete return structures. It is important to note that moving downstream a canal usually first serves land that is tributary (level one) and then begins serving higher level lands as new canals begin diverting. For example, the Otero Canal serves levels 1, 2 and 3 land, and the Rocky Ford Highline Canal serves levels 2, 3 and 4 lands. In contrast, the Rocky Ford Canal serves only level 1 land. This distinction between directly and indirectly tributary lands is used in this study to weight the impacts of the model variables on the stream flow.

The lengths and areas shown were measured from (scale 1:15840) aerial photos of Bent, Otero, and Crowley Counties that are available from the Soil Conservation Service [20, 21, 22] and from U.S. Geological Survey topographic maps [23] (scale 1:24000) for Pueblo County. The maps represent conditions in the counties at the following dates.

Pueblo — varies, 1960-1963

Crowley — 1962

Otero — 1966

Bent — 1962

The river miles shown for the Arkansas River were established by setting the river mile at the Pueblo-Otero County line equal to 100 and measuring relative upstream and downstream distances from the USGS and SCS maps. The tributaries are dimensioned upstream from their confluence with the Arkansas River. The irrigation canals are dimensioned downstream from their point of diversion.

The amount of detail was considerably greater in the SCS maps than in the USGS maps permitting a more accurate definition of phreatophyte and crop areas. However, the SCS maps also show areas that have been irrigated in the past but are no longer active, and the differentiation between active and inactive areas is difficult. Furthermore, small nonproductive areas, i.e. roads and homesteads, are included in the measured areas; but large nonproductive areas, i.e. towns, water bodies and undeveloped land are not included in the values shown.

The irrigated areas are also grouped with respect to the ditch providing the water. The sum of the acres irrigated by each ditch are compared with reported values [24, 25, 26] in Table 3. Some of the land

Table 3. Comparison of irrigated acre estimates

Name of diverter	Estimates of irrigated acres served by each diversion: This study	[24][a]	[25][a]	[26][a]	Est. acres not tributary to segment	Ratios of study estimates to others	Estimated acres used for model variables
Bessemer Ditch	20827	19500	20000	–	10380	1.07, 1.04	20827
West Pueblo Ditch[b]	400	400	–	–	400	1	400
Hamp Bell Ditch[b]	135	160	–	–	135	0.84	135
Riverside Dairy Ditch[b]	50	50	–	–	50	1	50
Booth Ditch[b]	1231	1400	1400	–	1231	0.88, 0.88	1231
Excelsior Ditch	2367	1583	2000	–	1213	1.50, 1.18	2367
Collier Ditch	745	643	–	–	0	1.16	745
Colorado Canal	31611	45000	43000	–	13389	–	31611
Rocky Ford Highline Canal	26027	22500	24000	21579	0	1.16, 1.08, 1.21	26027
Oxford Farmers Ditch	5892	5250	5000	–	0	1.12, 1.18	5892
Otero Canal	13969	6866	6000	–	0	2.03, 2.33	7965[c]
Catlin Canal	20332	18000	19960	–	0	1.13, 1.02	20332
Holbrook Lake Canal	8323	16000	16000	–	7677	–	8323
Rocky Ford Canal	9335	8000	8000	–	0	1.17, 1.17	9335
Fort Lyon Canal	31299	94000	94000	–	62701	–	31299
Las Animas Consolidated Ditch	7081	9482	6000	–	0	0.75, 1.18	7081

[a] Source references are listed in the Literature Cited section with the identification numbers shown in brackets.

[b] System not tributary to study segment.

[c] See text for method of revising Otero Canal estimate.

served by the Colorado and Holbrook Lake Canals are tributary to reservoirs, and a large portion of the Ft. Lyon system is tributary below the study segment. The crop areas shown for these three systems in Figure 8 do not contain these nontributary acres and, therefore, the study estimate cannot be compared with the total areas reported for each system. The ratios between the areas measured in this study and the other reported values are also shown in Table 3 for the remaining systems. The system areas in Pueblo County which were estimated from the USGS maps show a random variance between the estimates from this study and the other reported values with the estimate for the single large system being 4-7% higher in this study. The smaller system estimates ranged from 18% high to 16% low with one exception. The Excelsior Ditch estimate is 18% higher than one reported value and both values are considerably higher than the third reported value. Since only a portion of the Excelsior Ditch is in the study segment, this discrepancy was not resolved.

The study estimates derived from the aerial photos (the systems below the Collier Ditch in Table 3) range from 2% to 21% high with two exceptions. The study estimate for the Otero Canal is more than twice the reported values, and the discrepancy is primarily caused by the difficulty in differentiating abandoned cropland from active cropland in the aerial photographs. Several areas served by the Otero Canal have been removed from production in recent years because the systems water rights have been sold and transferred. The estimate of cropland area served by the Otero Canal was revised for this study by multiplying the Division of Water Resources (DWR) estimate by 1.16 which makes the

Otero estimate high by a percentage that is in the same range as the other estimates. The revised figure appears in the column of acreage estimates used in this study. The second discrepancy involves the Las Animas Consolidated system where the DWR estimate does not agree with this study's estimate or the other reported value. The Las Animas Town Ditch sold its water rights to the Rocky Ford Highline Canal a few years ago, and several thousand acres of this system were removed from production. Since this change is the likely cause of the discrepancy, the estimate prepared in this study is used because the Town Ditch has not been included.

The acreage estimates presented above are used to calculate variables representing patterns of water use along the segment and over time. Since the variables do not represent accurate estimates of water use, the use of estimates containing unproductive land will not affect the study adversely as long as the percentage of unproductive land is about the same for each system. The ratios in Table 3 indicate that this study's estimates for the systems affecting the study segment are clustered around an average overestimate of 13% if the three discrepancies are excluded. Therefore, the study estimates are assumed to include an equal percentage of unproductive land and are used to calculate variables in this study.

Surface runoff of irrigation water occurs as tail water from the fields which returns as either distributed or discrete flow or as waste water from the canals. Most of the canals have a waste connection near the beginning and at the end of the canal. These waste connections are shown in Figure 8.

Municipal and industrial dischargers, USGS gage stations, and reservoirs are also shown in Figure 8.

Four regulation activities affect the stream flow in the study segment.

1) The operation of the priority system affects the amount of natural flow in the river and tributaries.

2) Exchanges of stored water for natural flow affect the amount of water in some segments of the stream.

3) Imported water is transported to the user via the river and thus affects the stream flows above the downstream users.

4) The operation of John Martin Dam can affect the operation of the priority system above the dam.

The first activity has been discussed above, and the other three activities are discussed below.

Three of the irrigation systems, (1) the Colorado Canal, (2) the Holbrook Lake Canal, and (3) the Fort Lyon Canal system can utilize the exchange activity in providing water to irrigators in the upstream portions of their service areas. The exchange activity involves the simultaneous diversion of flow at an upstream point and the release of an appropriate amount of stored water downstream. For example, the Colorado Canal often diverts water at its headgate (river mile 76.5 in Figure 8) and also releases water from Lake Meredith through the Fort Lyon Storage Canal to the Arkansas River at river mile 121.5. This exchange practice affects the stream flow for a distance of 45 miles. The ratio of released to diverted water can be established by the Division Engineer who is responsible for the operation of the priority system, exchanges, and delivery of imported water and/or by agreement with other diverters in the area. Exchanges are permitted only in areas where no senior right can be adversely affected. The system acquires the stored water from surface runoff, irrigation return flows, natural flow diversions when their storage rights are in priority, and/or transmountain water imports. The Colorado Canal system uses Lake Henry and Lake Meredith, the Holbrook Lake Canal uses Dye Reservoir and Holbrook Reservoir, and the Fort Lyon system uses Horse Creek and Adobe Creek Reservoirs and several reservoirs located below the study segment. Whenever an exchange activity is operating, the diversion and discharge are measured and reported daily to the DWR.

Imported water is natural flow from another drainage basin that has been diverted across the basin divide. Such diversions, often called transmountain diversions, are regulated by the priority system in the basin of the water's natural occurrence, and the imported water is not subject to the priority system in the basin receiving the water. Imported water is usually stored in reservoirs near its entry point into the basin and is delivered to the user on demand. The water is transported in the natural river, and the delivery of a quantity of imported water is commonly termed a "river run". The user requests delivery from the Division Engineer who decrees the amount and time of discharge and diversion. The amount discharged is always greater than the amount diverted because the Division Engineer deducts for "transit losses". Previous and current transit loss studies have been discussed by Livingston [27]. Currently the transit loss is calculated as 0.07% per mile of travel, but Livingston's studies indicate the loss may be as high as 0.16% per mile. All gage data between the point of discharge and the point of diversion plus the diversion data will include the imported water. The effect of imported water is to increase the stream flow in affected segments in contrast to the depletion effect of exchanges as discussed above.

The Arkansas River basin receives imported water from several transmountain diversions, and the impacts of these diversions have been summarized in the Arkansas River Basin Water Quality Management Plan [28]. Of the three largest transmountain diversions, (1) the Homestake Project, (2) the Fryingpan-Arkansas Project, and (3) the Twin Lakes diversions, the latter two importers have an effect in the study segment. The Homestake project diverts water from the Colorado River basin to the Arkansas basin and then into the South Platte basin from the Arkansas River far above the study segment. A portion of this water reenters the Arkansas basin above Colorado Springs but this water is so thoroughly consumed that it has a negligible effect on the study segment.

The Twin Lakes imported water is mostly delivered to the Colorado Canal at river mile 76.5, but the Division Engineer has delivered water from Twin Lakes to several other diverters in the study segment. The Fryingpan-Arkansas Project (FAP) delivers imported water to several municipal and agricultural users including several users in the study segment. This project is being developed by the Bureau of Reclamation and is thoroughly discussed in the Draft Environmental Statement [29] for the FAP. Pueblo Reservoir, which is the upstream boundary of the schematic diagram in Figure 8, is part of the Fryingpan-Arkansas Project, and will soon become the source of imported water deliveries to downstream users. This reservoir may also be used in the future to store winter flows for use in the spring which will have a very large effect on the wintertime low flow regime in the study segment.

The fourth regulation activity affecting the surface flows in the study segment involves the operation of John Martin Reservoir which is located in Stretch 5 of the study segment. This reservoir was built by the U.S. Corps of Engineers in the 1940's for two primary purposes, (1) flood control and (2) the implementation of the Arkansas Compact [25, 26, 30, 31]. This latter purpose has dictated the operating rules of the reservoir. Each year is divided into a winter storage period (November 1 to March 31) and a summer storage period (April 1 to October 31) by the Compact. The stored water is apportioned between the Colorado and Kansas users, and the right of these users to demand water is severely limited during the winter storage period so that storage usually occurs. However, the summer storage restrictions permit demands far in excess of the supply so the reservoir is usually empty within a few weeks after the beginning of the summer storage period. Whenever the reservoir contains water, the Colorado users downstream cannot place a call on the river above the reservoir which can affect the operation of the priority system in the study segment. Whenever the reservoir is empty, all inflows are supposed to be passed through the reservoir during the summer storage period, and this is the condition found during the time period used to develop the water balance model in this study. The gage data at the downstream terminus of Stretch 5 is directly related to the operation of John Martin Dam and deviations from the principle of matching discharge to inflow will affect the fit of the model to the data for this stretch.

A comprehensive presentation of the study segment and adjacent areas would be massive because the water resource system is so complex. However, the above discussion, while not complete, does present the details necessary to support the development of variables for the water balance model as discussed below.

Time Interval Selection

The selection of the time elements required two decisions. First, a time increment between observations must be selected, and then one or more calendar time periods must be chosen. The unit time increment was chosen as one day for three principal reasons.

1) A large portion of the available data are reported as daily observations or averages, i.e. gaged flows, precipitation, and temperatures.

2) Low flows are usually stated in terms of minimum daily average flow for seven consecutive days, and using time increments longer than one day complicates the evaluation of these moving averages.

3) Using a time increment less than one day would increase the data handling task without markedly increasing the accuracy of the analysis.

Two calendar time periods were selected to provide the data base for the model parameters estimation. Both time periods were restricted to the critical low flow period of July through October. One time period was selected to represent a "wet year", and the other was selected to represent a "dry" year. The year 1973 fulfilled the first criteria, and 1974 fulfilled the second although October 1974 could not be included because the data had not been reported at the time of this study. The selection of recent data periods also means that the model will represent more recent basin conditions, and this feature helped in collecting some of the data. So the selected calendar time periods for data definition are:

July 1, 1973, to October 31, 1973

July 1, 1974, to September 30, 1974.

Development of Model Variables

The structure of the sample model is shown in Figure 9. The data base selected above provides a total of 1075 possible observations which may be reduced when time lagged variables are used. The independent variables are shown as subdivided into hydraulic and other variables. This subdivision permits the analysis of the parameters by individual stretches for hydraulic variables if the hydraulic factors differ markedly between stretches while all other variables whose parameters should be invariate between stretches can be analyzed over the whole segment. The hydraulic variables include QOUT, QDIV, QDIS, and ΔS, and the other variables include Distributed inflows and Stress-Aquifer Exchanges. The results of this subdivision element of the analysis are reported in the following chapter.

<table>
<tr><th rowspan="2"></th><th rowspan="2">STRETCH</th><th rowspan="2">OBSERVATION ID NUMBERS</th><th colspan="3">VARIABLES</th><th rowspan="2">DATA BASE YEAR</th></tr>
<tr><th>DEPENDENT</th><th>HYDRAULIC</th><th>OTHER</th></tr>
<tr><td rowspan="10">OBSERVATIONS</td><td rowspan="2">1</td><td>$1 \le N \le 123$</td><td rowspan="10">STRETCH INFLOW, QIN</td><td rowspan="10">DISCRETE FLOWS AND S VARIABLES</td><td rowspan="10">DISTRIBUTED FLOWS AND STREAM-AQUIFER EXCHANGES</td><td>1973</td></tr>
<tr><td>$124 \le N \le 215$</td><td>1974</td></tr>
<tr><td rowspan="2">2</td><td>$216 \le N \le 338$</td><td>1973</td></tr>
<tr><td>$339 \le N \le 430$</td><td>1974</td></tr>
<tr><td rowspan="2">3</td><td>$431 \le N \le 553$</td><td>1973</td></tr>
<tr><td>$554 \le N \le 645$</td><td>1974</td></tr>
<tr><td rowspan="2">4</td><td>$646 \le N \le 768$</td><td>1973</td></tr>
<tr><td>$769 \le N \le 860$</td><td>1974</td></tr>
<tr><td rowspan="2">5</td><td>$861 \le N \le 983$</td><td>1973</td></tr>
<tr><td>$984 \le N \le 1075$</td><td>1974</td></tr>
</table>

N = 1075 OBSERVATIONS

Figure 9. Sample model structure

Discrete flows

The dependent variable QIN and the independent variable QOUT were assigned the reported daily average flows [32, 33] for the upstream and downstream stretch boundary gages. As described earlier, the QOUT values for Stretch 5 will also be the QIN values for Stretch S + 1, e.g. QOUT for n = 1 will also be QIN for n = 216. The accuracy of the gage data cannot be well defined, but the potential for error is high. The stage-discharge rating curves, which are usually checked by USGS staff once or twice a month, were changed during the study period for three of the six stations[1]. The change for one station was caused by nearby construction altering the natural flow characteristics, and the other two gages were affected by natural changes in stream conditions. Such calibration changes are typical for gages on streams with alluvial stream beds like the Arkansas River in the study segment and are a source of potentially large error. Low flow measurements are especially subject to large potential error because the control section can change so easily in the sandy stream bed. Burkham and Dowdy [34] examined the accuracy of gage data on the Gila River in Arizona (the Gila River and the Arkansas River in the study segment have similar flow regimes and stream conditions) and reported a standard error in the 20-30% range for low summer flows at their two study gages. This high potential for error is considered in the discussion of the study results.

The QDIV variable was constructed by summing the reported daily diversions [24] within each stretch (see Figure 8) and using that sum

---

[1]Harold Petsch, Denver Region, U.S. Geological Survey, personal communication, July 1975.

as a single observation. Each day in the base data period produces five observations in the sample model. For example, July 1, 1973, produced observations n = 1, 216, 431, 646, and 861. Practically all diversions of any magnitude are measured with Parshall flumes. The largest diversions are also equipped with recorders. The data are collected, usually daily, by the local Water Commissioner who also sets the control gates that regulate the amount of water diverted. The diversion will vary during the day as the river stage rises and falls, but since the stage changes very little in low flow conditions, the diversions are assumed to be invariant during the day. During the study the author received a report that some staff gages are set high to provide flow measurements less than actual flow, but the report cannot be corroborated so no adjustment was made to any diversion data.

The variable QDIS was constructed in a manner similar to the QDIV construction with the exception that discharges (see Figure 8) replaced diversions; however, the construction of this variable was more complex because some of the discharges had to be estimated. The variable includes both gaged tributaries and municipal, industrial, and exchange discharges. The tributary data was taken from Division of Water Resources data [24]. Error potential for both sources is discussed above.

The study segment contains 18 known dischargers and discharge data are available [35] for only two of the discharges. Six discharges were not included in the study because the daily discharge was estimated at less than 25000 gallons (0.04 cfs) which is considerably less than the expected errors in the other flow data. Two power plant discharges also are not included as discrete discharges because they discharge into sanitary sewer systems and are therefore included in another discharge estimate. The Farmland Foods discharge near Rocky Ford was overlooked during the model construction resulting in a discrepancy of approximately 0.3 cfs [28] in QDIS for Stretch 3. The small potential error did not justify revising the model data when the error was discovered. The discharge of the Las Animas Fish Hatchery is estimated as 3-3.5 cfs by the Division of Water Resources [33]; however, this discharge is not included in the model because its influence on the main stream is considered to be negligible. The discharge occurs in Adobe Creek, 5.2 miles from the Arkansas River, in an intensely irrigated area; and the flow in Adobe Creek is often zero according to local residents. Evidently the hydrologic factors arising from irrigation activities, e.g. diversions and groundwater pumping, mask any impact from the fish hatchery discharge. The Adobe Creek area is included in the distributed inflow variables discussed below.

Estimates were prepared for the remaining eight discharges. The La Junta Sewage Treatment Plant (STP) discharge was estimated from monthly total flows that were measured in 1974 [28]. The values were taken from a smooth curve plotted through the available data points and rounded to the nearest 0.1 cfs. Data were also available for the Fort Lyon VA Hospital STP [28], but the data showed practically no seasonal daily variations. Weekend flows appear to be 70-80% of weekday flows, but all flows round off to 0.2 cfs. The American Crystal sugar plant discharge near Rocky Ford occurs only during the sugar beet refining season which begins about October 1 each year. This discharge was estimated as 7.7 cfs based on verbal contacts with plant personnel, and the estimate is in acceptable agreement with the 8.5 cfs discharge limit shown in American Crystal Sugar's NPDES permit [35]. The remaining five discharges were estimated [28] from NPDES permit limits, populations served, and hydraulic limitations of the treatment systems. The discharges range from 0.1 to 1.6 cfs. The techniques used in developing these QDIS estimates are prone to large percentage errors, but the impact of these potential errors on the model will not be as great as the errors discussed above because the discharges are considerably smaller than the diversions and boundary gage flows.

Distributed inflows

The Distributed Inflow variable(s) represents the surface runoff entering the tributaries and the main stem. This surface flow arises from three sources: (1) inflow from outside the study area, (2) precipitation, and (3) the application of excess irrigation water. Three variables were developed to simulate the effects of these surface flow sources with one variable representing the first two sources and two variables representing the effects of the last source.

No data have ever been collected to determine the surface inflows to the study area at the boundary of the irrigated lands. Visual observations by the author during the summer of 1975 indicated that none of the tributaries has any base flow during the low flow periods. Numerous contacts with area residents confirmed this impression so the model was constructed on the assumption that tributary inflow to the study area would occur only as intermittent runoff from precipitation events. This Surface Run Off from Precipitation, SROP, variable was constructed from the reported precipitation at a single weather station in each stretch. These five stations are known as:

1. Stretch 1 — Pueblo WSO AP — Index No. 6740
2. Stretch 2 — Fowler 1 SE — Index No. 3079
3. Stretch 3 — Rocky Ford 2 SE — Index No. 7167
4. Stretch 4 — La Junta FAA AP — Index No. 4720
5. Stretch 5 — Las Animas — Index No. 4834.

All five stations are in Division 1 of the Environmental Data Service. The stations used for stretches 2 and 3 are located near the center of the stretch, and the stations for stretches 4 and 5 are located at the western boundary of the stretch. The station for stretch 1 is about ten miles west of the stretch's western boundary. Most precipitation events in the study area move in an easterly direction, and this element influenced the selection of the stations. The daily precipitation amounts are summarized in Table 4. The largest daily event was 1.82 inches which occurred in stretch 3 on the same day (9-26-73) as the maximum event in stretches 2, 4 and 5. The other three events over one inch also occurred in 1973. The summary in Table 4 illustrates that precipitation events in the study area during the data base period were infrequent, ranging from 29 to 62 events out of a possible 215 events, and were low in magnitude with a large majority of events being less than 0.25 inches.

The SROP variable was developed in five different configurations. SROP is the string of daily observations with no transformations, and SROP1 is SROP lagged one day, SROP2 is SROP lagged two days. SROP3 and

Table 4. Summary of precipitation data (number of daily values)

Precipitation range (in/day)	Stretch number 1	2	3	4	5
0 - 0.25	47	25	28	40	21
0.25 - 0.50	8	8	5	2	4
0.50 - 1.00	6	2	2	3	1
1.00 - 2.00	1	1	1	1	3
> 2.00	0	0	0	0	0
Totals	62	36	36	46	29

SROP4 were constructed by adding portions of previous events to the daily values. The equations are:

$$SROP3 = SROP + 0.5\ SROP1 \tag{6}$$

$$SROP4 = SROP + 0.6\ SROP1 + 0.2\ SROP2 \tag{7}$$

The variables SROP, SROP1, and SROP2 were used conjunctively, but the variables SROP3 and SROP4 were never used conjunctively with any other SROP variable.

