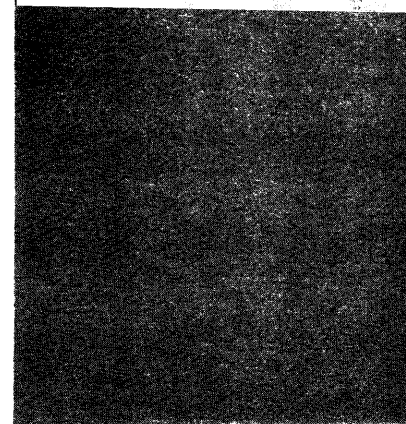


Daily Water and Sediment Discharges From Selected Rivers of the Eastern United States: A Time-Series Modeling Approach



United States
Geological
Survey
Water-Supply
Paper 2216



Daily Water and Sediment Discharges From Selected Rivers of the Eastern United States: A Time-Series Modeling Approach

By MICHAEL G. FITZGERALD and MICHAEL R. KARLINGER

UNITED STATES DEPARTMENT OF THE INTERIOR
JAMES G. WATT, Secretary

GEOLOGICAL SURVEY
Dallas L. Peck, Director



UNITED STATES GOVERNMENT PRINTING OFFICE: 1983

For sale by Distribution Branch
Text Products Section
U.S. Geological Survey
604 South Pickett Street
Alexandria, Virginia 22304

Library of Congress Cataloging in Publication Data

Fitzgerald, Michael G.

Daily water and sediment discharges from selected rivers of
the eastern United States.

(Geological Survey water-supply paper ; 2216)

Bibliography: p.

Supt. of Docs. no.: I 19.13:

1. Stream measurements—United States—Mathematical
models. 2. Sediment transport—United States—
Mathematical models. 3. Time-series analysis—
Mathematical models. I. Karlinger, Michael R.
II. Title. III. Series.

GB1216.F57 1982 551.48*3*0724 82-600163

CONTENTS

Abstract	1
Introduction	1
Acknowledgments	1
Review of previous work	1
Scope	2
Theory	2
Spectral analysis	2
Time-series models	5
Univariate models	5
Transfer-function models	6
Results	7
Data compilation	7
Spectral results	8
Time-series models	11
Univariate models	14
Transfer-function models	14
Accuracy of model prediction	19
Discussion	19
Summary	19
References	23
Metric conversion table	24

FIGURES

1. Map showing location of rivers included in study 3
- 2-14. Graphs showing
 2. Water-discharge (log transformed) spectra for the (A) Delaware River at Trenton, N.J., (B) Juniata River at Newport, Pa., and (C) Yadkin River at Yadkin College, N.C. 9
 3. Sediment-discharge (log transformed) spectra for the (A) Delaware River at Trenton, N.J., (B) Juniata River at Newport, Pa., and (C) Yadkin River at Yadkin College, N.C. 9
 4. Cross-correlation function for Delaware River at Trenton, N.J., log-transformed water and sediment discharge 10
 5. Cospectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J. 10
 6. Coherency spectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J. 11
 7. Phase spectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J. 12
 8. Gain function between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J. 12
 9. Autocorrelation function for water discharge, Tar River at Tarboro, N.C., log transformed and undifferenced 13
 10. Autocorrelation function for water discharge, Tar River at Tarboro, N.C., log transformed and differencing of order 1 13

11. Autocorrelation (*A*) and partial-autocorrelation (*B*) functions for water discharge (log transformed and first differenced), Tar River at Tarboro, N.C. 15
12. Autocorrelation (*A*) and partial-autocorrelation (*B*) functions for water discharge (log transformed and first differenced), Monocacy River at Frederick, Md. 16
13. Standardized step-response weights as a function of lag 21
14. Sediment concentration-water discharge relations for (*A*) Juniata River at Newport, Pa., and (*B*) Yadkin River at Yadkin College, N.C. 21

TABLES

1. Summary of drainage-basin characteristics 2
2. Summary of harmonic analysis 8
3. Cross-spectral characteristics between log-transformed daily discharge of water and sediment 9
4. Estimates of parameters for univariate models 17
5. Estimates of parameters for transfer-function model 18
6. Predictions of logarithms of storm-of-record discharge for Delaware River at Trenton, N.J., using transfer-function model 20
7. Frequency of suspended-sediment discharge 22

