



Time series analysis of irrigation return flow
by Michael Earl Nicklin

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Civil Engineering
Montana State University
© Copyright by Michael Earl Nicklin (1983)

Abstract:

The potential of time series analysis as a tool for quantifying irrigation return flows along a reach of the Beaverhead River near Dillon, Montana is investigated. Univariate time series analysis is used to determine the stochastic character of both the irrigation diversion series and the irrigation return flow series. Box and Jenkins' transfer function identification procedures are used to determine the nature of the diversion-return flow relationship. Analysis of the 1974 irrigation season yielded a transfer function relationship in which return flows are most strongly dependent upon irrigation diversions that were made 54 days earlier. The 1974 relationship also indicates that 53.5 percent of the irrigation season diversion volume returns to the Beaverhead reach during the same irrigation season.

Although the study demonstrates the potential of the transfer function methodology as a tool for return flow quantification, it also shows how that potential is severely limited by data inadequacies. Proper application of time series analysis to irrigation return flow systems requires reliable and sufficient records of daily stream flows and daily diversions.

TIME SERIES ANALYSIS OF IRRIGATION
RETURN FLOW

by

Michael Earl Nicklin

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Civil Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

August 1983

APPROVAL

of a thesis submitted by

Michael Earl Nicklin

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

8/10/83
Date

Richard L. Brustheim
Chairperson, Graduate Committee

Approved for the Major Department

8/8/83
Date

Theodore Williams
Head, Major Department

Approved for the College of Graduate Studies

22 Aug. 1983
Date

Michael Malone
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute copies of the dissertation in and from microfilm and the right to reproduce and distribute by abstract in any format."

Signature

Date

Michael E. Nicklen
8/17/83

ACKNOWLEDGMENTS

The writer wishes to express his appreciation to Dr. Richard L. Brustkern for his guidance and assistance during the preparation of this thesis. Further appreciation is extended to the other graduate committee members, most notably Ted Williams, and to the entire Civil Engineering and Engineering Mechanics staff at Montana State University for their support.

Special thanks is extended to the proficient typist of this report, Jean Julian.

Finally, extreme gratitude is expressed to the author's wife, Rebecca, for the encouragement, patience, and understanding given during this investigative effort.

TABLE OF CONTENTS

	Page
APPROVAL	ii
STATEMENT OF PERMISSION TO USE	iii
ACKNOWLEDGMENTS	vi
TABLE OF CONTENTS	v
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	x
INTRODUCTION	1
TIME SERIES ANALYSIS	3
Definition of Time Series	3
Methodologies for Time Series Analysis	5
Statistical Descriptions of a Univariate Time Series	6
Autocorrelation Function	7
Partial Autocorrelation Function	8
Stationarity	10
Models for Univariate Time Series	11
Moving Average Models	11
Autoregressive Models	14
Mixed Autoregressive and Moving Average Models	17
Autoregressive Integrated Moving Average Models	18
Identification of Univariate Time Series Models	20
Identification of Moving Average Models	21
Identification of Autoregressive Models	25
Identification of ARMA Models	26
Identification of ARIMA Models	26
Transfer Function Models for Bivariate Time Series Analysis	27
Linear Transfer Function Models	27
Cross Correlation Function	33
Identification of Transfer Function Models	34
Model Diagnostics	39

TABLE OF CONTENTS—Continued

	Page
SELECTED STUDY SITE.....	45
Study Area Description.....	45
Climate.....	45
Surface Water.....	45
Groundwater.....	48
Study Site Data.....	51
Preliminary Data Analysis.....	54
TIME SERIES ANALYSIS OF STUDY SITE AREA.....	62
Model Identification Procedure.....	62
First Transfer Function Model.....	71
Second Transfer Function Model.....	75
Third Transfer Function Model.....	78
Summary of Identified Models.....	81
SUMMARY AND CONCLUSIONS.....	85
BIBLIOGRAPHY.....	87
APPENDIX.....	89

LIST OF TABLES

Tables	Page
1. Yearly Component Values as Obtained from the Hydrographs in Figures	59
2. The Seasonal Returns Expressed as a Percentage of the Seasonal Diversions	60
3. The Time Displacement in Days Between the Centroid of the Area Under the Seasonal Diversion Hydrograph and the Centroid of the Area Under the Seasonal Return Flow Hydrograph (1966-1975)	61
4. Parameter Estimates of the ω , δ_1 , and ϕ_1 Coefficients at Various Lags.	72
5. Parameter Estimates of the ω_b , δ_1 , and θ_1 Coefficients at Various Lags.	77
6. Parameter Estimates of the ω , δ_1 , and ϕ_1 Coefficients at Various Lags.	79
7. A Summary of the Statistics for the Three Model Forms	82

LIST OF FIGURES

Figures	Page
1. A conceptual municipal water use series provides an example of a time series exhibiting a trend plus a random component	4
2. The stream discharge of the Beaverhead River at the Barretts gaging station provides an example of a time series exhibiting both periodic and random components.....	4
3. An example of an autocorrelation function.....	8
4. A typical partial autocorrelation function for a time series with a time dependency at lag 1 only	10
5. Theoretical autocorrelation and partial autocorrelation functions for moving average processes of order 1 and 2.....	22
6. Theoretical autocorrelation and partial autocorrelation functions for autoregressive processes of order 1 and 2.....	23
7. A linear system	28
8. Impulse response function plots and their corresponding rational form transfer functions for constants r , s , and b	31
9. The Beaverhead Valley between the Clark Canyon Reservoir and Twin Bridges	46
10. The East Bench Unit	47
11. The study area	49
12. A map of the geologic units and a conceptual cross-section of the geologic units along the Beaverhead River between the Barretts and Blaine gaging stations	50
13. Water table contours of the hydrogeologic units east of the Beaverhead River near Dillon, Mt.....	52
14. A conceptual representation of the irrigation and return flow process	53

Figures	Page
15. The inflow and outflow hydrographs for the years 1966 through 1975 for the Beaverhead River between the Barretts and Blaine gaging stations.	56
16. Diversion and return flow hydrographs for the Beaverhead River reach between the Barretts and Blaine gaging stations during the years 1966 through 1975	57
17. Mean weekly inflow and outflow hydrographs for the years of 1966 through 1975	58
18. The diversion and return flow hydrographs for the 1974 irrigation season	64
19. The autocorrelation and partial autocorrelation coefficients for the 1974 diversions, and the cross correlation coefficients between the 1974 diversions and returns.	66
20. The autocorrelation and partial autocorrelation coefficients for the 1974 differenced ($d = 1$) diversions.	68
21. The autocorrelation and partial correlation coefficients for the whitened diversions, and the cross correlation coefficients between the 1974 whitened diversions and the filtered returns.	69
22. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series.	73
23. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series.	76
24. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series.	80
25. Diversion series and return series.	82
26. Diversion series and return series.	83
27. The impulse response function for the 1974 irrigation season.	84

ABSTRACT

The potential of time series analysis as a tool for quantifying irrigation return flows along a reach of the Beaverhead River near Dillon, Montana is investigated. Univariate time series analysis is used to determine the stochastic character of both the irrigation diversion series and the irrigation return flow series. Box and Jenkins' transfer function identification procedures are used to determine the nature of the diversion-return flow relationship. Analysis of the 1974 irrigation season yielded a transfer function relationship in which return flows are most strongly dependent upon irrigation diversions that were made 54 days earlier. The 1974 relationship also indicates that 53.5 percent of the irrigation season diversion volume returns to the Beaverhead reach during the same irrigation season.

Although the study demonstrates the potential of the transfer function methodology as a tool for return flow quantification, it also shows how that potential is severely limited by data inadequacies. Proper application of time series analysis to irrigation return flow systems requires reliable and sufficient records of daily streamflows and daily diversions.

INTRODUCTION

In the semiarid West the major water use is agricultural irrigation. A good water management plan must therefore properly quantify those volumes of water that are diverted, utilized, and returned in connection with the irrigation operation. Quantification of these various volume components is complicated by the fact that although diversion data may be available, the diverted discharge often greatly exceeds crop consumptive requirements. Consequently, an unknown portion of the water eventually percolates into the groundwater system and finally back into the stream adjacent to the irrigated area.

A state-of-the-art study by Nicklin and Brustkern (1981) determined that most current mathematical methodologies dealing with irrigation return flow analysis center around research-oriented computer models. Typically these models employ either a finite difference or finite element numerical solution procedure. The major data needs for such return flow models are groundwater levels, hydraulic conductivities, storage coefficients, aquifer boundary information, and surface water measurements.

A mathematical technique known as time series analysis offers a potential analysis approach for return flow quantification. Time series analysis involves the development of mathematical descriptions for data sets (such as streamflow hydrographs) that arise sequentially over time. Although no applications of time series analysis to irrigation return flows could be identified in the literature, numerous applications to other hydrologic time series have been documented. The approach has been particularly useful in forecasting stochastic events such as streamflow. Salas et al. (1980) provide a good review of traditional hydro-

logic uses of time series. Time series analysis is also widely used in other areas such as control theory for machines and electronic circuits.

The time series characterization of irrigation return flows as herein presented requires only surface water input data, some of which is often readily available. Streamflow records, in particular, are routinely gathered at numerous locations and constitute a generally reliable network of time series observations. Successful application of time series analysis procedures to this data base provides information on the timing and quantity of return flows. Such information is critical for planners who must make decisions relative to additional irrigation diversions and the impacts of changing irrigation practices upon downstream users.

This study investigates the potential of time series analysis as a tool for quantifying irrigation flows.

TIME SERIES ANALYSIS

Definition of Time Series

A time series is a collection of numerical observations arranged according to their order of occurrence. Although the data may in general arise sequentially over time or space, this study is concerned only with data that arises sequentially over time. Time series data can be characterized as having (1) deterministic properties, (2) random properties or (3) both deterministic and random properties. A time series has strictly deterministic properties if it can be exactly resynthesized. On the other hand, a time series with any random properties cannot be exactly resynthesized. Owing to the various possible outcomes, it can only be described by utilizing probability theory. Most hydrologic time series exhibit both deterministic and random characteristics.

The deterministic component of a hydrologic time series may consist of either or both of the following elements:

1. Trends—such as linear or quadratic changes in the level of a hydrologic series over time (for example, Figure 1 shows a deterministic trend and a random component), and,
2. Periodicities—such as cyclical variations in the level of a hydrologic series due primarily to seasonal meteorological changes (for example, Figure 2 shows a deterministic periodicity and a random component).

Random (or stochastic) properties in hydrologic data are due primarily to the host of uncertainties related to the meteorological cycle. Annual hydrologic measurements such as total precipitation often approximate strictly random processes.

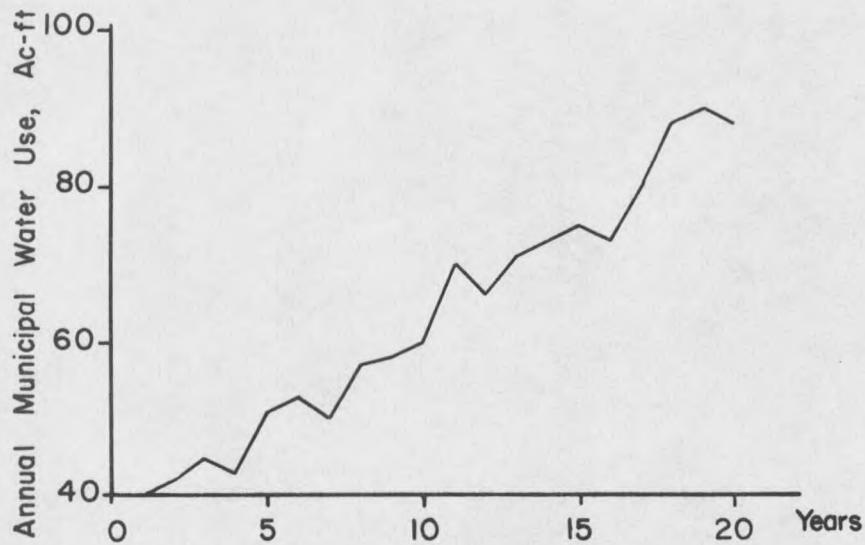


Figure 1. A conceptual municipal water use series provides an example of a time series exhibiting a trend plus a random component.

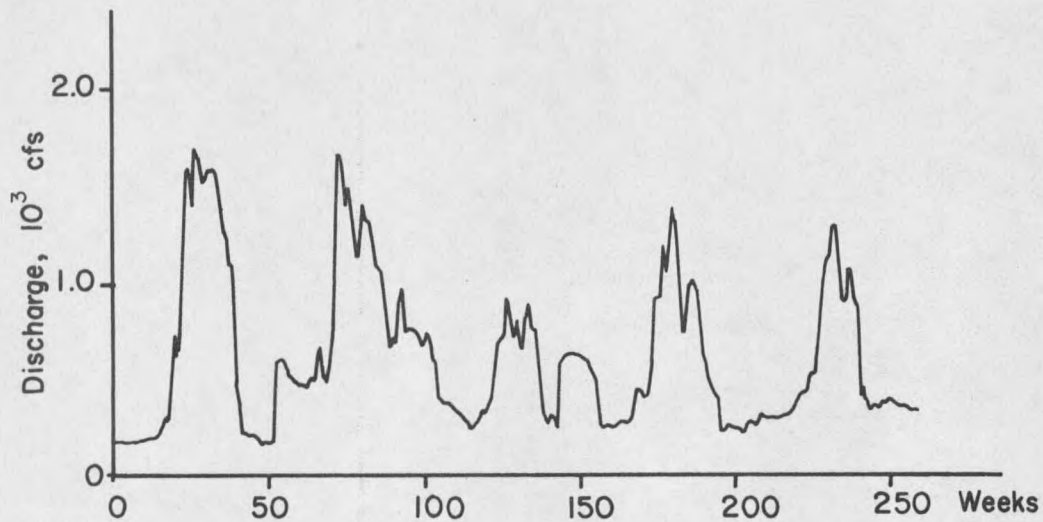


Figure 2. The stream discharge of the Beaverhead River at the Barretts gaging station provides an example of a time series exhibiting both periodic and random components.

Time series analysis generally requires that a mathematical description of the time series data be developed. Although time series processes may be continuous or discrete, a discrete time series description is generally employed (out of either necessity or convenience) to describe the hydrologic process. A discrete time series may evolve from a continuous time series process in two ways:

1. A discrete time series results from sampling a continuous time series at discrete intervals. For example, measurements of water levels in an observation water well at monthly increments result in a discrete time series from a continuous process of water level fluctuation.
2. A discrete time series also results from accumulating or averaging a continuous variable over discrete time intervals. A relevant example is provided by a record of streamflows. The raw data set is a continuous record of water level stages when viewed on a recorded strip chart. Using a stage-discharge relationship, the raw data set is converted to a discrete time series of mean daily flows.

Methodologies for Time Series Analysis

Two mathematical analyses have been particularly useful in the interpretation of time series data. One assesses the behavior of the data set in the time domain, time domain analysis, and the other assesses the behavior in the frequency domain, frequency domain (or spectral) analysis. This investigation concentrates on time domain analysis procedures.

In time domain analysis, a model of the time series is generally developed. Although the model identification and development processes, which will be discussed later, are mechanical and methodical, a physical understanding of the time series process is an invaluable aid in model identification.

Assessment of a single time series record is often termed univariate time series analysis. Assessment of the relationship between two dependent time series records is frequently termed transfer function analysis of a bivariate process. Statistical descriptions of univariate time series will be discussed in the section that follows. Later, transfer function methods pertinent to bivariate processes will be presented.

Statistical Descriptions of a Univariate Time Series

A time series sample record over a given interval is often termed a realization of the underlying process. The collection of all possible records over the given interval of time is usually designated as the ensemble (analogous to the sample space in probability theory). The analysis of a univariate time series sample (realization) provides information regarding interdependencies within the series.

The three most important parameters generally used to characterize a stationary (stationarity is defined later) stochastic time series are the mean, variance, and autocovariance. The expected value of a time series, $X(t)$, is its mean, μ

$$E[X(t)] = \mu \quad (1)$$

The expected value of the second moment about the mean is the time series variance, σ^2

$$E\{X(t) - \mu\}^2 = \sigma^2 \quad (2)$$

Furthermore, the expected value of the autocovariance between a time series, $X(t)$, and the same time series displaced τ time units, $X(t+\tau)$, is defined as

$$E\{X(t) - \mu\} \{X(t+\tau) - \mu\} = \gamma_\tau \quad (3)$$

Inspection of Equation (3) reveals that $\gamma_0 = \sigma^2$ when $\tau = 0$. The time series sample statistics for the mean, variance, and autocovariance are respectively

$$\hat{\mu} = \frac{1}{N} \sum_{t=1}^N x(t) \quad (4)$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{t=1}^N \left\{ x(t) - \hat{\mu} \right\}^2 \quad (5)$$

$$\hat{\gamma}_{\tau} = \frac{1}{N} \sum_{t=1}^{N-\tau} \left\{ x(t) - \hat{\mu} \right\} \left\{ x(t+\tau) - \hat{\mu} \right\} \quad (6)$$

The use of upper case and lower case letters will designate population and sample variables respectively for the remaining discussion.

Autocorrelation Function

Dividing Equation (3) by Equation (2) yields the autocorrelation coefficient, $\hat{\rho}_{\tau}$, between pairs of time series values that are separated by τ time units,

$$\rho_{\tau} = \frac{\gamma_{\tau}}{\sigma^2} \quad (7)$$

and as previously noted $\sigma^2 = \gamma_0$ for $\tau = 0$. Therefore

$$\rho_{\tau} = \frac{\gamma_{\tau}}{\gamma_0} \quad (8)$$

The sample estimate of the autocorrelation coefficient for time displacement τ is

$$\hat{\rho}_{\tau} = \frac{\hat{\gamma}_{\tau}}{\hat{\gamma}_0} \quad (9)$$

For a given time series sample, the collection of autocorrelation coefficients for all possible time displacements is referred to as the autocorrelation function. The graphical representation of the autocorrelation function is sometimes referred to as the correlogram. The autocovariance and therefore the autocorrelation function is a second order property of a time series. The mean is a first order property and the variance is of course a second order property. As can be seen from Equations (6) and (8), the autocorrelation function may range from 1 to -1 (negative values resulting from negative correlations) with a maximum of one always occurring at $\tau = 0$. The autocorrelation function is symmetric about zero.

The shape of the graphed autocorrelation function serves as a guide in choosing an appropriate statistical model for characterizing a time series data set. That is it identifies the similarities between the realization, $x(t)$, and the realization shifted τ units. For example, a realization with the strongest positive dependencies between the most closely spaced values and progressively weaker dependencies between more widely separated values will, for small lags, exhibit large autocorrelation coefficients, and, for large lags, exhibit small autocorrelation coefficients. Thus, the autocorrelation function will decay in this case as τ is increased. A typical autocorrelation function for such a dependence is shown in Figure 3. Because of the function's previously mentioned symmetry, it is customary to represent the autocorrelation function to the right of the origin only. For a realization with no significant time dependencies between neighboring values, the autocorrelation coefficient will be insignificant even at small lags.

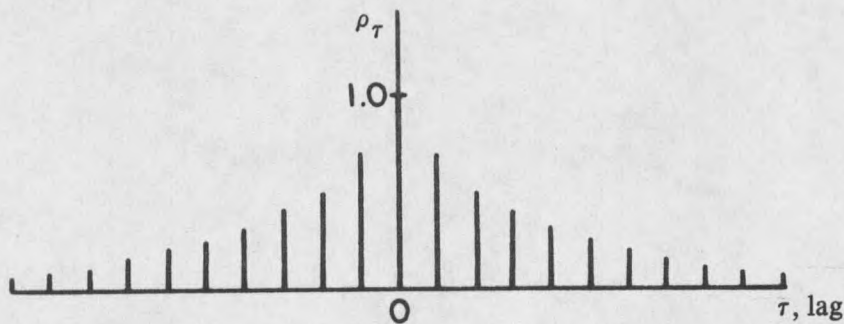


Figure 3. An example of an autocorrelation function.

Partial Autocorrelation Function

Another parameter useful in model identification is the partial autocorrelation coefficient. It is often important to measure the autocorrelation between the dependent variable, $X(t)$, and a particular independent variable, $X(t+\tau)$, with the effects of intervening variables removed. For example, if $X(t)$ is significantly autocorrelated to $X(t+1)$ and

$X(t+1)$ is significantly autocorrelated to $X(t+2)$, the result may well be a significant autocorrelation between $X(t)$ and $X(t+2)$ that is reflected in the autocorrelation coefficient ρ_2 . The partial autocorrelation coefficient removes the effect of the intervening variable $X(t+1)$ on the autocorrelation between $X(t)$ and $X(t+2)$, and thus provides an indication of the direct dependency between $X(t)$ and $X(t+2)$. More generally, the partial autocorrelation coefficient between $X(t)$ and $X(t+\tau)$ removes the effects of the intervening variables $X(t+1), X(t+2), \dots, X(t+\tau-1)$.

The partial autocorrelation coefficient for a realization and the same realization displaced τ time units, $\hat{\rho}_{\tau\tau}$, may be estimated from

$$\hat{\rho}_{\tau\tau} = \begin{cases} \hat{\rho}_1 & \text{if } \tau = 1 & \text{(no intervening variables)} \\ \frac{\hat{\rho}_\tau - \sum_{j=1}^{\tau-1} \hat{\rho}_{\tau-1,j} \hat{\rho}_{\tau-j}}{1 - \sum_{j=1}^{\tau-1} \hat{\rho}_{\tau-1,j} \hat{\rho}_j} & \text{if } \tau = 2, 3, \dots \end{cases} \quad (10)$$

where $\hat{\rho}_{\tau j} = \hat{\rho}_{\tau-1,j} - \hat{\rho}_{\tau\tau} \hat{\rho}_{\tau-1,\tau-j}$ for $j = 1, 2, \dots, \tau-1$

Note that $\hat{\rho}_\tau$ is the sample autocorrelation between the variable $x(t)$ and the variables displaced τ units, $x(t+\tau)$.

The collection of partial autocorrelation coefficients for all possible time displacements is referred to as the partial autocorrelation function. The shape of the graphed partial autocorrelation function is a guide in choosing a statistical model for a given time series. For example, if $X(t)$ is significantly autocorrelated with $X(t+1)$ but not with $X(t+2)$, then the partial autocorrelation function will indicate a significant partial autocorrelation only at a displacement of one as shown in Figure 4.

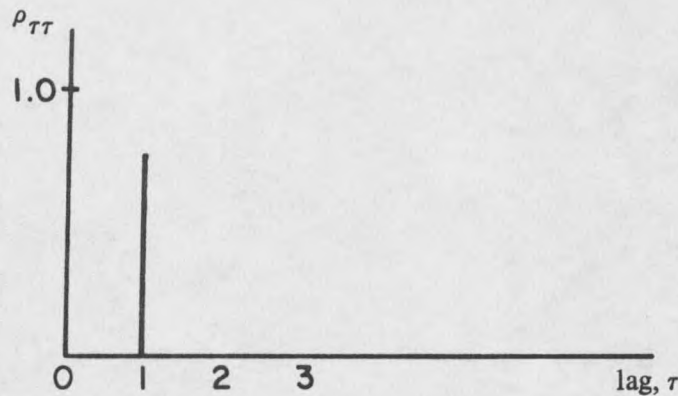


Figure 4. A typical partial autocorrelation function for a time series with a time dependency at lag 1 only.

Stationarity

A stochastic time series is stationary if its statistical properties do not change with time. Strict stationarity of a stochastic time series requires that statistical properties of all orders do not change over time. For example, strict stationarity for a normally distributed time series would require that the mean, variance, and autocovariance not change with time. For time series that are not normally distributed, and thus cannot be completely characterized by the mean, variance and autocovariance, strict stationarity requires constancy of the remaining characterization parameters.

Fortunately in most circumstances, strict stationarity is not a requirement for developing an adequate model. Stationarity of order two is usually considered sufficient. The normally distributed time series of the previous paragraph is completely characterized by the first and second order parameters of the mean, variance and autocovariance and thus is not only stationary to order two but also strictly stationary. For other distributions, constancy of the mean, variance and autocovariance would indicate the required second order stationarity but not strict stationarity.

The mean, variance, and autocovariance are statistics utilized in the determination of the autocorrelation function. Consequently, the autocorrelation function is useful in

assessing the stationarity of a time series data set. If the autocorrelation function fails to dampen or cut off quickly to zero, nonstationarity in the first and/or second order statistics is the usual cause. Another set of conditions implied by the concept of stationarity is described in a later section.

Stationarity of a time series data set is a requirement for developing a stochastic model of that data set. However, as will be shown later, there are effective ways to transform nonstationary data sets into stationary forms.

Models for Univariate Time Series

A group of models which often provides parsimonious representations of univariate time series are discussed in this section. The models described herein apply to either (1) stationary time series or (2) nonstationary time series transformable to stationary time series (the mechanics of this transformation are described later in this section). Model selection is based on the modeler's knowledge of the physical process and on his interpretation of the sample autocorrelation and partial autocorrelation functions of the process time series.

Moving Average Models

Based on an idea by Yule (1927), Box and Jenkins (1976) have developed a mathematical model that is useful in characterizing certain stochastic time series. The time series, $Z(t)$, is regarded as a set of successive highly dependent values which are generated from a series of independent and randomly distributed variables, $A(t)$, often termed white noise. By definition, the mean of a white noise process is zero and the variance is constant, or

$$E[A(t)] = \mu_A = 0 \quad (11)$$

$$E\left\{A(t) - \mu_A\right\}^2 = \sigma_A^2 = C \quad (12)$$

Furthermore, the autocovariance of a white noise series is zero,

$$E[A(t) A(t+\tau)] = 0 \quad (13)$$

for lags τ not equal to zero. The time series (and model) may be represented as

$$Z(t) = \mu + A(t) - \theta_1 A(t-1) - \theta_2 A(t-2) - \dots - \theta_q A(t-q) \quad (14)$$

The time series is thus dependent upon a "moving average" involving current and previous values from the white noise series. The previous values are weighted by the θ coefficients. The order of the moving average is indicated by q . A moving average model of order q is often abbreviated as MA(q).