Time lagged variables were used in the model to simulate the typical time delayed impacts of precipitation events on stream flow. The principle is illustrated in Figure 10 where a sample composite hydrograph has been constructed for three discrete identical precipitation events occurring on three consecutive days. Prior to the beginning of event 2 the hydrograph represents only the impact of event 1; however, following the beginning of event 2, the hydrograph represents the impacts of two or more events. For example, the stream flow at point C

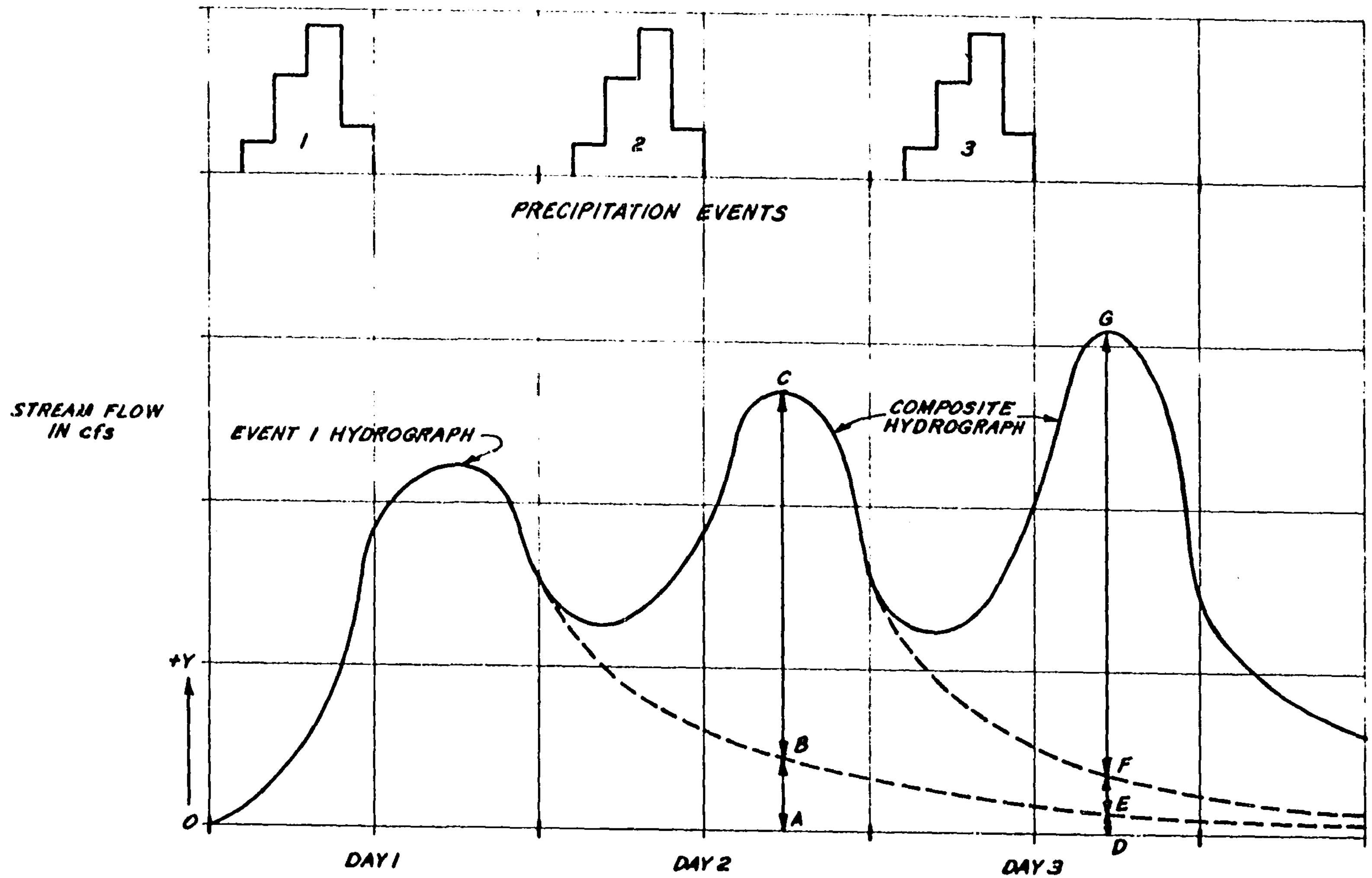


Figure 10. Composite hydrograph example

C is composed of segment AB which resulted from event 1 and segment BC which resulted from event 2. Similarly, flow G is composed of three segments, DE, EF, and FG, representing each of the three events. When the variables SROP, SROP1 and SROP2 are used in the model, the regression procedure will determine the relative impact of the individual events within the parameters. On the other hand, SROP3 and SROP4 establish the relative impacts of previous events before parameter estimation. The relative merits of these different configurations are discussed in the following chapter.

Two variables were constructed to represent surface runoff from irrigated cropland. The TRibutary Water Applied, TRWA, variable is associated with cropland that is directly tributary to the river or its side stream. Land tributary to return flow structures passing under canals is also included in the TRWA variable if the return flow is then tributary to a stream. The NonTRibutary Water Applied, NTRWA, variable is associated with cropland that is tributary to another irrigation system and will eventually reach the surface stream system after some time delay. Cropland tributary to streams or reservoirs above a gaging station is not included in either variable because the gage records serve as the inflow variables for those areas. The separation of this distributed inflow variable is based on the assumption that tributary areas will have a significantly larger impact on stream flow than the nontributary areas, and the validity of this assumption is discussed in the following chapter.

Both variables, TRWA and NTRWA, are constructed by the same method. The basic equation is:

$$TRWA_{st} \text{ (or } NTRWA_{st}) = \sum_{x=1}^{m} \frac{IA_{xs}*WD_{xt}*DE_x}{TA_x} \qquad (8)$$

where $TRWA_s$ (or $NTRWA_s$) is a distributed inflow variable for stretch s, IA_{xs} is the irrigated acres in stretch s (either tributary or nontributary) served by ditch x, WD_{xt} is the water diverted by ditch x on day t, DE_x is the ditch efficiency of ditch x, and TA_x is the total number of acres served by ditch x. The summation includes the m tributary (or nontributary) ditches in stretch s. The development of estimates for irrigated acres in each stretch and in the total system for each ditch has been presented above. The water diverted was taken from the daily diversion records [24] of each ditch, and this data source has also been discussed above. The ditch efficiency is an estimate of the portion of the water diverted at the headgate that will actually reach the cropland, and these estimates were provided by the Bureau of Reclamation's Fryingpan-Arkansas Project office[1]. The efficiencies used in this study averaged 55.2% and ranged from 42.6% to 64.1%.

The TRWA and NTRWA variables are similar to the QDIV variable since diversions are a basic part of both variables; however, the differences in the variables are large enough to justify using all of these variables in the model. The basic difference between the variables is that TRWA and NTRWA contain some diversions from upstream areas while QDIV contains some diversions that do not appear in the other two variables in that same stretch.

---

[1]P. A. Abbot, Fryingpan-Arkansas Project, U.S. Bureau of Reclamation, Pueblo, Colorado, personal communication, 1975.

The TRWA and NTRWA variables ignore the amount of groundwater applied to the croplands because the estimation of these quantities introduces too much potential error. Taylor and Luckey [36] estimate groundwater consumption in the Lower Arkansas River Valley is about 25% of the surface water use, and they also have estimated quantities used for each ditch using known well capacities and power consumption data. Estimating pumped quantities from power data is subject to large potential errors [37] for several reasons including lack of information about total dynamic head and well and pump efficiencies. Furthermore, power data are incomplete and difficult to obtain. Taylor and Luckey's estimates are in 30-day time increments and also could not be subdivided into stretches and tributary/nontributary areas for this study. These considerations are the basis for the decision not to directly include groundwater use in the model at this time. Another variable that is discussed later does introduce some elements of groundwater use, but the surface runoff impact is not included.

Another element that can also be evaluated in another study is the impact on the model of time lagging the diversion values, i.e. the surface runoff on day t may be more closely related to the water diverted on day t-1 and/or day t-2.

Stream-aquifer water exchanges

The Stream-Aquifer Water Exchange (SAWE) element of the model presented complex problems in developing variables. The basic approach to this element utilizes a qualitative mass balance analysis of the alluvial aquifer underlying the study segment. Inflow sources to the aquifer are upstream alluvial aquifers and deep percolation from streams, lakes, reservoirs, canals, and the soil moisture bank. Inflow from confined aquifers is assumed to be negligible because no shallow confined aquifers are known to exist near the study area. Major outflow sinks for the study segment aquifer include downstream alluvial aquifers, exfiltration to streams, direct plant use, and wells. The deep percolation from streams and the exfiltration to streams are the two basic units of the stream-aquifer water exchange element needed for the water balance model and cannot occur simultaneously. These two basic units are modeled as dependent variables in terms of some of the other sources and sinks.

The aquifer-aquifer exchanges and the deep percolation from lakes and reservoirs were assumed to be constant over the study period. This assumption is based on the knowledge that the John Martin and Dye Reservoirs which are on or near the main stem were empty during the study period, and the other major reservoirs are located far enough from the stream that impacts from declines in percolated water would not be felt in the surface system during the study period. The aquifer-aquifer exchanges were assumed as constant because groundwater flow across large interfaces typically changes very slowly. Since the study period is very short, the assumption of constant exchange should not introduce any large potential error.

The SAWE can be viewed as the sum of two parts, (1) a constant exchange and (2) a dynamic exchange. The constant element would be the result of constant inflow-outflow elements such as those described above. When the SAWE is added to the regression model, the constant element will become a part of the regression constant. So no variables were constructed to simulate the groundwater inflows and outflows that are assumed to be constant.

The deep percolation from the canals is a function of the amount of water in the canals. This quantity is estimated by the TRWA and NTRWA variables discussed above so no additional variable is constructed to account for this inflow to the aquifer.

The effect of direct plant use of groundwater is simulated by estimating the Evapo-Transpiration (E_t) of the PHreatophytes (ETPH) in the study segment. The phreatophyte evapotranspiration literature was reviewed prior to developing the estimation technique for this study. The term, evapotranspiration, refers to both the water evaporated from the soil and the water taken up and not returned by the plants for growth and respiration.

Phreatophytes are a group of plants that send their roots down into the saturated groundwater table or into the capillary fringe just above the groundwater surface [37]. Salt Cedar, Cottonwood, Willow and Salt Grass are common phreatophytes. Meinzer [38] identified the correlation between phreatophytes and accessible groundwater in 1927. Robinson [39] has identified most of the phreatophytes found in the U.S. and has provided detailed descriptions of the more common phreatophytes.

Phreatophytes consume large quantities of water through evapotranspiration, and researchers in the semiarid western states have been investigating potential water conservation through phreatophyte elimination for over forty years. In 1952, Robinson [40) estimated the water consumed by phreatophytes in the western states at 20-25 million acre-feet per year. The magnitude of the potential benefits from reducing this consumptive use has encouraged numerous research efforts into the realizable benefits and associated costs of phreatophyte reduction [41, 42, 43, 44, 45, 46]. Many of the early efforts have been summarized by a Select Committee on National Water Resources [47]. Several areas have been investigated to determine consumptive use levels in addition to those studies cited above [48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60].

The phreatophyte areas along the Arkansas River channel between Pueblo Reservoir and John Martin Reservoir have been estimated (see Figure 8) in this study to be approximately 38000 acres and composed of light and medium density growth. Blaney and Criddle [51] estimated the annual water consumption in Colorado for these two densities at 28 and 35 inches respectively which are less than most estimates for other areas. Therefore, the annual consumption in the Lower Arkansas Valley of the phreatophytes could be as high as 90000 to 110000 acre-feet.

The independent variables affecting phreatophyte water consumption include solar radiation, species (type and size), climate variables, temperature, relative humidity and wind. These relationships are quite similar to those discussed below for agricultural crops, but several additional factors must be included in the phreatophyte case.

The need for phreatophyte roots to reach the water table introduces the depth to water table measurement as an important variable. In general, the phreatophyte consumption is reported [49; and many others] to vary inversely with the depth to groundwater. When the groundwater table is near the land surface, the capillary fringe may intercept the surface and create a more direct atmosphere-groundwater connection which will greatly increase the apparent evapotranspiration.

Another variable affecting evapotranspiration by phreatophytes is the growth density. An artificial term, volume-density [42], has been used to estimate the combined areal density (units per unit area) and the vertical density (the vertical concentration of foliage). A volume density of 100% represents maximum concentration. The growth density can vary dramatically from very sparse to extremely dense, and this variability is difficult to measure.

Phreatophyte water consumption is also affected by the underlying soils. Agricultural crops usually are planted in prepared ground that is relatively smooth and homogeneous. Phreatophytes, on the other hand, often grow in several different soil conditions within one contiguous area. This soil variability is suspected as the main cause for some inconsistent results in a recent research project [41].

The consumption characteristics also vary among the species of phreatophytes as is illustrated in Table 5 (from [47]). The values have been taken from numerous reported and unreported projects of several Federal and State agencies. All of the values except those for Alder and Mesquite were developed by growing the plants in tanks. The high consumption rates of Saltcedar and Cottonwood are two to eight times greater than the relatively low rates of Salt grass, Greasewood, and Mesquite. When natural phreatophyte stands contain two or more species the estimation of consumptive use becomes even more complex because

Table 5. Annual rate of water use by some common phreatophytes in Western United States[a]

Plant	Annual rate including precipitation	Volume density	Depth to water	Locality
	Acre-feet per acre	Percent	Feet	
Alder	5.3	—	6	Santa Ana River, Calif.
Batamole or seepwillow	4.7	100	3-4	Safford Valley, Ariz.
Cottonwood	7.6-5.2	100	6	San Luis Rey River, Calif.
Do.	6.0	100	—	Safford Valley, Ariz.
Greasewood	2.5-0.08	—	10	Escalante Valley, Utah
Mesquite	3.3	100	—	Safford Valley, Ariz.
Sacaton	4.0-3.5	—	—	Pecos River Valley, N. Mex.
Saltcedar	5.5-4.7	—	1-7	Do.
Do.	9.2-7.3	100	1.5-5	Safford Valley, Ariz.
Saltgrass	4.1-1.1	—	2-4	Owens Valley, Calif.
Do.	2.9-1.1	—	0.3-2.1	Santa Ana, Calif.
Do.	2.3-1.1	—	2.0	San Luis Valley, Colo.
Do.	4.5	—	0.65	Carlsbad, N. Mex.
Do.	2.6	—	0.4-3.1	Isleta, N. Mex.
Do.	4.0-0.8	—	2.2	Los Griegos, N. Mex.
Do.	1.9	—	1.9-2.6	Mesilla Dam, N. Mex.
Do.	2.3-1.6	—	2.0	Escalante Valley, Utah
Do.	2.0	—	2.0	Vernal, Utah
Willow	4.4	—	1.1	Santa Ana, Calif.
Do.	2.5	—		Isleta, N. Mex.

[a]From Ref. [47].

heterogeneous phreatophyte growths have received limited research attention.

The amount of water available to the phreatophytes can affect the consumptive use, with abundant water increasing the rate of evapotranspiration [61]. The quality of the water supply also affects water use, with higher dissolved solids concentrations reducing the evapotranspiration rate [39]. These two factors plus volume density differences where densities are unreported are the most probable explanations for the variations seen in Table 5 for the same species at different sites but similar climates.

Several methods have been employed to field measure evapotranspiration. Gatewood, et al. [42] describe six of these methods, (1) tank studies, (2) transpiration wells, (3) seepage runs, (4) inflow-outflow (i.e. the water budget approach), (5) chloride increase, and (6) the slope seepage method. Weeks and Sorey [41] employed a water budget method utilizing finite difference solution techniques, and a current project is using a water budget approach [62]. All of these methods are subject to severe measurement error, and in some of the methods some variables must be ignored, assumed, and/or left uncontrolled. Due to this high probability of error, Gatewood, et al. [42] were quite pleased when the six methods they employed each yielded results within 20% of the mean for the project.

Blaney [43] first attempted to estimate phreatophyte evapotranspiration using the Blaney-Criddle format in 1952. The Blaney-Criddle equation for estimating evapotranspiration is:

$$U = K \sum^{m} \frac{t \times p}{100} \tag{9}$$

where U is evapotranspiration in inches of water for a complete growing season, K is a consumptive use coefficient, p is the percentage of annual daylight hours in each month of the growing season, t is the mean monthly temperature in ^{o}F, and m is the number of months in the growing period. The constant K is derived empirically and is a function of most of the factors discussed above. The units for K are inches of water per ^{o}F per growing season. The Blaney-Criddle equation is also given as:

$$u = \frac{k \times t \times p}{100} \tag{10}$$

where t and p remain as defined above, u is evapotranspiration in inches per month, and k is a consumptive use coefficient similar to K but derived monthly instead of seasonally.

The Blaney-Criddle equation is an attractive approach to E_t estimation because of its dependence on readily available data for its two variables providing reliable estimates of K (or k) are available. The method has been used extensively for estimating the consumptive water use of crops because the K (or k) values for each type of crop appears to be transferrable from area to area. Blaney [61] has reported several attempts to estimate E_t using phreatophyte consumptive use coefficients from similar areas, but the approach has met with limited acceptance. Rantz [49] refined the Blaney-Criddle coefficients to include the effect of the depth to the groundwater in addition to species and density effects, but the use of these refined coefficients can still introduce large potential errors. The broad diversity of phreatophyte mixtures of species, densities, and growth conditions makes the use of consumptive use coefficient estimates from other areas too risky.

However, transferring coefficients into this study area is not necessary because the work of Weeks and Sorey [41] was performed in this same area. The data developed in their work can be used to estimate consumptive use coefficients in the study area. Weeks and Sorey attempted to measure the evapotranspiration at four sites near (1) Boone, (2) Las Animas, (3) Lamar, and (4) Holly, Colorado. The project covered the years 1966-1969, and the evapotranspiration was estimated with a water budget technique involving the solution of the groundwater flow equations using finite difference arrays. The project purpose was to evaluate the analytical technique. The results at Boone and Holly were not acceptable to the authors and the cause of the poor results were felt to be a high water table at Boone and the heterogeneous soil conditions at Holly. This research yielded total annual evapotranspiration values of 29 inches in 1966 and 26 inches in 1968 at the Las Animas site, and values of 23 inches in 1966, 21 inches in 1968, and 30 inches in 1969 at the Lamar site. Considering the high potential error these results appear to agree with the Blaney-Criddle estimates presented earlier in this discussion, but the most significant aspect of these results is that the estimates are considerably below the estimates for other western areas. While the exact cause of this phenomena cannot be identified, the extremely poor quality of the groundwater in the Arkansas Valley area and the very acute shortage of water during the dry season are probably major reasons for these lower consumptive use rates. As a final comparison, the K rates given by Rantz [49] were used to estimate annual evapotranspiration using a 5-foot depth to groundwater dimension (this value is about the mean of the Weeks and Sorey [41] observations). This yielded an estimate of 53 inches per year if the vegetation is cottonwood and willow and 60 inches if the vegetation is saltcedar. The Arkansas Valley phreatophytes are actually a mixture of these three common plants. This comparison supports the advisability of developing K rates from the Weeks and Sorey work.

As a result of the above considerations, phreatophyte evapotranspiration was simulated for the study area in the variable ETPH by developing a unit estimate of E_t for each observation and multiplying by the acres of phreatophytes found in each stretch. The unit estimate was in terms of inches E_t per inch of seasonal E_t, and the parameter estimate developed in the regression procedure will account for the number of inches of E_t actually occurring during the study period.

Evapotranspiration estimates are usually prepared for annual, seasonal, or monthly time intervals, and time intervals of less than one month are considered too short to give the estimate acceptable accuracy. This causes a problem with the model because the selected model interval is one day. This problem was circumvented by computing daily unit estimates of the monthly unit E_t estimate and entering 1/31th of that estimate as the daily observation. The equation is:

$$u' = kp \times f \times \frac{1}{31} \tag{11}$$

where u' is 1/31 of the monthly unit E_t estimate in inches per inch of seasonal E_t, kp is a monthly unit consumptive use coefficient in inches per inch of seasonal E_t-^{o}F-% sunlight, and f is the consumptive use factor. The variable f is defined as:

$$f = \frac{t \times p}{100} \tag{12}$$

where f is in ^{o}F-% sunlight, t is the average daily temperature for each stretch in ^{o}F, and p is % of annual sunlight per month. This latter variable was constructed as a daily estimate of a monthly value using a daily estimate of sunlight hours for the study segment's latitude and then converting to % annual per month with the factors $\frac{3100}{4453}$ or $\frac{3000}{4453}$ depending on the number of days in the month. Average daily temperatures were not available for the weather stations in the study area so t was calculated as the average of the daily high and low values. The temperature data for each stretch were taken from the following station records [63].