SYMBOLS AND DEFINITIONS

		Units
a_t	Univariate-model, white-noise process	—
\hat{a}_t	Univariate-model residuals	—
$A(\omega)$	Amplitude gain	—
ACF	Autocorrelation function	—
AR	Autoregressive	—
B	Backward-shift operator	—
c_{xx}	Autocovariance for series x	—
c_{xy}	Cross-covariance between x and y	—
$C(\omega)$	Cospectrum	—
d	Order of differencing process	—
E	Expectation operator	—
e_t	White-noise process from transfer function of residuals	—
$G_{xx}(\omega)$	Power spectrum of x	—
$G_{yy}(\omega)$	Power spectrum of y	—
G_{yx}	Cross-spectrum between x and y	—
$H(\omega)$	Frequency-response function	—
k	Summation index	—
MA	Moving average	—
N	Length of data set	—
p	Order of autoregressive process	—
$PACF$	Partial-autocorrelation function	—
q	Order of moving-average process	—
Q	Portmanteau-autocorrelation statistic	—
$Q(\omega)$	Quadrature spectrum	—
R^2	Percent variance explained	—
r_{xx}	Autocorrelation coefficient of x	—
r_{xy}	Cross-correlation between x and y	—
S	Portmanteau cross-correlation statistic	—
s	Lag value	Day
t	Time	Day
V_k	Step-response weight	—
x_t	Daily water discharge	Cubic feet per second
\bar{x}	Mean of series x	Cubic feet per second
$X(\omega)$	Fast Fourier Transform of x	—
$X^*(\omega)$	Conjugate of $X(\omega)$	—
y_t	Daily sediment discharge	Tons per day
\bar{y}	Mean of series y	Tons per day
$Y(\omega)$	Fast Fourier Transform of y	—
$Y^*(\omega)$	Conjugate of $Y(\omega)$	—
z_t	A general random process	—
ϵ_t	Noise series from frequency-response model	—
$\gamma^2(\omega)$	Coherence	—
$\delta(B)$	Output operator	—
η_t	Transfer-function white-noise process	—
$\hat{\eta}_t$	Transfer-function residuals	—
θ	Moving-average parameter	—
$\theta(B)$	Moving-average operator	—
$\theta(\omega)$	Phase angle	Degrees
σ_w^2	Variance of water-discharge record	(Cubic feet per second) ²
σ_s^2	Variance of sediment-discharge record	(Cubic feet per second) ²
$v(B)$	Transfer function	—

v_k	Impulse-response weight	—
$\hat{v}(B)$	Transfer function for residuals	—
ϕ	Autoregressive parameter	—
$\phi(B)$	Autoregressive operator	—
χ^2	Chi-square	—
ω	Angular frequency	Radians per unit time
$\omega(B)$	Input operator	—

Daily Water and Sediment Discharges From Selected Rivers of the Eastern United States: A Time-Series Modeling Approach

By Michael G. Fitzgerald and Michael R. Karlinger

Abstract

Time-series models were constructed for analysis of daily runoff and sediment discharge data from selected rivers of the Eastern United States. Logarithmic transformation and first-order differencing of the data sets were necessary to produce second-order, stationary time series and remove seasonal trends.

Cyclic models accounted for less than 42 percent of the variance in the water series and 31 percent in the sediment series. Analysis of the apparent oscillations of given frequencies occurring in the data indicates that frequently occurring storms can account for as much as 50 percent of the variation in sediment discharge.

Components of the frequency analysis indicate that a linear representation is reasonable for the water-sediment system. Models that incorporate lagged water discharge as input prove superior to univariate techniques in modeling and prediction of sediment discharges. The random component of the models includes errors in measurement and model hypothesis and indicates no serial correlation. An index of sediment production within or between drainage basins can be calculated from model parameters.

INTRODUCTION

Temporal variations in water and sediment discharges are essential considerations in the design and operation of water-resource developments. However, the historic records of streamflow and sediment yield commonly exclude critical periods of flood or drought. Models that preserve these relevant statistics of the historical record make possible the prediction of system response to major floods or droughts and thus aid in development design.