Equation (14) can be more concisely written as

$$Z(t) = \mu + \theta^q(B) A(t) \quad (15)$$

where

$$\theta^q(B) = -\theta_0 B^0 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q$$

and by definition $\theta_0 = -1$. The symbol B is a backstep operator. It serves to transform $A(t)$ as follows

$$B^1 A(t) = A(t-1)$$

$$B^2 A(t) = A(t-2)$$

$$\vdots$$

$$B^q A(t) = A(t-q)$$

Another form of Equation (15) is

$$\tilde{Z}(t) = \theta^q(B) A(t) \quad (16)$$

where $\tilde{Z}(t) = Z(t) - \mu$. $\tilde{Z}(t)$ then represents the deviation of the time series from its mean, μ .

A time series that results from a moving average model of order q is always stationary. First order stationarity is easily demonstrated. The expected value of a white noise process equals zero or

$$E[A(t)] = \mu_A = 0$$

Thus, since $\theta^q(B)$ is a constant operator,

$$E[\theta^q(B) A(t)] = \theta^q(B) E[A(t)]$$

and consequently

$$E[\theta^q(B) A(t)] = 0.$$

Second order stationarity may be shown by forming the product of $\hat{Z}(t)\hat{Z}(t-\tau)$ and taking expectations term by term,

$$\begin{aligned} E[\hat{Z}(t)\hat{Z}(t-\tau)] &= E\left[\left\{1 - \theta_1 B^1 - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q\right\} A(t) \cdot \right. \\ &\quad \left.\left\{1 - \theta_1 B^1 - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q\right\} A(t-\tau)\right] \\ &= E\left[\left\{1 - \theta_1 B^1 - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q\right\} A(t) \cdot \right. \\ &\quad \left.\left\{B^\tau - \theta_1 B^{\tau+1} - \theta_2 B^{\tau+2} - \theta_3 B^{\tau+3} - \dots - \theta_q B^{\tau+q}\right\} A(t)\right] \quad (17) \end{aligned}$$

The expectation of the cross products involving the independently distributed white noise variables is zero at displacements $\tau \neq 0$ (recall Equation (13)). Thus for $\tau = 1$, for example,

$$E[\hat{Z}(t)\hat{Z}(t-1)] = \sigma_A^2 (-\theta_1 + \theta_1 \theta_2 + \theta_2 \theta_3 + \dots + \theta_{q-1} \theta_q)$$

Generalizing, the autocovariance is

$$\gamma_\tau = \sigma_A^2 \sum_{j=0}^{q-\tau} \theta_j \theta_{j+\tau} \quad (18)$$

Recall that $\theta_0 = -1$. When $\tau = 0$, Equation (17) yields

$$\gamma_0 = \sigma_A^2 (1 + \theta_1^2 + \theta_2^2 + \theta_3^2 + \dots + \theta_q^2)$$

and so the variance may be written

$$\sigma^2 = \gamma_0 = \sigma_A^2 \sum_{j=0}^q \theta_j^2 \quad (19)$$

Equations (18) and (19) demonstrate that the autocovariance and variance are independent of time and thus stationarity to order two of moving average processes is justified.

An important requirement for moving average processes is the requirement of invertibility. Although invertibility applies only to moving average processes, its importance is most easily demonstrated in the section that follows.

Autoregressive Models

A moving average process $\hat{Z}(t)$ may also be characterized in terms of a current white noise disturbance, $A(t)$, and all past time series observations of $\hat{Z}(t)$. For example, a moving average process of order one,

$$\hat{Z}(t) = (1 - \theta_1 B^1) A(t) \quad (20)$$

may be rewritten as

$$\frac{1}{1 - \theta_1 B^1} \hat{Z}(t) = A(t)$$

Dividing, yields

$$(1 + \theta_1 B^1 + \theta_1^2 B^2 + \theta_1^3 B^3 + \dots) \hat{Z}(t) = A(t) \quad (21)$$

Operating on $\hat{Z}(t)$ and rearranging gives

$$\hat{Z}(t) = -\theta_1 \hat{Z}(t-1) - \theta_1^2 \hat{Z}(t-2) - \dots + A(t) \quad (22)$$

This relationship is termed an autoregressive process as the variable $\hat{Z}(t)$ is regressed on past observations of itself. Thus a moving average process of order one may be transformed into an autoregressive process of infinite order. The moving average process is said to be invertible if the infinite series converges. As can be seen from the equation, convergence will occur only if $|\theta_1| < 1$. Consequently not all moving average processes of order one are invertible. If $|\theta_1| > 1$, the current deviation $\hat{Z}(t)$ depends upon events $\hat{Z}(t-1)$, $\hat{Z}(t-2)$, ...

with event weights increasing as events become more far removed in time. This is physically unreasonable and so invertibility is required.

Using Box and Jenkins' notation, Equation (21) may be written

$$\phi(B) \hat{Z}(t) = A(t) \quad (23)$$

where

$$\phi(B) = (1 + \theta_1 B^1 + \theta_1^2 B^2 + \dots).$$

Autoregressive processes may be more generally expressed as

$$\phi^p(B) \hat{Z}(t) = A(t) \quad (24)$$

where

$$\phi^p(B) = -\phi_0 B^0 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$$

and

$$\phi_0 = -1$$

Equation (24) may also be written as

$$\hat{Z}(t) = \phi_1 \hat{Z}(t-1) + \phi_2 \hat{Z}(t-2) + \dots + \phi_p \hat{Z}(t-p) + A(t) \quad (25)$$

The time series is thus dependent upon previous values of the time series weighted by the ϕ coefficients and the current white noise value $A(t)$. The order of the autoregressive process is indicated by p . An autoregressive process of order p is often abbreviated AR(p).

A comparison of Equation (20), an autoregressive form, and (23), a moving average form, indicates that it is parsimoniously unwise to select an autoregressive model to represent a process which could be sufficiently explained by a moving average model of low order (just as it will be shown parsimoniously unwise to use a moving average model to represent a process which is sufficiently modeled by an autoregressive model of low order).

The invertibility condition may be extended to a moving average process of any order. Consider a moving average of order q

$$\hat{Z}(t) = A(t) - \theta_1 A(t-1) - \dots - \theta_q A(t-q)$$

or

$$\hat{Z}(t) = \theta(B) A(t)$$

where $\theta(B) = \theta^q(B)$. It has already been shown that a first order moving average process is invertible if $|\theta_1| < 1$, or equivalently,

$$\pi(B) = (1 - \theta_1 B)^{-1} = \sum_{j=0}^{\infty} \theta_1^j B^j$$

converges. That is equivalent to saying that the root $B = \theta_1^{-1}$ of the equation $1 - \theta B = 0$ (the characteristic equation) lies outside of the unit circle (if $|\theta_1| < 1$ for invertibility, then $|1/\theta_1| > 1$). By similar reasoning, the invertibility condition for a MA(q) process requires that the roots of the characteristic equation

$$1 - \theta_1 B^1 - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q = 0$$

lie outside the unit circle for invertibility.

Autoregressive processes are always invertible because the series

$$\phi^p(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$$

is finite. However, an autoregressive process may or may not be stationary. Stationarity of an autoregressive process requires certain conditions on the ϕ coefficients in Equation (25).

To insure stationarity for an autoregressive process of order p,

$$\phi^p(B) \hat{Z}(t) = A(t)$$

where

$$\phi^p(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p$$

the roots of the characteristic equation

$$1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p = 0$$

must lie outside of the unit circle. For illustration of the importance of this requirement, an autoregressive process of order one, AR(1), may be written as

$$\hat{Z}(t) = \phi_1 \hat{Z}(t-1) + A(t) \quad (26)$$

or

$$(1 - \phi_1 B) \hat{Z}(t) = A(t)$$

Rearranging

$$\hat{Z}(t) = \frac{1}{1 - \phi_1 B} A(t)$$

and then dividing yields

$$\hat{Z}(t) = (1 + \phi_1 B^1 + \phi_1^2 B^2 + \dots) A(t) \quad (27)$$

The stationarity requirement for an AR(1) process is that the absolute value of ϕ_1 be less than one (or alternatively that the root of the characteristic equation $1 - \phi_1 B = 0$ be outside the unit circle). Note that when $|\phi_1| > 1$, the series is infinite and not convergent and the process is nonstationary.

A comparison of Equation (26), an autoregressive form, and (27), a moving average form, indicates that it is parsimoniously unwise to select a moving average model to represent a process which could be sufficiently explained by an autoregressive model of low order.

Mixed Autoregressive and Moving Average Models

Sometimes it is parsimoniously advantageous to use mixed autoregressive moving average models of order p and q (or ARMA(p, q)) to model certain processes. Such models often require fewer parameters for adequate description of stochastic processes. ARMA models with order $p \leq 2$ and $q \leq 2$ are adequate for many processes. An ARMA process (and model) with orders p and q may be written as

$$\begin{aligned}\hat{Z}(t) = & \phi_1 \hat{Z}(t-1) + \phi_2 \hat{Z}(t-2) + \dots + \phi_p \hat{Z}(t-p) \\ & + A(t) - \theta_1 A(t-1) - \theta_2 A(t-2) - \dots - \theta_q A(t-q)\end{aligned}\quad (28)$$

or, in abbreviated form

$$\phi^p(B) \hat{Z}(t) = \theta^q(B) A(t) \quad (29)$$

Rearranging,

$$\hat{Z}(t) = \frac{\theta^q(B)}{\phi^p(B)} A(t) \quad (30)$$

where $\hat{Z}(t)$ is the output from a white noise process, $a(t)$, which is weighted by the coefficients $\theta^q(B)/\phi^p(B)$. The coefficients thusly expressed are said to be in the rational form. The main advantage of the rational form is that fewer parameters are often required to sufficiently express the relationship between the white noise process and the stochastic process.

In order for processes to be properly represented by ARMA(p,q) models, the requirements of stationarity and invertibility must be met since both autoregressive and moving average coefficients are contained in these models.

Autoregressive Integrated Moving Average Models

Autoregressive integrated moving average models (ARIMA models) may be employed when the applicability of AR and ARMA models is restricted because of stationarity problems. ARIMA models are often useful when the mean of the process is nonstationary. For instance if the changes in the mean are uniform (trends), then a differencing transformation of the nonstationary time series will frequently provide a stationary time series so that meaningful model parameters can be determined. Following a successful stationary transformation, autoregressive and moving average parameters are estimated in the usual manner.

An ARIMA model's order is described as order p , d , and q . The orders p and q again refer to the mixed autoregressive and moving average orders. The order d refers to the differencing order which is described below. An ARIMA process (and model) with orders p , d , and q may be written as

$$\Theta^P(B) \hat{Z}(t) = \theta^q(B) A(t) \quad (31)$$

where $\Theta^P(B)$ is a nonstationary autoregressive operator. Defining

$$\begin{aligned} \Theta^P(B) &= \phi^P(B) \nabla^d \\ \phi^P(B) \nabla^d \hat{Z}(t) &= \theta^q(B) A(t) \end{aligned} \quad (32)$$

where, ∇^d , is a differencing operator that transforms the nonstationary process into a stationary form of the ARMA process and d is the number of differencing operations performed. The ∇^d operator serves to transform $\hat{Z}(t)$ as follows

$$\begin{aligned} \nabla^1 \hat{Z}(t) &= \hat{Z}(t) - \hat{Z}(t-1) \\ \nabla^2 \hat{Z}(t) &= \hat{Z}(t) - 2\hat{Z}(t-1) + \hat{Z}(t-2) \end{aligned}$$

For many situations stationarity is attained with $d \leq 2$. Equation (31) may be written

$$\phi^P(B) W(t) = \theta^q(B) A(t)$$

where

$$W(t) = \nabla^d \hat{Z}(t)$$

is the new differenced stationary series.

The parameters $\phi(B)$ and $\theta(B)$ are as previously described provided the differenced series is stationary and invertible. It is useful to note that MA(q), AR(p), and ARMA(p, q) processes (and models) are special cases of the more general ARIMA(p, d, q) processes (and models). Thus, if p and d are zero an ARIMA(p, d, q) model reduces to a MA(q) model, if d and q are zero the ARIMA(p, d, q) model reduces to an AR(p) model, or if d is zero the ARIMA(p, d, q) model reduces to an ARMA(p, q) model.

Summarizing, four basic stochastic models are available to characterize interdependencies within a time series:

1. Moving average models of order q , $MA(q)$,
2. Autoregressive models of order p , $AR(p)$,
3. Autoregressive moving average models of order p and q , $ARMA(p,q)$, and,
4. Autoregressive integrated moving average models of order p , d and q , $ARIMA(p,d,q)$.

Identification of Univariate Time Series Models

Four alternative linear stochastic models for approximating stochastic time series processes were described in the previous section. If possible, a parsimonious model should be chosen from these alternatives for representing a process. This selection is aided by a physical understanding of the process and by an assessment of the autocorrelation and partial autocorrelation functions of the process.

Recall that the autocorrelation function identifies the dependency between a time series value and earlier time series values. Recall also that the partial autocorrelation function identifies dependencies between particular time series values with the effects of intermediate time series values removed. Since a moving average process is equivalent to an autoregressive process of infinite order, the partial autocorrelation functions are of infinite extent and tail off. The autocorrelation function for a moving average process of order q cuts off at q since, as will be shown, only the autocorrelation coefficients $\rho_1, \rho_2, \dots, \rho_q$ as computed from Equation (9) are nonzero. Noting the correspondence between moving average and autoregressive processes (one is the inverse of the other), it is not surprising that the partial autocorrelation and autocorrelation function of an autoregressive process behave like the autocorrelation function and partial autocorrelation function respectively of a moving average process. Thus, for an autoregressive process of order p , the autocor-

relation function tails off and the partial autocorrelation function cuts off at lag p . Typical moving average and autoregressive functions for orders 1 and 2 are shown in Figures 5 and 6. For mixed autoregressive moving average processes both the autocorrelation and partial autocorrelation functions tail off. If the autocorrelation and partial autocorrelation functions tail off very slowly the cause may be due to process nonstationarity.

Thus a preliminary indication of the type and order of a tentative model can be determined from the behavior of the process autocorrelation and partial autocorrelation functions. Once the type and order of the model is tentatively identified, the autocorrelation function may be used to estimate the θ and/or ϕ parameters.

Identification of Moving Average Models

The time series autocorrelation function of a moving average process may be used to identify the order of the moving average process (and model) since the autocorrelation function will cut off after lag q (the moving average order). The autocovariance for lag τ and variance ($\tau = 0$) for a moving average process of order q are given by Equations (18) and (19) respectively,

$$\gamma_{\tau} = \sigma_A^2 \sum_{j=0}^{q-\tau} \theta_j \theta_{j+\tau} \quad \tau = 1, 2, \dots, q$$

$$\gamma_{\tau} = 0$$
(18)

and

$$\gamma_0 = \sigma_A^2 \sum_{j=0}^q \theta_j^2$$
(19)

Equation (8) indicates that the autocorrelation coefficient at displacement τ is

$$\rho_{\tau} = \frac{\gamma_{\tau}}{\gamma_0}$$
(8)

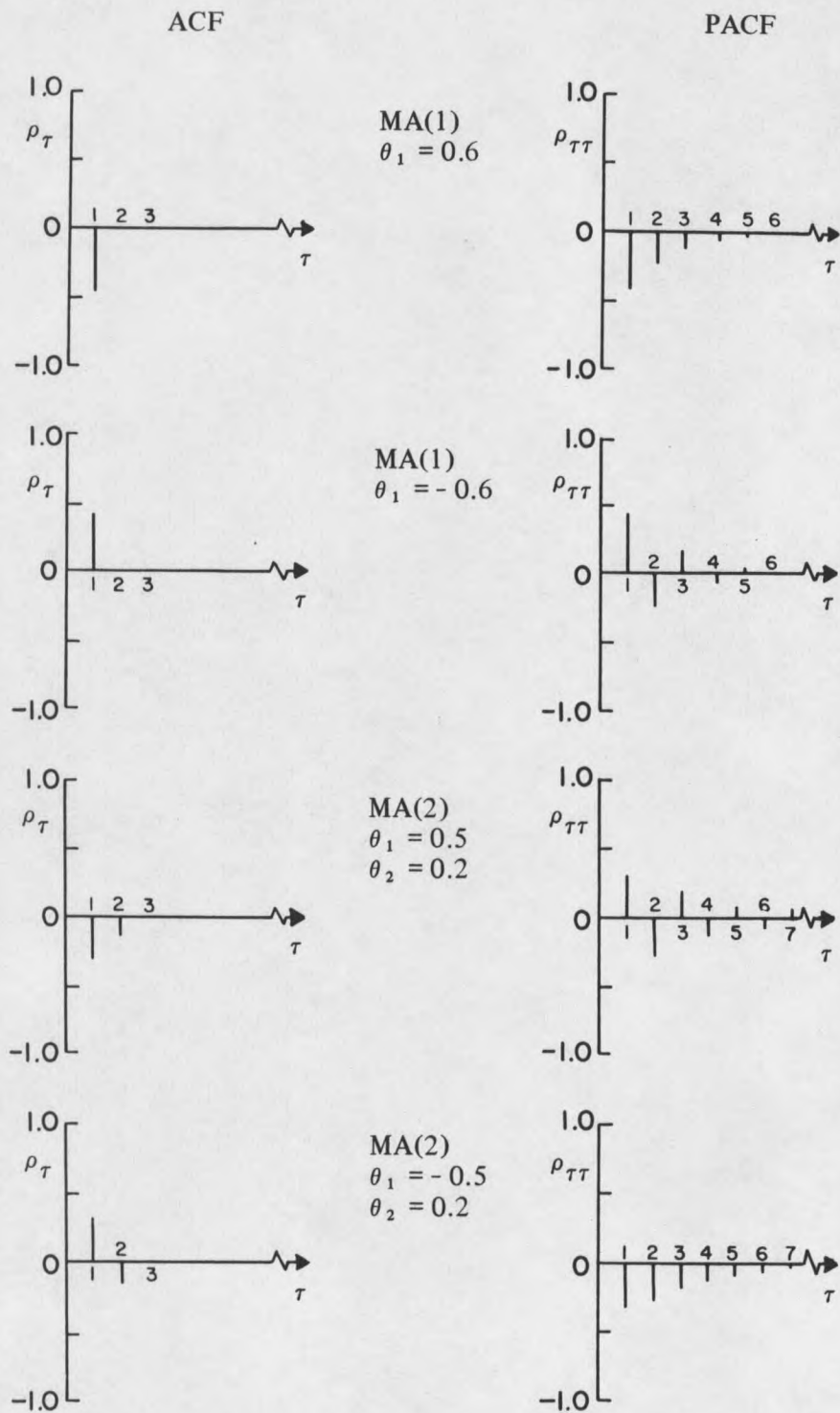


Figure 5. Theoretical autocorrelation and partial autocorrelation functions for moving average processes of order 1 and 2.

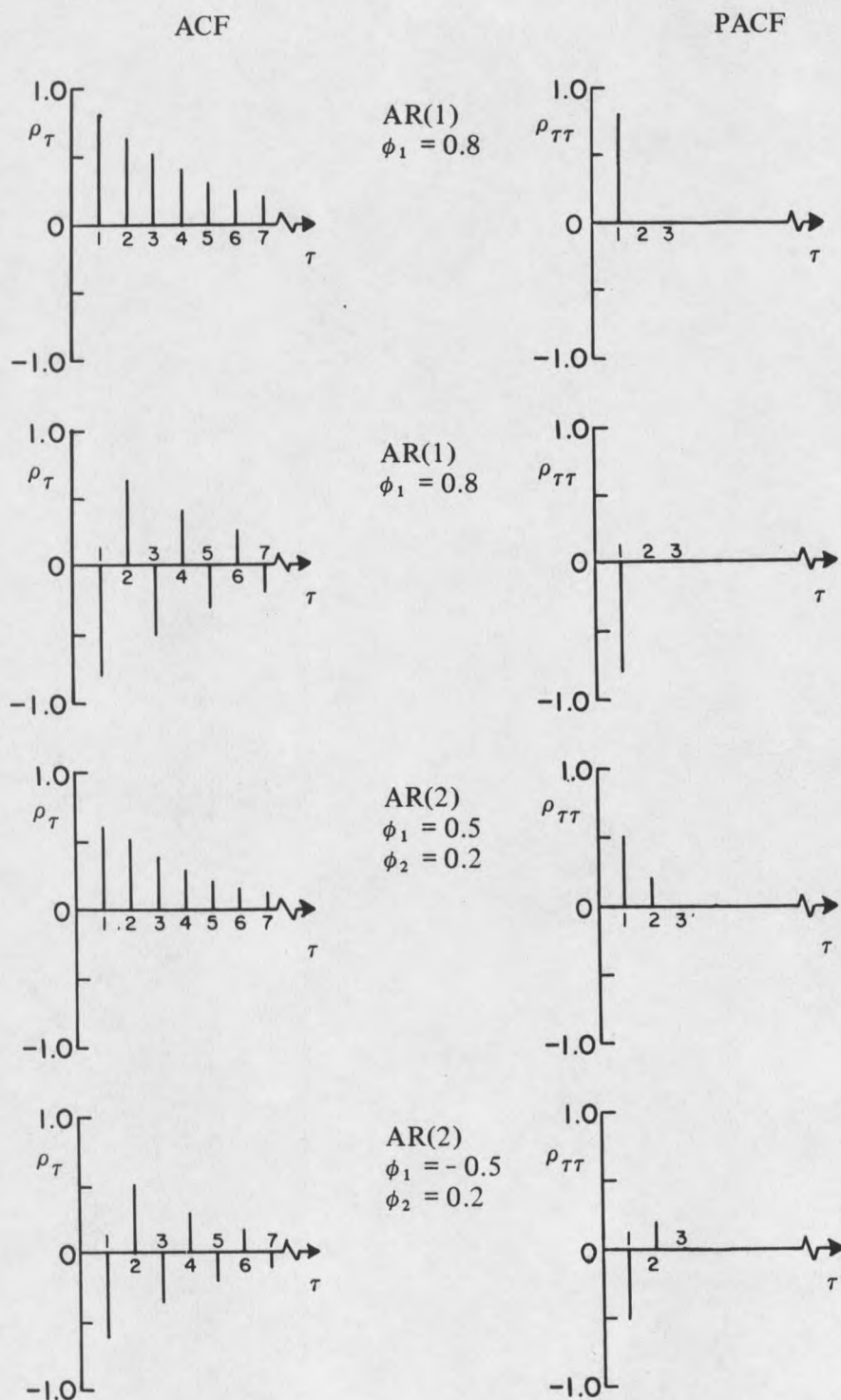


Figure 6. Theoretical autocorrelation and partial autocorrelation functions for autoregressive processes of order 1 and 2.

Thus

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1\theta_{\tau+1} + \theta_2\theta_{\tau+2} + \dots + \theta_{q-\tau}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases} \quad (33)$$

Sample estimates, $\hat{\theta}$, of the population moving average parameters, θ , can be made by replacing the population autocorrelation coefficients, ρ , in Equation (33) with sample autocorrelation coefficients, $\hat{\rho}$. The estimated autocorrelation coefficient, $\hat{\rho}_1$, for a MA(1) process is

$$\hat{\rho}_1 = \frac{-\hat{\theta}_1}{1 + \hat{\theta}_1^2} \quad (34)$$

Equation (34) is quadratic in terms of $\hat{\theta}_1$ and so $\hat{\theta}_1$ may be determined from

$$\hat{\theta}_1 = \frac{1}{2\hat{\rho}_1} \pm \left[\frac{1}{(2\hat{\rho}_1)^2} - 1 \right]^{\frac{1}{2}} \quad (35)$$

Equation (34) may be expressed as

$$\hat{\theta}_1^2 + \frac{\hat{\theta}_1}{\hat{\rho}_1} + 1 = 0 \quad (36)$$

The product of the roots is unity. If $\hat{\theta}_1$ is a solution then so will $1/\hat{\theta}_1$ be a root. Furthermore, if $\hat{\theta}_1$ satisfies the invertibility condition $|\theta_1| < 1$, the absolute value of the other root $|\theta_1^{-1}|$ will be greater than one and will not satisfy the invertibility condition. Thus only the invertible root of Equation (35) is selected.

Estimates of moving average parameters $\hat{\theta}_1$ for moving average models of order q are determined by developing q equations with q unknowns from Equation (33) and solving simultaneously. The invertibility requirements must be obeyed for moving average models of all orders.