Stretch 1 — Pueblo WSO AP — Index 6740

Stretch 2 — Pueblo WSO AP — Index 6740 and

Rocky Ford 2 SE — Index 7167

Stretch 3 — Rocky Ford 2 SE — Index 7167

Stretch 4 — La Junta FAA AP — Index 4720

Stretch 5 — Las Animas — Index 4834

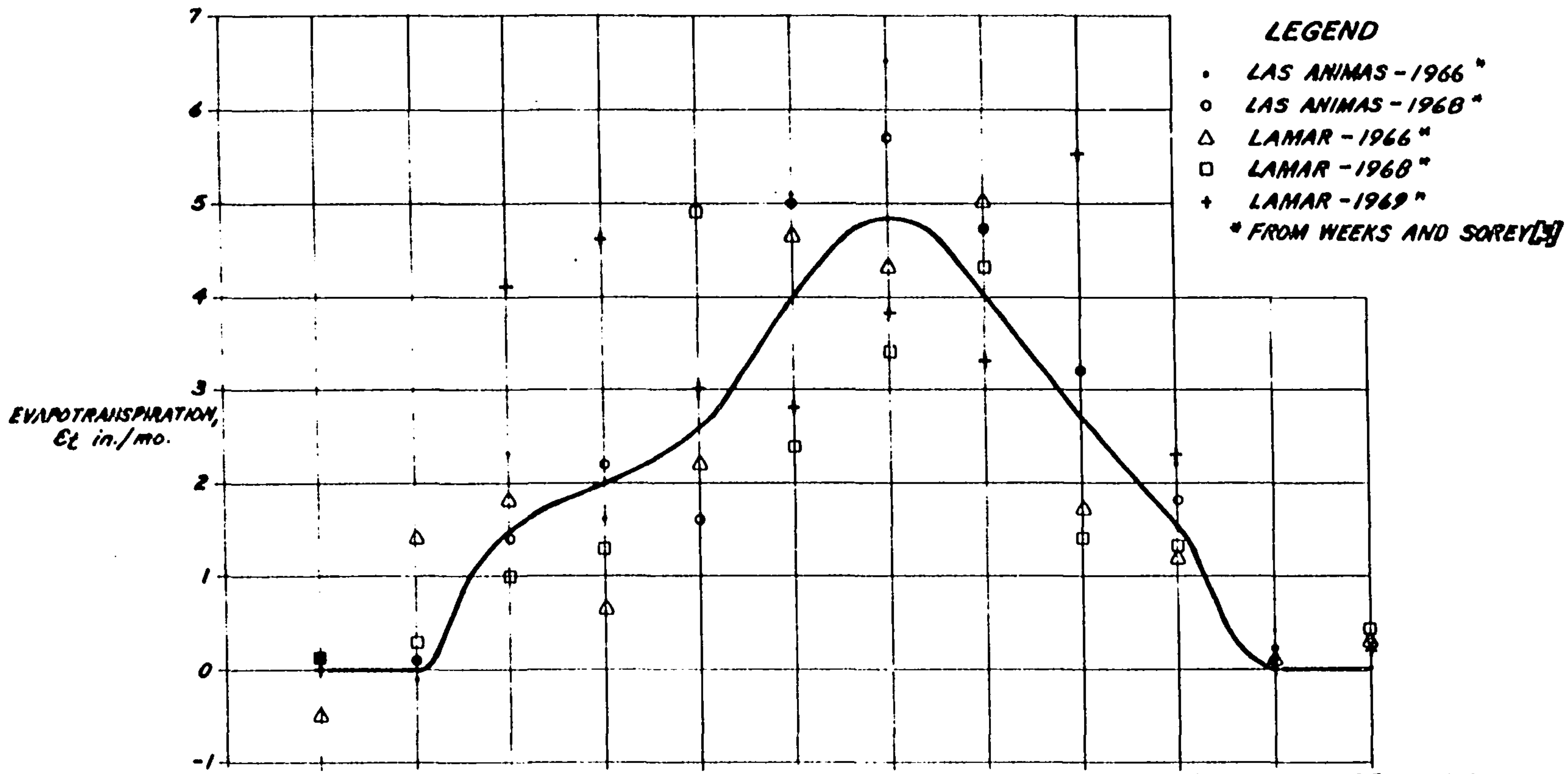
Two stations were averaged for stretch 2 because the stretch is midway between the stations. The other stretches either contain or are near the stations used.

The daily estimate unit monthly consumptive use coefficient k_p was calculated with the equation:

$$k_p = \frac{\hat{u}_d}{f_d} , \tag{13}$$

where $\hat{u}_d$ is a daily estimate of unit monthly evapotranspiration in inches per inch annual E_t and f_d is a daily estimate of a monthly consumptive use factor. The estimation of $\hat{u}_d$ utilized the monthly evapotranspiration data reported by Weeks and Sorey [41]. This data bank was gathered over the four-year period, 1966-1969, at two sites in the lower Arkansas River valley. The site near Las Animas is located between stretches 4 and 5 in the study segment and the other site near Lamar is about 20 miles downstream from the study segment. Data from both sites are plotted in Figure 11, and a smooth curve has been visually fitted to the data. Monthly values of E_t were read from the curve and are shown along the bottom of Figure 11. The monthly values are also represented as percentages of the sum of the monthly values below each month's estimate in Figure 11. These monthly percentages were then plotted in bar graph form in Figure 12, and a smooth curve was visually fitted so that the areas under both curves are approximately equal. Daily estimates of $\hat{u}_d$ were then read from the ordinate in Figure 12.

The estimation of f_d used a similar curve fitting technique. Normal monthly temperatures [63] and monthly values of p were used to calculate monthly consumptive use factors for five stations in and above the study segment. Seasonal consumptive use factors were calculated by summing the monthly values, and each seasonal value was compared with the average for the five stations. The results of these calculations and the values for T and p are shown in Table 6. Since the deviation from the mean among the five stations is quite small (1.34% maximum) a single consumptive use factor for the whole study segment was developed instead of developing factors for each stretch. The monthly factors were averaged for the five stations and the percentage of the seasonal factor was calculated for each month. These values



FITTED CURVE VALUES		JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.	ΣE_t
	E_t, in.	0	0	1.5	2.0	2.6	4.0	4.8	4.0	2.7	1.5	0	0	23.1
	% of ΣE_t	0	0	6.49	8.66	11.26	17.32	20.77	17.32	11.69	6.49	0	0	100

Figure 11. Monthly distribution of evapotranspiration by phreatophytes. Lower Arkansas River valley

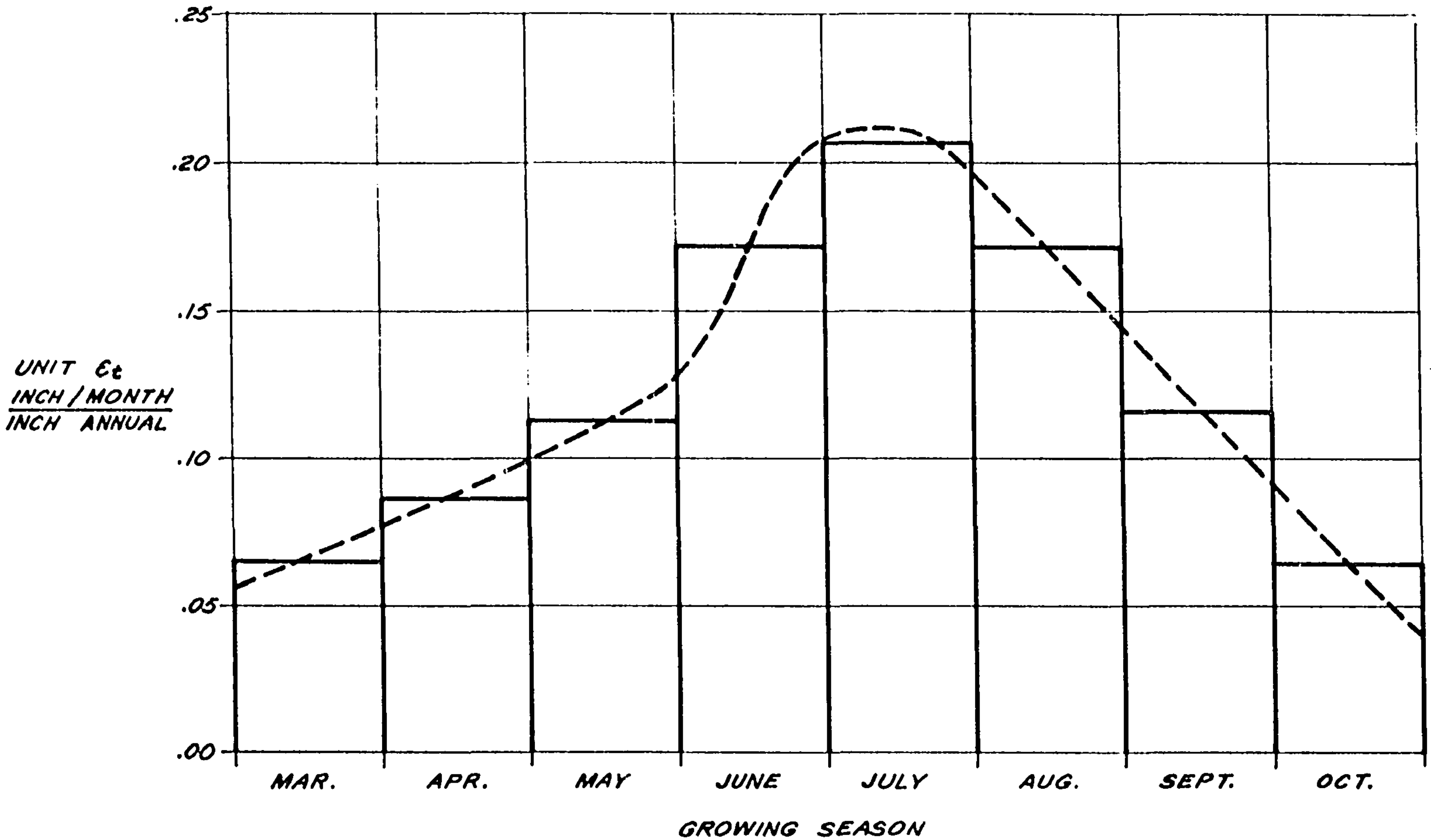


Figure 12. Estimation curve for daily unit evapotranspiration (E_t) by phreatophrytes

Table 6. Monthly consumptive use factors, for normal temperatures [63]

		Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.
Canon City	T_N	—	—	42.0	51.8	60.8	69.5	75.5	73.9	66.2	56.4
Index No. 1294	$\frac{TP}{100}$	—	—	3.50	4.61	6.03	6.92	7.63	7.00	5.55	4.40
Pueblo WSO	T_N	—	—	40.0	51.7	61.1	70.4	76.4	74.5	66.2	54.5
Index No. 6740	$\frac{TP}{100}$	—	—	3.34	4.60	6.06	7.00	7.72	7.06	5.55	4.25
Rocky Ford	T_N	—	—	41.0	52.9	62.1	71.5	76.5	74.6	66.3	58.4
Index No. 7167	$\frac{TP}{100}$	—	—	3.42	4.71	6.16	7.11	7.73	7.06	5.56	4.56
Las Animas	T_N	—	—	41.2	53.9	63.4	72.9	78.1	76.1	67.6	55.5
Index No. 4834	$\frac{TP}{100}$	—	—	3.44	4.80	6.29	7.25	7.89	7.21	5.66	4.33
John Martin Dam	T_N	—	—	41.3	53.4	63.1	72.8	78.3	76.3	67.8	55.8
Index No. 4388	$\frac{TP}{100}$	—	—	3.44	4.75	6.26	7.24	7.91	7.23	5.68	4.35
	P	—	—	8.34	8.90	9.92	9.95	10.10	9.47	8.38	7.80
Avg. $\frac{TP}{100}$		—	—	3.43	4.69	6.16	7.11	7.77	7.11	5.60	4.38
% $\Sigma \frac{TP}{100}$		—	—	7.42	10.14	13.32	15.37	16.80	15.37	12.11	9.47

Table 6. Continued

		Nov.	Dec.	$\Sigma \frac{TP}{100}$	% of avg.
Canon City	T_N	—	—	—	
Index No. 1294	$\frac{txp}{100}$	—	—	45.63	98.66
Pueblo WSO	T_N	—	—	—	
Index No. 6740	$\frac{TP}{100}$	—	—	45.57	98.53
Rocky Ford	T_N	—	—	—	
Index No. 7167	$\frac{TP}{100}$	—	—	46.30	100.11
Las Animas	T_N	—	—	—	
Index No. 4834	$\frac{TP}{100}$	—	—	46.86	101.32
John Martin Dam	T_N	—	—	—	
Index No. 4388	$\frac{TP}{100}$	—	—	46.87	101.34
	P	—	—	—	—
Avg. $\frac{TP}{100}$		—	—	46.25	100
% $\Sigma \frac{TP}{100}$		—	—	100	

are also shown in Table 6. The monthly average factors were then plotted in bar graph form in Figure 13. A smoothed curve was visually fitted to the bar curve so that the areas under each curve are approximately equal, and daily estimates of f_d were read from the ordinate in Figure 13.

Numerous estimates were read from the smooth curves in Figures 12 and 13 and k_p estimates were calculated with equation 10. The resulting smooth curve is shown in Figure 14. Values of k_p were then read from the ordinate in Figure 14 and used for estimation of ETPH observations using equation 11. The sign of the ETPH observations is always positive, and since ETPH represents an outflow from the surface system, the sign of the parameter should also be positive.

The inflow to the aquifer from soil moisture bank deep percolation and the outflow through the wells were modeled conjunctively. The combining of these two variables is based on their inherent interrelationship because of the structure of the surface water, cropping, and groundwater use regime. Whenever the consumptive water needs of the crops exceeds the quantity of surface water available, the soil moisture bank declines. This negative soil moisture stress reduces the deep percolation to the aquifer, and the farmer will normally increase his groundwater use. As a result, negative soil moisture stress produces an outflow increase from the groundwater aquifer. Conversely, when the surface water supply exceeds the consumptive needs of the crops, the use of groundwater declines and the moisture content of the soil bank will increase resulting in more deep percolation. This positive soil moisture stress results in a decrease in outflow and an

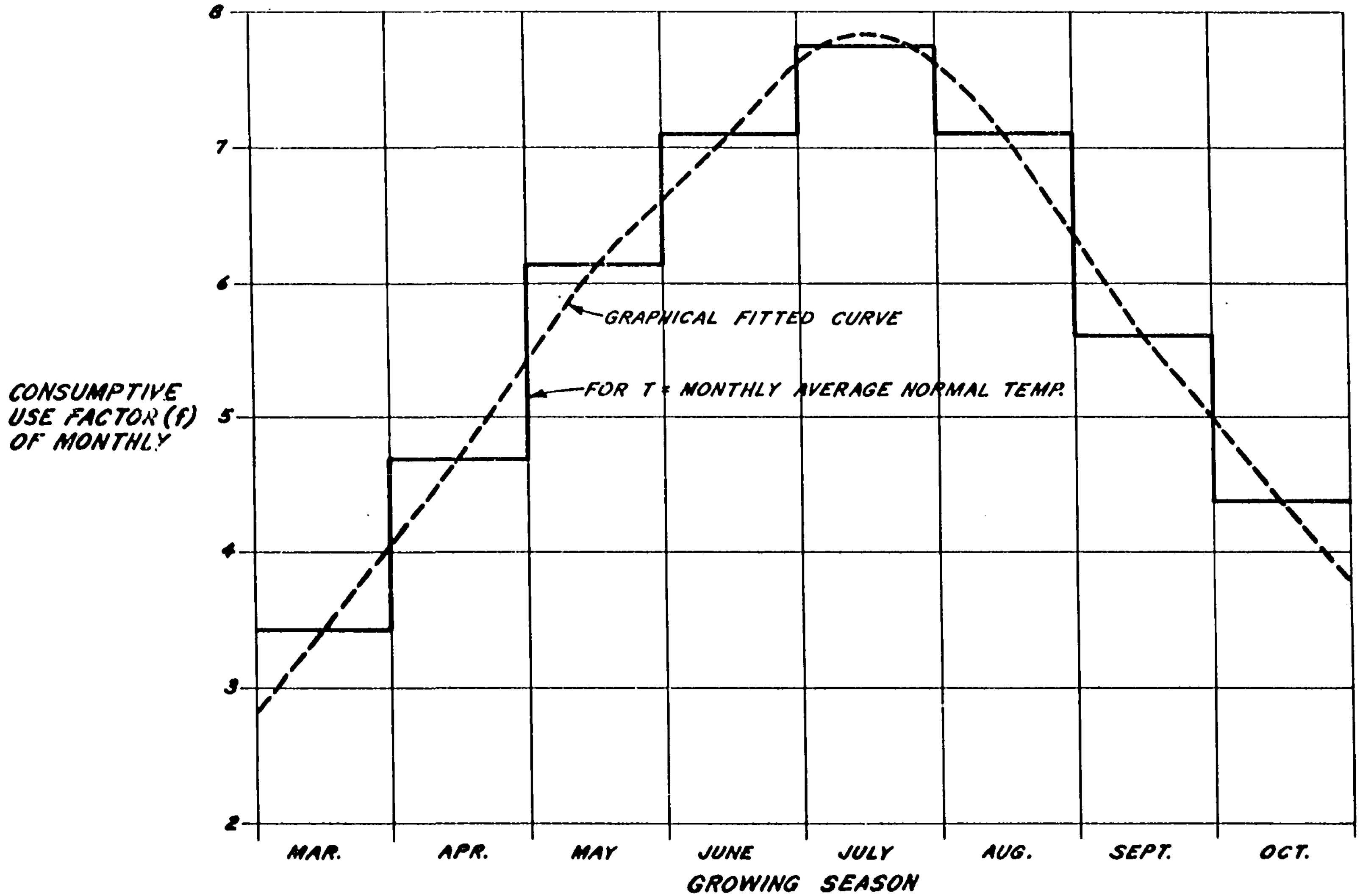


Figure 13. Estimation curve for monthly normal consumptive use factor (f)

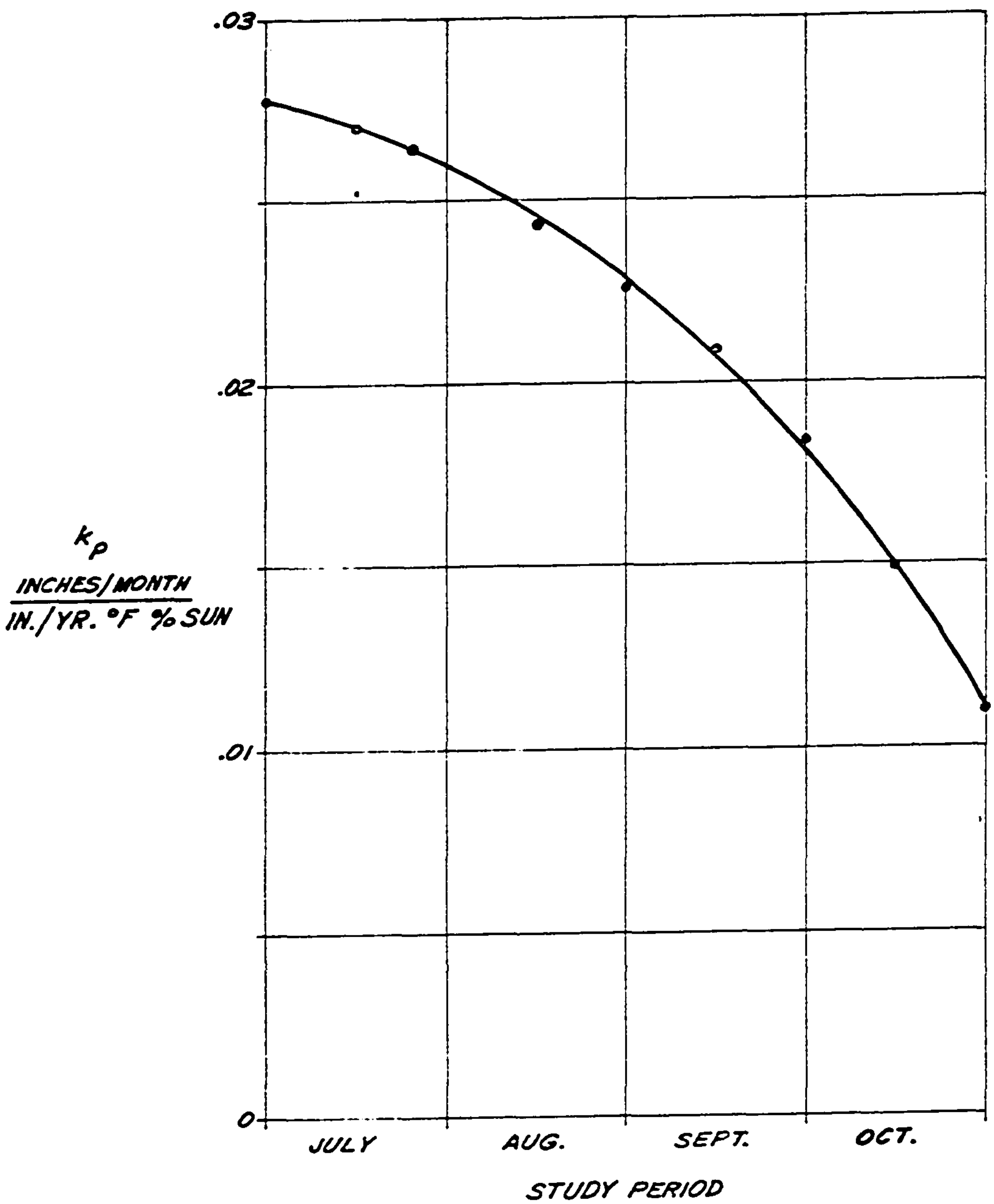


Figure 14. Estimation curve for k_p

inflow to the aquifer. A soil moisture stress concept was used to conjunctively model the deep percolation inflow and well use outflow variables.

Two activities are the principal elements in the soil stress variable: (1) the amount of surface water applied to the crops, and (2) the consumptive use demands of the crops. The first element has been modeled above in the variables TRWA and NTRWA, and the modeling of the latter element is described below.

Consumptive use and evapotranspiration are usually considered synonymous [64], and include water used by plants for tissue and transpiration and the evaporation of soil moisture and intercepted or standing water. Interest in estimation of crop water needs has existed since near the beginning of agriculture, and reported U.S. research on this subject began appearing in the late 1800's when irrigated agriculture began to develop in the western part of the country. Research interest has continued to grow for about a hundred years, and today consumptive use is still a very popular research topic. As a result of this broad and growing research interest, the literature is now quite large and contains numerous models for estimating consumptive use. However, the literature does not contain a conclusive basis for choosing the consumptive use model to be used in this study.

The factors affecting consumptive use are numerous and include (1) climatic variables, (2) crop variables, and (3) geologic and hydrologic variables. Munson [65] lists 38 variables (or groups of variables) that form the bulk of these three variable groups. This large group of causative factors is a primary reason for the large number of consumptive use models. Several different treatments of these variables are illustrated in the comparison of consumptive use models in Table 7. The seventeen models chosen for this comparison include those models most commonly researched and applied but do not represent a complete list of all consumptive use models.

The comparison in Table 7 shows that consumptive use models do not all predict the same quantity. Two models predict lake evaporation which is sometimes also accepted as an estimate of evapotranspiration. Some models predict specially defined quantities, e.g. Model ID #16. The majority of the models predict potential evapotranspiration (E_{tp}) which is the amount of water that will be consumptively used during a unit time interval when an adequate water supply is available. Some of the potential evapotranspiration estimates are specific to a reference crop (E_{tpr}), e.g. alfalfa or grass, while other estimates represent an average for several crops. The potential evapotranspiration estimates can be converted to other specific crop E_{tpc} estimates using the equation

$$E_{tpc} = K_{co} E_{tp} \text{ (or } E_{tpr}\text{)}, \tag{14}$$

where K_{co} is a crop coefficient (see [64], pp. 87-89). The crop coefficient typically increases from the season's start until the crop foliage is well developed and then remains constant until the crop matures. The E_{tp} estimates [64] can be converted to evapotranspiration (Et) estimates for conditions where a scarce water supply limits consumptive use by reducing K_{co} with the equation,

Table 7. Comparison of consumptive use models

Model ID #	Model author and references	Predicted quantity[a]	Time interval[b]	Independent variables: Δ[c]	Independent variables: Atmospheric pressure	Independent variables: Specific heat
1	Penman [66]	E_{tpr}	D	X	X	X
2	Kohler, Nordenson, & Fox [67]	LE	D	X	X	
3	vanBavel, Businger [68, 69]	E_{tp}	H, D	X	X	X
4	Makink [70]	E_{tpr}	10 days	X	X	X
5	Christiansen [71]	E_{tp}	M			
6a	Jensen-Haise [72]					
6b	Jensen [73]	E_{tpr}	5 days			
6c	Jensen, Robb, & Franzoy [74]					
7	Stephens [75]	E_{tpr}	10 days			
8	Lane [76]	LE	M			
9	Christianson & Hargreaves [77, 78]	E_{tp}	M			
10	Papadakis [79]	RE_t	M			
11	Hamon [80]	E_{tp}	D			
12	Blaney-Criddle [81, 82, 83, 84]	E_{tpc}	A, M			
13	Thornthwaite [85]	E_{tp}	M			
14	Lowry-Johnson [86]	E_{tp}	A			
15	Munson [65]	E_{tp}	M			
16	Olivier [87]	BWR	M			
17	Behnke-Maxey [88]	E_{tp}	M			

[a] LE = lake evaporation, E_{tp} = potential evapotranspiration, E_{tpc} = potential evapotranspiration for for crop c, E = evaporation, E_{tpr} = potential evapotranspiration for a specific reference crop Y, BWR = basic water requirements.

[b] H = hourly, D = daily, M = monthly, A = annual (or seasonal). Time periods shown are the most commonly recommended intervals.

[c] Δ = slope of the saturation vapor pressure-temperature curve.

Table 7. Continued

Model ID #	Predicted quantity[a]	Time interval[b]	Independent variables: Latent heat-vaporization	Net radiation	Solar radiation	Wind velocity	Vapor pressure	Temperature
1	E_{tpr}	D	X	X		X	X	X
2	LE	D		X		X	X	
3	E_{tp}	H, D	X	X		X	X	X
4	E_{tpr}	10 days	X		X			
5	E_{tp}	M			X	X		X
6a								
6b	E_{tpr}	5 days			X		X	X
6c								
7	E_{tpr}	10 days			X			X
8	LE	M			X			X
9	E_{tp}	M				X		X
10	RE_t	M					X	
11	E_{tp}	D						X
12	E_{tpc}	A, M						X
13	E_{tp}	M						X
14	E_{tp}	A						X
15	E_{tp}	M						X
16	BWR	M						X
17	E_{tp}	M						X

Table 7. Continued

Model ID #	Predicted quantity[a]	Time interval[b]	Independent variables: Humidity	Evaporation	Sunlight	Elevation	Water requirement constant[d]	Air density	PE index[e]	Heat index[f]
1	E_{tpr}	D								
2	LE	D								
3	E_{tp}	H, D				X		X		
4	E_{tpr}	10 days								
5	E_{tp}	M	X							
6a										
6b	E_{tpr}	5 days				X				
6c										
7	E_{tpr}	10 days								
8	LE	M								
9	E_{tp}	M	X	X	X					
10	RE_t	M		X						
11	E_{tp}	D	X		X					
12	E_{tpc}	A, M			X					
13	E_{tp}	M			X					X
14	E_{tp}	A								
15	E_{tp}	M							X	
16	BWR	M	X				X			
17	E_{tp}	M					X			

[d]Water requirement constant includes the effects of latitude and temperature on radiation.

[e]PE index = a statistical index derived from precipitation and evaporation data.

[f]Heat index = an index based on temperature that was developed for U.S. East Coast conditions.

$$K = K_{co} \frac{\ln(AW + 1)}{\ln(101)} , \tag{15}$$

where AW is the percentage of the potential soil moisture bank that is available to the crops. However, the Blaney-Criddle model predicts potential evapotranspiration for a number of individual crops using crop coefficients that have been developed solely for this model, and therefore does not use the crop coefficients, K_{co}, discussed above.

Consumptive use estimates have been developed for a broad variety of applications, and this variety has caused the wide range of time intervals used in the estimation models. For example, the shorter interval models, i.e. hourly, daily or 5 and 10 days, are used for operating irrigation systems, and the longer interval models are used for sizing water resource projects, negotiation of compacts and treaties, and water rights litigation and adjudication. A daily interval model is preferred for this study, but these shorter interval models require a larger amount of input data. Typically, the application of the more complex shorter interval models has been severely restricted because of these data limitations.

The climatic data [63] available for one or more stations in the study area include:

1. Precipitation — daily, monthly, annual, normal
2. Temperature, air — daily, monthly, annual, normal
3. Evaporation — daily, monthly
4. Wind — daily, monthly
5. Degree days — monthly, seasonal, normal
6. Relative humidity — 4 days per month.

An effort was made to locate either net or solar radiation data for the study area, but no radiation data could be found. Helton [31] also encountered this data gap, and he overcame the problem by correlating the data in the area with data from a station near Akron, Colorado, which is approximately 150 air miles north of the area. Even though Helton achieved a high correlation, the method introduces another source of high potential error, and the magnitude of this potential error is difficult to assess. A choice between synthesizing data as Helton did or using a different model should be based on the criteria of minimizing the potential error, but this evaluation was not possible in this situation. Therefore, the models using a radiation variable were arbitrarily not considered further for use in this study. This data gap eliminates eight of the models listed in Table 7 including five (ID's 1, 2, 3, 5, and 6) models that have received considerable research attention.

Of the remaining nine models, four (ID's 13, 15, 16 and 17) were eliminated because they use special variables, e.g. a water requirement constant, that were developed for areas quite dissimilar to the study area. The Lowry-Johnson model (ID 14) was eliminated because the available data can support one of the more complex models. The Hamon (ID 11) and Papadakis (ID 10) models have received very limited research and application compared to the Blaney-Criddle (ID 12) and Christiansen and Hargreaves (ID 9) models, so these latter two models were given the most comprehensive consideration for this study. Cruff and Thompson [89] have compared six of the models (ID's 2, 8, 11, 12, 13, and 14) in Table 7, and their results are summarized in Table 8. The Blaney-Criddle model appears to yield more accurate results than all but the Kohler, et al. model (ID 2) which supports the choice of the Blaney-Criddle model over the Hamon model. The ASCE study group [64] has compared a large number of models including the two models most seriously considered for this study, and the results indicate that the Christiansen and Hargreaves model is more accurate. However, the ASCE report also points out that this model may be difficult to calibrate. The Christiansen and Hargreaves model also requires humidity data which is only partially available for the study area.

The Blaney-Criddle model has been applied more times than any other model shown in Table 7 [43, 46, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 61, 64, 81, 82, 86, 90, 91], and many of these applications have been in the Western United States. The Soil Conservation Service of the U.S. Department of Agriculture has adopted the Blaney-Criddle model for consumptive use estimations, and has refined and developed the model extensively. The current SCS techniques are well documented in the SCS Technical Release No. 21 [81] which includes crop coefficient estimation curves for 25 crops, and this wealth of documentation was a primary factor in the selection of the SCS version of the Blaney-Criddle model for the consumptive use estimates in this study. Other factors that contributed to this selection were the limited amount of humidity data, potential problems in estimating E_{tp} for specific crops using the Christiansen and Hargreaves model, and the difference in computational effort required. The sacrifice in estimation accuracy was considered acceptable because the regression procedure can compensate

Table 8. Comparison of E_t estimation accuracy for selected models[a]

Method	Blaney-Criddle # Sta	Blaney-Criddle % Err	Kohler, et al. # Sta	Kohler, et al. % Err	Thornthwaite # Sta	Thornthwaite % Err	Lowry-Johnson # Sta	Lowry-Johnson % Err	Hamon # Sta	Hamon % Err	Lane # Sta	Lane % Err
Calendar year												
Arid climate	9	-31	1	-1	9	-54	7	-48	9	-49	7	-8
Modified arid climate	4	+5	3	+14	4	-29	4	-21	4	-17	4	+41
Subhumid climate	10	-4	3	-1	10	-33	10	-22	10	-35	9	+22
Total sta/aug % E	23	-13	7	+5	23	-41	21	-30	23	-38	20	+16
Growing season												
Arid climate	10	-26	1	-1	10	-46	8	-46	10	-44	8	-5
Modified arid climate	4	+11	3	+11	4	-18	4	-22	4	-11	4	+41
Subhumid climate	11	-1	3	-3	11	-31	11	-19	11	-34	10	+23
Total sta/aug % E	25	-9	7	+3	25	-35	25	-29	25	-35	22	+16

[a]From Cruff and Thompson [89].

for a consistent error in the estimate when the regression parameters are estimated.

The SCS modified Blaney-Criddle model used in this study is:

$$u = kf \tag{16}$$

where u is monthly consumptive use of a specific crop in inches, and k is an empirical consumptive use crop coefficient. The monthly consumptive use factor is defined by equation 12 above. The variables u and f are similar for the consumptive use estimates for both crops and phreatophytes (see above discussion of the construction of ETPH). The temperature data and the consumptive use factor f were developed for the crop consumptive use estimate by the same method used in estimating phreatophyte evapotranspiration for the variable ETPH as discussed above.

However, the factor k is defined differently for the crop phreatophyte estimates. In the crop consumptive use estimates, the crop coefficient is defined as:

$$k = k_t k_c, \tag{17}$$

where k_t is mean monthly air temperature in ^oF. The growth stage crop coefficient, k_c, was read from SCS curves [82], and each k_c is for a specific crop. The curves plot k_c versus calendar date for the rest of the crops. Growing seasons were assumed for several crops in the study area, and the assumed seasons are shown in Table 9. The assumptions are based on several oral contacts with residents in the study area and data provided by the Statistical Reporting Service of

Table 9. Assumed crop growing seasons

Crop	Beginning date	End date
Field corn (grain)	May 11	Oct. 8
Dry beans	June 16	Oct. 1
Sugar beets	May 15	Nov. 5
Field corn (silage)	May 11	Oct. 8
Spring grain	April 23	Aug. 1
Grain sorghum	June 10	Dec. 1
Small vegetables	May 1	Sept. 1
Winter wheat	—	July 20

the U.S. Department of Agriculture[1]. The curves provided daily estimates of monthly k_c values, and the coefficients were not modified to reflect any impact from an insufficient water supply.

Two consumptive use variables were constructed for each stretch: (1) CON1 was constructed for tributary areas and (2) CON2 was constructed for nontributary areas with the distinction between tributary and nontributary being identical with the method used in constructing TRWA and NTRWA as discussed above. Each consumptive use variable is the summation of the individual crop consumptive use estimates within its appropriate land area. The variables were constructed in units of cfs to be compatible with the TRWA and NTRWA variables.

[1]Statistical Reporting Service Office, U.S. Department of Agriculture, Denver, Colorado, personal communication, 1975.

The use of individual crop coefficients, k_c, required the disaggregation of tributary and nontributary areas into crop acreages for each area. This disaggregation was accomplished in several steps which began with the estimation of the percent of irrigated acres in each county containing the study area allocated to each crop. These percentages were estimated from data found in Colorado Agricultural Statistics [84] for the years 1966-1970 and 1972. The percentage of total irrigated acres in each county allocated to each crop was calculated for each of the six years of data. The percentages were then averaged for each county and are presented in Table 10. These average percentages were used to estimate the percentage of acres in each irrigation system allocated to each crop. Since the Colorado Canal is the only large irrigation system in Crowley County, the percentages for that county were assigned directly to that system. In the other three counties the percentages estimated for each system were modified according to the general strength of each system's water rights. The systems with higher priority rights, e.g. the Rocky Ford Canal, were assigned larger proportions of the higher consumptive use crops, e.g. corn. Irrigation systems were given the same percentages in all counties. The results of this estimation procedure for the systems in the study area are presented in Table 11. Barley, oats, and spring wheat were combined in the spring grains category, and potatoes was included in the vegetable category. Some estimates were influenced by additional available information. For example, the Rocky Ford, Catlin, and Oxford Farmers systems are known to produce a large proportion of vegetables so most of the Otero County vegetable

Table 10. Percentage of total irrigated acres by crop

Crop	County: Pueblo	Crowley	Otero	Bent
Spring wheat	0	0	0.1	0
Winter wheat	3.2	1.4	2.6	8.1
Corn-grain	17.5	9.6	20.0	6.1
Corn-silage	7.3	8.2	12.2	3.5
Barley	1.8	0.3	1.9	1.0
Sorghum-grain	4.8	20.0	7.9	32.3
Dry beans	8.5	2.4	1.1	0
Sugar beets	0.5	0.3	1.5	0
Oats	0.7	0.3	1.2	0.7
Hay	51.5	57.5	46.3	47.1
Potatoes	0	0	0.6	0
Vegetables	4.2	0	4.6	1.2
Total	100.0	100.0	100.0	100.0

production was assigned to these three systems. The Rocky Ford Hiline and Oxford Farmers systems were also influenced by special reports [26, 92] on those systems. Some crops, e.g. melons, do not appear in the analysis because the percentages for those crops rounded off to zero.

The crop acreage estimates for tributary and nontributary areas in each stretch were then calculated by multiplying the percentage of a particular crop in a particular system by the total number of

Table 11. Percentages of total system irrigated acres by crop

Crop	Irrigation systems						
	Bessemer Ditch	Excelsior Ditch	Collier Canal	Colorado Canal	R. F. Hiline Canal	Oxford Far. Ditch	Otero Canal
Spring grain	2.5	2.5	2.5	0.6	1.5	0.4	8.0
Winter wheat	3.2	3.2	3.2	1.4	1.3	0.5	6.6
Corn-grain	17.5	17.5	17.5	9.6	18.9	22.3	19.0
Corn-silage	7.3	7.3	7.3	8.2	13.8	12.0	10.5
Sorghum-grain	4.8	4.8	4.8	20.0	5.4	0.8	10.8
Dry beans	8.5	8.5	8.5	2.4	2.0	1.4	2.4
Sugar beets	0.5	0.5	0.5	0.3	1.0	1.7	1.5
Hay	51.5	51.5	51.5	57.5	52.2	51.1	41.2
Vegetables	4.2	4.2	4.2	0	3.9	9.8	0
Totals	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 11. Continued

Crop	Irrigation systems: Catlin Canal	Holbrook Lake Canal	Rocky Ford Canal	Ft. Lyon Canal	Las Animas Consolidated	Highland Canal	A. J. Anderson
Spring grain	0	8.0	0	2.8	1.0	1.7	3.2
Winter wheat	0	6.6	0	7.8	8.1	8.1	2.6
Corn-grain	20.2	19.0	23.3	8.4	6.1	6.1	20.0
Corn-silage	11.2	10.5	12.9	4.7	3.5	3.5	12.2
Sorghum-grain	11.6	10.8	0	28.5	32.3	32.3	7.9
Dry beans	0	2.4	0	0.4	0	0	1.1
Sugar beets	1.6	1.5	1.9	0.3	0	0	1.5
Hay	44.5	41.2	51.0	46.1	47.1	47.1	46.3
Vegetables	10.9	0	10.9	1.0	1.2	1.2	5.2
Totals	100.0	100.0	100.0	100.0	100.0	100.0	100.0

acres irrigated by that system in that area. The consumptive use variables were then calculated with the equation:

$$\text{Consumptive Use (CON1 or CON2)} = \sum^{K}\sum^{L} Ac_{kl} \times (k_c)_{li} \times (k_t)_i \times (f)_i \times 0.042, \qquad (18)$$

where Ac_{kl} is the acres of crop l served by system c, $(kc)_{li}$ is the K_c estimate for crop l on observation i, and $(k_t)_i$ and $(f)_i$ are the values of k_t and f for observation i. The individual estimates are summed for K systems and L crops. The constant 0.042 converts acre-inches per day to cfs.

The variables TRWA, NTRWA, CON1, and CON2 were then used to construct the variables. Tributary area SOil STress, TSOST, and Non-Tributary area SOil STress, NTSOST were not used independently in the water balance model.

The equations:

$$\text{TSOST} = \text{CON1} - \text{TRWA, and}$$

$$\text{NTSOST} = \text{CON2} - \text{NTRWA} \qquad (19)$$

were used initially. Whenever TSOST and NTSOST are positive, the variables are simulating a condition where the SAWE element is exchanging water from the stream to the aquifer which would be a positive quantity in the water balance model. Therefore, the expected sign of the water balance model parameter is positive when the dominant element in TSOST and NTSOST in the groundwater withdrawal effect. When the parameter sign is negative the dominant effect will be the surface runoff from the applied irrigation water and/or groundwater exfiltration.

Inspection of the consumptive use and surface water applied variables revealed that the latter never exceeded the former so that TSOST and NTSOST were always positive. While this constant relationship may actually exist in the study area, it also may result from overestimation of the consumptive use variables. Assuming that the soil stress variable should be both positive and negative, two additional sets of soil stress variables were prepared. One set was calculated with the equations:

$$\mathrm{TSOST} = \frac{\mathrm{CON1}}{\overline{\mathrm{CON1}}} - \frac{\mathrm{TRWA}}{\overline{\mathrm{TRWA}}}\text{ , and}$$

$$\mathrm{NTSOST} = \frac{\mathrm{CON2}}{\overline{\mathrm{CON2}}} - \frac{\mathrm{NTRWA}}{\overline{\mathrm{NTRWA}}}\text{ ,} \tag{20}$$

where $\overline{\mathrm{CON1}}$, $\overline{\mathrm{CON2}}$, $\overline{\mathrm{TRWA}}$, and $\overline{\mathrm{NTRWA}}$ are the mean values of each respective variable. The second set was calculated with the equations:

$$\mathrm{TSOST} = \frac{\mathrm{CON1} - \overline{\mathrm{CON1}}}{\mathrm{SD(CON1)}} - \frac{\mathrm{TRWA} - \overline{\mathrm{TRWA}}}{\mathrm{SD(TRWA)}}\text{ , and}$$

$$\mathrm{NTSOST} = \frac{\mathrm{CON2} - \overline{\mathrm{CON2}}}{\mathrm{SD(CON2)}} - \frac{\mathrm{NTRWA} - \overline{\mathrm{NTRWA}}}{\mathrm{SD(NTRWA)}}\text{ ,} \tag{21}$$

where SD(CON1, CON2, TRWA, or NTRWA) is the standard deviation for the respective variables.

The relative contributions of these three sets of soil stress variables were evaluated during the study.

The Δs variable

The Δs element of the water balance model was first modeled by estimating the daily change in cross-sectional area along the stream axis in each stretch [93]. The change in the vertical dimension was calculated with the stage-discharge relationship for the two boundary gages. This approach was abandoned for several reasons. First, the calculation of stage at the upstream station requires knowledge of the QIN value which means that an independent variable in the model contains an element of the dependent variable, and this method of constructing Δs does not permit the elimination of this unacceptable condition. This first approach was also abandoned because the submodel was unstable at very low flows. Furthermore, this approach assumed a constant stream surface area which is a difficult assumption to accept over the range of flows found in the observations.

The second approach to the Δs variable resembles a "black box" technique and was suggested by Sauer's [93] unit-response method of flow routing. The approach assumes that the regression procedure will estimate a relationship between Δs and an impulse(s) that triggers the Δs function, i.e. the "black box", if the impulse(s) are included in the model as independent variables. The change in flow between consecutive observations is selected as the impulse variable on the assumption that an increase in an inflow or a decrease in an outflow will increase the storage activity and conversely the opposite changes in flow magnitude will decrease the storage activity.