Identification of Autoregressive Models

The time series partial autocorrelation function of an autoregressive process may be used to identify the order of the autoregressive process (and model) since the partial autocorrelation function will cut off after lag p (the autoregressive process order). Once the order of an autoregressive model has been identified, the Yule-Walker equations (Yule, 1927) can be used to obtain estimates for the ϕ parameters. These equations may be developed by multiplying the expression for an autoregressive process of order p ,

$$\tilde{Z}(t) = \phi_1 \tilde{Z}(t-1) + \phi_2 \tilde{Z}(t-2) + \dots + \phi_p \tilde{Z}(t-p) + A(t)$$

by the same process lagged τ intervals, $\tilde{Z}(t-\tau)$. The product may be written

$$\begin{aligned} \tilde{Z}(t) \tilde{Z}(t-\tau) &= \phi_1 \tilde{Z}(t-1) \tilde{Z}(t-\tau) + \phi_2 \tilde{Z}(t-2) \tilde{Z}(t-\tau) + \\ &\dots + \phi_p \tilde{Z}(t-p) \tilde{Z}(t-\tau) + A(t) \tilde{Z}(t-\tau) \end{aligned} \quad (37)$$

Taking expectations for each term yields a series of autocovariance terms (recall Equation (3))

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \phi_3 \gamma_{\tau-3} + \dots + \phi_p \gamma_{\tau-p} \quad (38)$$

Dividing through by γ_0

$$\frac{\gamma_\tau}{\gamma_0} = \phi_1 \frac{\gamma_{\tau-1}}{\gamma_0} + \phi_2 \frac{\gamma_{\tau-2}}{\gamma_0} + \phi_3 \frac{\gamma_{\tau-3}}{\gamma_0} + \dots + \phi_p \frac{\gamma_{\tau-p}}{\gamma_0}$$

yields a series of autocorrelation terms (recall Equation (10)),

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \phi_3 \rho_{\tau-3} + \dots + \phi_p \rho_{\tau-p} \quad (39)$$

Equations may be developed from Equation (39) to determine the ϕ 's when the autocorrelation coefficients, ρ_τ , are known.

For a first order autoregressive model, AR(1),

$$\hat{\rho}_1 = \hat{\phi}_1$$

where $\hat{\rho}_1$ is estimated from the time series (recall $\rho_0 = \rho_{\tau-1} = 1$). For an AR(2) model,

$$\hat{\rho}_1 = \hat{\phi}_1 + \hat{\phi}_2 \hat{\rho}_{-1} = \hat{\phi}_1 + \hat{\phi}_2 \hat{\rho}_1$$

$$\hat{\rho}_2 = \hat{\phi}_1 \hat{\rho}_1 + \hat{\phi}_2$$

($\hat{\rho}_{-1} = \hat{\rho}_1$ due to the symmetry of the autocorrelation function). Solving these last two equations for $\hat{\phi}_1$ and $\hat{\phi}_2$ in terms of $\hat{\rho}_1$ and $\hat{\rho}_2$ yields

$$\hat{\phi}_1 = \frac{\hat{\rho}_1 (1 - \hat{\rho}_2)}{1 - \hat{\rho}_1^2}$$

$$\hat{\phi}_2 = \frac{\hat{\rho}_2 - \hat{\rho}_1^2}{1 - \hat{\rho}_1^2}$$

Similarly, for a p order autoregressive process, p equations involving p unknown ϕ coefficients can be determined. These equations can be solved for the ϕ coefficients once the ρ values are estimated from the time series sample.

Identification of ARMA Models

Mixed autoregressive and moving average processes (and models) of order p and q (ARMA(p, q)) are indicated by a decaying autocorrelation function and a decaying partial autocorrelation function. The ARMA processes are just combinations of the AR and MA processes and the computation of the parameters ϕ and θ from the autocorrelation function is similar to the procedures defined in the two previous sections on moving average and autoregressive models. The invertibility and stationarity requirements mentioned in the two previous sections must also be considered in this mixed process.

A more detailed discussion of parameter estimation for ARMA(p, q) models is given in Box and Jenkins (1976).

Identification of ARIMA Models

Autoregressive integrated moving average processes (and models) are suggested by an autocorrelation function which decays very slowly and a partial autocorrelation function which also decays very slowly. When nonstationarity is indicated by these slowly decaying

functions, differencing d times may transform the nonstationary series into a stationary form. Provided the differencing for stationarity is successful, an autoregressive, moving average, or mixed autoregressive moving average model is then identified for that differenced time series as previously described in this section.

A more detailed discussion of parameter estimation for ARIMA(p,d,q) models is provided in Box and Jenkins (1976).

Transfer Function Models for Bivariate Time Series Analysis

The discussion to this point has dealt primarily with univariate time series and their mathematical descriptions. In certain hydrologic problems one of these series may constitute an input to a physical system wherein the series is transformed and subsequently output as a second univariate time series. A study of the input and output series may reveal the nature of the transformation. The following sections describe some of the basic concepts involved in the analysis of two such series:

Linear Transfer Function Models

A model which approximates the relationship between a linearly related input and output time series is called a linear transfer function model. A conceptual linear system is shown in Figure 7. The linear system acts as a mechanism which takes the input, modifies it, and transfers it to an output (this mechanism is often termed a filter in time series literature). Input modifications may include (1) weighting, (2) delaying, or (3) weighting and delaying to form the output, $Y(t)$. For example, a simple transfer function model might be

$$Y(t) = W[X(t-1)]$$

wherein the output $Y(t)$ is the response to a delayed input $X(t-1)$ that is weighted by the W coefficient. In natural systems, the transfer relationship is generally more complex with



Figure 7. A linear system.

the output $Y(t)$ often a function of previous weighted output values and current and previous weighted input values. Thus a more general transfer function model is

$$\delta^r(B) Y(t) = [\omega^s(B) X(t)] B^b \quad (40)$$

where

$$\delta^r(B) = -\delta_0 B^0 - \delta_1 B^1 - \delta_2 B^2 - \dots - \delta_r B^r$$

and

$$\omega^s(B) = \omega_0 B^0 - \omega_1 B^1 - \omega_2 B^2 - \dots - \omega_s B^s$$

The coefficient δ_0 is defined as -1. The model described by Equation (40) is not unlike the process relationships of the previous section that related a univariate time series to a white noise series. The weighting coefficients, ω and δ , are analogous to the θ and ϕ weighting coefficients employed in the univariate time series description. The previously mentioned delaying action of the system is accomplished through the backstep operator B^b which effectively provides a delay of b intervals. The order of the transfer function model is defined by the integers r and s .

Equation (40) may be rearranged to obtain the "rational" form

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b \quad (41)$$

The parameter $\omega^s(B)/\delta^r(B)$ in the rational form constitutes the so-called "transfer function" which describes the bivariate relationship between the input and output time series.

A better appreciation of the physical relationship between the two univariate time series $X(t)$ and $Y(t)$ is obtained by carrying out the division as indicated by the rational form parameter, $\omega^s(B)/\delta^r(B)$. The character of the resulting equation can best be illustrated by considering an example. Suppose that $r=1$, $s=0$ and $b=3$ so that the rational form becomes

$$Y(t) = \left[\frac{\omega_0}{1 - \delta_1 B} X(t) \right] B^3$$

Suppose further that the weighting coefficients are $\omega_0 = 0.5$ and $\delta_1 = 0.5$, then

$$Y(t) = \left[\frac{0.5}{1 - 0.5B} X(t) \right] B^3$$

Dividing and carrying out the backstep operation yields the "impulse response" function model for this particular case,

$$Y(t) = 0.5 X(t-3) + 0.25 X(t-4) + 0.125 X(t-5) + \dots$$

where the output $Y(t)$ is a response to an infinite series of input impulses, $X(t)$, whose weights decay exponentially. The weights decay in this manner, and the finite series converges, when $|\delta_1| < 1$ as can be seen from carrying out the previously mentioned division. It should be noted that the example is not completely general in that the weights do not necessarily decay exponentially but rather depend upon the δ and ω values. A general expression involving the impulse response weights can be similarly obtained from Equation (41) and written as

$$Y(t) = v_0 X(t) + v_1 X(t-1) + v_2 X(t-2) + \dots$$

or

$$Y(t) = v(B) X(t) \tag{42}$$

where

$$v(B) = \sum_{j=0}^{\infty} v_j B^j$$

By definition, $v_j = 0$ for $j < b$. The collection of weights designated by $v(B)$ is termed the impulse response function. A plot of the impulse response coefficients v_j often provides an indication of the character of the transfer function in the more parsimonious rational form. Figure 8 illustrates several impulse response function plots and their corresponding rational form transfer functions for constants r , s , and b .

The sequence of terms in Equation (42) may be finite or infinite depending upon the δ values. Stability requires the sequence to be finite, or infinite and convergent. The stability requirement for bivariate time series (and transfer function models) is analogous to the stationarity requirement for univariate time series (and models). Although the terms stationarity and stability are sometimes used synonymously, the term stability will be herein used with regard to the bivariate time series (and transfer function models) only. To insure stability for a transfer function model of the form

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b$$

the roots of the characteristic equation

$$\delta^r(B) = 1 - \delta_1 B^1 - \delta_2 B^2 - \dots - \delta_r B^r = 0$$

must lie outside the unit circle. For example, given a bivariate process where $r=1$, the characteristic equation is

$$1 - \delta_1 B^1 = 0$$

and the root of this equation $B=\delta^{-1}$ must be outside of the unit circle or equivalently the absolute value of δ is less than one.

A quantity that is useful for defining the steady state response of $Y(t)$ to $X(t)$ is termed the steady state gain. The steady state gain, g , expressed in terms of the bivariate process (and model) parameters, is the sum of the response function weights

$$g = v_0 + v_1 + v_2 + \dots$$

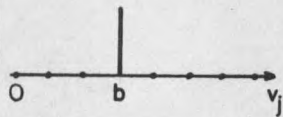
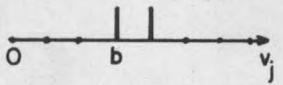
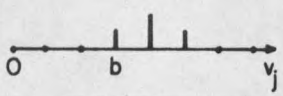
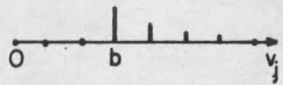
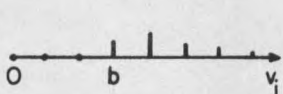
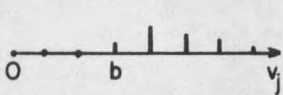
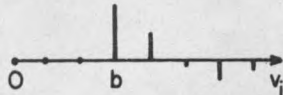
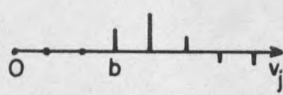
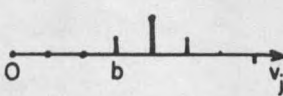
Model Order (r, s, b)	Model Form	Response Function $v(B)$
0,0,3	$Y(t) = X(t)B^3$	
0,1,3	$Y(t) = [(.5 + .5B) X(t)] B^3$	
0,2,3	$Y(t) = [(.25 + .5B + .25B^2) X(t)] B^3$	
1,0,3	$Y(t) = \left[\frac{.5}{1 - .5B} X(t) \right] B^3$	
1,1,3	$Y(t) = \left[\left(\frac{.25 + .25B}{1 - .5B} \right) X(t) \right] B^3$	
1,2,3	$Y(t) = \left[\left(\frac{.125 + .25B + .125B^2}{1 - .5B} \right) X(t) \right] B^3$	
2,0,3	$Y(t) = \left[\left(\frac{.8}{1 - .6B + .4B^2} \right) X(t) \right] B^3$	
2,1,3	$Y(t) = \left[\left(\frac{.4 + .4B}{1 - .6B + .4B^2} \right) X(t) \right] B^3$	
2,2,3	$Y(t) = \left[\left(\frac{.2 + .4B + .2B^2}{1 - .6B + .4B^2} \right) X(t) \right] B^3$	

Figure 8. Impulse response function plots and their corresponding rational form transfer functions for constants r, s, and b (modified from Box and Jenkins, 1976).

or

$$g = \sum_{j=0}^{\infty} v_j$$

Equivalently

$$g = \frac{\omega_0 - \omega_1 - \omega_2 - \dots - \omega_s}{1 - \delta_1 - \delta_2 - \dots - \delta_r} \quad (43)$$

The steady state gain may be interpreted in the following way. If an input $X(t)$ were to remain constant for an extended period, the steady state output would be a constant fraction g (the steady state gain) of the input.

The output $Y(t)$ of Equation (41) is the deterministic response to an input series, $X(t)$. A more general form of the transfer function relationship considers the output time series $Y(t)$ to be a response to both an input time series, $X(t)$, and a noise or error input series, $N(t)$,

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

The series $N(t)$ is the remaining structure of $Y(t)$ not explained by the transfer function relationship between $X(t)$ and $Y(t)$. It may be the result of measurement errors in the input and output series, the neglect of secondary factors that impact the input-output relationship, the inadequacy of the linear bivariate process approximation, or some combination of these things. $N(t)$ may be characterized as a moving average, an autoregressive, a mixed autoregressive moving average, or an autoregressive integrated moving average process as previously discussed.

The input $X(t)$ and the output $Y(t)$ may exhibit deterministic (and thus nonstationary) and/or random characteristics. Unlike the previously discussed models for univariate processes whose applicabilities were limited to stationary time series data sets (or data sets

transformable to stationary sets), transfer function models may be utilized to identify relationships between pairs of univariate time series whether the individual univariate series are stationary or not.

Cross Correlation Function

The cross correlation function for two time series is analogous to the previously discussed autocorrelation function for single time series. The primary difference between these functions is that the autocorrelation function is used to identify the relationships within a time series (univariate analysis), whereas, the cross correlation function is used to identify the relationships between two separate time series (bivariate analysis). The cross correlation function may be helpful in identifying a tentative transfer function model form (Equation (41)).

The cross correlation function is defined in terms of the cross covariance statistic. The cross covariance which is analogous to the autocovariance (Equation (3)) may be determined from

$$\gamma_{\tau_{xy}} = E[\{X(t) - \mu_x\} \{Y(t+\tau) - \mu_y\}] \quad (45)$$

The variances of $X(t)$ and $Y(t)$ are from Equation (2)

$$\begin{aligned} \sigma_x^2 &= \gamma_{0_{xx}} = E[\{X(t) - \mu_x\}^2] \\ \sigma_y^2 &= \gamma_{0_{yy}} = E[\{Y(t) - \mu_y\}^2] \end{aligned} \quad (46)$$

where μ_x and μ_y are the mean values of the $X(t)$ and $Y(t)$ time series respectively. The cross correlation coefficient between $X(t)$ and $Y(t)$ is analogous to the autocorrelation coefficient and is defined as

$$\rho_{\tau_{xy}} = \frac{\gamma_{\tau_{xy}}}{\sigma_x \sigma_y} \quad (47)$$

An estimate of the cross covariance at lag τ is

$$\frac{1}{N} \sum_{t=1}^{N-\tau} \{x(t) - \bar{x}\} \{y(t+\tau) - \bar{y}\}, \quad \tau = 0, 1, 2, \dots \quad (48)$$

$$\hat{\gamma}_{xy}(\tau) = \frac{1}{N} \sum_{t=1}^{N+\tau} \{y(t) - \bar{y}\} \{x(t-\tau) - \bar{x}\}, \quad \tau = 0, -1, -2, \dots \quad (49)$$

where N is the sample size, and $\bar{x}$ and $\bar{y}$ are the means of the sample $x(t)$ and $y(t)$ time series. An estimate of the variance $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ may be obtained from Equation (5)

$$\hat{\sigma}_x^2 = \hat{\gamma}_{0xx} = \frac{1}{N} \sum_{t=1}^N (x(t) - \bar{x})^2 \quad (50)$$

$$\hat{\sigma}_y^2 = \hat{\gamma}_{0yy} = \frac{1}{N} \sum_{t=1}^N (y(t) - \bar{y})^2$$

The estimated cross correlation coefficient at lag τ is

$$\hat{\rho}_{\tau xy} = \frac{\hat{\gamma}_{\tau xy}}{\hat{\sigma}_x \hat{\sigma}_y} \quad (51)$$

The cross correlation function is often presented as a graphical collection of cross correlation coefficients for various τ lags. Unlike the autocorrelation function, the cross correlation function is not necessarily symmetric about lag zero. For most real physical processes, the negative lag cross correlation function is of no concern since the present output cannot respond to future inputs. Therefore, the remaining discussion will be concerned with the positive cross correlation function only.

Identification of Transfer Function Models

A number of transfer function model forms that are useful for approximating bivariate time series processes were presented in Figure 8. Those model forms provide parsimoni-

ous representations for many linear bivariate processes. The following paragraphs discuss model identification procedures pertinent to transfer function models.

A cross correlation function determined from system input and output time series samples is often useful in the preliminary identification of an appropriate transfer function model (Equation (41)). The value of the cross correlation function as a preliminary identifier stems from its relationship to the impulse function $v(B)$ which is in turn related to the δ and ω coefficients of the transfer function in Equation (41). Model identification is practical only if the input series is whitened and the output series is filtered prior to cross correlation function determination (Box and Jenkins, 1976). The whitening and filtering are accomplished by first characterizing the input time series with an appropriate ARIMA-(p,d,q) model. Recall that the ARIMA model is general and reduces to an MA, AR, or ARMA model when one of the model orders (p,d,q) or combination of orders is zero. An ARIMA model (Equation (32)) for the input series may be written

$$X(t) = \frac{\theta_x^q(B) \alpha(t)}{\phi_x^p(B) \nabla^d} \quad (52)$$

where $\alpha(t)$ denotes a white noise input series and the x subscripts identify the θ and ϕ coefficients as input series, $X(t)$, model coefficients. By rearranging Equation (52), a white noise series, or a "whitened" time series, $\alpha(t)$, is obtained from the input series $X(t)$,

$$\alpha(t) = \frac{\phi_x^p(B) \nabla^d}{\theta_x^q(B)} X(t) \quad (53)$$

In whitening an ARIMA process time series, the moving average, autoregressive and nonstationary character of the time series is essentially removed.

A "filtered" output series $\beta(t)$ is obtained by using the input whitening parameter, $\phi_x^p(B) \nabla^d / \theta_x^q(B)$, to operate on the output time series $Y(t)$,

$$\beta(t) = \frac{\phi_X^p(B) \nabla^d}{\theta_X^q(B)} Y(t) \quad (54)$$

This operation filters out the moving average, autoregressive, and nonstationary character of the output series that is attributable to the input series. The output series of Equation (54), $Y(t)$, may also be expressed as a response to the input series, $X(t)$. Recall the impulse response function form of a bivariate process wherein

$$Y(t) = v(B) X(t) \quad (42)$$

Adding a noise component, $N(t)$, to account for the $Y(t)$ structure not attributable to $X(t)$ (see previous discussion), Equation (42) becomes

$$Y(t) = v(B) X(t) + N(t) \quad (55)$$

Substitution of Equation (55) into Equation (54) yields

$$\beta(t) = \frac{\phi_X^p(B) \nabla^d}{\theta_X^q(B)} [v(B) X(t) + N(t)]$$

Substitution of Equation (53) into this equation gives

$$\beta(t) = v(B) \alpha(t) + \frac{\phi_X^p(B) \nabla^d}{\theta_X^q(B)} N(t) \quad (56)$$

Multiplying Equation (56) by $\alpha(t+\tau)$ and taking expectations produces the cross covariance at lag τ between the whitened input series $\alpha(t)$ and the filtered output series $\beta(t)$.

$$\gamma_{\tau\alpha\beta} = E[\alpha(t-\tau)\beta(t)] = E\left\{\alpha(t-\tau)\left[v(B)\alpha(t) + \frac{\phi_X^p(B) \nabla^d}{\theta_X^q(B)} N(t)\right]\right\}$$

Assuming that $N(t)$ and $X(t)$ are uncorrelated and thus that their cross covariance values for all τ are zero

$$\gamma_{\tau\alpha\beta} = E\left\{\alpha(t-\tau)[v(B)\alpha(t)]\right\} = E\left\{\alpha(t-\tau)[v_0\alpha(t) + v_1\alpha(t-1) + v_2\alpha(t-2) + \dots]\right\}$$

Expanding the right side of this equation results in a series of autocovariance terms for the white noise series $\alpha(t)$. Since the autocovariance of a white noise process is zero except for lag zero,

$$\gamma_{\tau\alpha\beta} = E \left\{ \alpha(t-\tau) v_{\tau} \alpha(t-\tau) \right\} = v_{\tau} \gamma_0 = v_{\tau} \sigma_{\alpha}^2 \quad (57)$$

Rearranging Equation (57) and multiplying the numerator and denominator by σ_{β} yields

$$v_{\tau} = \frac{\gamma_{\tau\alpha\beta}}{\sigma_{\alpha}^2} \frac{\sigma_{\beta}}{\sigma_{\beta}} = \frac{\gamma_{\tau\alpha\beta}}{\sigma_{\alpha}\sigma_{\beta}} \frac{\sigma_{\beta}}{\sigma_{\alpha}} = \rho_{\tau\alpha\beta} \frac{\sigma_{\beta}}{\sigma_{\alpha}} \quad (58)$$

Thus the previously mentioned relationship between the impulse response coefficients v_{τ} and the cross correlation coefficients, $\rho_{\tau\alpha\beta}$, is established.

The impulse response coefficients v_{τ} suggested by the cross correlation coefficient relationship of Equation (58) are in turn related to the δ and ω coefficients of the transfer function model (Equation (41)). Recall again the impulse response function model form for a bivariate process,

$$Y(t) = v(B) X(t) = \sum_{j=0}^{\infty} v_j B^j X(t) \quad (42)$$

and the transfer function model form for a bivariate process

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b \quad (41)$$

The identity formed by equating the model forms of Equations (42) and (41)

$$\sum_{j=0}^{\infty} v_j B^j X(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b$$

may be rearranged to form

$$\delta^r(B) \sum_{j=0}^{\infty} v_j B^j X(t) = [\omega^s(B) X(t)] B^b$$

Recalling that

$$\delta^T(B) = 1 - \delta_1 B^1 - \delta_2 B^2 - \dots - \delta_r B^r$$

$$\sum_{j=0}^{\infty} v_j B^j = v_0 + v_1 B^1 + v_2 B^2 + \dots$$

and

$$\omega^S(B) = \omega_0 - \omega_1 B^1 - \omega_2 B^2 - \dots - \omega_s B^s$$

the equation may be written

$$(1 - \delta_1 B^1 - \delta_2 B^2 - \dots - \delta_r B^r) (v_0 + v_1 B^1 + v_2 B^2 + \dots) =$$

$$(\omega_0 - \omega_1 B^1 - \omega_2 B^2 - \dots - \omega_s B^s) B^b$$

Expanding and equating B^j operators gives rise to the following series of equations.

$$\begin{aligned} v_j &= 0 & j < b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + \omega_0 & j = b \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} - \omega_{j-b} & j = b+1, b+2, \dots, b+s \\ v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} & j > b+s \end{aligned} \quad (59)$$

The simultaneous solution of these equations establishes a relationship between the v coefficients and the δ and ω coefficients. This relationship plus the previously developed relationship between the v coefficients and the cross correlation function provide the mechanism by which the cross correlation function can be used to identify δ and ω coefficients for a transfer function model. Impulse response function coefficients, v_{τ} , are estimated from Equation (58) wherein the cross correlation coefficient, $\rho_{\tau_{\alpha\beta}}$, and variances, σ_{α} and σ_{β} , are sample estimates from the whitened input and filtered output. The v_{τ} coefficients thus obtained are substituted into the appropriate expressions of Equation (59) and the resulting relationships are solved simultaneously for the δ and ω values.

In addition to estimating δ and ω values, the cross correlation coefficients may be useful in estimating transfer function model order, r and s , and lag b . The cross correlation function (the graphical plot of the collection of cross correlation coefficients) ideally corresponds to the shape of the impulse response function as evidenced by Equation (58). The shape of the cross correlation function thus suggests the shape of the impulse response function which in turn may be compared qualitatively to transfer functions for models of known order and lag. Recall that Figure 8 illustrates some characteristic impulse response function shapes for certain common transfer function orders and lags.

Cross correlation functions may also be developed from the unmodified input and output time series (unwhitened input and unfiltered output) or from differenced input and output series. These cross correlation functions do not provide reliable evidence of the form of the impulse response function (and therefore the order of the transfer function) but may be useful in certain circumstances for identifying the lag parameter, b , of the backstep operator, B^b .

Model Diagnostics

An important part of statistical modeling is the application of diagnostic checks. The autocorrelation, partial autocorrelation, and cross correlation functions are not only used in time series model identification, but are also used in testing the adequacy of the estimated model. The discussion that follows considers the diagnostics used in transfer function model analysis.

With reference to Equation (44),

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

if the estimated transfer function component and the estimated noise component adequately explain the process output, then the difference between the estimated and observed

output values, the residuals, $a(t)$, should be white noise. If the residuals do not constitute white noise, the model is inadequate in that the residual series exhibits a significant dependence in its structure. The nature of this dependence may be detected by utilizing the autocorrelation, partial autocorrelation, and cross correlation functions. The autocorrelation and partial autocorrelation functions of the residuals, and the cross correlation function between the input and residuals are determined and plotted. If the plots reveal no significant autocorrelation, partial autocorrelation or cross correlation coefficients, the residuals constitute white noise and the model is adequate. If significant residual autocorrelation and/or partial autocorrelation coefficients exist with no significant cross correlation coefficients, then the model inadequacy is likely in the noise component of Equation (44). If significant cross correlation coefficients between the input and residuals exist, then significant autocorrelations and/or partial autocorrelations will usually be present (Box and Jenkins, 1976). A significant cross correlation coefficient indicates an inadequacy in the transfer function component. The associated significant autocorrelation and/or partial autocorrelation coefficients may or may not indicate an inadequacy in the noise component.

There are two common methods of testing the residuals to assess the adequacy of the transfer function model. One method employs a "t-like" statistic. The other employs a chi-square test of fit (the portmanteau test, Box and Pierce, 1970). The "t-like" statistic test will be discussed first.