The approach used in the water balance model assumes a linear relationship between Δs and $\Delta Q/\Delta t$ and is written as:

$$\Delta s_t = (P_1)(\Delta Q/\Delta t)_m + (P_2)(\Delta Q/\Delta t)_{m-1}$$
$$+ (P_3)(\Delta Q/\Delta t)_{m-2} \; .... + (P_n)(\Delta Q/\Delta t)_{m-n+1} \qquad (22)$$

where $\Delta Q/\Delta t$ is the change in an inflow or outflow in a one day period, m denotes the period immediately preceding the observation, and m-1 is the period immediately preceding the previous day's observation, with a total of n $\Delta Q/\Delta t$ quantities affecting Δs for observation t. The quantities P_1, P_2, P_3, ..., P_n represent the parameters that define the linear relationship, and these quantities must be estimated. The signs of the ΔQIN and ΔQDIS variable parameters should be positive, and the signs of the ΔQOUT and ΔQDIV variable parameters should be negative.

The $\Delta Q/\Delta t$ quantities were estimated with the equation:

$$(\Delta Q/\Delta t)_{m-n} = Q_{m-n} - Q_{m-n-1} \qquad (23)$$

where n = 0, 1, 2 .... N. Various values of n were evaluated in the study to select an optimum value of N, and these results are discussed in the next chapter. For discussion purposes, a value of 2 has been assumed for N.

Substituting water balance variables into equation 23, the Δs variable becomes:

$$\Delta s = (P_{1s})(QIN_t - QIN_{t-1}) + P_{2s}(QIN_{t-1} - QIN_{t-2})$$
$$+ (P_{3s})(QOUT_t - QOUT_{t-1}) + (P_{4s})(QOUT_{t-1} - QOUT_{t-2})$$
$$+ (P_{5s})(QDIV_t - QDIV_{t-1}) + (P_{6s})(QDIV_{t-1} - QDIV_{t-2})$$

$$+ (P_{7s})(QDIS_t - QDIS_{t-1}) + (P_{8s})(QDIS_{t-1} - QDIS_{t-2}), \tag{24}$$

where the subscript s denotes the stretch. The number of variables could be five times as large as shown if no stretches are combined. Rearranging equation 24 produces:

$$\begin{aligned} \Delta s = {} & (P_{1s})(QIN_t) + (P_{2s} - P_{1s})(QIN_{t-1}) + (-P_{2s})(QIN_{t-2}) \\ & + (P_{3s})(QOUT_t) + (P_{4s} - P_{3s})(QOUT_{t-1}) + (-P_{4s})(QOUT_{t-2}) \\ & + (P_{5s})(QDIV_t) + (P_{6s} - P_{5s})(QDIV_{t-1}) + (-P_{6s})(QDIV_{t-2}) \\ & + (P_{7s})(QDIS_t) + (P_{8s} - P_{7s})(QDIS_{t-1}) + (-P_{8s})(QDIS_{t-2}) \end{aligned} \tag{25}$$

Based on equation 25, the Δs variable in the water balance model was simulated by time lagging the variables QIN, QOUT, QDIV, and QDIS. The construction of these variables has been discussed above. Equation 25 also shows that the parameters estimated in the regression procedure are combinations of the model parameters, e.g.

$$B_{xs} = P_{2s} - P_{1s} \tag{26}$$

for the model parameter of QIN_{t-1}, where B_{xs} is the estimated regression coefficient for the xth regression variable in stretch s.

The relationship $\Delta s = f(\Delta Q/\Delta t)$ may not be linear as assumed, and this possibility was considered before the decision was made to accept the linear assumption for this study. A model that is nonlinear in the variables but linear in the parameters, e.g.

$$\Delta s = \sum_{i}^{I} P_i (\Delta Q/\Delta t)^a \tag{27}$$

where a represents a vector of exponents of I length, adds considerable effort and complexity to the computation and analysis. For example, the terms .... + $P_{1s}(QIN_t - QIN_{t-1}) + P_{2s}(QIN_{t-1} - QIN_{t-2})$ .... become .... + $(P_{1s})(QIN_t) + (P_{2s} - P_{1s})(QIN_{t-1}) + (P_{2s})(QIN_{t-2})$ after rearrangement, but the terms ... + $P_{1s}(QIN_t - QIN_{t-1})^2 + P_{2s}(QIN_{t-1} - QIN_{t-2})^2$ + ... become .... + $(P_{1s})(QIN_t)^2 - (2P_{1s})(QIN_t)\ (QIN_{t-1}) + (P_{1s} + P_{2s})(QIN_{t-1})^2 - (2P_{2s})(QIN_{t-1})\ (QIN_{t-2}) + (P_{2s})(QIN_{t-2})^2$ + ... on rearrangement. Two additional variables, the cross-product terms, have been added for each pair of Δs model variables, and this single change could add 40 variables to the water balance model. An additional problem is created in the analysis of the dependent variable. In the example above, QIN_t is the dependent variable in the water balance model so terms containing QIN_t must be moved to the left-hand term in the model. As a result, the dependent variable becomes $QIN_t - P_{1s}(QIN_t)^2 + (2P_{1s})(QIN_t)\ (QIN_{t-2})$ which causes considerable additional complexity in the analysis of the regression results. These two considerations are the basis for accepting the linear assumptions for this study. Other versions of the Δs model can be investigated later if the water balance model appears to warrant further research and additional funds can be secured.

Two additional variations in constructing Δs were investigated in this study. Both variations were designed to remove model error caused by unequal stretch lengths and the uneven distribution of discharges and diversions along the length of each stretch. The unequal stretch lengths caused the impact of a change in QIN, QDIV, or QDIS to show in QOUT at differing time intervals so that in the shorter stretches the effect appeared in the same observation and in the longer stretches it often was not felt until the next day. As a result, the effect of the different stretch lengths will be absorbed by the Δs variables. A variable, $QOUT_{t+1}$, was constructed with the QOUT values from the observation following observation t and was added to the water balance model. Similar variables for QDIV were also evaluated. This addition of the t+1 variables adds a $\Delta Q/\Delta t$ term for period m+1, i.e. the period between today and tomorrow, and does not create a parameter estimation problem since the term can be calculated from the data. However, the term does cause a problem when the model is used for prediction purposes if the future flow cannot be predicted. In the low flow conditions, the flow can be assumed to be constant over a multiday period, and the diversion and discharge patterns are also unlikely to change on a daily basis. In other words, the addition of special Δs variables, e.g. the t+1 variables, to improve the accuracy of parameter estimation is reasonable when the model is to be used for low flow predictions because Δs is assumed to be zero in low flow conditions which eliminates the estimation problem associated with the special variables.

The uneven distribution of diversions and discharges along each stretch also causes inconsistent responses in the data. Two sets of variables were evaluated as methods of alleviating this problem. The variables were constructed by regrouping the individual values of QDIV and QDIS on the basis of time of impact at the downstream stretch boundary. For example, assuming the water velocity in the stream is ten miles per day, the diversion or discharge values occurring in the lower ten miles of the stretch for day t were combined with the day t-1 values for diversions or discharges occurring between ten and 20 miles above the lower boundary. Day t-2 values would be used for discharges or diversions occurring between 20 and 30 miles, and the resulting heterogeneous variable would be used with observation t. One set of these variables, QDIV1 and QDIS1, was developed with an assumed velocity of 15 miles per day, and a second set, QDIV2 and QDIS2, was developed with an assumed velocity of ten miles per day. These variable sets were substituted for the Δs variables described above, and the effect on the water balance model was evaluated.

The Regression Model

After the construction of the above variables, the water balance model can be written as:

$$\begin{aligned} QIN_t = {} & P_{0s} + P_{1s}(QOUT_t) + P_{2s}(QDIV_t) + P_{3s}(QDIS_t) + P_{4s}(SROP) \\ & + P_{5s}(TSOST) + P_{6s}(NTSOST) + P_{7s}(ETPH) \\ & + P_{8s}(QIN_t - QIN_{t-1}) + P_{9s}(QIN_{t-1} - QIN_{t-2}) \\ & + P_{10s}(QOUT_{t+1} - QOUT_t) + P_{11s}(QOUT_t - QOUT_{t-1}) \\ & + P_{12s}(QOUT_{t-1} - QOUT)_{t-2} + P_{13s}(QDIV_{t+1} - QDIV_t) \\ & + P_{14s}(QDIV_t - QDIV_{t-1}) + P_{15s}(QDIV_{t-1} - QDIV_{t-2}) \\ & + P_{16s}(QDIS_t - QDIS_{t-1}) + P_{17s}(QDIS_{t-1} - QDIS_{t-2}) \end{aligned} \tag{28}$$

assuming n = 2 for the Δs terms (P_{8s} through P_{17s}). The term P_{1s} is the parameter estimate for each variable and P_{0s} represents any constant effects not included in the independent variables. Moving the dependent variable terms to the left of the equation and rearranging yields the equation:

$$
\begin{aligned}
QIN_t(1 - P_{8s}) = {} & P_0 + (P_{1s} - P_{10s} + P_{11s})(QOUT_t) \\
& + (P_{2s} + P_{13s} + P_{14s})(QDIV_t) + (P_{3s} + P_{16s})(QDIS_t) \\
& + P_{4s}(SROP) + P_{5s}(TSOST) + P_{6s}(NTSOST) + P_{7s}(ETPH) \\
& + (P_{9s} - P_{8s})(QIN_{t-1}) + (-P_{9s})(QIN_{t-2}) + P_{10s}(QOUT_{t+1}) \\
& + (P_{12s} - P_{11s})(QOUT_{t-1}) + (- P_{12s})(QOUT_{t-2}) + P_{13s}(QDIV_{t+1}) \\
& + (P_{15s} - P_{14s})(QDIV_{t-1}) + (- P_{15s})(QOUT_{t-2}) \\
& + (P_{17s} - P_{16s})(QDIS_{t-1}) + (- P_{17s})(QDIS_{t-2}) \qquad (29)
\end{aligned}
$$

When a regression is performed on the constructed dependent and independent variables (assuming n = 2 for the Δs variables), the procedure estimates the B_i parameters in the model:

$$
\begin{aligned}
QIN = {} & B_0 + B_{1s}(QOUT_t) + B_{2s}(QDIV_t) + B_{3s}(QDIS_t) + B_{4s}(SROP) \\
& + B_{5s}(TSOST) + B_{6s}(NTSOST) + B_{7s}(ETPH) + B_{8s}(QIN_{t-1}) \\
& + B_{9s}(QIN_{t-2}) + B_{10s}(QOUT_{t+1}) + B_{11s}(QOUT_{t-1}) + B_{12s}(QOUT_{t-2}) \\
& + B_{13s}(QDIV_{t+1}) + B_{14s}(QDIV_{t-1}) + B_{15s}(QDIV_{t-2})
\end{aligned}
$$

$$+ B_{16s}(QDIS_{t-1}) + B_{17s}(QDIS_{t-2}) + \text{an error term} \tag{30}$$

where

$$B_0 = \frac{P_{0s}}{1 - P_{8s}}$$

$$B_{1s} = \frac{P_{1s} - P_{10s} + P_{11s}}{1 - P_{8s}}$$

$$B_{2s} = \frac{P_{2s} - P_{13s} + P_{14s}}{1 - P_{8s}}$$

$$B_{3s} = \frac{P_{3s} + P_{16s}}{1 - P_{8s}}$$

$$B_{is} = \frac{P_{is}}{1 - P_{8s}} \quad \text{for } i = 4, 5, 6, 7, 10, 13$$

$$B_{is} = \frac{P_{i+1,s} - P_{is}}{1 - P_{8s}} \quad \text{for } i = 8, 11, 14, 16$$

$$B_{is} = \frac{- P_{is}}{1 - P_{8s}} \quad \text{for } i = 9, 12, 15, 17 \tag{31}$$

and the error term is equal to the expected error divided by $1 - P_{8s}$. During the study the best values of n were determined for the QIN, QOUT, QDIV, and QDIS elements of the Δs variable separately, and any n value different from 2 will change the interpretation of the B_i estimates from those given above. The results discussion in the following chapter expands on this impact point when the P_{is} values are calculated from the B estimates using equations similar to the above.

Analytical Procedures

The basic regression methods used in this study have been presented clearly by Draper and Smith [94]. Ordinary multiple linear regression and stepwise regression procedures[1] were both used but the latter procedure was used sparingly because its special features did not benefit the study enough to justify the additional computation expense.

The computer work was accomplished at two different locations on two different machines. The early work which was performed during the summer of 1975 while the author resided in Ames, Iowa, was done on IBM 370/158 and IBM 360/65 computers that are operated as a large single computer by the Iowa State University Computation Center. This early work involved two distinct elements: (1) a data management effort and (2) the initial regression runs. The data were assembled, and the variables constructed with Fortran IV programs written by the author. The basic data bank contained approximately 28000 pieces of data, and the data management programs used 170 K of computer core. The data management element produced a large card deck containing the basic data which served as input to a program called "Models". This program constructed the variables discussed above and then transferred the calculated variables to the regression procedure. The software used at ISU for regression was the Statistical Analysis System (SAS) [95] which was recommended by the Iowa State Statistical Laboratory.

---

[1]In stepwise regression the variables are added to the model in steps (one per step) and entering variables are chosen or rejected by the procedure on the basis of predetermined criteria.

The latter stage of the computational effort was accomplished with a Control Data Corporation (CDC) Model 6400 computer located at Colorado State University (CSU), Fort Collins, Colorado. This machine was accessed from remote terminals located in Denver, Colorado. This change in computer facilities was necessary because the author returned to his permanent residence and employment. This hardware required a few minor modifications in the Fortran software and a change to the Statistical Package for the Social Sciences (SPSS) [96], software that is available at the CSU facility. Most of the regression work in this study was accomplished with the CSU equipment.

Both SAS and SPSS produce numerous regression statistics, and several of these statistics are discussed below because they were used extensively in the study. These key statistics are:

1) R^2 — the square of the multiple correlation coefficient,
2) SD — the standard deviation of the residuals,
3) F — the F statistic,
4)
5) confidence intervals,
6) the van-Neumann ratio,
7) the Durbin-Watson test, and
8) the Z statistic.

The square of the multiple correlation coefficient, R^2, is a measure of the portion of variation about the mean found in the observed data that is explained by the regression model. The equation for R^2 is:

$$R^2 = \frac{\sum^{N} (\hat{Y}_i - \overline{Y})^2}{\sum^{N} (Y_i - \overline{Y})^2} \tag{32}$$

where $\hat{Y}_i$ is the value of the dependent variable predicted by the regression model from the independent variable in observation i, $\overline{Y}$ is the mean observed value of the dependent variable, Y_i is the observed value of the dependent variable for observation i, and the summations include all N observations. This statistic is often used as the basic measure of a regression model's value, and this practice can be misleading. The statistic states the explanation variation as a fraction of the total variation with the unexplained variations in the dependent variable being the $1 - R^2$ portion of the total variation. When the total variation is very large, this unexplained portion can also be a large quantity even when the R^2 value is quite large. A data transformation often reduces the total variation so that the unexplained variation is also reduced even when the R^2 value is reduced. Therefore, if error minimization is a basic criterion in evaluating a model, the maximization of R^2 will not guarantee that this basic criterion is optimized. Furthermore, the addition of variables automatically increases R^2, and as the number of variables approaches the number of observations, R^2 approaches 1. This statistic would then indicate that the model was quite effective when it actually was a trivial solution.

The standard deviation, SD, of the residuals is a better measure of the unexplained error in the regression model. The residuals are calculated as:

$$e_i = \hat{Y}_i - Y_i \ , \tag{33}$$

and the standard deviation is calculated as:

$$SD = \left[\frac{\Sigma e_i^2}{N - k - 1}\right]^{1/2} \tag{34}$$

where k is the number of independent variables. This method of calculating SD assumes the e_i are a sample from a random normally distributed population of residual errors E_i with variance σ^2, but if the residuals contain a nonrandom element or an element not distributed normally, the SD^2 will be larger than σ^2. The minimization of the SD was used as the basic criteria in selecting variables for inclusion in the water balance model, and considerable effort was required to minimize the effects of nonrandomness and nonnormality on this statistic.

The F statistic is used to measure the effectiveness of the regression equation or the individual parameters estimates by the regression procedure. The statistic is calculated with the equation:

$$F = \frac{\dfrac{\sum^{N} (\hat{Y}_i - \bar{Y})^2}{k}}{SD^2} \tag{35}$$

for the overall regression equation. This F ratio can also be calculated with the equation:

$$F = \frac{\dfrac{R^2}{k}}{\dfrac{1 - R^2}{N - k - 1}} \ . \tag{36}$$

In other words, the ratio represents the relationship of explained to unexplained variation with weights assigned for the numbers of variables and observations. The F ratio is a better method of evaluating the significance of the regression model because it accounts for the effect of adding variables on R^2. In equation 36 the F ratio becomes zero when the number of variables exceeds the number of observations by one. The F ratio is also calculated for each B estimate which are assumed to be estimates of the true regression coefficients. The ratio is calculated with the equation:

$$F_i = \frac{SSinc_i}{SD^2} , \tag{37}$$

where $SSinc_i$ is the increase in the sum of squares, $\sum^{N}(\hat{Y}_i - \bar{Y})^2$, explained by the regression equation when variable i is added to the model.

The F statistics are assumed to follow the F distribution and are used in the significance calculations. The F distribution is defined in terms of two parameters, the degrees of freedom. In the regression case, these parameters are k and N - k - 1 for the test of the overall model, and 1 and N - k - 1 for the individual parameter estimates. The regression software used in this study computes the 100 $(1 - \alpha)\%$ confidence level associated with a theoretical F value equal to the calculated F ratio and reports the value of α as "significance". This statistic represents the probability that the calculated F ratio could be exceeded given the degrees of freedom associated with the F ratio. Low significance values for the regression equation represent a high probability that R^2 is greater than 0, and low values for the parameter estimates indicate that the B estimate is greater than 0.

The T statistic is also reported for the individual B estimates and is equal to the square root of the F ratio. The standard error of each B estimate is calculated as:

$$\mathrm{SEofB}_i = \frac{B_i}{T} \,. \tag{38}$$

The standard error is then used to construct a confidence interval using the Student t distribution using the equation:

$$\mathrm{maxmin}\ B_i = B_i \pm (T_{N-k-1,1-\alpha})(\mathrm{SE\ of\ } B_i) \tag{39}$$

where $t_{N-k-1,1-\alpha}$ is the dimension on both sides of the mean that encloses 100(1-α)% of the area under the t cumulative density curve for N-k-1 degrees of freedom. The confidence intervals in this study were calculated at the 95% level ($\alpha = 0.05$).

Both regression software packages output the residuals, e_i, for all observations; and these residuals were analyzed for (1) the presence of autocorrelation, (2) correlation between the residuals and the dependent or independent variable(s) or the observations, (3) nonrandomness in the residual and (4) outliers. Time series data contain a nonrandom element because any single daily observation is often correlated to one or more preceding values. This element is called autocorrelation, serial correlation, or autoregression, and the presence of autocorrelation in the data usually results in autocorrelated residuals which violates the critical randomness assumption underlying the regression procedure. A common autocorrelation model is written as:

$$X_t = \rho_1 X_{t-1} + \rho_2 X_{t-2} \cdots\cdot + \rho_n X_{t-n} + E \tag{40}$$

where $-1 < \rho_i < 1$, and E represents the random error. This Markovian model is of order n, often used with order equal to 1. The autocorrelation is positive if ρ_n is greater than zero and is negative if ρ_n is less than zero. The specific effects of autocorrelation, methods of testing for it, and methods of removing its effects by transforming the basic data are well known today [95]. The use of regression on autocorrelated data causes the variances of the parameter estimates to be larger than the variances that would be obtained from independent data, but the estimates of the variances will be less than the actual variances. As a result, the significance tests and the confidence interval estimates will be invalid. In addition, the model's predictions will also underestimate the potential error. For these reasons the removal of autocorrelation effects is desirable.

The SPSS software calculates two statistics that are used to test for the presence of autocorrelation in the residuals. The von-Neumann Ratio [97] is calculated with the equation:

$$\frac{\delta^2}{S^2} = \frac{\dfrac{\sum_{t=2}^{W} (e_t - e_{t-1})^2}{N - 1}}{\dfrac{\sum_{t=1}^{N} (e_t - \bar{e})^2}{N}} \tag{41}$$

where δ^2/S^2 is the von-Neumann ratio, e_t and e_{t-1} are the residuals calculated for observations t and t-1 respectively, and N is the number of observations. When the residuals are independent, the ratio is normally distributed and the expected value is $\frac{2N}{N - 1}$. For large values of N the expected value is approximately 2.

The Durbin-Watson "d" statistic [98] is the second measure of autocorrelation provided by SPSS. The statistic is calculated with the equation:

$$d = \frac{\sum_{t=2}^{N} (e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2} \qquad (42)$$

where d is the Durbin-Watson statistic.

The two statistics are quite similar, and the interrelationship is defined as:

$$d = \frac{\delta^2}{S^2} \left(\frac{N - 1}{N}\right). \qquad (43)$$

In this study the $\frac{N - 1}{N}$ term is 0.999 so the two statistics should be nearly equal. Both statistics have been used to test the hypothesis that autocorrelation is zero. The von-Neumann ratio is used with the normal distribution to establish a confidence level for acceptance of the hypothesis. Durbin and Watson [98] found that the hypothesis could be accepted if d exceeded an upper limit d_u and could be rejected if d did not exceed a lower limit d_1 but the values between d_1 and d_u were inconclusive. Both statistics show positive autocorrelation when the values are significantly less than 2 and negative autocorrelation when the values are significantly above 2.