The significance of a residual series autocorrelation coefficient, $\hat{\rho}_T$, may be determined with the so-called "t-like" statistic

$$t_{\hat{\rho}_T} = \frac{\hat{\rho}_T - \rho_T}{S_{\hat{\rho}_T}}$$

where $\hat{\rho}_\tau$ is the sample autocorrelation coefficient for lag τ , ρ_τ is the mean of the sampling distribution of the autocorrelation coefficients, and $S_{\hat{\rho}_\tau}$ is the sample estimate of the standard deviation (standard error) of the sampling distribution of the autocorrelation coefficients. It is assumed that the residuals are normally distributed. Since the mean of the sampling distribution of the autocorrelation coefficients is zero for a white noise process,

$$t_{\hat{\rho}_\tau} = \frac{\hat{\rho}_\tau}{S_{\hat{\rho}_\tau}} \quad (60)$$

where the standard deviation, $S_{\hat{\rho}_\tau}$, may be obtained from the Bartlett approximation (Anderson, 1976),

$$S_{\hat{\rho}_\tau} = \left[\frac{1}{N} \left(1 + 2 \sum_{j=1}^{\tau-1} \hat{\rho}_j^2 \right) \right]^{\frac{1}{2}}$$

which reduces to

$$S_{\hat{\rho}_\tau} = \left[\frac{1}{N} \right]^{\frac{1}{2}}$$

for a white noise process. The distribution of the "t-like" statistic is similar to the "student t" distribution and, like the "student t" approaches a normal distribution for large sample size. Because of this similarity, the "t-like" statistic is often used as a t statistic (Bowerman and O'Connell, 1979). Use of the "t-like" statistic in this manner presumes that the autocorrelation coefficients of the residual white noise series are normally distributed about zero. The number of effective observations for the residual series (and also in this case, the degrees of freedom ν) is $n = N - d - p - s - b$ where N is the number of observations. Recall that d defines the number of differencing operations on the data, p defines the order of the autoregressive $\phi(B)$, s defines the order of the transfer function parameter $\omega(B)$, and b describes the order of the backstep operator B^b .

The "t-like" statistic for the residual series partial autocorrelation coefficient is similarly defined as

$$t_{\hat{\rho}_{\tau\tau}} = \frac{\hat{\rho}_{\tau\tau}}{1/\sqrt{n}} \quad (61)$$

where the standard error, $1/\sqrt{n}$, is based upon the Quenouille approximation (Anderson, 1976) and n is as previously defined for the autocorrelation coefficient.

The "t-like" statistic for the cross correlation coefficient between the input and the residuals $a(t)$ is

$$t_{\hat{\rho}_{\tau_{xa}}} = \hat{\rho}_{\tau_{xa}} / S_{\hat{\rho}_{\tau_{xa}}} \quad (62)$$

where the standard error $S_{\hat{\rho}_{\tau_{xa}}}$ (also based upon a Bartlett approximation) is

$$S_{\hat{\rho}_{\tau_{xa}}} = \frac{1}{\sqrt{m}}$$

and $m = N - \tau - d - p - s - b$ is the number of effective observations (and also in this case, the degrees of freedom ν).

A rough rule of thumb often used to gage the significance of these autocorrelation, partial autocorrelation, and cross correlation coefficients is that if

$$|t_{\hat{\rho}_{\tau}}| = \left| \frac{\hat{\rho}_{\tau}}{S_{\hat{\rho}_{\tau}}} \right| \leq 2, \text{ then } \hat{\rho}_{\tau} = 0,$$

and if

$$|t_{\hat{\rho}_{\tau\tau}}| = \left| \frac{\hat{\rho}_{\tau\tau}}{1/\sqrt{n}} \right| \leq 2, \text{ then } \hat{\rho}_{\tau\tau} = 0,$$

and if

$$|t_{\hat{\rho}_{\tau_{xa}}} = \left| \frac{\hat{\rho}_{\tau_{xa}}}{S_{\hat{\rho}_{\tau_{xa}}}} \right| \leq 2, \text{ then } \hat{\rho}_{\tau_{xa}} = 0, \quad (63)$$

This rule of thumb is based upon the fact that the $t(.05, \nu)$ value (95 percent significance level for a two-tailed table) is approximately 2 for 50 degrees of freedom. This corresponds to the number of observations (and approximately the number of effective observations or degrees of freedom recommended by Box and Jenkins for transfer function analysis).

As previously mentioned, a second test is commonly employed in judging model inadequacy. It is the "portmanteau" lack of fit test which gages model inadequacy by testing the first K autocorrelation coefficients as a group, and the first M cross correlation coefficients as a group. The statistic

$$Q = \frac{(\hat{\rho}_1 - \mu_\rho)^2 + (\hat{\rho}_2 - \mu_\rho)^2 + \dots + (\hat{\rho}_\tau - \mu_\rho)^2 + \dots + (\hat{\rho}_K - \mu_\rho)^2}{\sigma^2}$$

has a chi-square distribution if the autocorrelation coefficients, $\hat{\rho}_\tau$, from the white noise residual series are normally distributed. Since the mean, μ_ρ , of a white noise process is zero by definition,

$$Q = \frac{\hat{\rho}_1^2 + \hat{\rho}_2^2 + \dots + \hat{\rho}_\tau^2 + \dots + \hat{\rho}_K^2}{\sigma^2} = \frac{\sum_{\tau=1}^K \hat{\rho}_\tau^2}{\sigma^2}$$

has a chi-square distribution also. A sample estimate of the variance, σ^2 , of the autocorrelation coefficient population is provided by the previously mentioned Bartlett approximation,

$$S^2 = \frac{1}{n}$$

where $n = N-d-s-b-p$ is the number of effective observations. Substitution into the previous expressions yields the statistic,

$$Q = n \sum_{\tau=1}^K \hat{\rho}_{\tau}^2 \quad (64)$$

which is used in the "portmanteau" lack of fit test. The statistic has a chi-square distribution with $\nu = K-p-q$ degrees of freedom. Recall that q is the order of the moving average operator, $\theta(B)$. For the first M cross correlation coefficients, the "portmanteau" test statistic can be similarly defined as

$$S = n \sum_{\tau=0}^M \hat{\rho}_{\tau_{\bar{x}a}}^2 \quad (65)$$

where $n = N-d-s-b-p$ is the number of effective of observations. The S statistic also has an approximate chi-square distribution with $\nu = M-r-s$ degrees of freedom. Recall that r is the order of the transfer function parameter, $\delta(B)$. The model is not considered inadequate if the Q and S statistics are less than the chi-square significance points for the appropriate ν degree of freedom.

In summary, a significant residual series autocorrelation structure or cross correlation structure, as evidenced by the "t-like" or "portmanteau" statistics, indicates that the model is inadequate.

SELECTED STUDY SITE

The application of the time series methodology to the irrigation diversion-return flow problem requires an appropriate area possessing sufficient data. A reach of the Beaverhead River near Dillon in Beaverhead County, Montana was selected for this study (Figure 9). Although certain data deficiencies arose during the course of the study, surface and groundwater data for the area are relatively abundant.

Study Area Description

Some of the assumptions employed in the irrigation diversion-return flow analysis to be described later are dependent upon the geohydrologic features existing at the study area. Therefore, a description of the study area climate, surface water, and groundwater systems follows.

Climate

A semiarid climate dominates the area. The average annual rainfall at the Dillon Airport weather station measures 9.5 inches. The average annual temperature for January is 20°F and for July it is 66°F. The May through September average temperature is 59°F.

Surface Water

Several stream gaging stations (refer to Figure 10) provide discharge information for the Beaverhead River near Dillon. There are five gaging stations between the Clark Canyon Reservoir and Twin Bridges that have been active at various times. The tributary discharges of Blacktail Deer Creek and Rattlesnake Creek have also been monitored periodically in

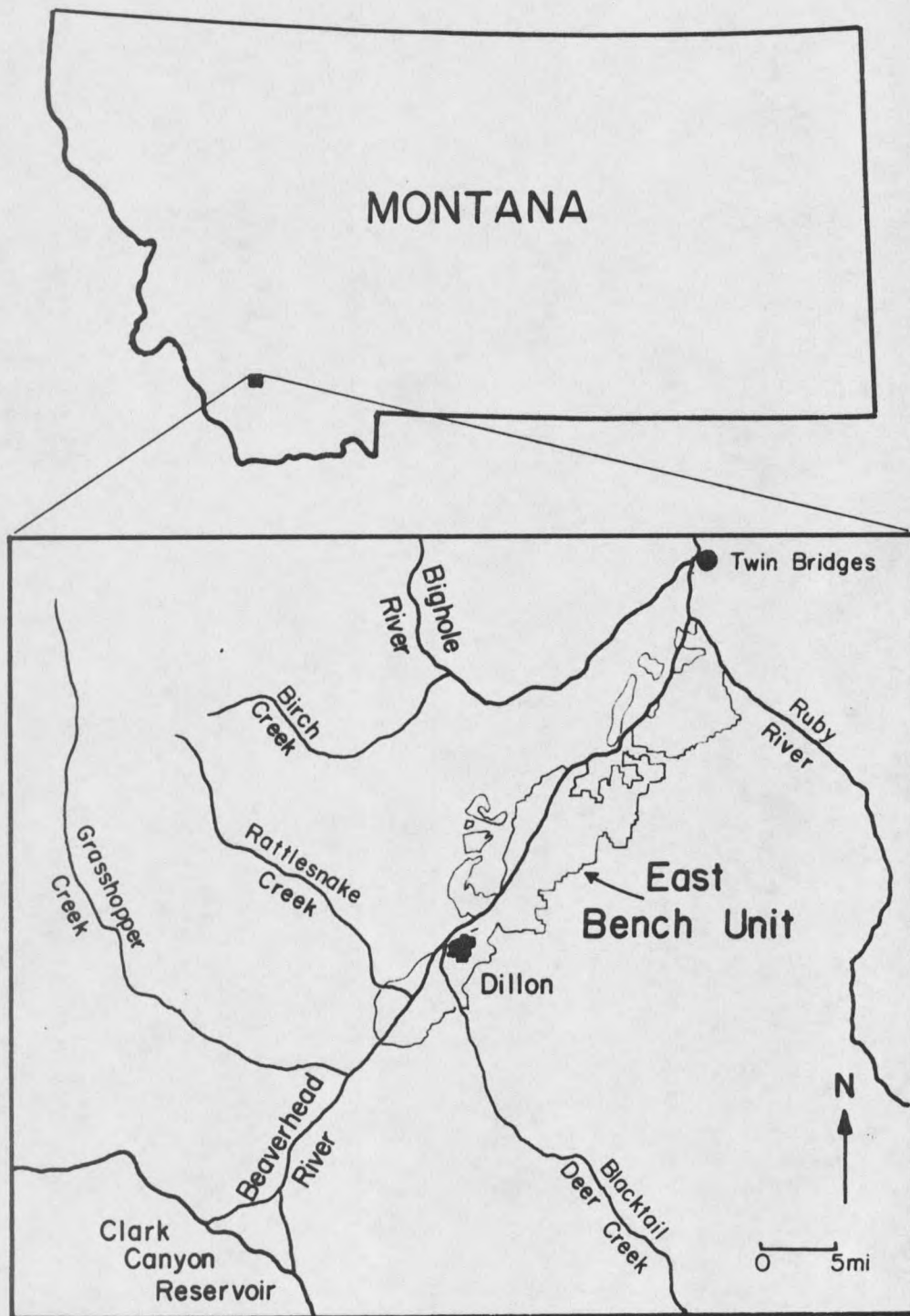


Figure 9. The Beaverhead Valley between the Clark Canyon Reservoir and Twin Bridges (after Nicklin and Brustkern, 1981).

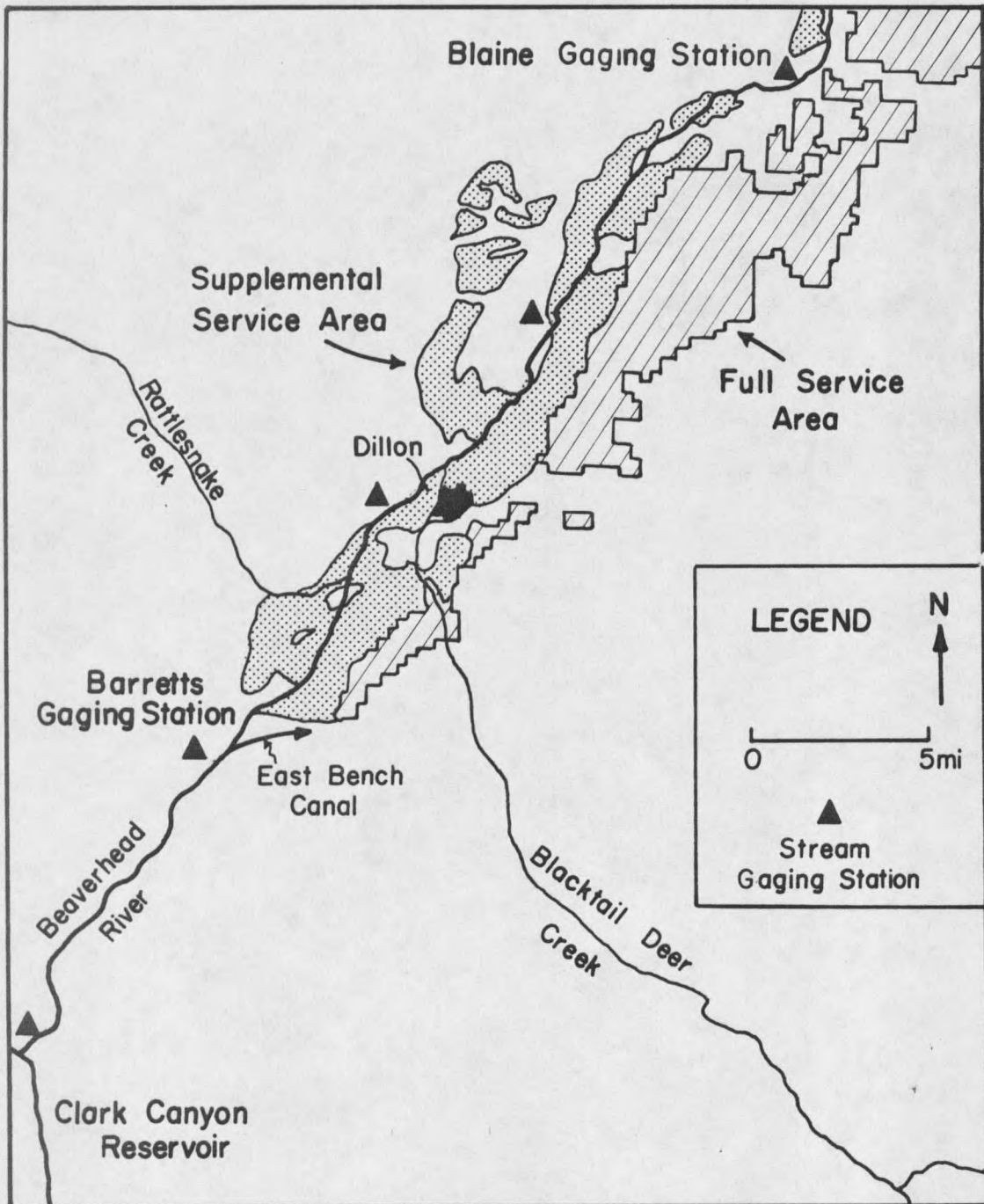


Figure 10. The East Bench Unit (after Nicklin and Brustkern, 1981).

the past. More importantly, the diversions in two of the major irrigation canals, the East Bench Canal and the Clark Canyon Canal, are monitored (Figure 11).

The Clark Canyon Reservoir was completed in 1965 as a multipurpose facility. Part of its function is to supply irrigation water for the East Bench Unit (which consists of the full service area and supplemental service area shown in Figure 10). Full service water is supplied to 21,800 acres of benchland east of Dillon and supplemental water is supplied to 28,000 acres of valley land. Only a portion of this 49,800 acre total, approximately 33,400 acres, lies inside the study area (Figure 11). Tributary inflows irrigate an additional 8,000 acres within the study area. Thus, a total of approximately 41,400 acres of bench and bottom land constitute the irrigated acreage within the study area.

Groundwater

The two geohydrologic units of primary concern for this project are the consolidated Tertiary sedimentary deposits and the unconsolidated Quaternary alluvial deposits (refer to Figure 12). The unconsolidated alluvial deposits comprise the floodplain and consist of coarse sand and gravel intermixed with scattered lenses of silt and clay. The underlying Tertiary sedimentary unit is made up of finer material including lenticular layers of clay, silt, fine to coarse sand, and some gravel. The thickness of the alluvium is approximately 200 feet. The Tertiary unit's thickness is not documented, but is estimated to exceed 1000 feet (Betz, 1967).

The alluvial portion of the geohydrologic system is highly permeable as evidenced by the reported transmissivity of 660,000 gpd/ft for the Dillon city well. Though less permeable, the Tertiary system underlying the Bench areas provides sufficient water for domestic and stock purposes (Betz, 1967).

Although a few irrigation wells exist in the study area, their impact on the hydrologic system is believed to be minimal. Most of the water development in the area relies upon

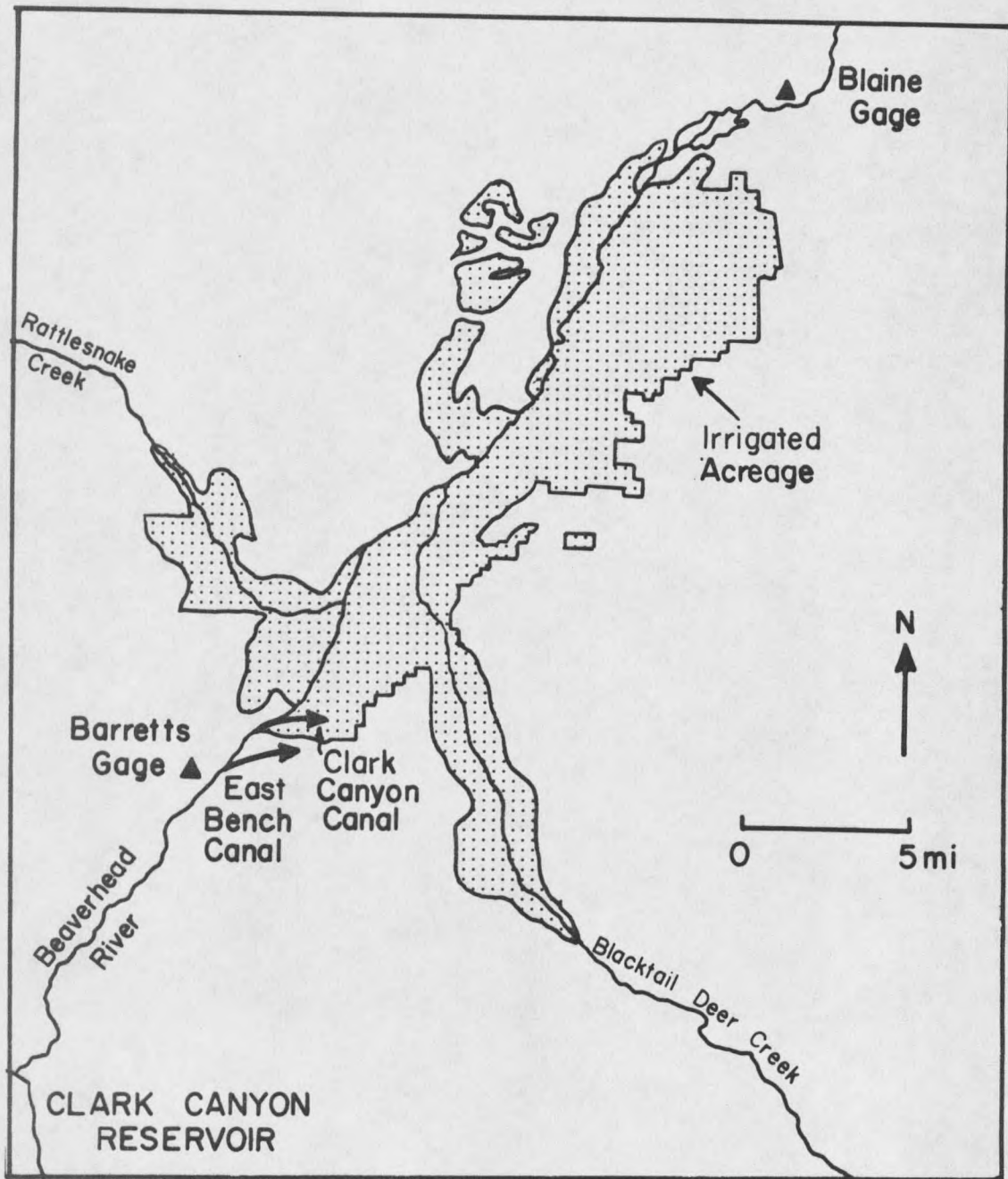


Figure 11. The study area.

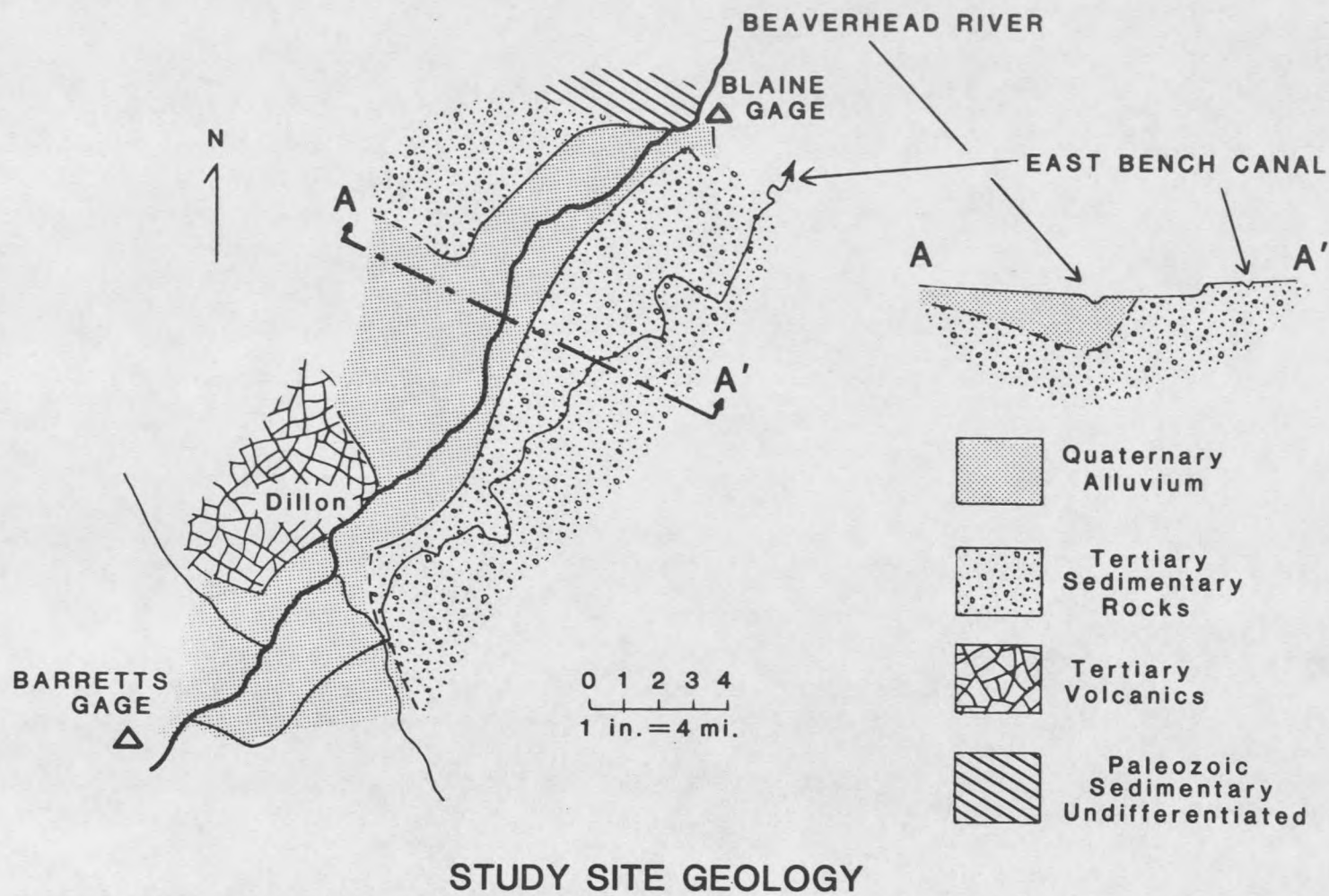


Figure 12. A map of the geologic units and a conceptual cross-section of the geologic units along the Beaverhead River between the Barretts and Blaine gaging stations (modified from Ross, 1955).

the relatively abundant surface water supplies. This limited groundwater development reduces the complexity of the geohydrologic system and enhances the prospects of assessing the magnitude and timing of irrigation return flow quantities that reach the Beaverhead River via the groundwater system.

The groundwater hydraulic gradient indicates that the Beaverhead River is an effluent stream (refer to Figure 13). This fact provides support for a later assumption that the unconsumed diversion eventually returns to the channel and becomes part of the surface discharge in the studied reach of the Beaverhead River.

The Beaverhead hydrologic system between the Barretts and Blaine (published as Near Twin Bridges) gaging stations is schematically illustrated in Figure 14. The figure shows how excess irrigation applications contribute to outflow via the return flow process. The spatial and temporal distribution of irrigation return flow is dependent upon irrigation activity and the return system (the aquifer system and, to a much smaller degree, the surface return flow system). The time distributions of diversion and return flow constitute two time series that are linked by this return flow system. This study attempts to link these two time series through a bivariate transfer function model. Reliable and adequate hydrologic data are obviously a prerequisite to the successful development of such a model.

Study Site Data

The Beaverhead system is of course much more complex than the schematic of Figure 14 indicates and consequently it is necessary to consider more hydrologic measurements than are suggested by the figure.

The surface inflow $I(t)$ of Figure 14 consists of Beaverhead inflow at the Barretts gage plus the surface runoff and tributary inflow between the Barretts and Blaine gages. Daily discharge records at the Barretts gage were obtained from the United States Geologi-

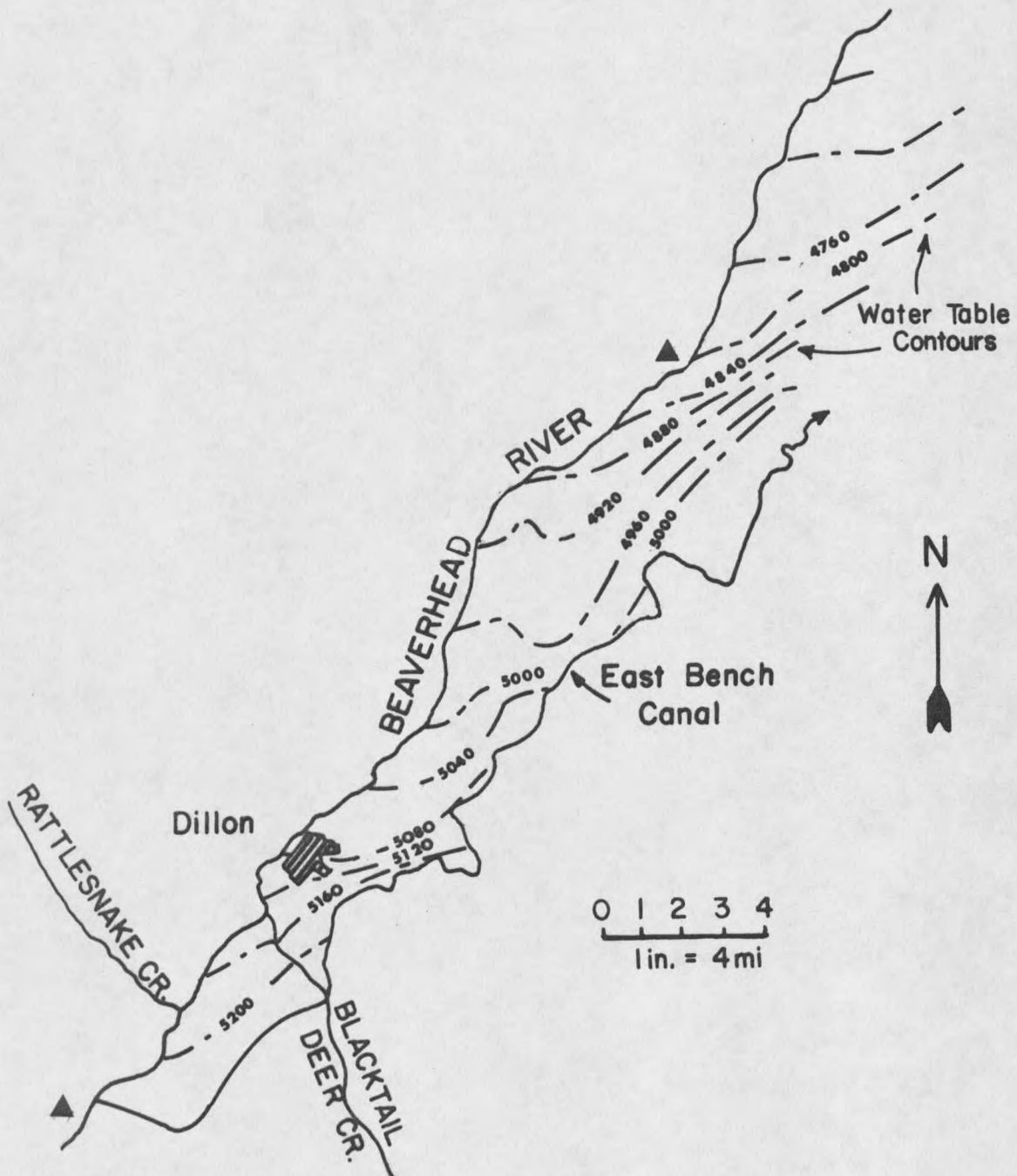


Figure 13. Water table contours of the hydrogeologic units east of the Beaverhead River near Dillon, Mt. (modified from Botz, 1967).

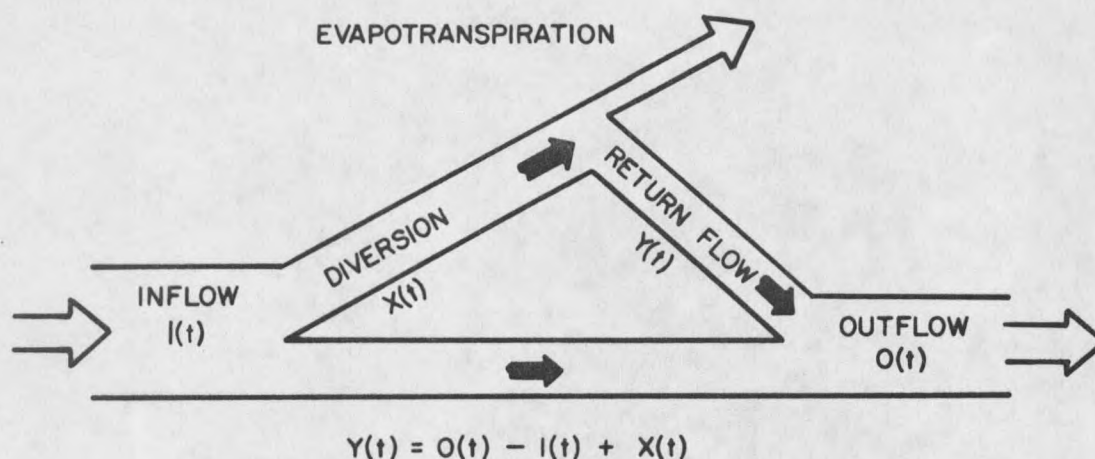


Figure 14. A conceptual representation of the irrigation and return flow process.

cal Survey (USGS) Water Supply Papers for the period of interest. Natural surface runoff was assumed to be negligible in view of the relatively small study area (Figure 11) and limited precipitation. Records (USGS) of tributary inflow for Blacktail Deer Creek and Rattlesnake Creek, the two significant tributaries within the study reach, were used where available. Missing tributary inflows were estimated from linear regression relationships that were established between each of the tributaries and nearby Birch Creek whose record spans the period of interest. Birch Creek, a tributary to the Big Hole River, is hydrologically very similar to the Blacktail Deer and Rattlesnake tributaries (see further comments, Nicklin and Brustkern, 1983).

The surface outflow $O(t)$ of Figure 14 includes Beaverhead outflow at the Blaine gage and irrigation diversion bypass (irrigation flow diverted within but applied below the study site). Daily discharge records for the Blaine gage were obtained from the USGS for the period of record. Irrigation diversion bypass was estimated by proportioning the known total diversion on the basis of the irrigated areas served within and without the study area.

The diversion $X(t)$ of Figure 14 includes all diversion within the study reach. Diversion records for the major canals, the East Bench and Clark Canyon Canals (Figure 11), were

obtained from the East Bench Irrigation District Office. Records for certain other diversions within the reach, though not continuous, were also obtained from this office. Incomplete or missing diversion records were estimated by assuming that irrigation application rates were uniform over the study area, and thus unmeasured diversion could be determined from measured diversion by applying water on the same per unit area basis. Return flow $Y(t)$ in Figure 14 was computed from the balance equation for the system

$$Y(t) = O(t) - I(t) + X(t) \quad (66)$$

A more detailed discussion of the time series data base, its development and limitations, is presented by Nicklin and Brustkern (1983).

Preliminary Data Analysis

In an attempt to gain a better physical understanding of the area's hydrologic processes, and in particular the return flow process, certain preliminary data analyses were carried out. The results of these analyses are described in the following paragraphs.

A set of hydrographs were developed to show the study area inflow and outflow (measured and estimated) for the years 1966 through 1975 (Figure 15). As previously described, the inflow hydrograph, $I(t)$, represents total inflow to the study area and the outflow hydrograph, $O(t)$, represents the total outflow from the area. A second set of hydrographs showing total diversion (measured and estimated) and return flow (computed from balance equation) was also determined and plotted (Figure 16). The figure shows some inconsistencies (such as negative or anomalously small quantities) in the estimated return discharge at the beginning and end of some seasons. These apparent inconsistencies are very probably related to the fact that diversion is sometimes monitored less rigorously at the season's beginning when the water supply is plentiful and at the season's end when the demand is small. Less rigorous diversion accounting results in return flow anomalies

since computed return flow is dependent in part on irrigation diversion through Equation (66). The significance of the diversion and return flow activities is apparent in both Figures 15 and 16. For example, a comparison of the inflow and outflow rates during the irrigation and nonirrigation seasons reveals an outflow hydrograph that is markedly different from the inflow hydrograph as a result of these activities.

The mean annual diversion volume (measured and estimated) along the stream reach between the Barretts and Blaine stations for the years 1966 through 1975 was 130,970 cfs-days. The mean irrigation season inflow and outflow volumes for the same years were respectively 132,940 and 72,840 cfs-days. Since the mean annual diversion volume was approximately equal to the irrigation season mean inflow volume to the area, it is obvious that significant return flow to the stream occurred during the irrigation season. Furthermore, only a small percentage, less than 5 percent of the return, is believed to be surface return (this percentage is based upon records made available from the East Bench Irrigation District by Dick Kennedy, 1981). Thus, groundwater return flow is the major component of flow in the Beaverhead River reach during the irrigation season.

Figure 17 is a conceptual representation of the various components of the area flow system. The figure shows the mean weekly inflow and mean weekly outflow for the years 1966 through 1975. It is assumed that irrigation season effective precipitation constitutes an insignificant source of water for the study area (for the years of interest, the mean seasonal precipitation between the dates of May 1 and September 30 was approximately 6.6 inches at the Dillon Airport). It is also assumed that within the limited reach under consideration, groundwater recharge to the Beaverhead River from sources other than irrigation return flow is insignificant. It is believed that these assumptions are reasonable for the analysis herein described.

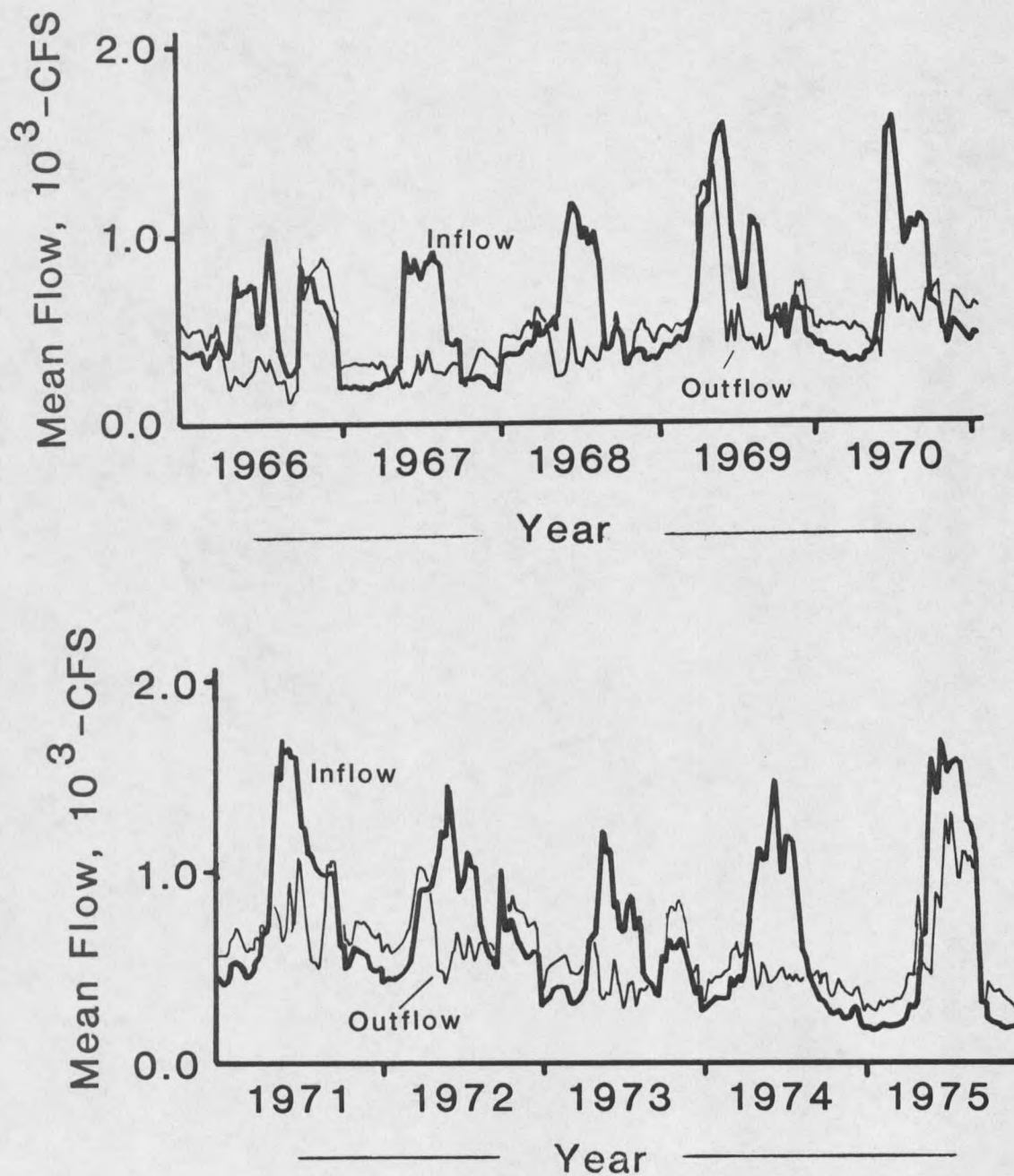


Figure 15. The inflow and outflow hydrographs for the years 1966 through 1975 for the Beaverhead River between the Barretts and Blaine gaging stations (mean weekly flows).

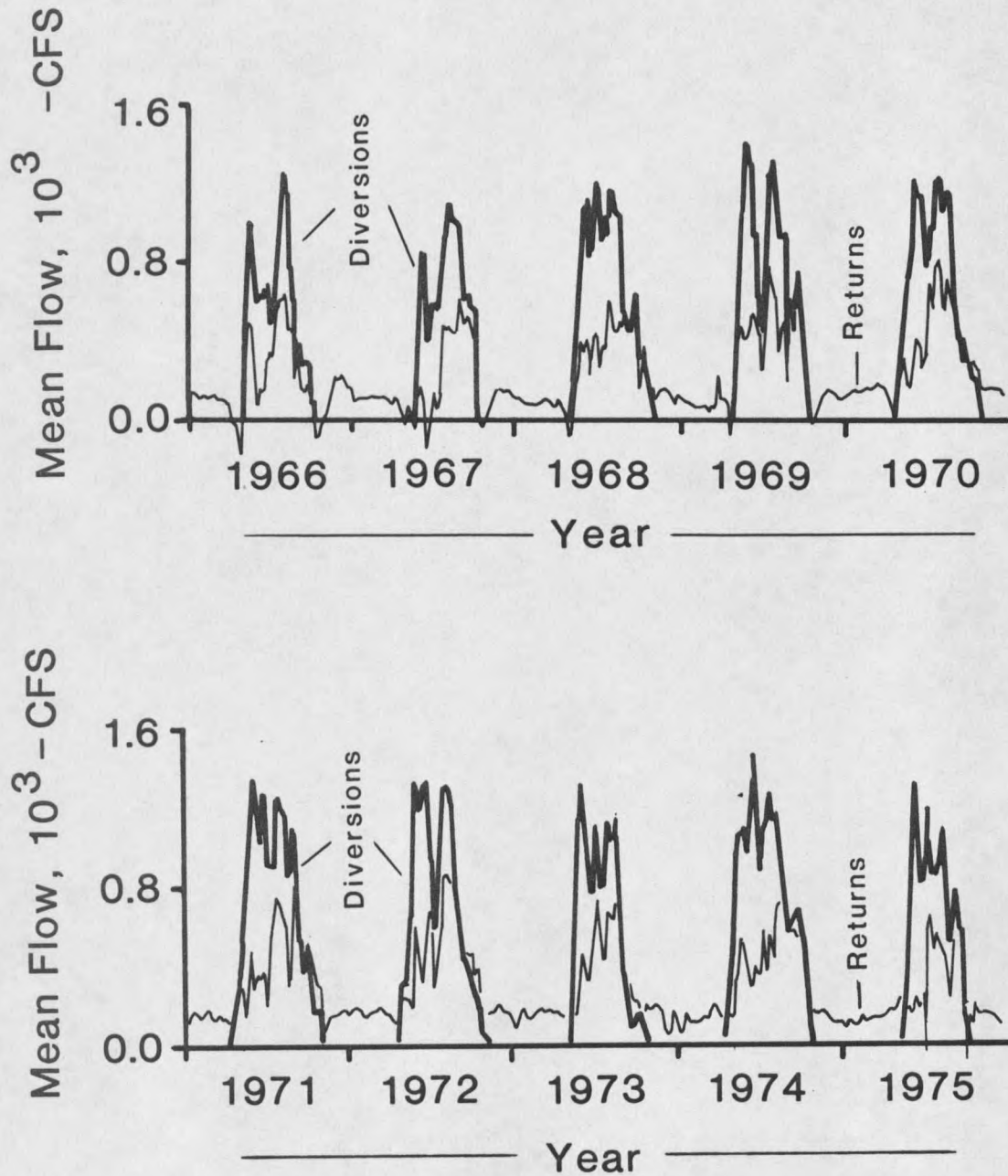


Figure 16. Diversion and return flow hydrographs for the Beaverhead River reach between the Barretts and Blaine gaging stations during the years 1966 through 1975 (mean weekly flow).

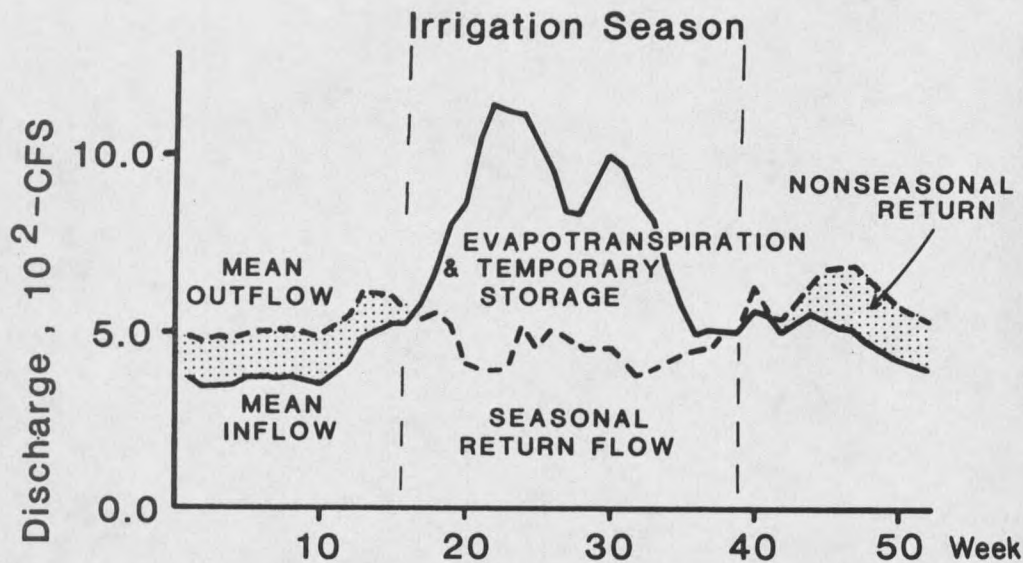


Figure 17. Mean weekly inflow and outflow hydrographs for the years of 1966 through 1975.

The inflow and outflow hydrographs of Figure 17 can be separated into a number of components. The total area below the inflow hydrograph for the irrigation season represents that season's total inflow volume to the area. This area also approximates total seasonal diversion as was previously pointed out. Thus the area below the irrigation season outflow hydrograph represents seasonal return flow volume (including both groundwater return flow and the much less significant surface return flow). The importance of return flow in maintaining irrigation season discharge in the study reach is apparent. The area between the hydrographs for the irrigation season represents both evapotranspiration and temporary groundwater storage. The area between the hydrographs for the nonirrigation season represents nonseasonal groundwater return flow from this temporary storage.

It should be reemphasized that the conceptualization of Figure 17 involves certain assumptions and is based upon mean discharge values. However, the qualitative insight it provides is helpful in understanding the irrigation diversion-return flow system.

Since irrigation return flow is derived from irrigation diversion, a relationship between these quantities should be identifiable. The total annual diversion volumes (areas under the annual diversion hydrographs of Figure 16), IA, were regressed with the total annual return flow volumes (areas under the annual return flow hydrographs of Figure 16), RA, for the years 1966 through 1975. Table 1 summarizes these annual volumes (along with similar annual volumes obtained from Figure 15). Utilizing the figures in this table, the regression equation,

$$RA = -14.6 + 0.85IA \quad (67)$$

was determined. The sample correlation coefficient, $r = 0.92$, suggests a strong cause-effect relationship between diversion and return. The units for Equation (67) are 1000 cfs-days. It should be noted that the regression relationship and correlation coefficient are impacted to a small but unknown degree by the fact that the irrigation return flow rate from a given year's application decreases exponentially with time and thus, there is a limited carry-over from year to year.

Table 1. Yearly Component Values as Obtained from the Hydrographs in Figures.

Year	1000 cfs-days				
	Inflow Irrigation Season	Outflow Irrigation Season	Diversion Annual	Return Annual	Return/Diversion (Percent)
1966	92.65	53.07	101.91	76.01	74.6
1967	76.78	44.61	100.42	62.78	62.5
1968	115.18	69.61	134.29	87.88	65.4
1969	131.82	81.53	141.71	105.81	74.7
1970	152.82	82.50	138.22	111.55	80.7
1971	183.23	84.69	151.30	111.18	73.5
1972	165.93	98.20	148.05	124.13	83.8
1973	116.93	66.92	119.18	89.43	75.0
1974	129.31	88.04	161.95	115.12	71.1
1975	164.73	51.67	112.67	83.12	73.8
Total	1329.38	720.84	1309.70	967.01	
Mean	132.94	72.08	130.97	96.70	73.8

A yearly comparison of the total annual diversion and return flow volumes show that return flow as a percentage of diversion varies between 62.5 and 84 percent with a mean annual return percentage of 73.8 (refer to Table 1). This percentage may be slightly high because the method of return flow computation unavoidably includes some limited groundwater flow from sources other than the listed diversions. For purposes of comparison, the Soil Conservation Service (1978) estimates a gross diversion of 1.7×10^6 acre feet and a return flow of 1.3×10^6 acre feet annually for Beaverhead County. The SCS 76.4 percent-age compares favorably with the values of Table 1.

A consideration of the irrigation season diversion and return (as opposed to the annual diversion and return of the previous paragraph) provides insight about the timing of return flow. Return flow occurring between the date of the first recorded diversion and September 1 (approximate end of the irrigation season) is referred to as seasonal return flow. Return that does not occur within this interval is termed nonseasonal return flow. Although somewhat arbitrary, the difference in designation is important as seasonal return flow is generally available for reuse by downstream irrigators whereas nonseasonal return is not. Table 2 summarizes the seasonal diversion and return flow volumes for the years 1966 through 1975.

Table 2. The Seasonal Returns Expressed as a Percentage of the Seasonal Diversions.

Year	1000 cfs-days		Seasonal Return Percentage
	Seasonal Diversions	Seasonal Returns	
1966	93.29	48.19	51.7
1967	81.68	28.63	35.1
1968	111.81	46.37	41.5
1969	117.14	57.70	49.3
1970	116.62	57.41	49.2
1972	139.48	74.27	53.2
1973	89.97	55.58	61.8
1974	139.01	57.75	41.5
1975	104.28	35.77	34.3
Total	1123.62	521.87	
Mean	112.36	52.19	46.4

If seasonal return is expressed as a percentage of seasonal diversion, the percentage varies from 34.3 to 61.8 percent with a mean percentage of 46.4 (Table 2). This variability in percentage is quite probably because of differences in irrigation timing from year to year. For example, unusually warm dry early irrigation season weather with its attendant high early season irrigation activity would result in increased return flow percentages within the irrigation season.

A rough indication of return flow delay may be obtained by determining the time displacement between the centroid of the area under the seasonal diversion hydrograph and the centroid of the area under the seasonal return flow hydrograph (Figure 16). The average displacement was calculated as 61 days (see Table 3). This figure represents in a sense the time delay between diversion and return flow. The delay is of course based on aggregate figures, whereas the time delay for a particular diversion depends on application timing, point of application, local aquifer properties, etc.

Table 3. The Time Displacement in Days Between the Centroid of the Area Under the Seasonal Diversion Hydrograph and the Centroid of the Area Under the Seasonal Return Flow Hydrograph (1966-1975).

Year	Centroid Difference (Days)
1966	62.6
1967	62.0
1968	52.9
1969	54.4
1970	63.0
1971	61.6
1972	53.6
1973	60.2
1974	64.1
1975	78.2
Mean	61.3

TIME SERIES ANALYSIS OF STUDY SITE AREA

In the Beaverhead hydrologic system the input time series (irrigation diversion) is transformed to the output time series (return flow). In this section the Box and Jenkins' transfer function methodology is employed to identify the nature of this transformation process.