Autocorrelation was removed from the residuals of several versions of the model in the following manner. A first order model was assumed to be:

$$e_t = \rho e_{t-1} + E, \tag{44}$$

where the E term represents random error. Substituting

$$e_t = Y_t - X_tB, \text{ and}$$
$$e_{t-1} = Y_{t-1}X_{t-1}B \tag{45}$$

into equation 45 and rearranging yields

$$Y_t - \rho Y_{t-1} = [X_t - \rho X_{t-1}]B + E, \tag{46}$$

where Y is the observed dependent variable values for observations t and t-1, X is the vector of independent variable for the same observations, and B is estimated regression coefficients. This transformed model, called an error model, will contain only a random error term if the correct error model has been assumed. The optimum value of ρ was found by performing several regressions with a range of values for ρ. The SD was then plotted versus the ρ values and the points connected with a smooth curve. The plotted curves resembled parabolas, and the ρ value that minimized the SD was selected. A regression was then performed using the optimum ρ to find the B estimates.

The residuals were also examined to determine if the variance of the residuals was independent of the variables and the observations. This nonconstant variance imperfection in a regression model, called heteroscedasticity, causes the variances associated with the

B estimates to be larger than the variances associated with a constant residual variance when ordinary least squares procedures are used in the regression, and introduces a larger potential error into the prediction applications of the model. These effects are similar to the effects of autocorrelation that is discussed above. Johnston [99] offers two possible remedial actions, (1) use a generalized least squares procedure or (2) transform the data with a model of the relationships involving the residuals and the variables or observations and then use ordinary least squares for the regression. This latter method was used in this study because software for a generalized least squares solution was not available.

A nonconstant variance is often visible in plots of residuals versus a variable or observations whenever a relationship exists. Correlation analysis can also be used when the presence of nonconstant variance is suspected but not obvious. Sometimes the residuals are correlated to several variables and the observations because of correlations between variables and/or observations. In this situation, the construction of a model using one of the highly correlated variables will remove all correlations. Finally, interpretation of the regression results on the transformed data must treat the transformed variables as if no transformation had been performed. For example, the independent variable 1/(the error model) is still treated as the regression constant.

The SPSS software also produces a "Z" statistic that can be used to analyze the residual of nonrandomness. This statistic is based on the analysis of runs procedure. A run is defined as any sequence of residuals of the same sign, and any time the signs of two consecutive residuals are different, a run is terminated and a new run is begun. A run can contain one or more residuals. If the residuals are random, the number of runs in any sequence of residuals approaches an expected mean value, and the distribution of runs will be normally distributed. Draper and Smith [94] have presented this concept quite clearly.

The SPSS software calculates the number of runs u occurring in the residual string, the expected number of runs μ, and the expected standard deviation of the runs distribution. These latter two quantities are calculated with the equations:

$$\mu = \frac{2n_1 n_2}{n_1 + n_2} + 1, \text{ and} \qquad (47)$$

$$\sigma = \left(\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(N_1 + n_2)^2 (n_1 + n_2 - 1)}\right. \qquad (48)$$

where n_1 is the number of positive signs and n_2 is the number of negative signs found in the residuals sequence. The Z statistic is then calculated as:

$$Z = \frac{(u - \mu + 1/2)}{\sigma} \qquad (49)$$

This statistic is the unit normal deviate of the runs and is used with the normal cumulative distribution to test the hypothesis that nonrandomness is zero.

Draper and Smith [94] point out that this procedure is only valid for runs produced by independent observations, and this condition is not true for most of the independent variables used in this study.

Draper and Smith also note that the effect of nonrandom data is small unless the ratio, $\frac{N - k - 1}{N}$, is small. In this study that ratio is approximately 0.95 so this residual correlation effect should be small.

The runs were also analyzed for outliers which are usually defined as residuals larger than 2 SD's. These outliers usually represent aberrations in the data or unique problems in the model structure. Since these problems are often unique to each outlier, no general approach to this type of analysis is possible.

RESULTS

This discussion of the results will include the findings in the areas of independent variable selection, model transformations, and evaluation of the model.

Evaluation of Independent Variables SROP, Soil Stress, and Δs

Three elements in the independent variables were resolved during the study, and these elements include the selection of the optimum SROP, Soil Stress, and Δs variables configurations.

The SROP variable

The relative effectiveness of the SROP configurations is illustrated in Table 12.

Table 12. Comparison of SROP configurations

Variable configuration	Run ID	B estimate	95% confidence interval	R^2 change	Significance, α
SROP	C	2.9060	-49.2388 to 55.0508	0.00000	0.913
SROP1	C	4.1186	-47.7526 to 55.9898	0.00000	0.874
SROP2	C	-11.4439	-63.0241 to 40.1363	0.00001	0.668
SROP3	G	33.6162	-13.1705 to 80.4029	0.00000	0.159
SROP4	F	53.8247	8.7679 to 98.8815	0.00025	0.021
SROP4	M	46.4344	4.0404 to 88.8284	0.00099	0.032

Runs F and G used the identical set of independents with the exception that G used QDIV1 and F used QDIV. A comparison of the B estimates, confidence intervals, and significance values of variables other than

SROP indicated that the runs are essentially identical and provide a valid comparison of SROP3 and SROP4. Runs C and M provide a good comparison between SROP, SROP1, SROP2, and SROP4. The results indicate that the SROP variables contribute very little to the effectiveness of the model with SROP4 contributing a maximum 0.099% additional explanation of the total variance. The variables SROP, SROP1, and SROP2 were never significant at the 95% confidence level, and SROP4 was more significant than SROP3 except for one run which included only QIN values less than 50 cfs. The confidence intervals for SROP, SROP1, SROP2, and SROP3 contain the value zero which is typical of variables with little significance. The SROP4 configuration was chosen for use in later model runs because the variable was the only SROP variable to show any significance; however, this may be a questionable criteria because the sign of the B estimates is the opposite of the expected sign. The overall significance of the SROP variable is discussed in the model evaluation section.

The soil stress variable

The three methods of calculating TSOST and NTSOST were compared in two different runs. Run 1A used the basic variable group QOUT, QDIV, QDIS, SROP4, and ETPH with the QDIV1/QDIS1 version of the Δs variables. Run 1B was identical except the QDIV2/QDIS2 version of the Δs variables was used. The results of these runs are shown in Table 13. The SD of the error terms is minimized in both runs with the difference of means method (equation 20) of calculating TSOST and NTSOST, and the same equations also produced the highest R^2 values.

Table 13. Comparison of TSOST/NTSOST variables

Variable configuration	Run ID	Variable ID	Sig. of B	SD resids.	R^2
Arithmetic difference-	1A	TSOST	0.067	120.8758	0.95695
		NTSOST	0.042		
equations 19	1B	TSOST	0.108	120.3608	0.95761
		NTSOST	0.074		
Difference of means	1A	TSOST	0.001	119.72259	0.95777
		NTSOST	0.070		
equations 20	1B	TSOST	0.003	119.21625	0.95841
		NTSOST	0.055		
Difference of means in	1A	TSOST	0.003	120.7227	0.95706
SD units equations 21		NTSOST	0.449		

Furthermore, the B estimates for the difference of means method are more significant, and the significance level is more consistent between the runs when the difference of means configuration is used. Therefore, equations 20 were used for TSOST and NTSOST.

The effect on the model of using TSOST and NTSOST instead of TRWA, NTRWA, CON1, and CON2 was evaluated with parallel regressions of the variables combined with QOUT, QDIV, QDIS, SROP4, and ETPH. The results of this analysis is shown in Table 14.

Table 14. Comparison of TSOST/NTSOST with TRWA/NTRWA/CON1/CON2

Variables	SD	R^2	F
TRWA/NTRWA/ CON1/CON2	130.8747	0.95161	2327.17
TSOST/NTSOST	131.5351	0.95103	2960.30

The lower SD and the higher R^2 values of the four variable groups might be used as a basis for selecting that group. However, these two statistics are affected favorably by adding variables to the model so the small differences between the variable groups in these two statistics could be the result of the two additional variables in the four variable group. The F ratio for the regressions appears more favorable to the TSOST/NTSOST group. In addition, the B estimates for the variables CON1, CON2, and ETPH were not significant at the 95% level when the four variable group was used, but all variables except the constant were significant when TSOST and NTSOST were used. Theoretically the TSOST/NTSOST variable group is more desirable because consumptive use and irrigation water application effects should not be separable. Since the comparison of these two variable groups do not conclusively favor either group, the TSOST and NTSOST were used in the model because of their theoretical appeal.

The Δs variable

The relative effects of the three configurations in the change in storage variable were evaluated in several ways. Early in the study stepwise regression was used, and this procedure was allowed to select both the QDIV1/QDIS1 and the QDIV2/QDIS2 variable groups. The procedure chose QDIV2 first, and the F ratios for QDIV2 and QDIV1 were 30.09 and 8.88 respectively. QDIS1 was also chosen first, and the QDIS1 and QDIS2 F ratios were 4.38 and 1.93 respectively. Since the QDIV variables were both more significant than the QDIS variables, the QDIV2/QDIS2 group was used initially for the model development. At that point in the study the development of the Δs variables was trying to apply the unit hydrograph routing technique of Sauer [93], and when the theoretical basis discussed in the preceding chapter finally evolved, the time lagged QDIV and QDIS variables were arbitrarily chosen for the continuing model development. Late in the study this arbitrary decision and the questionable basis of the choice between QDIV1/QDIS1 and QDIV2/QDIS2 were reexamined with three parallel regressions, each using a set of Δs variables plus the basic group of QOUT, QDIV, QDIS, SROP4, TSOST, NTSOST, and ETPH. Lagged QIN variables were also used in each regression. The results of this analysis are shown in Table 15. An n value of two was used for the lagging of the QIN and QDIV related variables, and an n value of 3 was used for the QDIS related variables.

Table 15. Comparison of Δs variables

Variable configuration	SD	R^2	F	Number of cases
QDIV/QDIS	124.4962	0.95483	1595.87	1072
QDIV1/QDIS1	120.6790	0.95613	1588.02	1035
QDIV2/QDIS2	123.3745	0.95387	1476.94	1015

The three regressions do not represent identical conditions because the number of cases used in each run are different. The time lagging of variables typically removes cases from the set of observations, and the available observations for this analysis was expected to be 1035. The time lagging procedures available in SPSS are a little complex, and the undesirable inclusion or exclusion of cases seen in the Table 15 results occurred several times during the study. Since this analysis of the Δs variables was performed after the study funds were exhausted, a rerun of the regressions was not performed because the expected benefits did not appear to justify the cost. The results in Table 15 favor the QDIV1/QDIS1 configuration in minimizing SD and maximizing R^2. The results also favor the QDIV/QDIS configuration if maximizing the significance of the regression is the criterion. The QDIV/QDIS regression included 37 additional cases, and these surplus inclusions all occurred during the early days of the study data period. These observations included the highest observed values of QIN, QOUT, and QDIV and are generally associated with the larger residuals. This element would increase SD and decrease R^2 so the apparent differences between the QDIV/QDIS and QDIV1/QDIS1 variable group is not as large as the results indicate. However, a similar conclusion is not justified in comparing QDIV1/QDIS/ and ODIV2/QDIS2. The twenty cases missing from the QDIV2/QDIS2 regression would probably increase the SD and decrease the R^2 statistics if they were included because they also are associated with the larger residual segment of the data. So the differences in the three key statistics can be used as the basis for preferring the QDIV1/QDIS1 variables over the QDIV2/QDIS2 variables. As a result of this analysis, any results discussed below that use the QDIV2/QDIS2 version of the Δs variables can probably be improved by substituting the QDIV/QDIS or the QDIV1/QDIS1 configurations, and results using the QDIV/QDIS version may be improved if the QDIV1/QDIS1 version is substituted. This element of the water balance model is a good subject for further study.

Structural and Variable Transformations

The transformations of the model structure were performed to equalize the variance, eliminate autocorrelation, and eliminate the error caused by hydraulic differences between the stretches.

Variance equalization

Visual analysis of the plots of residuals versus observations showed a considerable change in residual variance between the observations early in the study data period and those occurring late in the period. This change in the variance occurred in all five stretches. An analysis of the data indicated that the variance change also corresponded to the changing magnitude of the variables QIN, QOUT, and QDIV.

Two models were used to equalize the variance in the manner discussed in the preceding chapters and both models used the relationships between the residuals and the flow variables QIN and QOUT. The simplest model involved multiplying all variables by the quantity 1/QOUT, and this model was used during the middle of the study to equalize the variance in a model containing the QDIV2/QDIS2 version of the Δs variable plus two experimental variables $QOUT^2$ and $QOUT \cdot QDIV$. An evaluation of the effectiveness of this equalization model is not possible because no comparison regressions were performed. The experimental variables were dropped from the model because they were incompatible with the underlying theory and another variance equalization model was later selected for the developed model. The 1/QOUT model reduced the total sum of squares, the R^2, and the F ratio considerably as expected and reduced the Z statistic from about -18 to about -12, which still represents a number of runs about 12 standard deviations less than expected from a random normal population. The transformed model produced a SD of 1.47833.

The second variance equalization model was chosen because the variability of QIN is expressed in terms that are related directly to QIN. The variance equalization model is

$$\frac{1}{\text{EHAT}} = \left(\frac{1}{a + b \ * \ (\text{QINHAT})^{c}}\right)^{1/2} \tag{50}$$

where EHAT is a predicted residual, QINHAT is the dependent variable predicted for each observation by the regression model prior to variance equalization, and a, b, and c are constants. The values of a, b, and c were estimated with a regression procedure using the residuals and the predicted QIN from the unequalized model for the dependent and independent variables. The regression equation is:

$$e^{2} = B_{0} + B_{1} * (\text{QINHAT})^{c}, \tag{51}$$

where B_0 and B_1 are regression estimates of a and b, and e^2 is the square of the observed residuals. The value of c was found by making repeated regression runs at different values of c and then selecting the c value producing the B_0 estimate closest to zero so that the expected variability about an estimated QIN of zero will be minimized. This minimization criterion required the acceptance of a SD 0.07% larger than the minimum obtainable and reduced the R^2 value 1% below optimum. Neither of these compromises are considered as sufficient reason to abandon the minimization criterion. This procedure produced the following model:

$$e^2 = 4.21 + 5.47 \times (\text{QINHAT})^{1.165} \tag{52}$$

This regression produced the following statistics.

$SD = 31,225.0549$

$R^2 = 0.11100$

$F = 129.23$

$\text{Sig.} = 0.000$

$\frac{\delta^2}{S^2} = 1.51156$

$d = 1.51010$

$Z = -11.43321$

The F ratio and the significance level indicate that this regression equation does explain a portion of the residual variance, but the SD and the R^2 values show a large portion of the variation in the residuals is not explained. The Durbin-Watson and von-Neumann statistics indicate positive autocorrelation, and the Z statistic indicates the residuals of this equation are not normally distributed.

The impact of variance equalization on the water balance model was slight. The Durbin-Watson and von-Neumann statistics moved slightly closer to the value of 2 but both statistics were already very close to that optimum value. The Z statistic moved 0.3 of a standard deviation in the wrong direction. Visual examination of the residual plots revealed some equalization in all stretches, but the effect was obvious only in stretches 1 and 5. The effect also appeared more pronounced in the 1973 data blocks. These observations suggest that variance equalization models be developed for each stretch and data time block. But time and money limitations precluded the pursuit of this possibility in this study, and this element of the variance equalization transformation will be a good subject for future study.

The inclusion of a variance equalization step in developing a water balance model for low flow estimation is desirable because it permits the estimation of errors that are specific for the low flow range. This concept is illustrated in Figure 15 where the 95% confidence limits have been plotted for the variance equalized model. The curves in Figure 15 were constructed by assuming that QINHAT is equal to the QIN predicted by the model, and the model residuals are assumed to be normally distributed for the calculation of the confidence interval. The limits were calculated with the equation:

$$QIN_l = QIN_p \pm (t_{95\%})[4.21 + 5.47(QINHAT)^{1.165}]^{0.5} \ (SE_e), \tag{53}$$

where QIN_l is an upper or lower confidence limit for QIN, QIN_p is the value predicted by the variance equalized model, t is the number of standard deviations either side of the central value that includes 95% of the area under the cumulative normal distribution curve, QINHAT is assumed equal to QIN_p, and SD_e is the standard deviation of the residuals in the variance equalized model. This latter value was

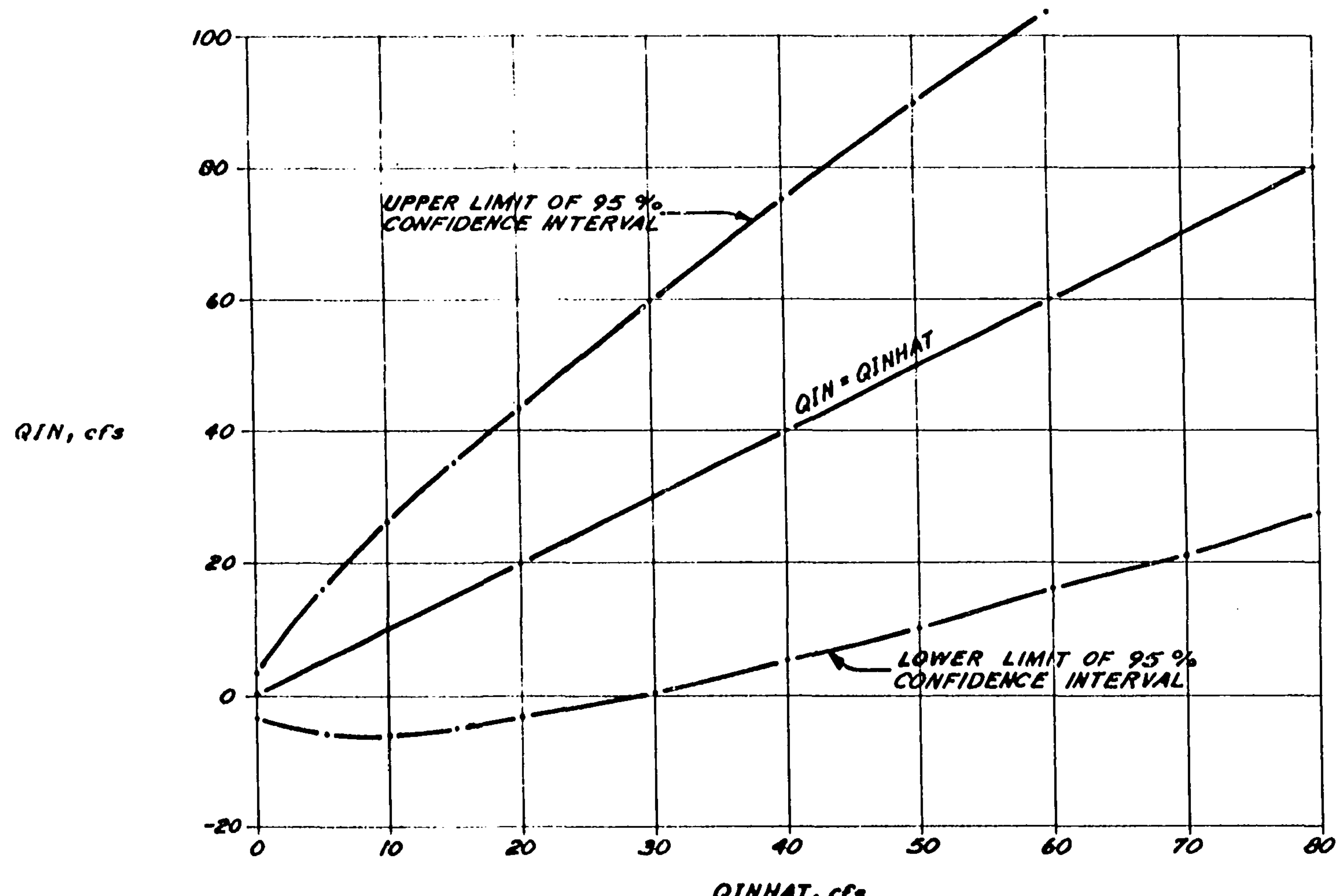


Figure 15. 95% confidence limits in low flow range for the variance equalized regression model

found to be 1.05743, and t equals 1.65. Without the variance equalization the relationship between error and the size of QIN could not be estimated, and the 95% confidence interval would be $\pm$ 149.5 cfs at all levels of QIN. Such a large error is intuitively known to be grossly overstated for the low flows as Figure 15 illustrates.

The lower confidence limit in Figure 15 can also be used to illustrate a reason for preferring this second variance equalization model to the 1/QOUT model. The lower limit equals zero at a QINHAT of approximately 30 cfs. Assuming that the QIN/QOUT ratio equals the $\overline{\text{QIN}}/\overline{\text{QOUT}}$ ratio (= 1.45), the OOUT associated with a QIN of 30 cfs would be 20.7 cfs. Using the SD from the 1/QOUT results discussed above, the 95% confidence interval for QIN would be $\pm$ 50.5 cfs which is considerably larger than the $\pm$ 30 cfs predicted using the QINHAT model. This comparison inspires more confidence in the latter model even though the underlying assumptions of normality are unfulfilled and do not support a decisive evaluation.

Autocorrelation elimination

Several versions of the model exhibited strong positive autocorrelation during the study, and the autocorrelation was removed with a first order model as discussed in the preceding chapter. A ρ of 0.57 was found to be optimum for a regression model using the QDIV2/QDIS2 Δs variables and containing no disaggregation of the hydraulic variables (see next discussion section on disaggregation). The procedure increased the Durbin-Watson statistic from 1.39854 to 2.10389 and increased the von-Neumann ratio from 1.39991 to 2.10595.

These statistics indicate that the procedure was effective in removing autocorrelation. The effectiveness of this procedure is also indicated by the increase in Z from - 18.31651 to - 9.98187, although the latter value still indicates a strong nonnormal element in the residuals.