Model Identification Procedure

The BMDP statistical software from the Department of Biomathematics at UCLA (1981) provides the time series program to carry out the Box and Jenkins analysis. The BMDP program uses a Gauss-Marquardt method (Gausshaus) to calculate least squares estimates for the δ , ω , ϕ , and θ coefficients by minimizing the sum of squares function $S(\hat{\delta}, \hat{\omega}, \hat{\phi}, \hat{\theta})$

$$S = \sum_{t=1}^n (y(t) - \hat{y}(t))^2$$

where n is the number of residuals, $y(t)$ is the observed return, and $\hat{y}(t)$ is the estimated return. The software requires initial estimates of the δ , ω , ϕ , and θ coefficients and successively adjusts each parameter to converge on a final set of coefficients that minimizes the sum of squares function $S(\hat{\delta}, \hat{\omega}, \hat{\phi}, \hat{\theta})$.

The capability of the software was tested by modeling certain hypothetical systems. Each hypothetical system consisted of an assumed input (diversion) series and an output (return flow) series that was generated from that input series through a known (assumed) transfer function relationship. These series pairs were then analyzed using the BMDP software to determine the Box and Jenkins' transfer function. The software's ability to iden-

tify the known transfer function relationships from these hypothetical systems verified its time series analysis capability.

Although diversion activity is limited to the irrigation season, return flow continues through the full year and into the following irrigation season. Because of the year-to-year activity in the return flow process, initial analyses considered full year and even multi-year time series records of diversion and return flow. These initial analyses were unsuccessful primarily because of the discontinuous character of the diversion series. Subsequent, and more fruitful analyses, concentrated on diversion and return flow series pairs that spanned only the irrigation season, generally from mid-April to mid-October.

The seasonal time series were divided into daily and 3-day increments at various stages of the study. Subdivision of the series into 3-day increments produced about 60 observations for each of the years of interest. This number exceeds the minimum of 50 observations required for proper application of the methodology (Box and Jenkins, 1976). The use of 3-day time increments tended to produce better results than did the use of daily increments. This is quite probably due to the smoothing that occurs when averaging the daily data over 3-days (for example, the negative anomalies that were seen in Figure 16 are more prominent when the data are plotted for shorter time increments).

The time series analysis was most successfully applied to the 1974 irrigation season (April 21 through October 23), and it is the analysis of that season's data that is described in this section. The relative success of the 1974 analysis is believed to be attributable to that year's more reliable and complete time series data base. The 1974 irrigation season diversion (measured and estimated) and return data (calculated from the balance equation, Equation (66)) were analyzed using 3-day time increments (62 observations). The daily discharge data for 1974 are listed in the Appendix. The diversion and return hydrographs are shown in Figure 18.

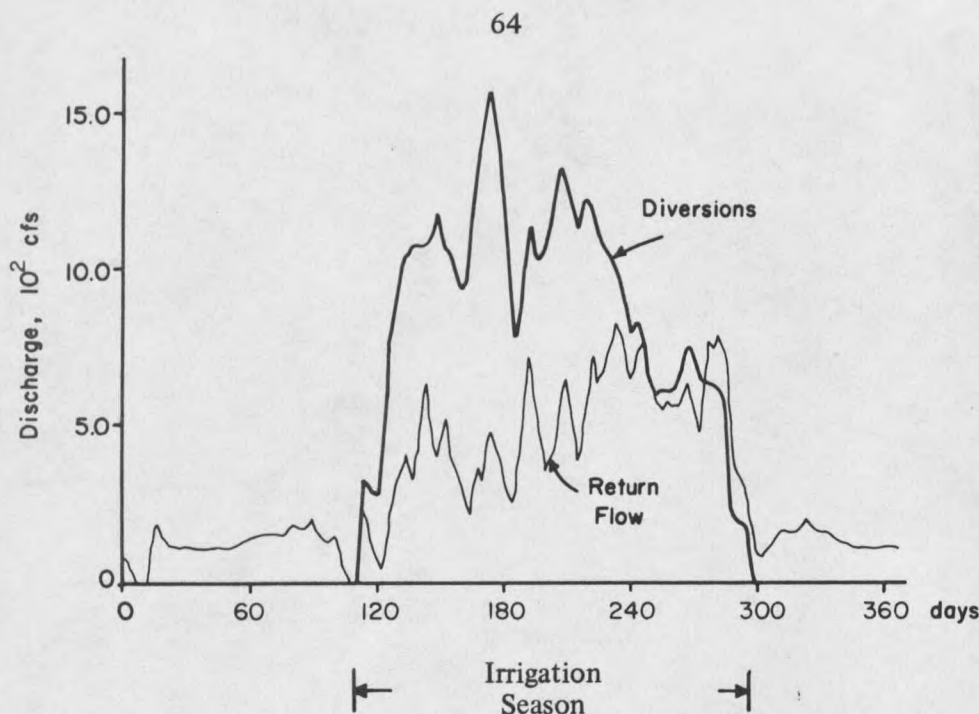


Figure 18. The diversion and return flow hydrographs for the 1974 irrigation season.

Recall again the bivariate transfer function model of Equation (44)

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

where $X(t)$ represents an input time series (the diversions), $N(t)$ represents an input noise series, and $Y(t)$ represents an output time series (the return flow). In order to model the physical relationship between the diversion and return flow, the transfer function parameters, δ and ω , and the lag parameter, b , must be determined from the two time series (diversion and return). The noise component, which accounts for the character of the return flow series that is not explained by the transfer function, must also be determined from these two series. The noise may be the result of measurement errors in the diversion and return series, the neglect of secondary factors that impact the diversion-return flow relationship, and/or the inadequacy of the transfer function approximation.

The previously described procedure for transfer function model identification requires that the input series be whitened and that the output series be filtered. These procedures require the identification of an appropriate ARIMA(p,d,q) model to characterize the input series. Recall that the AR(p), MA(q), and ARMA(p,q) models are special cases of the ARIMA(p,d,q). Following the input model identification, the ARIMA input series is whitened (Equation (53)) and the whitening parameter is then used to filter the output (return) series (Equation (54)). The whitened cross correlation function is then determined for the whitened diversion, $\alpha(t)$, and the filtered return, $\beta(t)$. The shape of this function often suggests the shape of the impulse response function which in turn may be compared qualitatively to transfer functions for models of known order and lag such as those illustrated in Figure 8.

The character (MA, AR, etc.) of a time series and its stationarity, or lack thereof, is often suggested by the appearance of the series autocorrelation and partial autocorrelation functions. In a like manner, the cross correlation function often suggests something of the nature of the relationship (the transfer function) between two time series. Figure 19 shows the autocorrelation and partial autocorrelation functions for the 1974 diversion time series (unwhitened). It also shows the cross correlation function for the diversion (unwhitened) and return flow (unfiltered) series of the same year. Nonstationarity of the diversion series is suggested by the fact that both the autocorrelation and partial autocorrelation functions tail off very slowly. The stronger cross correlation coefficients in and around lag $\tau = 20$ of the cross correlation function indicate an approximate time delay of 3τ or $3 \times 20 = 60$ days between irrigation application and significant return flow response.

The determination of an appropriate ARIMA model

$$X(t) = \frac{\theta_x^q(B)}{\phi_x^p(B) \nabla^d} \alpha(t) \quad (52)$$

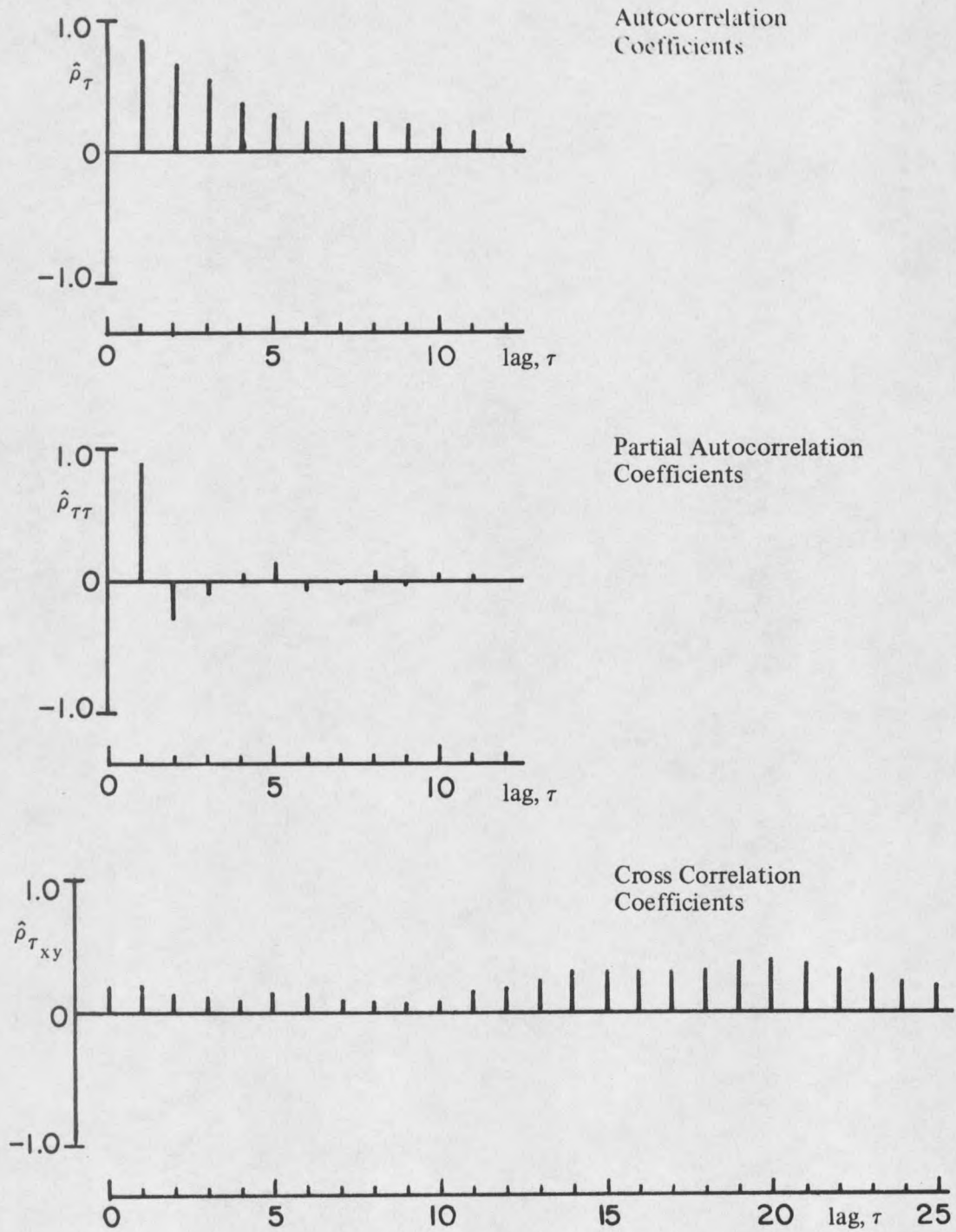


Figure 19. The autocorrelation and partial autocorrelation coefficients for the 1974 diversions, and the cross correlation coefficients between the 1974 diversions and returns.

for a nonstationary input series (the first step in whitening) is facilitated by differencing the time series and then determining, plotting and inspecting the various correlation coefficients from the differenced series. The autocorrelation and partial autocorrelation functions for the differenced ($d=1$) diversion series are plotted in Figure 20. The autocorrelation function shows a significant autocorrelation only at lag $\tau = 1$ indicating process stationarity was achieved by the differencing operation. The abrupt decrease in autocorrelation coefficient significance after lag $\tau = 1$ combined with the gradual tailing off of the partial autocorrelation function suggests that the appropriate ARIMA input model includes a moving average parameter ($q=1$) as well as the previously identified differencing parameter ($d=1$).

The diversion series was whitened with the following ARIMA model

$$\alpha(t) = \frac{1}{1 + 0.51B} x(t)$$

where $p=0$, $d=1$, and $q=1$. Using the diversion whitening parameter ($1/1 + 0.51B$), the return flow series was filtered. The cross correlation function for the whitened diversion and the filtered return flow series is shown in Figure 21 along with the whitened diversion series autocorrelation and partial autocorrelation functions. The plotted cross correlation function indicates a significant input and output relationship at lag zero only. Recall that the preliminary data analysis indicated a probable time delay between diversion and return flow of approximately 61 days. Though a transfer function parameter at lag zero is reasonable (owing to probable surface return flows within the 3-day time increment), the high cross correlation at this lag is probably more directly the result of the return flow computation procedure. Recall that return flow is computed with the balance equation (Equation (66)) and thus is dependent upon the diversions (measured and estimated during the same time increment).

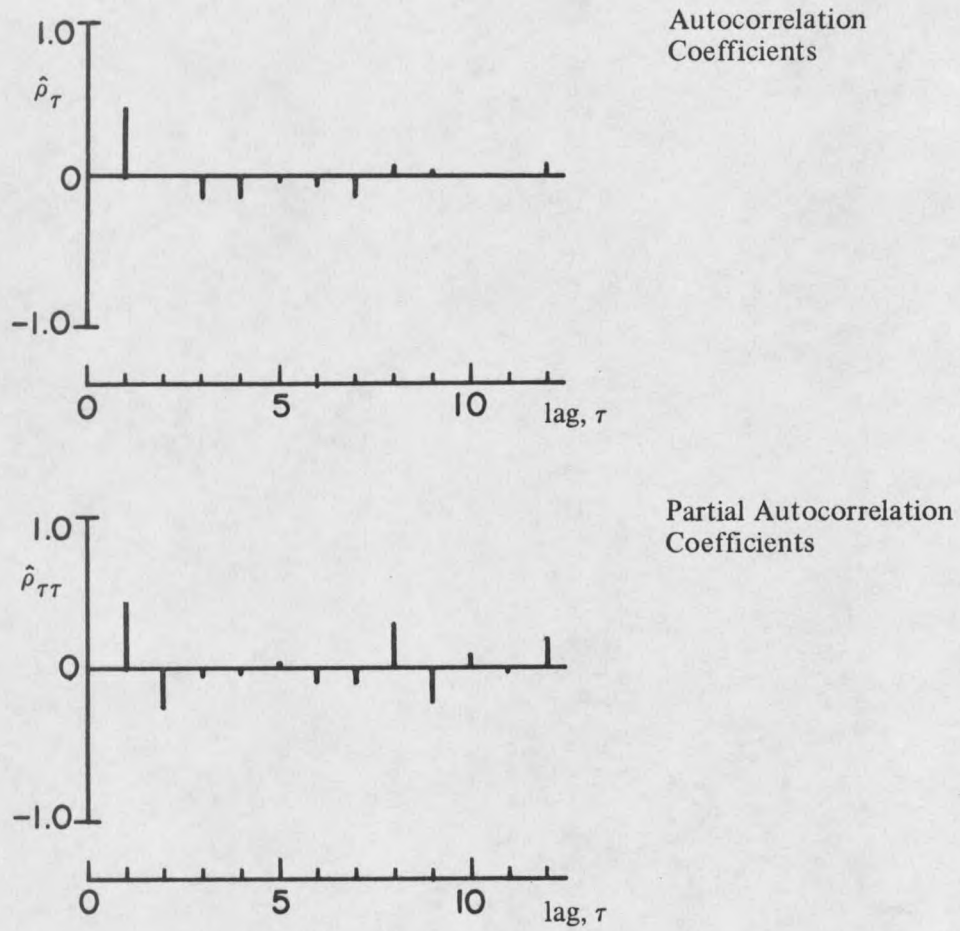


Figure 20. The autocorrelation and partial autocorrelation coefficients for the 1974 differenced ($d = 1$) diversions.

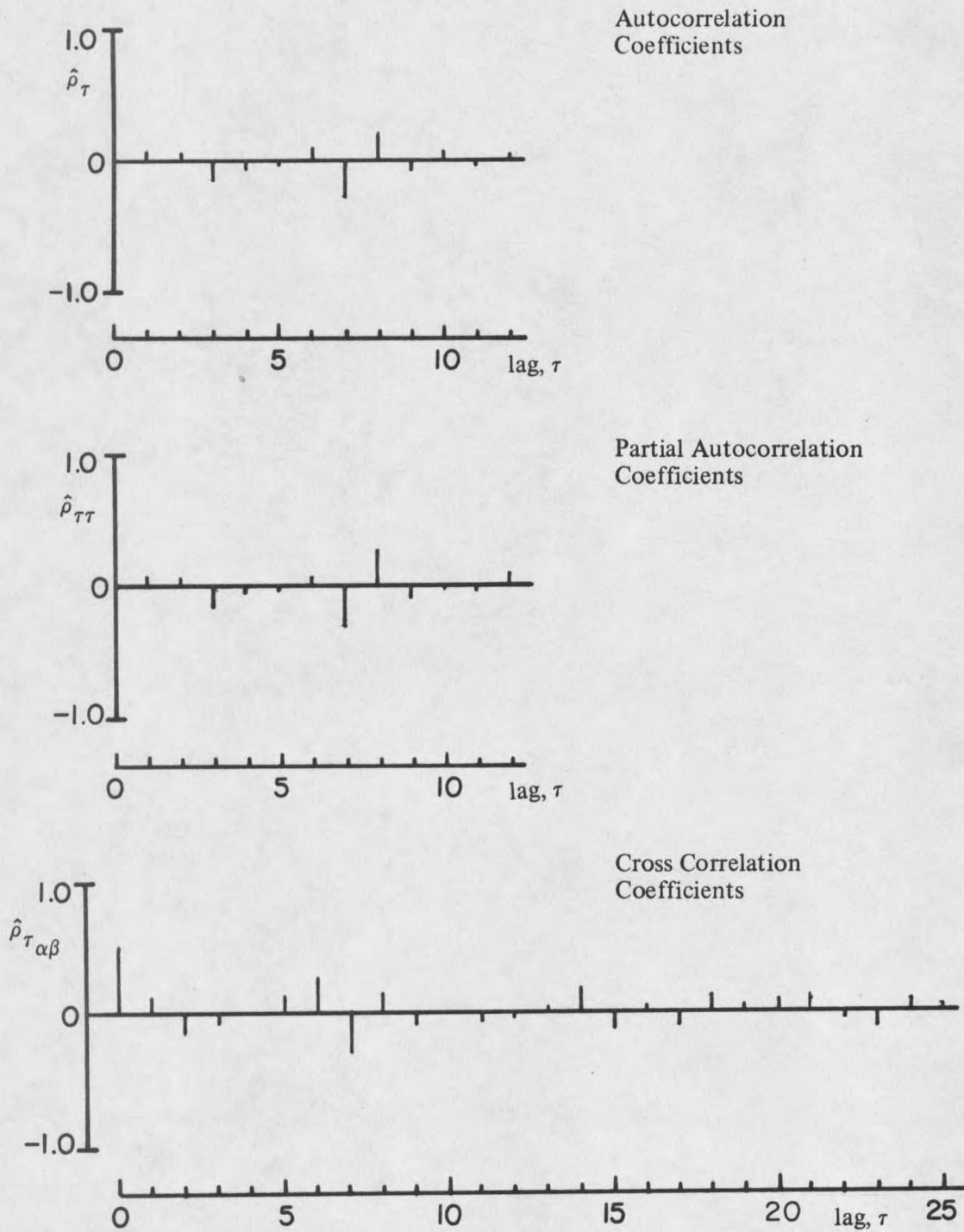


Figure 21. The autocorrelation and partial autocorrelation coefficients for the whitened diversions, and the cross correlation coefficients between the 1974 whitened diversions and the filtered returns.

Because of the failure of the whitening and filtering procedure to identify the anticipated model (input-output lag on the order of 61 days), an alternative trial and error identification procedure was employed in subsequent analyses. According to the literature, the use of a trial and error procedure is a common necessity because of real world data peculiarities. Although the cross correlation function between the unmodified input and output (Figure 19) does not provide reliable evidence of the form of the impulse response function (and therefore the order of the transfer function), it does suggest the magnitude of the lag parameter b . A trial and error procedure, which relied upon this cross correlation function indication of b was applied in the following manner:

1. The approximate lag b was determined from the cross correlation function for the sample diversion and return ($b=17$ to 22 in Figure 19);
2. These estimated lags were incorporated in parsimonious forms (orders r , s , p , and $q \leq 2$) of the transfer function relationship (initial estimates of pertinent δ , ω , ϕ , and θ coefficients were provided for BMDP least squares determination of these coefficients);
3. Model diagnostics were applied to these various trial models to test for model adequacy.

Although theoretically not as efficient or mathematically sound as the whitening and filtering model identification methods, this trial and error procedure was effective in identifying transfer function relationships between the diversion series and the return flow series. Three of the more promising 1974-season transfer function models that were identified by this procedure are herein described in the following paragraphs.

First Transfer Function Model

A parsimonious form of the transfer function relationship

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

results when $r=1$ and $s=0$ in the transfer function component and when the noise $N(t)$ is an autoregressive series with $p=1$ and $q=0$. The transfer function model then becomes

$$Y(t) = \left[\frac{\omega_b}{1-\delta_1 B} X(t) \right] B^b + \frac{A(t)}{1-\phi_1 B} \quad (68)$$

Using this model and varying lag b from 17 to 22, the BMDP algorithm produced least squares estimates of the coefficients ω_b , δ_1 , and ϕ_1 for each assumed lag. Within the constraints dictated by these lag and order assumptions, the model

$$Y(t) = \left[\frac{0.130}{1-0.754B} X(t) \right] B^{18} + \frac{A(t)}{1-0.584B} \quad (69)$$

resulted in a minimum residual series variance. Table 4 lists model parameter estimates and resultant residual series variances (mean square errors) estimates for three different assumed lag parameter values. Also listed are the standard errors of the parameter estimates and the associated t-ratios (parameter estimate/standard error) under the assumption that residuals are normally distributed. The parameters are all significant at the 95% level of confidence as indicated by t-ratios ($t_{0.05, 40} \approx 2$ for two-tailed table).

The residual series, which is obtained by subtracting the predicted returns, $\hat{y}(t)$, from the measured returns, $y(t)$,

$$a(t) = y(t) - \hat{y}(t)$$

consists of 44 observations. The smaller residual series size results from the fact that the first predicted return is produced when $t = 19$ since

Table 4. Parameter Estimates of the ω , δ_1 , and ϕ_1 Coefficients at Various Lags (first model—Equation (68)). The Degrees of Freedom $\nu = N-r-2s-b-2p-q-1$.

Model Parameter	Parameter Estimate	Standard Error	T-Ratio	Degrees Freedom	Mean Square Error
ϕ_1	0.5911	0.1280	4.62	41	12,873
ω_{17}	0.1003	0.0546	1.84		
δ_1	0.8081	0.1022	7.91		
ϕ_1	0.5841	0.1392	4.20	40	12,619
ω_{18}	0.1299	0.0632	2.06		
δ_1	0.7540	0.1178	6.40		
ϕ_1	0.5881	0.1421	4.14	39	13,111
ω_{19}	0.1502	0.0705	2.13		
δ_1	0.7155	0.1318	5.43		

$$x(t) B^{18} = x(t-18) = x(19-18) = x(1)$$

and $x(1)$ is the first observed diversion. The number of observations in the residual series is thus $62-18=44$ (the 1974 season had 62 observations). In general, the first predicted return is produced when $t = b+s+1$. Recall that $b=18$ and $s=0$ in this model.

The first 12 autocorrelation and partial autocorrelation coefficients for the residual series are shown in Figure 22 along with the first 25 cross correlation coefficients for the diversion series and the residual series. Each of the correlation coefficient plots indicates a bounding tolerance band. The tolerance bands represent the plus and minus two standard error line and thus delineate the "t-like" statistic (recall Equations (60), (61), (62), and (63)). Correlation coefficients that extend beyond these bands indicate a significant dependence in the residual series structure and an inadequacy in the model (significance in the "t-like" statistic).

Significant correlations are apparent on the partial autocorrelation coefficient plot at lag two and on the cross correlation coefficient plot at lag zero. Recall from an earlier section that a significant cross correlation coefficient generally indicates inadequacy in the

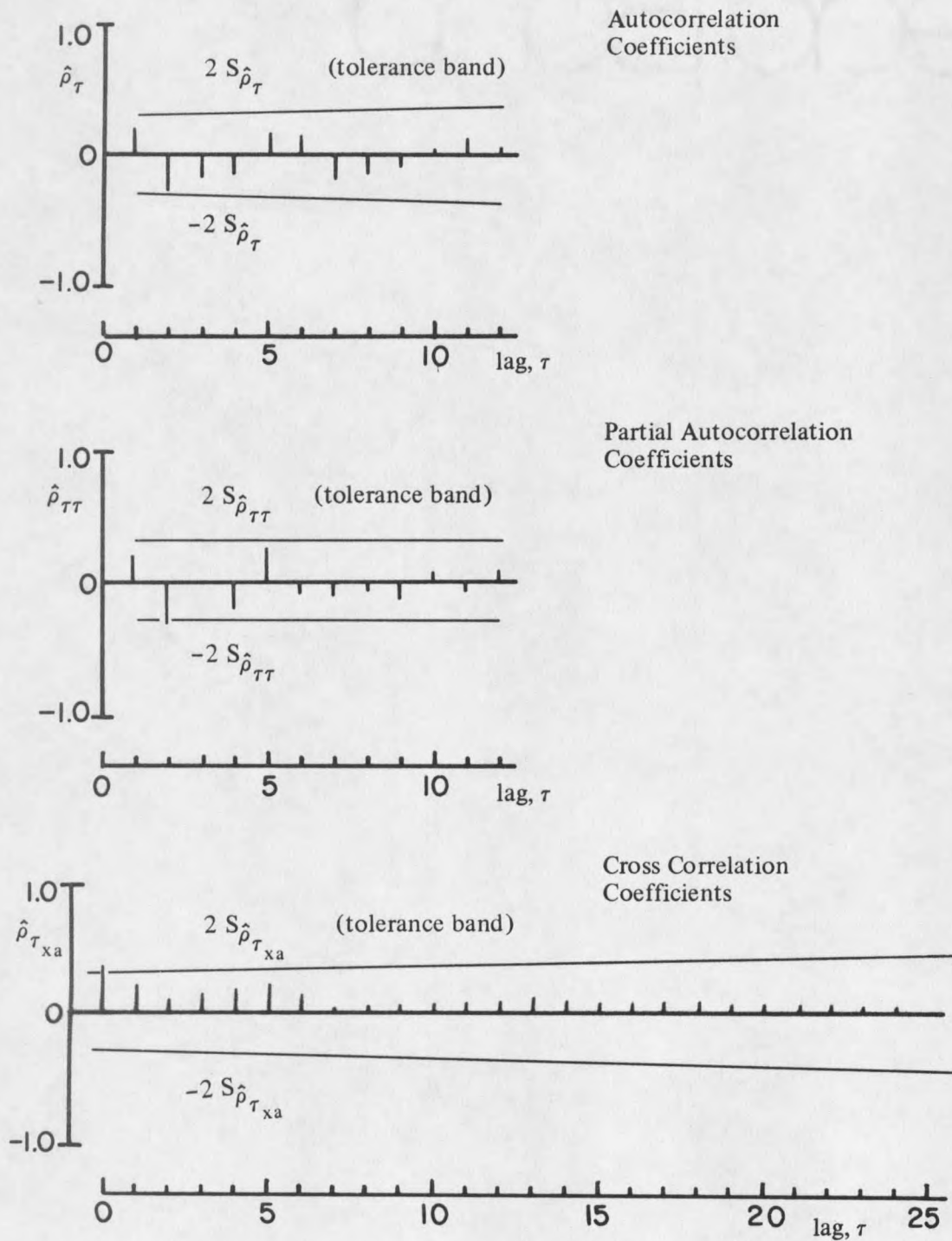


Figure 22. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series (first model form Equation (69)).

transfer function component of Equation (44). In addition, the somewhat slower decay of the residual partial autocorrelation function (when compared to the autocorrelation function) suggests a possible moving average MA(1) component in the residuals and hence an inadequacy in the noise component. The high lag zero cross correlation coefficient ($\hat{\rho}_{\tau_{xa}} = 0.373$) suggests the need for an additional coefficient $\omega_{\tau} = \omega_0$ in the transfer function.

The possible model inadequacy can be further checked with the previously described "portmanteau" statistic. The "portmanteau" statistic for the first 12 autocorrelation coefficients of the residuals is

$$Q = n \sum_{j=1}^K \hat{\rho}_j^2 = 43 \sum_{j=1}^{12} \hat{\rho}_j^2 = 12.47$$

where the number of effective observations $n = N-d-b-s-p = 62-18-0-1 = 43$. The chi-square values for $\nu = K-p-q = 12-1-0 = 11$ degrees of freedom and $\alpha = 5\%$ ($\chi^2_{0.05,11}$) is 19.7. Since $Q < \chi^2$, the model is not judged inadequate. The "portmanteau" statistic for the first 25 cross correlation coefficients is

$$S = n \sum_{j=0}^M \hat{\rho}_{j_{xa}}^2 = 43 \sum_{j=1}^{25} \hat{\rho}_{j_{xa}}^2 = 17.50$$

where the number of effective observations is again 43. The chi-square value for $\nu = M-r-s = 25-1-0 = 24$ and $\alpha = 5\%$ ($\chi^2_{0.05,24}$) is 36.4. Since $S < \chi^2$, the model is not judged inadequate.

Although the previous model is adequate in terms of the "portmanteau" test, the model is not adequate in terms of the "t-like" statistic (as delineated by the tolerance bands in Figure 22). The correlation coefficients indicated the possible need for a moving average parameter in the noise component and lag zero, ω_0 , coefficient in the transfer function component.

Second Transfer Function Model

A second parsimonious form of the transfer function relationship

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

results when $r=1$ and $s=0$ for the transfer function component and when the noise $N(t)$ is a moving average series with $p=0$ and $q=1$. The transfer function model then becomes

$$Y(t) = \left[\frac{\omega_b X(t)}{1 - \delta_1 B} \right] B^b + (1 - \theta_1 B) A(t) \quad (70)$$

Using this model and varying lag b from 17 and 22, the BMDP algorithm produced least squares estimates of the coefficients ω_b , δ_1 , θ_1 for each lag. Within the constraints dictated by these assumptions, the model

$$Y(t) = \left[\frac{0.134 X(t)}{1 - 0.749 B} \right] B^{18} + (1 + 0.778 B) A(t) \quad (71)$$

resulted in the minimum residual series variance. Table 5 lists residual series variances (mean square errors) and the parameter estimates for three different assumed lag parameter values. Note that the residual series variance is less than that of the first model indicating an improved model. Also listed are the standard errors of the parameter estimates and the associated t -ratios. The parameters are all significant at the 95% level of confidence as indicated by their t -ratios.

The first 12 autocorrelation and partial autocorrelation coefficients (and tolerance bands) for the residual series are shown in Figure 23 along with the first 25 cross correlation coefficients (and tolerance bands) for the diversion series and the residual series. No significant residual series correlations (" t -like" statistic) are apparent on either the autocorrelation or partial autocorrelation coefficient plot indicating the noise component of the model is adequate. This absence of significant correlation coefficients represents an

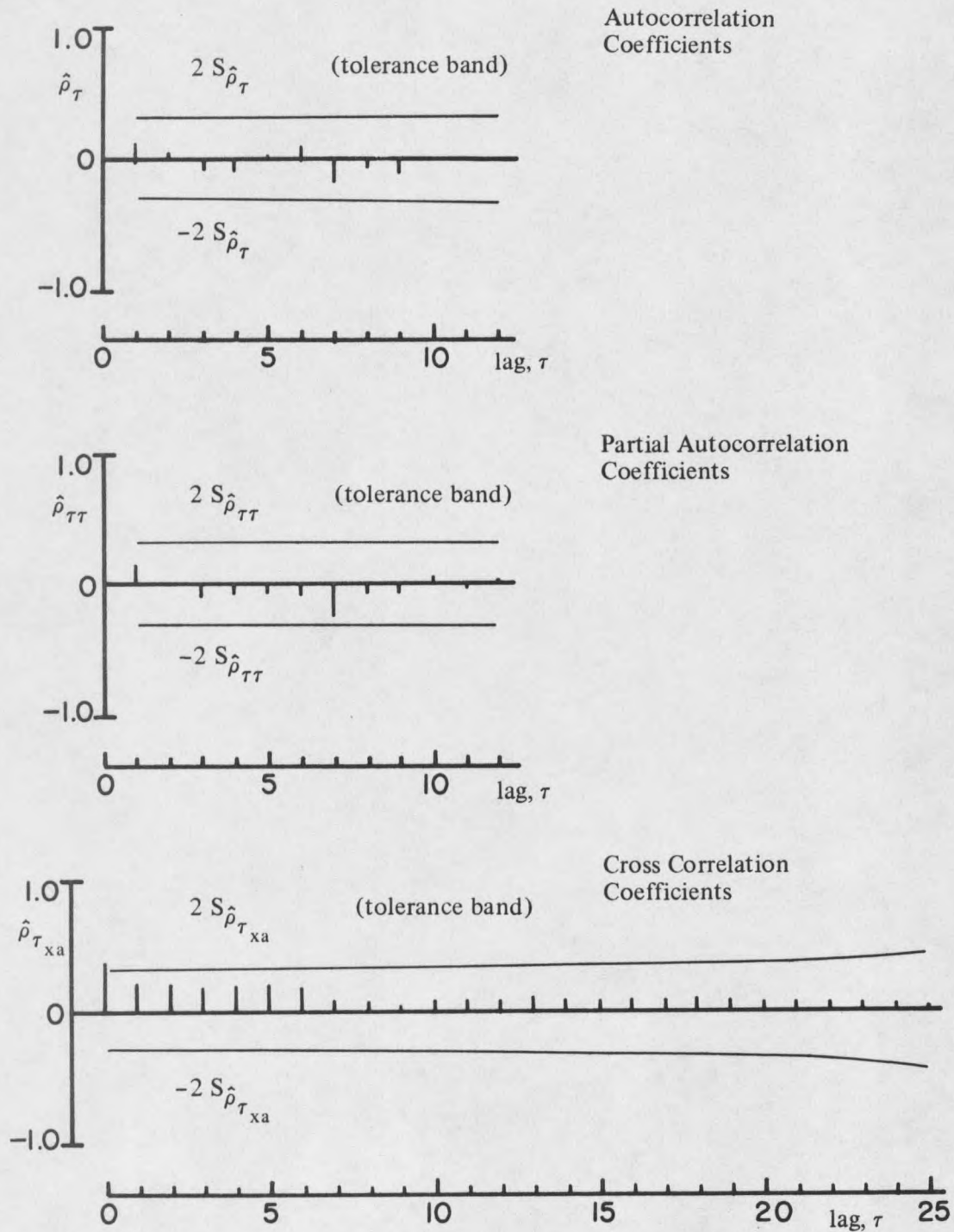


Figure 23. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series (second model form Equation (71)).

Table 5. Parameter Estimates of the ω_b , δ_1 , and θ_1 Coefficients at Various Lags (second model—Equation (70)). The Degrees of Freedom $\nu = N-r-2s-b-2p-q-1$.

Model Parameter	Parameter Estimate	Standard Error	T-Ratio	Degrees Freedom	Mean Square Error
θ_1	-0.7233	0.0927	-7.80	42	10,834
ω_{17}	0.1080	0.0438	2.46		
δ_1	0.7964	0.0813	9.80		
θ_1	-0.7775	0.0927	-8.39	41	10,081
ω_{18}	0.1344	0.0523	2.57		
δ_1	0.7489	0.0965	7.76		
θ_1	-0.7477	0.0952	-7.86	40	10,638
ω_{19}	0.1457	0.0567	2.57		
δ_1	0.7270	0.1053	6.91		

improvement over the first model. A significant residual cross correlation ($\hat{\rho}_{0_{xa}} = 0.356$) persists at lag zero, and the possible need for an additional coefficient $\omega_{\tau} = \omega_0$ in the transfer function remains.

The "portmanteau" statistic for the first 12 autocorrelation coefficients of the residuals is

$$Q = 44 \sum_{j=1}^{12} \hat{\rho}_j^2 = 5.10$$

where the number of effective observations is 44. The chi-square value for $\nu = 11$ and $\alpha = 5\%$ ($\chi^2_{.05,11}$) is 19.7. Since $S < \chi^2$, the model is not judged inadequate. The "portmanteau" statistic for the first 25 cross correlation coefficients is

$$S = 44 \sum_{j=0}^{25} \hat{\rho}_{j_{xa}}^2 = 20.68$$

where the number of effective observations is again 44. The chi-square value for $\nu = 24$ and $\alpha = 5\%$ ($\chi^2_{.05,24}$) is 36.4. Since $S < \chi^2$, the model is not judged inadequate.

Although the reduced mean square error suggests that the second model form is a substantial improvement over the second model, the possible need for a lag zero transfer function parameter remains. This lag parameter is incorporated in the final model which is described below.

Third Transfer Function Model

A parsimonious form of the transfer function relationship

$$Y(t) = \left[\frac{\omega^s(B)}{\delta^r(B)} X(t) \right] B^b + N(t) \quad (44)$$

results when $r=1$ and $s=1$ for the transfer function component and when the noise component $N(t)$ is a moving average series with $p=0$ and $q=1$. The transfer function model then becomes

$$Y(t) = \left[\frac{\omega_0 + \omega_b B^b}{1 - \delta_1 B} \right] X(t) + (1 - \theta_1 B) A(t) \quad (72)$$

Using this model and varying lag b from 17 to 22, the BMDP algorithm produced least squares estimates of the coefficients ω_0 , ω_b , δ_1 , and θ_1 for each lag. Within the constraints dictated by these assumptions, the model

$$Y(t) = \left[\frac{0.029 + 0.124 B^{18}}{1 - 0.722 B} \right] X(t) + (1 + 0.772 B) A(t) \quad (73)$$

resulted in a minimum residual series variance. Table 6 lists residual series variances (mean square errors) and the parameter estimates for three different assumed lag parameter values. Note that the residual series variance is only slightly less than that of the second model. Also listed are the standard errors of the parameter estimates and the associated t -ratios. All the parameters except ω_0 are significant at the 95% level of confidence as indicated by their t -ratios.

Table 6. Parameter Estimates of the ω , δ_1 , and ϕ_1 Coefficients at Various Lags (third model—Equation (72)). The Degrees of Freedom $\nu = N-r-s-b-2p-q-1$.

Model Parameter	Parameter Estimate	Standard Error	T-Ratio	Degrees Freedom	Mean Square Error
θ_1	-0.7209	0.0935	-7.71	41	10,752
ω_0	0.0176	0.0169	1.04		
ω_{17}	0.0959	0.0490	1.96		
δ_1	0.7898	0.0956	8.26		
θ_1	-0.7717	0.0956	-8.07	40	9,783
ω_0	0.0294	0.0225	1.31		
ω_{18}	0.1236	0.0602	2.05		
δ_1	0.7216	0.1246	5.49		
θ_1	-0.7369	0.0986	-7.48	39	9,951
ω_0	0.0418	0.0277	1.51		
ω_{19}	0.1371	0.659	2.08		
δ_1	0.6769	0.1435	4.72		

The first 12 autocorrelation and partial autocorrelation coefficients (and tolerance bands) for the residual series are shown in Figure 24 along with the first 25 cross correlation coefficients (and tolerance bands) for the diversion series and the residual series. No significant residuals are evident on these plots suggesting that both the noise and transfer function components are adequately explained by their respective parameters.

The "portmanteau" statistic for the first 12 autocorrelation coefficients of the residuals is

$$Q = 43 \sum_{j=1}^{12} \hat{\rho}_j^2 = 4.56$$

where the number of effective observations is 43. The chi-square value for $\nu = 11$ degrees of freedom and $\alpha = 5\%$ ($\chi_{0.05,11}^2$) is 19.7. Since $Q < \chi^2$, the model is not judged inadequate. The "portmanteau" statistic for the first 25 cross correlation coefficients is

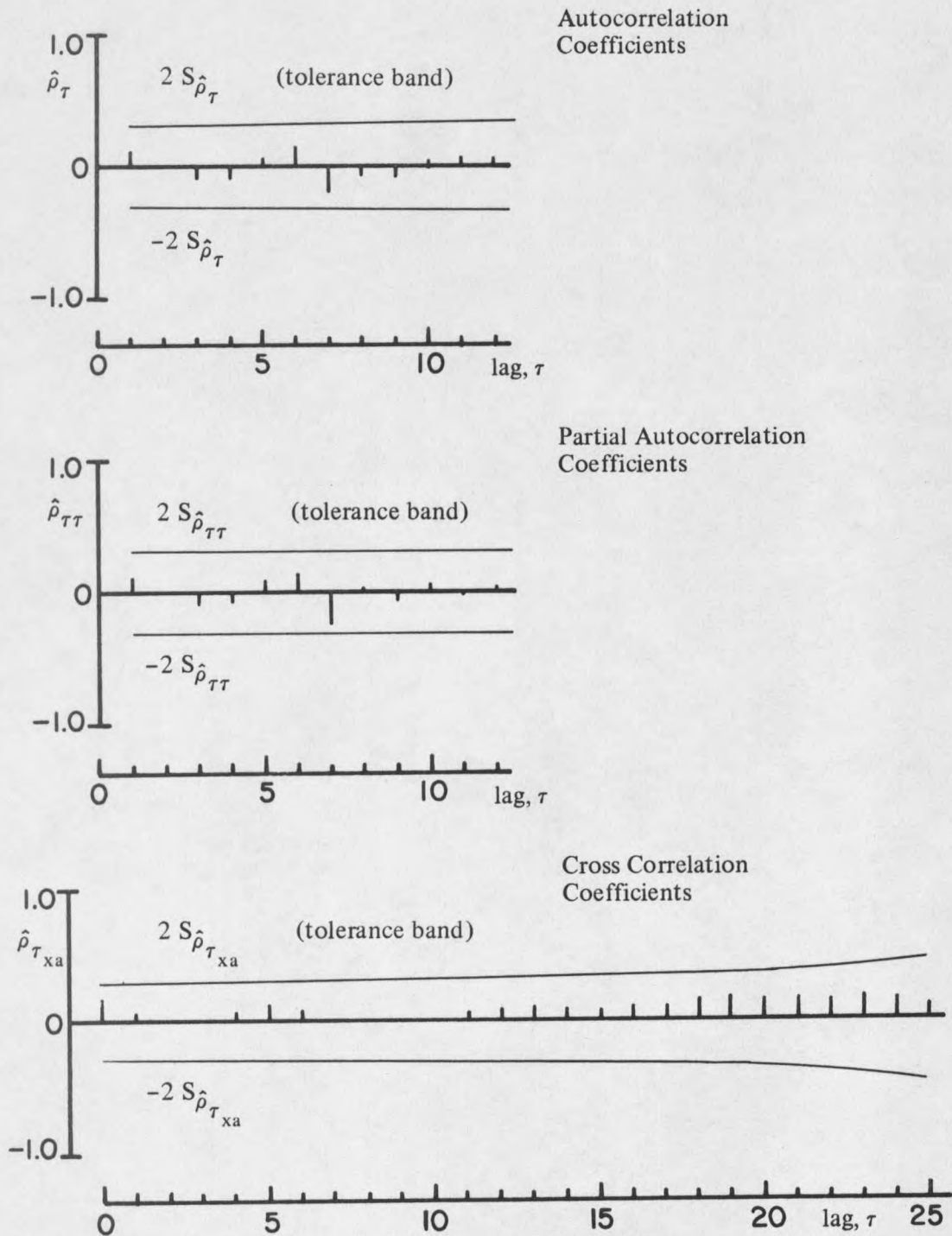


Figure 24. The autocorrelation and partial autocorrelation coefficients for the residual series, and the cross correlation coefficients between the diversion and residual series (third model form Equation (73)).

$$S = 43 \sum_{j=1}^{25} \hat{\rho}_{j_{xa}}^2 = 10.36$$

where the number of effective observations is again 43. The chi-square value for $\nu = 23$ and $\alpha = 5\%$ ($\chi^2_{.05,23}$) is 35.2. Since $S < \chi^2$, the model is not judged inadequate.

Summary of Identified Models

Table 7 summarizes each model's parameters (and standard errors), residual variance, and "portmanteau" statistics. This table combined with the residual correlation coefficient plots (Figures 23 and 24) provides a basis for comparison of the last two models. A comparison of the residual series diagnostics and the variance, σ_a^2 , for the three transfer function models (Equations (69), (71), and (73)) suggests that both the second and third transfer function models (fitted with moving average noise components) provide substantially better fits to the diversion and return data than does the first transfer function model (fitted with an autoregressive noise component). The third model (additionally fitted with a lag zero transfer function parameter, ω_0) produced only a slight further reduction in residual series variance. It did however reduce the lag zero cross correlation coefficient to a nonsignificant level.

The 1974 season estimated return from the second transfer function model is shown in Figure 25 along with the observed return (balance equation). This figure also shows the diversions for that year. Figure 26 shows the same diversion and return series along with an estimated return series from the third transfer functional model.

An improved physical understanding of the diversion-return flow relationship may be obtained by altering the form of the selected transfer function model. For example the form of the seasonal model,

Table 7. A Summary of the Statistics for the Three Model Forms.

Model	Parameter Estimates	Standard Error	Residual Variance	Q	S
First	$\phi_1 = 0.5841$	0.1392	12,619	12.47	17.50
	$\omega_{18} = 0.1299$	0.0632			
	$\phi_1 = 0.7540$	0.1178			
Second	$\phi_1 = -0.7775$	0.0927	10,081	5.10	20.68
	$\omega_{18} = 0.1344$	0.0523			
	$\delta_1 = 0.7489$	0.0965			
Third	$\phi_1 = -0.7717$	0.0956	9,783	4.56	10.36
	$\omega_0 = 0.0294$	0.0225			
	$\omega_{18} = 0.1236$	0.0602			
	$\delta_1 = 0.7216$	0.1246			

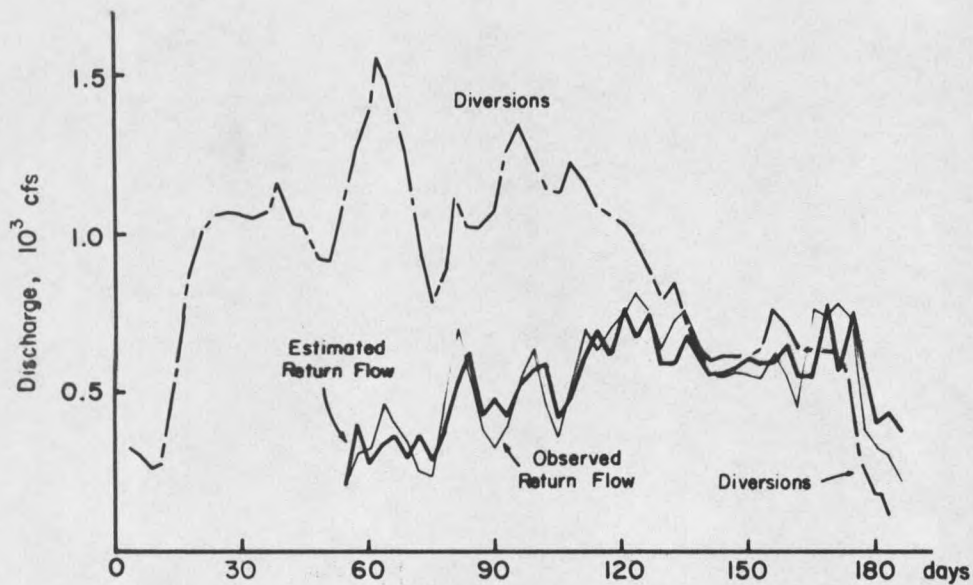


Figure 25. Diversion series and return series (observed and estimated by second model—Equation (71)).

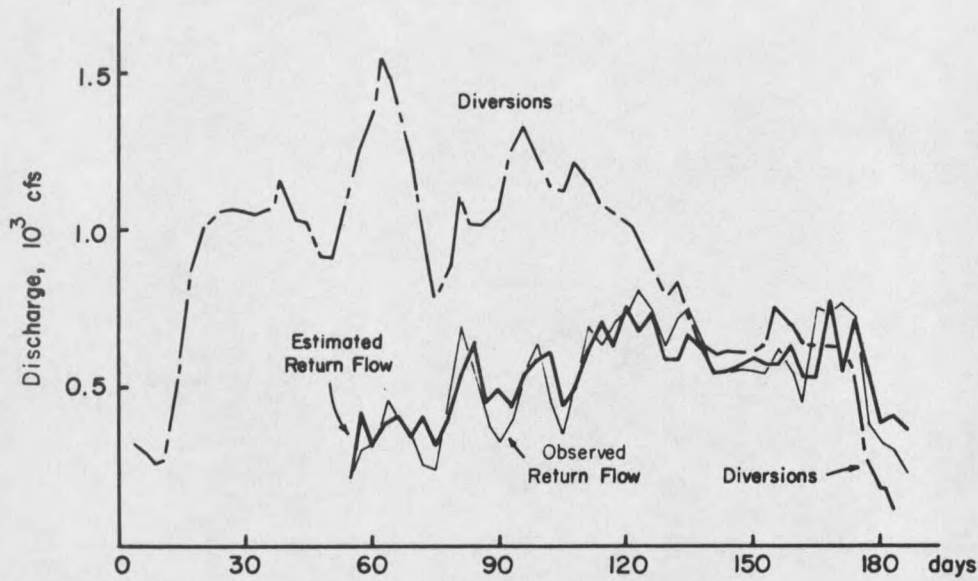


Figure 26. Diversion series and return series (observed and estimated by third model—Equation (73)).

$$Y(t) = \left[\frac{0.134}{1-0.749B} X(t) \right] B^{18}$$

may be modified by performing the division implied in the transfer function component of the model to obtain the impulse response function weights

$$\frac{0.134B^{18}}{1-0.749B} = 0.134B^{18} + 0.101B^{19} + 0.075B^{20} + \dots$$

Multiplying by $X(t)$ and carrying out the backstep operation results in a series of $X(t)$'s weighted by the impulse response function

$$\frac{0.134B^{18}}{1-0.749B} X(t) = 0.134 X(t-18) + 0.10 X(t-19) + 0.075 X(t-20) + \dots$$

or more concisely

$$\frac{0.134B^{18}}{1-0.749B} X(t) = 0.134 \sum_{j=0}^{\infty} (0.749)^j X(t-18-j)$$

Figure 27 shows the impulse response function weights that characterize the relationship between the diversion series and the return flow series. This relationship indicates that the

return flow responds most strongly to the diversion that is applied 54 days (18 3-day increments) earlier. The response to earlier inputs diminishes progressively as the decaying response function indicates.