The final model, however, did not show any significant autocorrelation in the two key statistics, $\frac{\delta^2}{S^2}$ and d, as reported in the variance equalization discussion. Nevertheless, an autocorrelation elimination procedure was performed using a first order error model, and a maximum SD reduction of 0.06% was possible with a ρ of - 0.07. These results confirm the d and $\frac{\delta^2}{S^2}$ statistics. Since the final model is the only version that did not contain considerable autocorrelation and the same version is the only model with some of the hydraulic variables disaggregated by stretch, a functional relationship between autocorrelation and combining hydraulic effects may exist. However, an examination of the residual plots reveals several residual sequences that are quite obviously autocorrelated and leads to the hypothesis that the lack of autocorrelation in the overall model may result from a coincidental combination of offsetting autocorrelation effects. These observations suggest another subject for future study.

Disaggregation of hydraulic variables

The purpose of combining all five stretches in each variable was to increase the strength of the statistical analysis by increasing the number of observations and to force similar effects to yield a single parameter estimate. Regressions were also performed on the data for each stretch, and a comparison of the B estimates suggested that some of the hydraulic variable parameters varied considerably among the stretches. As a result, a procedure was developed to provide separate parameter estimates for the hydraulic variables that can theoretically vary significantly among the stretches.

The procedure consisted of two steps. First, a regression was performed on the complete model with the variables similating the discrete inflows and outflows disaggregated by stretches. Then the stretches with similar parameter estimates were combined within each variable. Several variables were aggregated back into a single variable, but three key variables were each divided into three stretch groups. Stretches were combined whenever the 95% confidence intervals for the B estimates overlapped although a few small overlaps were ignored. A subjective decision was made in a few cases where two estimates overlapped a third estimate but not each other. The CSU computer could not perform the step one regression because the model was too large so the 2 and 3 day lagged variables were eliminated for that run and were then grouped in the same manner as the 1 day lagged variables. The results of this regression and the final stretch groupings are shown in Table 16. Some variables do not appear in Table 16 for a few stretches because the variable observations were all zero in that stretch. In addition, eight of the disaggregated variables were eliminated from the model because the significance of their B estimates exceeded 0.500. Three disaggregated variables with B estimates significances above 0.500 were inadvertently retained, and one of these variables became significant at the 99% confidence level in the final model. As a result the decision to eliminate the

Table 16. Hydraulic variable stretch groupings

Variable	Str. No.	B estimate	Sig.	95% conf. int. Lower limit	95% conf. int. Upper limit	Grouping by str. No.
$QOUT_t$	1	0.4965	0.000	0.4322	0.5608	
	2	1.0719	0.000	0.9749	1.1690	1-3
	3	0.3512	0.000	0.1843	0.5182	2-4
	4	0.8408	0.000	0.5260	1.1557	5
	5	-0.0665	0.585	-0.3051	0.1722	
$QOUT_{t+1}$	1	0.2279	0.000	0.1799	0.2758	
	2	0.1410	0.000	0.0732	0.2088	1-2-5
	3	0.6345	0.000	0.5019	0.7671	3-4
	4	0.4752	0.000	0.2395	0.7108	
	5	0.2812	0.003	0.0964	0.4660	
$QOUT_{t-1}$	1	-0.2923	0.000	-0.3800	-0.2045	
	2	-0.2687	0.000	-0.3941	-0.1434	1-2-3-4
	3	-0.3840	0.000	-0.5273	-0.2407	
	4	-0.4479	0.001	-0.7001	-0.1956	
	5	-0.0028	0.977	-0.1888	0.1833	
$QDIV_t$	1	0.6197	0.001	0.2479	0.9915	
	2	-0.0333	0.872	-0.4383	0.3716	1-2-3
	3	0.4312	0.000	0.3122	0.5503	
	4	-0.1523	0.806	-1.3689	1.0643	
$QDIV_{t+1}$	1	0.3618	0.002	0.1356	0.5880	
	2	0.3749	0.027	0.0427	0.7071	1-2-3-4
	3	0.2899	0.000	0.1971	0.3827	
	4	0.8225	0.102	-0.1634	1.8083	
$QDIV_{t-1}$	1	-0.2992	0.017	-0.5444	-0.0540	
	2	-0.3580	0.025	-0.6702	-0.0458	1-2-3
	3	-0.4025	0.000	-0.5083	-0.2968	
	4	-0.2768	0.570	-1.2334	0.6798	

Table 16. Continued

Variable	Str. No.	B estimate	Sig.	95% conf. int. Lower limit	Upper limit	Grouping by str. No.
$QDIS_t$	2	-1.9149	0.060	-3.9126	0.0827	
	3	-1.2418	0.000	-1.5639	-0.9197	2-3
	4	-12.3389	0.945	-455.3033	424.6256	
	5	-0.0224	0.690	-0.1324	0.0876	
$QDIS_{t-1}$	2	-0.5103	0.594	-2.3880	1.3673	
	3	0.7529	0.000	0.4316	1.0741	2-3
	4	5.0399	0.982	-433.2653	443.3451	5
	5	-0.6493	0.197	-0.1746	0.0360	
QIN_{t-1}	1	0.4114	0.000	0.2757	0.5471	
	2	0.1405	0.007	0.0377	0.2434	1
	3	0.5207	0.000	0.4286	0.6128	2-4
	4	0.1152	0.230	-0.0731	0.3034	3-5
	5	0.6730	0.000	0.4887	0.8573	
SROP4	All	2.5908	0.880	—	—	All
TSOST	All	-16.4978	0.012	—	—	All
NTSOST	All	-21.3925	0.001	—	—	All
ETPH	All	1.8166	0.000	—	—	All
Constant	All	2.0727	0.779	—	—	All

variables with insignificant B estimates now appears to be a mistake.

As a result of this disaggregation, the final model contained 27 independent variables and a constant. The effect on the model of regrouping the hydraulic variables was quite obvious. Key statistics from three regressions are presented in Table 17. The first regression included QOUT, QDIV, QDIS, SROP4, TSOST, NTSOST, ETPH, the t+1 variables, and the QDIV/QDIS version of the Δs variables. The second regression contained the completely disaggregated variables as shown in Table 15, and the third regression used the regrouped variables shown in Table 16 with the 2-day lagged variables included. A 3-day lagged QDIS variable was also included for stretch 5. The decline in the SD from regression 1 to regression 2 is evidence of a significant improvement in the model, and the F ratio remained very large even though it was reduced considerably. Furthermore, the regrouping of the variables recovered most of the F ratio decline with a very small increase in the SD. The fluctuations in the F ratio are strongly influenced by the large increases and decreases in the number of independent variables caused by the disaggregation and regrouping. The effects on the Z statistic is also quite large, and the improvement in the Z statistic attributed to the regrouping after the disaggregation was not expected.

Table 17. Statistical effects of regrouping the hydraulic variables

Statistic	Regression 1 No disaggregation	Regression 2 Complete disaggregation	Regression 3 Regrouped variables
R^2	0.95483	0.97689	0.97583
SD	124.496	89.831	90.584
F	1595.86	960.91	1568.26
$\frac{\delta^2}{S^2}$	1.49558	2.09955	2.07258
d	1.49419	2.09754	2.07058
Z	-18.15434	-10.24438	-8.38209

Evaluation of the Models

The regression model

The regression with the equalized variance but with no autocorrelation removal was selected as the final model for discussion instead of the regression with the autocorrelation removed because this final procedure provided essentially no improvement in the model but introduced additional complexity. The B estimates and their significance levels for this model are presented in Table 18.

The B estimates for the QOUT, QDIV, and lagged QIN variables all show high significance levels with the exception of most of the 2-day lagged variables; however, the B estimates for most of the QDIS variables appear to be zero. This lack of effectiveness for QDIS can be traced to several causes. The variable has no effect in stretch 1 because

Table 18. B estimates for final regression model

Variable	Stretch 1		Stretch 2		Stretch 3		Stretch 4		Stretch 5	
	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.
Const.	2.9184	0.013	2.9184	0.013	2.9184	0.013	2.9184	0.013	2.9184	0.013
$QOUT_{t+1}$	0.3547	0.000	0.3547	0.000	0.4347	0.000	0.4347	0.000	0.3547	0.000
$QOUT_{t}$	0.3043	0.000	0.7071	0.000	0.3043	0.000	0.7071	0.000	-0.1035	0.010
$QOUT_{t-1}$	-0.4266	0.000	-0.4266	0.000	-0.4266	0.000	-0.4266	0.000	0	–
$QDIV_{t+1}$	0.4543	0.000	0.4543	0.000	0.4543	0.000	0.4543	0.000	0	–
$QDIV_{t}$	0.2138	0.012	0.2138	0.012	0.2138	0.012	0	–	0	–
$QDIV_{t-1}$	-0.3961	0.000	-0.3961	0.000	-0.3961	0.000	0	–	0	–
$QDIV_{t-2}$	0.0243	0.684	0.0243	0.684	0.0243	0.684	0	–	0	–
$QDIS_{t}$	0	–	-0.3350	0.021	-0.3350	0.021	0	–	0	–
$QDIS_{t-1}$	0	–	0.0746	0.768	0.0746	0.768	0	–	-0.0287	0.448
$QDIS_{t-2}$	0	–	0.2068	0.280	0.2068	0.280	0	–	-0.0418	0.272
$QDIS_{t-3}$	0	–	0	–	0	–	0	–	-0.0209	0.516
QIN_{t-1}	0.6740	0.000	0.2526	0.000	0.6335	0.000	0.2526	0.000	0.6335	0.000
QIN_{t-2}	0.0489	0.415	0.0717	0.011	0.0447	0.275	0.0717	0.011	0.0447	0.275
SROP4	2.4732	0.656	2.4732	0.656	2.4732	0.656	2.4732	0.656	2.4732	0.656

Table 18. Continued

Variable	Stretch 1		Stretch 2		Stretch 3		Stretch 4		Stretch 5	
	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.	B Est.	Sig.
TSOST	-4.8048	0.010	-4.8048	0.010	-4.8048	0.010	-4.8048	0.010	-4.8048	0.010
NTSOST	-1.6377	0.610	-1.6377	0.610	-1.6377	0.610	-1.6377	0.610	-1.6377	0.610
ETPH	0.0747	0.446	0.0747	0.446	0.0747	0.446	0.0747	0.446	0.0747	0.446

it is always zero, and its values in stretch 4 show practically no variation between observations because the variable consists mostly of small discharges that were estimated from very little data. The largest variations and observations of QDIS occur in stretch 5 where the Purgatoire River joins the Arkansas River, and the lack of effectiveness for QDIS in this stretch was unexpected. An examination of the data found several inconsistencies in the passage of large surges through the stretch. The inconsistencies involve the time delay between the surge inflows and outflows and variations in the shape and size of the outflow hydrograph. Since these flows all pass through John Martin Reservoir, the probable cause of the problem is the deviation in the operation of the reservoir from the operating principle of passing all flows through the reservoir without delay. Even though the reservoir contents were reported as zero throughout the study data period, this analysis indicates that some short-term storage was probably occurring.

The B estimate for SROP4 was insignificant in many versions of the model so the ineffectiveness of this variable in this final model is expected. Evidently the precipitation that occurred during the study period became part of the soil moisture bank or percolated into the groundwater. Since the recorded precipitation contains few large events, this conclusion seems reasonable for this study area and period, but an SROP4 variable must be included in any water budget model during parameter estimation because this variable is always a potential model element. Further evaluation of the SROP4 variable can be accomplished in future research using a different study area or period that contains a more active surface runoff element.

The B estimates for TSOST were insignificant in most versions of the models, and this variable increases the explained variation by about 0.2% of the total variation. The impact of this single variable on the SD was not determined. The effectiveness of the variable appears to be related to an outflow element because the sign of the estimate is negative. This outflow element is either a good simulation of the SAWE model variable or the diversion data element of the TSOST variable is acting as an additional version of QDIV. The variable NTSOST contains diversion data in the same manner as TSOST, but the B estimates for this variable are not significant. Therefore, the reason for the significance of TSOST is probably its role as a simulator of SAWE.

The B estimate for NTSOST and ETPH were usually significant during the study and were usually insignificant for the regression constant. However, these conditions were reversed after the application of the variance equalization procedures. Both of the variables simulate slow acting distributed inflow and SAWE effects. The exchange of NTSOST and ETPH for the constant may indicate that these effects change so slowly during the short study period that a constant is a more effective simulator. If this hypothesis is confirmed in later studies, the effort required to develop a water balance model will be greatly reduced by eliminating the need to estimate phreatophyte and nontributary area consumptive use.

The significance of the B estimates was used to select the appropriate n value for the Δs variables. A value of was selected for QDIV and QIN. A value of 0 is indicated for QDIS which is not surprising because the QDIS observations are usually one or two magnitudes less than the other discrete flow variables. Johnston [99] suggests other methods of selecting n which could be investigated later.

The incremental improvements and changes in the model attributable to the major elements of the water balance were measured with a sequence of regressions, and the results are shown in Table 19.

Table 19. Relative model impacts of major variable groups

Regression No.	Independent variables	SD	R^2	d	Z
1	QOUT, QDIV, QDIS	136.43	0.94712	1.08134	-19.81517
2	Above plus distributed inflows and SAWE	131.54	0.95103	1.11597	-20.01930
3	Above plus Δs variables	120.68	0.95613	1.43731	-18.74832
4	Above with t+1 and regrouped hydraulic variables	90.58	0.97583	2.07058	-8.38209

Since the Δs and regrouped hydraulic variables have provided the most improvement in the model, these two variable groups probably are the best areas to investigate for further improvement in the model. Several subjects for future study in those two areas are suggested in this dissertation.

Prior to variance equalization, the model produced 53 residuals greater than 2 standard deviations, called outliers, representing 5.11% of the residuals. These values were reduced to 43 and 4.15% respectively by the variance equalization procedure. No authority could be found that provided norms for assessing the significance of these statistics, but subjectively they appear high. The data in the observations associated with each outlier were examined for clues to any effects not properly simulated in the model. Many pairs of outliers occurred in two consecutive observations which involved a surge in an inflow or outflow. The model appeared to react too slowly producing first a large positive residual and then a large negative residual. This problem is most likely related to the linear assumptions underlying the Δs variable. As discussed earlier, a polynomial or nonlinear model for this variable group creates more complex computation problems and has been left for future study. The examination of outliers was responsible for the inclusion of the t+1 variables, but the effect of this new variable was masked by the simultaneous inclusion of the regrouped hydraulic variables. Since the B estimates for these variables are significant and several of the outliers that indicated a need for the t+1 variables have been reduced this addition to the model was retained.

The SD of 90.58 in the most effective regression (#4) is about 19.1% of the mean value for QIN. This latter value is called the coefficient of variability, and usually falls in the 20-25% range for hydrologic studies. This comparison coupled with the large R^2 and the apparent lack of autocorrelation associated with the final regression model encourages the conditional acceptance of the water balance model as a part of a low flow estimation technique for arid and semiarid areas. Unconditional acceptance will require the elimination of the apparent nonnormal element in the residuals and a more complete evaluation of the potential error in the model.

The water balance model

The water balance model parameters were calculated from the B estimates in Table 17. All estimates with a significance below the 90% confidence level were assumed to equal zero, and the regression model then became:

$$\begin{aligned} QIN = {} & B_0 + B_{1s}(QOUT_t) + B_{2s}(QDIV_t) + B_{3s}(QDIS_t) \\ & + B_{4s}(TSOST) + B_{5s}(QIN_{t-1}) + B_{6s}(QIN_{t-2}) \\ & + B_{7s}(QOUT_{t+1}) + B_{8s}(QOUT_{t-1}) + B_{9s}(QDIV_{t+1}) \\ & + B_{10s}(QDIS_{t-1}). \end{aligned} \tag{54}$$

Equations in terms of P_{is} and B_{is} were developed in the same manner as shown in the preceding chapter and were rearranged to produce the following set of equations.

$$P_{5s} = \frac{B_{5s} + B_{6s}}{B_{5s} + B_{6s} - 1} \tag{55a}$$

$$P_{is} = B_{is}(1 - P_{5s}) \text{ for } i = 0, 3, 4, 7, 9 \tag{55b}$$

$$P_{is} = (B_{1s} + B_{7s} + B_{8s})(1 - P_{5s}) \tag{55c}$$

$$P_{2s} = (B_{2s} + B_{9s} + B_{10s})(1 - P_{5s}) \tag{55d}$$

$$P_{is} = B_{is}(P_{5s} - 1) \text{ for } i = 6, 8, 10 \tag{55e}$$

These equations were used to calculate the parameter estimates shown in Table 20.

The sign of the constant parameter, P_0, indicates that the variable is simulating an outflow from the stream. A constant understating of the diversion flows could be the source of this outflow element, but the loss of streamflow to the groundwater aquifer as discussed above is the most probable source of this element. If further research confirms that aquifer demand is the true source of this element, this water balance model technique will become a new approach to the estimation of regional consumptive use.

The sign of the TSOST variable indicates that this variable is simulating a distributed inflow which is either ungaged tributary inflow to the study area, surface runoff from irrigation tail water, and/or waste irrigation ditch water. Considering the arid conditions during the study data period and the associated high demand for irrigation water, the surface runoff is the most probable source of this element.

The signs of the discrete flow variables, QOUT, QDIV, and QDIS, are correct with the exception of QOUT in stretch 5. The problems caused by the John Martin Reservoir that were discussed above are the probable cause of this sign problem. The B estimate for this variable was insignificant when the hydraulic variables were completely disaggregated but became significant after recombining the variables and equalizing the variance. This behavior was not seen in the other

Table 20. Parameter estimates for water balance model

Independent variable	Parameter	Parameter estimates Stretch 1	Stretch 2	Stretch 3	Stretch 4	Stretch 5	Units
Constant	P_0	8.952	4.319	7.963	4.319	7.963	cfs
$QOUT_t$	P_1	0.713	0.940	0.852	1.058	-0.479	None
$QDIV_t$	P_2	0.834	0.403	0.742	0.086	0	None
$QDIS_t$	P_3	0	-0.496	-0.914	0	0	None
TSOST	P_4	-14.739	-7.111	-13.110	-7.111	-13.110	cfs
QIN_t-QIN_{t-1}	P_5	-2.067	-0.480	-1.729	-0.480	-1.729	None
$QIN_{t-1}-QIN_{t-2}$	P_6	0	-0.106	0	-0.106	0	None
$QOUT_{t+1}-QOUT_t$	P_7	1.088	0.525	1.186	0.643	0.968	None
$QOUT_t-QOUT_{t-1}$	P_8	1.309	0.631	1.164	0.631	1.164	None
$QDIV_{t+1}-QDIV_t$	P_9	1.394	0.672	1.240	0.672	0	None
$QDIV_t-QDIV_{t-1}$	P_{10}	1.215	0.586	1.081	0	0	None

discrete flow variables and is considered another symptom of a discontinuity in the functioning of the stretch.

The signs of the Δs variables are consistent, i.e. variables that share the same sign were expected to have like signs; but the signs are opposite from the expected signs. The accuracy of the parameter estimate calculations and a review of the assumptions underlying the Δs variables confirmed that the signs are wrong. If the Δs variables are rearranged as follows $(n = 1)$

$$X = (QIN_t - QOUT_t - QDIV_t + QDIS_t)$$

$$- (QIN_{t-1} - QOUT_{t-1} - QDIV_{t-1} + QDIS_{t-1}) \qquad (56)$$

The sign reversal can be explained. The signs of $\Delta QOUT$ and $\Delta QDIV$ were reversed in the rearrangement because the effects of these two variables are opposite to the effects of ΔQIN and $\Delta QDIS$. Rearranging the basic hydrologic water budget model leaves:

$$\Delta s = \Sigma \text{ inflow} - \Sigma \text{ outflow}. \qquad (57)$$

Each of the two quantities in the parentheses in equation 52 are good approximations of the right side of equation 53 so equation 52 really estimates the quantity $\Delta s_t - \Delta s_{t-1}$. Since the data used for the Q variables are in terms of cubic feet per second the Δs quantity becomes $\Delta s/\Delta t$; and, furthermore, the difference $(\Delta s/\Delta t)_t - (\Delta s/\Delta t)_{t-1}$ is actually

$$\frac{\frac{\Delta s_t}{\Delta t} - \frac{\Delta s_{t-1}}{\Delta t}}{\Delta t}$$

since the t - (t-1) increment is also Δt. This quantity is an integer approximation of the second derivative of storage with respect to time. In other words the Δs variable estimator is the rate of change in the daily rate of storage, and this quantity can be either positive or negative. The sign reversal means that the rate of change is the rate of storage is negative instead of positive as was implicitly assumed in the model structure. This result is not surprising because the rate of storage is related to the magnitude of the streamflow, and the streamflow magnitude typically declines during the study data period. This study finds this trend component to be a significant part of the water balance model so this element of the Δs variable should be included in the development of new models, but future research should investigate the value of an additional variable to simulate the magnitude of the change in storage.

SUMMARY AND CONCLUSIONS

Summary

This study began with the investigation of potential problems in the estimation of low flows for water quality management in arid and semiarid river basins. The special conditions created in these basins by the application of water law based on the Prior Appropriation Doctrine introduce a new role for low flow estimation procedures in that purchase of water rights and relocation of diversions could become water quality management alternatives. This economic implication produces a sizable research subject. The same special conditions also suggest a new low flow estimation technique involving (1) the establishment of the active diversions in a regulated stream segment by frequency analysis of call dates or a gaged flow that can be correlated with the call date history and (2) the estimation of the desired low flow with a water balance model using the diversions identified in step 1 (see pages 26-28) as outflows. The investigation of step 1 is left for future study, and this study concentrated on developing a method for constructing the water balance model.