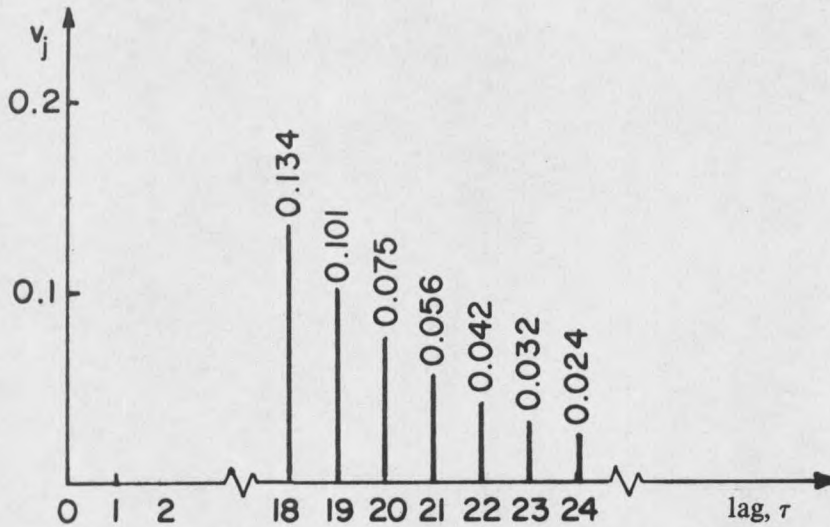


Figure 27. The impulse response function for the 1974 irrigation season.

The steady state gain

$$\frac{0.134}{1-0.749} = 0.535$$

suggests that 53.5 percent of the diversion returns to the stream as seasonal return flow. For comparison, the percentage of diversions returning was 54.4 within the 1974 irrigation season (April 21 through October 23).

SUMMARY AND CONCLUSIONS

The irrigation diversion-return flow system along a reach of the Beaverhead River, near Dillon, Montana was analyzed utilizing Box and Jenkins' transfer function methodology. Although the recommended whitening and filtering procedure failed to identify the nature of the transfer relationship between the diversion and return flow series, an alternative trial and error procedure did identify transfer function relationships for the 1974 irrigation season. These relationships are physically quite reasonable. For example, the transfer function model

$$Y(t) = \left[\frac{0.134 X(t)}{1-0.749B} \right] B^{18} + (1 + 0.778B) A(t) \quad (71)$$

appears to approximate both the timing and the amount of seasonal return flow. The return flow response of the model is most strongly dependent upon the irrigation diversions that occur 54 days earlier. The model yields a seasonal return flow volume that is 53.5 percent of the seasonal diversion. Both of these values agreed quite well with the estimates of return timing and amount that were based on hydrograph analysis.

Although the study demonstrated the potential of transfer function methodology as a tool for return flow quantification, it also showed how that potential is severely limited by data inadequacies. Incomplete discharge records (diversions and tributaries) necessitate estimates which in turn impact the transfer function model form. Discharge measurement errors, if present, are particularly troublesome since in addition to their obvious direct impact on the model, they impact it in a second indirect way through the computed return flow. Measurement errors in inflow, diversion, and outflow are transmitted via the balance equation to the return flow.

Transfer function analysis, as with any statistical methodology, is most meaningful when a large sample is evaluated. The analysis of only irrigation season returns limits the sample size considerably (about 60 three-day increments at the study site). A return flow response lag on the order of 18 three-day increments further reduces the effective sample size in that only $60 - 18 = 42$ values are available to establish a cross correlation relationship. In view of this sample size constraint, the possibility of full year or multi-year analyses should be further investigated.

Since the quantification of irrigation season return flows by time series analysis, or by any other means, is contingent upon the existence of a reliable and complete data base, consideration should be given to the establishment of a hydrologic monitoring program at a carefully selected site. The data base resulting from several years of comprehensive monitoring of diversions and streamflows at such a site would be an invaluable aid in understanding and quantifying the return flow process via transfer function analysis.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Anderson, O. D., 1976. Time Series Analysis and Forecasting. Butterworths. London, England.
- Botz, M. K., 1967. Hydrogeology of the East Bench Irrigation Unit, Madison and Beaverhead Counties, Montana. Montana Bureau of Mines and Geology. Montana College of Mineral Science and Technology. Butte.
- Bowerman, B. L., and R. T. O'Connell, 1979. Time Series and Forecasting: An Applied Approach. Duxbury. North Scituate, Massachusetts.
- Box, G. E. P., and G. M. Jenkins, 1976. Time Series Analysis Forecasting and Control. Revised Edition. Holden-Day. San Francisco, California.
- Box, G. E. P., and D. A. Pierce, 1970. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. Jour. Amer. Stat. Assoc., 64:1509.
- Kennedy, R., 1981. Irrigation Diversion Discharge Data Obtained by Personal Communication from the East Bench Irrigation District. U.S. Bureau of Reclamation. Dillon, Montana.
- Nicklin, M. E., and R. L. Brustkern, 1981. Assessment of Methodology to Quantify Irrigation Return Flows. Montana Water Resources Report no. 114. Montana Water Resources Research Center, Montana State University. Bozeman.
- Nicklin, M. E., and R. L. Brustkern, 1983. Assessment of Time Series as a Methodology to Quantify Irrigation Return Flows. Montana Water Resources Report no. 136. Montana Water Resources Research Center, Montana State University. Bozeman.
- Ross, C. P. et al., 1955. Geologic Map of Montana. Montana Bureau of Mines and Geology. Montana College of Mineral Science and Technology. Butte.
- Salas, J. D. et al., 1980. Applied Modeling of Hydrologic Time Series. Water Resources Publications. Littleton, Colorado.
- U.S. Department of Agriculture, 1978. Water Conservation and Salvage Report for Montana. Soil Conservation Service. Bozeman, Montana.
- University of California, 1981. BMDP Statistical Software. Department of Biomathematics. University of California Press. Los Angeles.
- Yule, G. U., 1927. On the Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers. Phil. Trans., A226:227.

APPENDIX

The Barretts Gaging Station daily discharge data (cfs) for 1974 as published by the United States Geological Survey (read from left to right).

220.	210.	220.	220.	215.	200.	200.	210.	220.	230.	255.	180.
181.	181.	187.	205.	253.	255.	248.	247.	243.	245.	245.	246.
245.	245.	249.	252.	270.	273.	276.	274.	273.	273.	279.	295.
295.	295.	294.	291.	291.	290.	290.	289.	289.	289.	292.	292.
293.	293.	293.	289.	286.	289.	287.	286.	286.	286.	281.	286.
290.	305.	295.	294.	291.	287.	286.	283.	278.	280.	287.	302.
297.	296.	358.	525.	601.	393.	321.	320.	319.	308.	304.	308.
326.	418.	479.	399.	377.	377.	366.	356.	347.	345.	347.	354.
354.	355.	360.	387.	388.	385.	366.	355.	359.	372.	393.	482.
462.	622.	610.	613.	717.	735.	740.	743.	739.	722.	742.	806.
800.	794.	788.	788.	788.	794.	788.	794.	836.	848.	878.	925.
946.	939.	925.	960.	1010.	1040.	1100.	1120.	1020.	918.	812.	872.
953.	946.	953.	981.	1020.	1010.	878.	776.	770.	764.	758.	836.
836.	830.	752.	698.	770.	878.	953.	994.	1020.	988.	1040.	1060.
1120.	1110.	1120.	1180.	1200.	1210.	1200.	1130.	1160.	1150.	1040.	974.
946.	891.	806.	788.	776.	728.	704.	692.	692.	668.	674.	704.
716.	728.	794.	818.	848.	866.	905.	981.	1050.	1070.	1080.	1100.
1100.	1040.	953.	905.	872.	842.	854.	946.	960.	998.	999.	1000.
1010.	953.	923.	844.	776.	747.	740.	737.	727.	698.	655.	649.
635.	619.	607.	557.	532.	500.	492.	486.	487.	486.	487.	501.
524.	507.	473.	446.	446.	438.	405.	384.	366.	354.	354.	370.
391.	392.	391.	391.	391.	390.	389.	388.	386.	389.	406.	434.
433.	434.	454.	427.	416.	414.	399.	400.	413.	253.	261.	257.
257.	253.	257.	261.	261.	245.	233.	233.	233.	237.	237.	237.
241.	237.	201.	190.	190.	190.	190.	190.	190.	190.	190.	190.
190.	194.	194.	194.	197.	194.	184.	178.	187.	190.	190.	190.
190.	194.	197.	197.	197.	194.	190.	190.	190.	190.	190.	172.
178.	184.	184.	184.	184.	184.	184.	184.	187.	190.	194.	190.
187.	187.	179.	176.	184.	186.	184.	182.	184.	187.	187.	186.
183.	185.	197.	238.	235.	225.	230.	237.	237.	236.	236.	230.
227.	233.	234.	233.	229.							

APPENDIX

The Blaine Gaging Station daily discharge data (cfs) for 1974 as published by the United States Geological Survey (read from left to right).

327.	333.	330.	300.	280.	270.	250.	230.	220.	210.	206.	219.
283.	346.	391.	442.	508.	466.	456.	418.	406.	391.	400.	403.
403.	403.	400.	400.	412.	421.	430.	433.	430.	430.	430.	442.
445.	445.	445.	442.	442.	445.	442.	442.	445.	445.	449.	452.
452.	452.	452.	442.	439.	439.	442.	445.	456.	449.	456.	466.
494.	473.	484.	480.	487.	480.	484.	470.	463.	463.	484.	515.
508.	494.	526.	668.	800.	772.	575.	536.	550.	529.	522.	529.
543.	656.	697.	707.	626.	609.	594.	579.	556.	546.	539.	479.
532.	518.	540.	754.	554.	582.	571.	543.	529.	526.	519.	568.
603.	636.	720.	724.	728.	772.	668.	636.	648.	652.	620.	628.
620.	557.	517.	471.	347.	275.	265.	222.	185.	167.	178.	177.
197.	187.	175.	124.	143.	173.	204.	389.	459.	416.	312.	238.
238.	241.	217.	181.	176.	257.	323.	297.	273.	243.	217.	200.
189.	187.	198.	165.	144.	151.	164.	174.	164.	148.	117.	117.
126.	121.	111.	107.	129.	165.	170.	149.	141.	161.	183.	190.
180.	200.	198.	245.	304.	275.	292.	291.	277.	262.	245.	241.
264.	250.	229.	219.	175.	160.	157.	155.	146.	142.	140.	141.
146.	157.	164.	163.	156.	146.	140.	126.	113.	115.	131.	124.
132.	136.	165.	182.	220.	218.	175.	162.	174.	194.	195.	189.
193.	182.	206.	225.	239.	247.	248.	230.	242.	220.	227.	241.
255.	309.	339.	357.	384.	405.	355.	336.	320.	281.	264.	258.
259.	250.	263.	280.	275.	257.	255.	228.	231.	225.	226.	215.
220.	209.	195.	183.	171.	158.	152.	146.	142.	319.	317.	311.
305.	305.	314.	349.	364.	442.	439.	421.	394.	388.	388.	403.
427.	430.	415.	370.	370.	358.	349.	346.	343.	346.	337.	325.
328.	325.	322.	331.	352.	364.	364.	352.	358.	403.	408.	417.
424.	449.	425.	430.	436.	424.	410.	414.	428.	428.	437.	409.
387.	391.	378.	379.	372.	368.	363.	363.	365.	374.	386.	394.
378.	368.	354.	340.	341.	345.	337.	336.	333.	342.	342.	339.
332.	337.	337.	374.	394.	381.	389.	392.	395.	391.	390.	380.
378.	386.	380.	382.	336.							

The estimated Rattlesnake Creek daily discharge data (cfs) for 1974, based upon the regression relationship $RS = 0.6258 + 0.5441 \times BIR$ where RS and BIR designate respectively the Rattlesnake and Birch Creek discharges (read from left to right).

3.	3.	3.	3.	3.	3.	3.	3.	3.	3.	3.
4.	3.	4.	4.	4.	4.	4.	4.	4.	4.	4.
4.	4.	4.	4.	4.	4.	4.	4.	4.	4.	4.
3.	4.	3.	3.	4.	4.	4.	4.	4.	3.	4.
4.	3.	3.	3.	3.	3.	3.	4.	4.	4.	4.
6.	7.	6.	4.	5.	4.	4.	4.	4.	4.	4.
4.	4.	4.	10.	4.	4.	4.	12.	16.	15.	14.
13.	11.	9.	8.	8.	9.	10.	12.	10.	9.	9.
10.	10.	12.	11.	11.	10.	15.	17.	19.	19.	17.
15.	14.	33.	45.	49.	42.	37.	38.	45.	52.	54.
13.	19.	33.	31.	33.	47.	62.	84.	93.	97.	99.
46.	39.	103.	102.	90.	98.	83.	78.	82.	77.	75.
109.	102.	48.	45.	44.	43.	43.	42.	37.	33.	35.
62.	57.	36.	37.	34.	31.	28.	27.	25.	29.	28.
32.	38.	23.	22.	21.	21.	21.	21.	21.	20.	20.
28.	25.	21.	21.	20.	19.	19.	17.	17.	14.	8.
20.	20.	7.	7.	6.	6.	6.	6.	5.	4.	4.
7.	7.	5.	5.	5.	5.	5.	5.	5.	4.	4.
5.	5.	5.	5.	5.	4.	4.	4.	4.	4.	4.
5.	4.	6.	4.	8.	7.	7.	7.	6.	6.	6.
7.	6.	6.	5.	5.	5.	5.	5.	5.	5.	6.
6.	6.	5.	6.	6.	4.	3.	2.	2.	4.	3.
7.	9.	8.	10.	8.	8.	8.	7.	6.	6.	3.
3.	3.	2.	2.	2.	2.	2.	2.	3.	3.	3.
5.	4.	4.	3.	3.	3.	3.	3.	3.	3.	3.
3.	3.	3.	3.	4.	4.	4.	4.	4.	4.	4.

The estimated Blacktail Deer Creek daily discharge data (cfs) for 1974 is based upon the polynomial relationship $BLDC = 15.04 + 4.45 \times BIR - 0.03096 \times BIR^2 + 0.00006973 \times BIR^3$ where BLDC and BIR respectively designate Blacktail and Birch Creek discharge (read from left to right).

32.	32.	32.	32.	32.	32.	32.	32.	37.	37.	37.	37.
32.	37.	41.	45.	45.	41.	45.	45.	41.	45.	45.	45.
45.	45.	41.	41.	41.	41.	41.	41.	41.	41.	41.	41.
41.	41.	41.	41.	41.	41.	41.	41.	41.	41.	41.	41.
37.	41.	37.	37.	37.	37.	37.	37.	41.	41.	37.	41.
45.	37.	37.	37.	37.	37.	37.	37.	41.	41.	41.	41.
37.	41.	57.	86.	57.	45.	41.	45.	45.	45.	45.	45.
53.	60.	53.	45.	49.	45.	41.	41.	41.	41.	41.	41.
41.	41.	45.	49.	45.	45.	45.	105.	120.	114.	111.	105.
99.	89.	75.	71.	71.	79.	86.	92.	82.	75.	75.	75.
82.	82.	79.	79.	89.	89.	82.	125.	133.	143.	131.	125.
114.	105.	95.	89.	89.	86.	82.	82.	79.	75.	75.	82.
102.	133.	196.	210.	215.	205.	194.	198.	210.	218.	223.	219.
211.	199.	186.	178.	186.	212.	223.	221.	219.	219.	219.	220.
224.	221.	221.	221.	220.	220.	221.	222.	221.	222.	223.	223.
223.	221.	214.	210.	208.	207.	206.	205.	194.	184.	190.	191.
181.	198.	193.	195.	189.	178.	169.	165.	159.	174.	173.	171.
169.	159.	152.	148.	143.	143.	143.	143.	141.	141.	138.	138.
136.	136.	143.	143.	138.	133.	131.	125.	123.	120.	108.	71.
64.	60.	60.	60.	57.	53.	53.	53.	49.	49.	45.	45.
49.	49.	49.	49.	49.	49.	49.	49.	49.	49.	45.	45.
49.	49.	49.	49.	49.	45.	45.	45.	45.	45.	45.	45.
45.	45.	41.	41.	41.	41.	53.	53.	49.	57.	57.	57.
60.	57.	57.	60.	68.	64.	60.	60.	57.	53.	57.	57.
53.	53.	53.	49.	49.	49.	49.	49.	49.	49.	53.	53.
53.	53.	49.	53.	53.	45.	37.	28.	28.	41.	41.	57.
64.	75.	71.	82.	71.	68.	68.	64.	53.	53.	49.	32.
32.	32.	28.	28.	28.	28.	28.	28.	32.	37.	37.	49.
49.	45.	41.	37.	37.	37.	32.	32.	37.	37.	37.	37.
37.	45.	53.	53.	41.	41.	45.	45.	49.	45.	45.	41.
41.	37.	37.	37.	37.	37.						

The East Bench Canal daily discharge data (cfs) for 1974 as obtained from the East Bench Irrigation District office in Dillon, Montana (read from left to right).

[illegible]

The Clark Canyon Canal discharge data (cfs) for 1974 as obtained from the East Bench Irrigation District office in Dillon, Montana (read from left to right).

[illegible]

The estimated total daily diversion (cfs) for 1974 is based upon the approximation $DIV = 0.456 \times EB + 7.4 \times CC$ where DIV, EB, and CC designate respectively daily diversion, East Bench, and Clark Canyon discharges (read from left to right).

[illegible]

The estimated total daily return (cfs) for 1974, based upon $RET = DIV - INF + OUT$ where RET, DIV, INF, and OUT designate respectively daily return, daily diversion, daily inflow, and daily outflow (read from left to right).

72.	88.	75.	45.	30.	35.	15.	-15.	-40.	-60.	-89.	-1.
67.	125.	159.	188.	206.	166.	159.	122.	118.	97.	106.	108.
109.	109.	106.	103.	97.	103.	109.	114.	112.	112.	106.	102.
105.	105.	106.	106.	106.	110.	107.	108.	111.	111.	112.	115.
119.	114.	119.	113.	108.	105.	115.	114.	125.	123.	131.	135.
155.	128.	149.	146.	156.	153.	158.	147.	140.	138.	152.	168.
171.	153.	105.	47.	136.	330.	209.	167.	182.	172.	169.	172.
159.	171.	160.	259.	195.	183.	183.	178.	164.	156.	147.	80.
133.	118.	131.	113.	159.	148.	156.	69.	34.	25.	0.	-33.
30.	-86.	26.	426.	190.	321.	269.	123.	75.	190.	70.	30.
36.	0.	40.	120.	107.	157.	254.	280.	349.	320.	348.	335.
373.	394.	412.	334.	303.	299.	287.	452.	620.	671.	686.	557.
457.	456.	345.	347.	344.	433.	542.	549.	512.	454.	442.	384.
321.	282.	364.	391.	305.	175.	135.	247.	287.	399.	317.	313.
268.	402.	403.	485.	492.	482.	390.	442.	404.	435.	408.	275.
219.	278.	299.	355.	207.	180.	352.	492.	577.	677.	740.	709.
657.	524.	452.	444.	376.	358.	336.	312.	344.	391.	412.	411.
454.	565.	653.	615.	645.	669.	574.	430.	394.	354.	380.	368.
387.	592.	627.	701.	749.	669.	640.	633.	632.	665.	717.	747.
743.	739.	781.	820.	818.	828.	822.	736.	714.	619.	606.	673.
670.	741.	779.	782.	733.	761.	655.	633.	617.	583.	563.	555.
552.	546.	560.	579.	577.	565.	574.	558.	560.	553.	545.	560.
633.	652.	617.	637.	549.	489.	480.	473.	446.	776.	769.	760.
738.	747.	747.	758.	745.	843.	851.	773.	600.	412.	385.	377.
310.	317.	338.	309.	309.	297.	288.	285.	99.	102.	89.	77.
80.	73.	74.	79.	97.	121.	140.	144.	141.	168.	173.	164.
163.	171.	148.	141.	159.	154.	144.	153.	180.	180.	193.	202.
174.	172.	164.	165.	158.	154.	149.	149.	143.	144.	152.	150.
137.	132.	130.	124.	117.	119.	118.	119.	109.	115.	115.	113.
109.	98.	82.	78.	114.	111.	110.	106.	104.	106.	105.	105.
106.	113.	106.	109.	67.							

The estimated total daily inflow (cfs) for 1974, calculated from the sum of the Barretts, Rattlesnake and Blacktail Deer daily discharges (read from left to right).

255.	245.	255.	255.	250.	235.	235.	245.	260.	270.	295.	220.
216.	221.	232.	254.	302.	300.	297.	296.	288.	294.	294.	295.
294.	294.	294.	297.	315.	318.	321.	319.	318.	318.	324.	340.
340.	340.	339.	336.	336.	335.	335.	334.	334.	334.	337.	337.
333.	338.	333.	329.	331.	334.	327.	331.	331.	326.	325.	331.
339.	345.	335.	334.	331.	327.	326.	323.	323.	325.	332.	347.
337.	341.	421.	621.	664.	442.	366.	369.	368.	357.	353.	357.
384.	485.	537.	448.	431.	426.	411.	401.	392.	390.	392.	399.
399.	400.	409.	441.	437.	434.	415.	474.	495.	501.	519.	601.
573.	722.	694.	693.	797.	823.	836.	847.	831.	806.	826.	890.
892.	886.	876.	876.	888.	913.	914.	937.	988.	1013.	1027.	1068.
1075.	1058.	1033.	1060.	1110.	1136.	1192.	1212.	1108.	1002.	896.	964.
1068.	1098.	1187.	1236.	1284.	1257.	1108.	1012.	1025.	1034.	1043.	1109.
1093.	1067.	971.	906.	989.	1137.	1238.	1299.	1332.	1305.	1355.	1378.
1453.	1433.	1444.	1503.	1510.	1528.	1504.	1430.	1464.	1450.	1336.	1272.
1231.	1169.	1068.	1043.	1027.	978.	952.	939.	922.	885.	899.	931.
929.	964.	1023.	1050.	1071.	1074.	1102.	1173.	1234.	1274.	1281.	1299.
1297.	1224.	1129.	1075.	1037.	1007.	1019.	1111.	1122.	1160.	1158.	1159.
1166.	1109.	1088.	1009.	935.	899.	889.	880.	867.	834.	777.	729.
706.	686.	674.	624.	595.	558.	550.	544.	541.	540.	536.	550.
578.	561.	527.	500.	500.	492.	459.	438.	420.	408.	403.	419.
445.	446.	445.	445.	445.	439.	438.	437.	435.	438.	455.	483.
482.	483.	499.	472.	461.	459.	457.	458.	467.	316.	324.	320.
324.	316.	320.	328.	337.	316.	300.	300.	296.	295.	300.	300.
299.	295.	259.	244.	244.	244.	244.	244.	244.	244.	248.	248.
248.	252.	248.	252.	255.	243.	224.	208.	217.	235.	235.	253.
261.	278.	277.	289.	277.	270.	266.	261.	248.	248.	244.	207.
213.	219.	214.	214.	214.	214.	214.	214.	222.	230.	234.	244.
241.	236.	224.	216.	224.	226.	219.	217.	224.	227.	227.	226.
223.	234.	255.	296.	280.	270.	279.	286.	291.	285.	285.	275.
272.	273.	274.	273.	269.							

The estimated total daily outflow (cfs) for 1974, calculated from the approximation
 $OUT = BLA + 0.544 \times EB + 0.2CC$ where OUT, BLA, EB, and CC designate respectively the
outflow, Blaine, East Bench, and Clark Canyon discharges (read from left to right).

327.	333.	330.	300.	280.	270.	250.	230.	220.	210.	206.	219.
283.	346.	391.	442.	508.	466.	456.	418.	406.	391.	400.	403.
403.	403.	400.	400.	412.	421.	430.	433.	430.	430.	430.	442.
445.	445.	445.	442.	442.	445.	442.	442.	445.	445.	449.	452.
452.	452.	452.	442.	439.	439.	442.	445.	456.	449.	456.	466.
494.	473.	484.	480.	487.	480.	484.	470.	463.	463.	484.	515.
508.	494.	526.	668.	800.	772.	575.	536.	550.	529.	522.	529.
543.	656.	697.	707.	626.	609.	594.	579.	556.	546.	539.	479.
532.	518.	540.	554.	596.	582.	571.	543.	529.	526.	519.	568.
603.	636.	720.	734.	735.	782.	713.	668.	674.	690.	656.	669.
665.	614.	577.	536.	436.	373.	386.	366.	344.	333.	353.	368.
398.	391.	378.	327.	346.	375.	414.	599.	665.	624.	523.	456.
458.	472.	448.	417.	423.	505.	569.	532.	508.	480.	458.	447.
435.	422.	423.	387.	376.	381.	409.	436.	429.	420.	388.	399.
411.	412.	403.	402.	424.	460.	459.	434.	427.	450.	466.	464.
450.	457.	440.	479.	527.	441.	508.	502.	489.	477.	468.	471.
495.	483.	466.	468.	426.	411.	413.	422.	425.	423.	426.	428.
433.	447.	454.	451.	447.	435.	427.	412.	396.	400.	416.	411.
421.	428.	449.	462.	489.	477.	436.	421.	426.	444.	437.	425.
420.	410.	433.	432.	429.	426.	423.	400.	411.	388.	383.	394.
402.	456.	479.	498.	523.	543.	488.	456.	435.	395.	378.	380.
393.	386.	399.	417.	414.	396.	399.	378.	379.	374.	379.	370.
377.	366.	350.	333.	316.	300.	292.	286.	278.	448.	448.	438.
426.	426.	432.	459.	463.	541.	535.	488.	407.	396.	396.	410.
432.	435.	420.	375.	375.	363.	354.	351.	343.	346.	337.	325.
328.	325.	322.	331.	352.	364.	364.	352.	358.	403.	408.	417.
424.	449.	425.	430.	436.	424.	410.	414.	428.	428.	437.	409.
387.	391.	378.	379.	372.	368.	363.	363.	365.	374.	386.	394.
378.	368.	354.	340.	341.	345.	337.	336.	333.	342.	342.	339.
332.	332.	337.	374.	394.	381.	389.	392.	395.	391.	390.	380.
378.	386.	380.	382.	336.							

D378
N514
cop.2