The ultimate development and acceptance of the estimation process presented in this dissertation will introduce the following activities into the planning and management functions in the physical and economic areas of water quality. Application of the step 1 process will produce a call date for a regulated area. The call date will define a set of diversions for a regional flow regime that is comparable to the 7-day, 10-year low flow regime in unregulated basins. Simultaneous application of step 2 in the estimation process will define a water balance model for the area. Some adjustments to the significant distributed inflow and outflow variables are required to account for special low flow conditions, e.g. temperatures for TSOST would have to be selected when the time of occurrence of the low flow is unknown. The desired critical low flow (QIN) is calculated using the appropriate diversions and associated discharges derived as a result of the step 1 analysis, using equation 54. This low flow value would then be used in a water quality model to determine required treatment levels, and a treatment cost model would provide a measure of the cost to achieve the desired water quality benefits without modifying the critical low flow (CLF). A planner or manager can then use the tools developed in the economic study to determine if social benefit/cost considerations make the modification of the critical low flow a more desirable alternative. For example, a model of the water rights market developed in the economic (and legal-institutional) study would yield cost estimates or water quality tradeoffs required to alter the existing water allocation system for various levels of critical low flow modification. This information could then be fed into a cost optimization model that would tradeoff treatment cost with water right modification costs, and/or the results of the water rights market analysis would serve as input to a regional water allocation model for use in a normative planning analysis.

The following conclusions have been drawn from the results of this study.

Conclusions

1. This study has found a need for a new low flow estimation technique for arid and semiarid basins because (1) the basic assumption of the common frequency analysis estimation method are seriously violated in the dry arid/semiarid basins and (2) the potentially rapid changes in the low flow process found in these dry basins are not reflected quickly enough in the frequency method, especially if the available data string is long.

2. The study has shown a potentially strong tie among low flow estimation for water quality management and the arid-semiarid basin economies if a responsive estimation method is developed and accepted. Interaction mechanisms are defined and a two phase research program is developed. The first phase will develop predictive tools for a second normative phase. Cost models are defined for the predictive study area, and an allocation model is defined for the normative phase.

3. The study results justify a conditional acceptance of the basic hypothesis that a hydrologic budget model is a suitable water balance model for step 2 of the proposed estimation method. Considerable improvement in the model's accuracy was demonstrated during the study, and additional improvement areas have been found during the final evaluations of the study. These potential areas are listed below with the conclusion on additional research.

4. The selection of multiple regression as the analytical technique provided several useful model development tools and should be used in future water balance models for this basin analysis.

5. The results indicate that slowly changing hydrologic variables, e.g. the consumptive use of phreatophytes and crops in nontributary areas can be effectively modeled by a constant parameter when the model time period contains four or less months. This conclusion may also extend to periods greater than four months because this study did not define the time period limits. Further investigation of this conclusion is warranted because this hydrologic budget approach might lead to a new method of estimating regional consumptive use of crops and phreatophytes.

6. This study demonstrated that for this sample study area and period the surface runoff from precipitation was not significant in the water balance model. This conclusion, however, cannot be applied universally to the water budget model approach. In this study, the only configuration of the surface runoff variable to show any functional relationship was a 2-day lagged compounded variable that simulated a runoff hydrograph. This observation may be useful in guiding future water budget model attempts.

7. The value of removing heteroscedasticity and autocorrelation was demonstrated during the model development even though the final model did not benefit substantially from either of these analytical techniques. Since these two problems are typical of many hydrologic time series data strings the value of these tools should be measured in similar future studies of water quantity or quality.

8. The study also demonstrated that the presence of a reservoir, even a "dry" reservoir, within a water balance model can seriously affect

the calibration accuracy. Evidently some random storage and release must occur even when the reservoir is supposed to be passing all inflows directly through the reservoir.

9. The study has also broadly defined a needed and companion research project into the frequency component of the proposed low flow estimation method. This project will further develop step 1 of the two step method.

10. The results of this study have defined the following research needs to further refine the accuracy of the water balance element.

a. Add an additional Δs variable to the model. A suggested form is $P_{\Delta s}$(QOUT + QDIV + QDIS) with a $P_{\Delta s} \cdot$QIN term moved to the dependent variable side of the equation and appropriate interpretations applied to the B estimates.

b. The relative value of the QDIV/QDIS and ODIV1/QDIS1 configurations should be evaluated.

c. The role of variance equalization models should be evaluated further, especially the value of using dissimilar models for different stretches or data blocks. In addition the value of QOUT, QDIV, or QOUT-QDIV models should be compared with the YHAT model.

d. The value of using different autocorrelation models for each stretch should be investigated.

e. One or more years of additional variables should be developed and used to verify the calibration results reported herein.

11. Finally, this study has shown that water budget models can be developed economically. This study was accomplished by a single researcher with a limited budget in about six months. With this experience, future models can be developed in even shorter periods by water quality management agencies and consultants.

ACKNOWLEDGMENTS

The work reported in this dissertation could not have been accomplished without the help of numerous people and institutions, and the support, encouragement, and assistance so willingly given by all concerned was very much appreciated. Several persons should receive special recognition for filling major roles in the author's education and research program. Drs. Merwin D. Dougal and Neil E. Harl, serving as co-major professors, have contributed substantial portions of their time to the direction of the author's graduate study program, with assistance from the balance of the study committee: Drs. E. Robert Baumann, James C. Young, Vincent A. Sposito, and John L. Cleasby, with T. Al Austin substituting during the current year while Dr. Cleasby was on sabbatical leave. Special advisory services were provided by Dr. Wayne Fuller of the Iowa State University Department of Statistics. Dr. Norman Evans, Director of the Environmental Resources Center at Colorado State University, coordinated the use of the CSU computational facilities, in cooperation with the Iowa State Water Resources Research Institute.

Financial support for the academic portion of the author's study program leading to this research was provided by an Environmental Protection Agency Fellowship, No. 910089. The financial support for computer usage was provided by the Department of Civil Engineering and the Engineering Research Institute of Iowa State University and the Iowa State Water Resources Research Institute, the latter being supported in part by the Office of Water Research and Technology, U.S. Department of the Interior.

Recognition in the nontechnical area is extended first to Dr. Charles Oulman of the sanitary engineering faculty in the Department of Civil Engineering at Iowa State University whose persistent encouragement led to the author's final perseverance to the task of completing this work; and second and most important, to my wife, Sharon Hubly, who made numerous contributions and sacrifices during the course of this research effort.

LITERATURE CITED

1. "Federal Water Pollution Control Act of 1965." U.S. Congress, 79 U.S. Statutes at Large 903, 1965.

2. "1972 Amendments to the Federal Water Pollution Control Act." U.S. Congress, PL92-500, 1972.

3. J. E. McKee and H. W. Wolf. "Water quality criteria." 2nd ed. State Water Quality Control Board Pub. 3-A. State of California, Sacramento, Calif., 1963.

4. H. W. Wolf. "Quality standards for natural waters; a historical background." Proceedings; National symposium on quality standards for natural waters, Pp. 27-31, University of Michigan, Ann Arbor, Michigan, 1966.

5. H. C. Riggs. "Frequency curves." In _Techniques of Water Resources Investigations of the United States Geological Survey_, Book, 4, Chapter A2. U.S. Government, Washington, D.C., 1968.

6. H. C. Riggs. "Low flow investigations." In _Techniques of Water Resources Investigations of the United States Geological Survey_, Book 4, Chapter B1. U.S. Government, Washington, D.C., 1968.

7. "Technical Report No. 1, Low-Flow Frequency Analysis." U.S. Army Engineers District, U.S. Corps of Engineers, Tulsa, Oklahoma, Civil Works Investigations, Project CW-154, 1964.

8. V. Yevjevich. _Probability and Statistics in Hydrology_. Water Resources Publications, Fort Collins, Colorado, 1972.

9. V. T. Chow. _Handbook of Applied Hydrology_. McGraw-Hill Book Co., New York, 1964.

10. E. G. Gumbel. _Statistics of Extremes_. Columbia University Press, New York, 1958.

11. E. G. Gumbel. "Statistical theory of droughts." _Proceedings of the American Society of Civil Engineers_ 80: 439, 1954.

12. "Streamflow statistics." U.S. Geological Survey, National Center, Reston, Virginia, Water Resources Division, Computer Program A969, 1974.

13. D. M. Thomas and M. A. Benson. "Generalization of streamflow characteristics from drainage-basin characteristics." U.S. Geological Survey, Water Supply Paper No. 1975, 1970.

14. John F. Orsborn. "Determining streamflows from geomorphic parameters." Journal of Irrigation and Drainage Division 100: IR4, Dec. 1974.

15. John F. Orsborn. "A geomorphic method for estimating low flows." Paper presented at the ASCE Annual Meeting, Denver, Colorado, Nov. 1975.

16. Wells A. Hutchins. Water Rights Laws in the Nineteen Western United States, vols. I, II, and III. Economic Research Service, U.S. Department of Agriculture, Miscellaneous Publication No. 1206, 1974.

17. Richard R. Luckey and Russell K. Livingston. "Reservoir release routing model for the Upper Arkansas River basin of Colorado." Colorado Water Conservation Board, State of Colorado, Denver, Colorado, Circular No. 27, 1975.

18. Water Resources Council. Water and related land resources, establishment of principles and standards of planning." Federal Register 38: 174, Sept. 10, 1973.

19. Frederich S. Hillier and Gerald J. Lieberman. Introduction to Operations Research. Holden-Day, Inc., San Francisco, California, 1967.

20. "Scil Survey, Bent County, Colorado." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., Jan. 1971.

21. "Soil survey, Otero County, Colorado." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., March 1972.

22. "Soil survey, Crowley County, Colorado." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., March 1972.

23. Topographic maps of Colorado. U.S. Geological Survey, U.S. Government Printing Office, Washington, D.C.,

24. "Daily diversion records, Division II." Division of Water Resources, State of Colorado, Denver, Colorado, 1973 and 1974.

25. W. W. Wheeler and Associates, and Woodward-Clyde and Associates. "Water legislation investigations for the Arkansas River basin in Colorado." Colorado Water Conservation Board, State of Colorado, Denver, Colorado, Sept. 1968.

26. "Water users census, summary — High Line Canal Company — 1972." U.S. Bureau of Reclamation, U.S. Government Printing Office, Washington, D.C., Unpublished report, Fryingpan-Arkansas Project, Pueblo, Colorado, 1972.

27. Russell K. Livingston. "Transit losses and travel times for reservoir releases, Upper Arkansas River basin, Colorado." Colorado Water Conservation Board, State of Colorado, Denver, Colorado, Circular No. 20, 1973.

28. "Arkansas River Basin Water Quality Management Plan." Water Quality Control Division, Department of Health, State of Colorado, Denver, Colorado, 1975.

29. "Draft Environmental Statement: Fryingpan-Arkansas Project." U.S. Bureau of Reclamation, U.S. Government Printing Office, Washington, D.C., March 1974.

30. "Arkansas River Compact." Colorado Water Conservation Board, State of Colorado, Denver, Colorado, 1948.

31. D. D. Helton. "An optimization study of John Martin Reservoir operation." M.S. Thesis, University of Colorado, Boulder, Colorado, 1972.

32. "Water Resources Data for Colorado; Part I — Surface Water Records," U.S. Geological Survey, U.S. Government Printing Office, Washington, D.C., 1973 and 1974.

33. "Reservoir reports." Division of Water Resources, State of Colorado, Denver, Colorado, 1968-1974.

34. D. E. Burkham and D. R. Dowdy. "Error analysis of streamflow data for an alluvial stream." U.S. Geological Survey, Professional Paper No. 655-C, 1970.

35. Discharge permit No. CO-00701. National Pollutant Discharge Elimination System, Environmental Protection Agency, U.S. Government Printing Office, Washington, D.C., 1975.

36. James O. Taylor and Richard R. Luckey. "Water-management studies of a stream-aquifer system, Arkansas River valley, Colorado." Groundwater 12: 1, January-February 1974.

37. O. E. Meinzer. "Outline of groundwater hydrology with definitions." U.S. Geological Survey, Water Supply Paper No. 494, 1923.

38. O. E. Meinzer. "Plants as indicators of groundwater." U.S. Geological Survey, Water Supply Paper No. 577, 1927.

39. W. F. Robinson. "Phreatophytes." U.S. Geological Survey, Water Supply Paper No. 1423, 1958.

40. T. W. Robinson. "Phreatophytes and their relation to water in western United States." Trans. of the American Geophysical Union 33: 57-61, 1952.

41. E. P. Weeks and M. L. Sorey. "Use of finite-difference arrays of observation wells to estimate evapotranspiration from groundwater in the Arkansas River valley, Colorado." U.S. Geological Survey, Water Supply Paper No. 2029-C, 1973.

42. J. S. Gatewood, W. T. Robinson, B. R. Colby, J. D. Hem, and L. C. Halpenny. "Use of water by bottom-land vegetation in Lower Safford Valley, Arizona." U.S. Geological Survey, Water Supply Paper No. 1103, 1950.

43. H. F. Blaney. "Determining evapotranspiration by phreatophytes from climatological data." <u>Trans. of the American Geophysical Union</u> 33: 1, 1952.

44. E. G. Dunford and P. W. Fletcher. "Effect of removal of stream-bank vegetation upon water yield." <u>Trans. of the American Geophysical Union</u> 28: 1, 1947.

45. S. F. Turner and H. E. Skibitzke. "Use of water by phreatophytes in 2,000 foot channel between Granite Reef and Gillespie Dams, Maricopa County, Arizona." <u>Trans. of the American Geophysical Union</u> 33: 1, 1952.

46. A. A. Young and H. F. Blaney. "Use of water by native vegetation." Department of Public Works, State of California, Sacramento, California, Division of Water Resources Bulletin 50, 1942.

47. "Water resources activities in the United States — evapotranspiration reduction." Select Committee on National Water Resources, U.S. Senate, Committee Print No. 21, Feb. 1960.

48. H. F. Blaney and W. D. Criddle. "Consumptive use and irrigation water requirements of crops in Colorado." Soil Conservation Service, Report-Region 6, 1949.

49. S. E. Rantz. "A suggested method for estimating evapotranspiration by native phreatophytes." U.S. Geological Survey, Professional Paper 600-D, 1968.

50. W. C. Barrett and C. H. Milligan. "Consumptive water use and requirements in the Colorado River area of Utah." Utah Agricultural Experiment Station, Special Report 8, 1953.

51. H. F. Blaney and W. D. Criddle. "Consumptive use of water in the irrigated areas of the Upper Colorado River basin (provisional)." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., 1949.

52. H. F. Blaney and P. A. Ewing. "Irrigation practices and consumptive use of water in the Salinas Valley, California." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., 1946.

53. H. F. Blaney, P. A. Ewing, O. W. Isrealson, and others. "Part III — Water utilization." National Resources Commission, Regional plan, part VI — Upper Rio Grande, 1938.

54. H. F. Blaney, K. V. Morin, and W. D. Criddle. "Consumptive water use and requirements — the Pecos River joint investigation reports of the participating agencies." National Resources Planning Board, U.S. Government Printing Office, Washington, D.C., 1942.

55. H. F. Blaney and Homer Stockwell. "Irrigation investigations, San Fernando Valley, California." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., 1941.

56. H. F. Blaney, C. A. Taylor, H. G. Nickle, and A. A. Young. "Water losses under natural conditions from wet areas in southern California." Department of Public Works, State of California, Sacramento, California, Division of Water Resources Bulletin 44, 1933.

57. H. F. Blaney, A. A. Young, and C. A. Taylor. "Rainfall penetration and consumptive use of water in Santa Ana River valley and coastal plain." Department of Public Works, State of California, Sacramento, California, Division of Water Resources Bulletin 33, 1930.

58. W. D. Criddle and D. F. Peterson, Jr. "Colorado River area investigations — a progress report on consumptive water use and requirements." State Engineer, State of Utah, Salt Lake City, Utah, 1949.

59. D. C. Muckel and H. F. Blaney. "Utilization of the waters of the Lower San Luis Rey Valley, San Diego County, California." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., 1945.

60. D. C. Muckel and H. F. Blaney. "Water losses in the Santa Ana River Canyon below Prado Dam, California." Soil Conservation Service, U.S. Government Printing Office, Washington, D.C., 1946.

61. Harry F. Blaney. "Consumptive use and water waste by phreatophytes." Journal of the Irrigation and Drainage Division 87: IR3, Sept. 1961.

62. I. S. McQueen and R. F. Miller. "Soil-moisture and energy relationships associated with riparian vegetation near San Carlos, Arizona." U.S. Geological Survey, Water Supply Paper No. 655-E, 1972.

63. "Climatological Data." Environmental Data Service, U.S. Government Printing Office, Washington, D.C., 1973 and 1974.

64. Marvin E. Jensen, ed. Consumptive Use of Water and Irrigation Water Requirements. ASCE, New York, 1975.

65. W. C. Munson. "Methods for estimating consumptive use of water for agriculture." Journal of the Irrigation and Drainage Division 86: IR4, Dec. 1960.

66. H. L. Penman. "Natural evaporation from open water, bare soil, and grass." Proc. of the Royal Society 193, April 1948.

67. M. A. Kohler, T. J. Nordenson, and W. E. Fox. "Evaporation from pans and lakes." U.S. Weather Bureau, Research Paper 38, 1955.

68. C. H. M. vanBavel. "Potential evaporation: The combination concept and its experimental verification." Water Resources Research 2: 3, 1966.

69. J. A. Businger. "Some remarks on Penman's equations for the evapotranspiration." Netherlands Jour. Agricultural Sciences 4: 77, 1956.

70. G. F. Makink. "Testing the Penman formula by means of lysimeters." Journal of Inst. Water Engineering 11: 3, 1957.

71. J. E. Christiansen. "Estimating pan evaporation and evapotranspiration from climatic data." In Methods for Estimating Evapotranspiration. ASCE, New York, 1966.

72. M. E. Jensen and H. R. Haise. "Estimating evapotranspiration from solar radiation." Journal of the Irrigation and Drainage Division 96: IR4, Dec. 1963.

73. M. E. Jensen. "Plant and irrigation water requirements." In Sprinkler Irrigation, 3rd ed. Edited by C. H. Pair, Sprinkler Irrigation Association, Washington, D.C., 1969.

74. M. E. Jensen, D. C. N. Robb, and C. E. Franzoy. "Scheduling irrigations using climate-soil data." Journal of the Irrigation and Drainage Division 96: IR1, March 1970.

75. J. C. Stephens. Disc. of "Estimating evapotranspiration from isolation." Journal of the Hydraulic Division, 91: HY5, 1965.

76. R. K. Lane. "Estimating evaporation from isolation." Journal of the Hydraulics Division 90: HY5, Sept. 1964.

77. W. C. Munson. "Estimating consumptive use by the Munson P.E.-Index method." In Methods for Estimating Evapotranspiration. ASCE, New York, 1966.

78. J. E. Christiansen and G. H. Hargreaves. "Irrigation requirements from evaporation." Trans. of the International Commission on Irrigation and Drainage 3, 1969.

79. J. Papadakis. Climates of the World and their Agricultural Potentialities. Buenos Aires, Argentina, 1966.

80. W. R. Hamon. "Estimating potential evapotranspiration." Journal of the Hydraulics Division 87: HY3, May 1961.

81. "Irrigation water requirements." Soil Conservation Service, Technical Release No. 21, 1970.

82. H. F. Blaney and W. D. Criddle. "Determining water requirements in irrigated areas from climatological and irrigation data." Soil Conservation Service, Technical Paper No. 96, 1950.

83. H. F. Blaney and W. D. Criddle. "Determining consumptive use and irrigation water requirements." Soil Conservation Service, Technical Bulletin No. 1275, 1962.

84. "Colorado agricultural statistics, 1973 preliminary, 1972 final." Colorado Crop and Livestock Reporting Service, State of Colorado, Denver, Colorado, July 1974.

85. C. W. Thornthwaite. "An approach toward a rational classification of climate." The Geological Review 38: 1, 1948.

86. R. L. Lowry and A. F. Johnson. "Consumptive use of water for agriculture." Trans. of the American Society of Civil Engineers 107, 1942.

87. H. Olivier. Irrigation and Water Resources Engineering. Edward Arnold, London, England, 1972.

88. J. J. Behnke and G. B. Maxey. "An empirical method for estimating monthly potential evapotranspiration in Nevada." Journal of Hydrology 8, 1969.

89. R. W. Cruff and T. H. Thompson. "A comparison of methods of estimating potential evapotranspiration from climatological data in arid and subhumid environments." U.S. Geological Survey, Water Supply Paper No. 1839-M, 1967.

90. H. F. Blaney and K. V. Morin. "Evaporation and consumptive use of water formulae." Trans. of the American Geophysical Union 23: 1, 1942.

91. W. D. Criddle. "Methods of computing consumptive use of water." Journal of the Irrigation and Drainage Division 84: IR1, January 1958.

92. "Progress report. Oxford Farmers Ditch Company system investigation. Irrigation seasons 1968 to 1970." Colorado Water Conservation Board, State of Colorado, Denver, Colorado, Unpublished report, 1971.

93. Vernon B. Sauer. "Unit-response method of open-channel flow routing." Journal of the Hydraulic Division 99: HY1, 1973.

94. N. R. Draper and H. Smith. Applied Regression Analysis. John Wiley and Sons, New York, 1966.

95. Jolayne Service. A User's Guide to the Statistical Analysis System. Sparks Press, Raleigh, N.C., 1972.

96. Norman H. Nie, C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Brent. SPSS-Statistical Package for the Social Sciences, 2nd ed. McGraw-Hill Book Co., New York, 1975.

97. J. von-Neumann. "Distribution of the ratio of the mean square successive difference to the variance." Ann. Math. Statist. 12, 1941.

98. J. Durbin and G. S. Watson. "Testing for serial correlation in least-squares regression." Biometrika 37 and 38, 1950 and 1951.

99. J. Johnston. Econometric Methods, 2nd ed. McGraw-Hill Book Co., New York, 1972.