Where long-term water and sediment records are available, time series and systems concepts offer a natural approach to the analysis, modeling, and simulation of the relations between runoff and sediment yield. Time-series techniques preserve the mean, variance, and autocorrelation structure of historical records; these techniques are simple and require a minimum number of parameters to represent adequately the process for most practical applications.

This paper presents a method for constructing discrete sediment-yield models that incorporate the input-output relationships between hydrologic series. Long-term records of water and sediment discharge for seven rivers in the Eastern United States were analyzed as weakly stationary time series, and the results are discussed in terms of the relationships between statistical properties of the series and physical characteristics of the drainage basin.

Acknowledgments

This study was part of the U.S. Geological Survey Water Resources Division research program on sedimentary processes in alluvial channels. The writers express their deep appreciation for the counsel and encouragement of Robert H. Meade, whose own work in fluvial processes provided the impetus for this work. Special thanks are due Brent Troutman, Carl Nordin, and William Emmett for their useful suggestions and stimulating discussions, and to Margaret Roberts and Herb Stevens for their programming assistance.

Review of Previous Work

Time-series techniques have been used in the analysis and planning of water-resource developments for several years. Matalas (1967) summarizes time-series modeling of streamflow and emphasizes the models' utility in synthesizing hydrologic records of any desired length for evaluation and optimization of development-design variables. In addition to their utility in synthesizing hydrologic records for long-term planning, models also can be useful for short-term predictions. Delleur and others (1976) provide a good discussion of hydrologic time-series modeling with respect to time resolution and the conditions under which prediction and simulation can best be achieved.

Rodriguez-Iturbe and Nordin (1968, 1974) used the causal relationship between water discharge and sediment discharge in a bivariate-modeling technique to examine the prediction of sediment discharges from a given drainage basin. Though the majority of their work

was confined to frequency-domain analysis of the hydrologic system, it started time-series modeling efforts in sediment hydrology. Sharma and Dickinson (1980) used transfer-function techniques to model sediment yield as a function of water discharge and related the step-response weights of the impulse-response function to drainage-basin characteristics. The study presented here applies the systems concepts described by Sharma and Dickinson (1980), with the additional aspects of extended record lengths and expanded spatial coverage to address the relations of model properties to hydrologic characteristics.

Scope

Published records of daily water and sediment discharge (U.S. Geol. Survey, 1956–70; 1971–79) for seven rivers in the Eastern United States (fig. 1 and table 1) were used in this study. These rivers were selected for several reasons. Because all sampling stations are located in the Middle Atlantic States, there is a gross homogeneity in physiographic variables such as climate, topography, and soil type. This homogeneity permits qualitative comparisons of model specifications and parameter values to differences in such factors as land-use practices and in-stream regulation. Qualitative studies of sediment yield, sediment storage, and land-use practices have been made for several of these rivers (Meade, 1982), providing a basis for these comparisons. Finally, the records used are among the longest available for simultaneous measurements of water and sediment discharge, ranging from a minimum of 2,048 daily values to a maximum of 8,192 daily values. For all seven rivers, the available records are longer than the records used for this study. The statistics in the last four columns of table 1 apply to the period of record used and not to the entire period of available record.

Spectral and cross-spectral techniques were applied to all seven rivers for initial examination of the relation between water and sediment discharges. In addition to the application of the spectral techniques, univariate time-series models were identified for each of the water and sediment-discharge series. Causal relations between members of each of the seven pairs of models were investigated, using the lagged cross-correlation values between the appropriate univariate-model residuals. Properties of the cross-correlations together with the univariate-model parameters subsequently permitted estimation of transfer-function models of the water-sediment systems for each of the rivers.

THEORY

Spectral Analysis

Spectral and cross-spectral techniques are exploratory tools for system identification, providing information on series oscillations and indicating interrelations between series. The theory is well described in the literature (Jenkins and Watts, 1968; Otnes and Enochson, 1972; Bloomfield, 1976); only estimation procedures will be outlined here.

Let x_t be a finite, discrete, stationary time series of Length N and mean μ . The average lagged product of x_t , that is, $E[(x_t - \mu)(x_{t+s} - \mu)]$, is called the autocovariance function and is estimated by

$$c_{xx}(s) = \frac{1}{N} \sum_{t=1}^{N-s} (x_t - \bar{x})(x_{t+s} - \bar{x}), s=0, 1, \dots, N-1 \quad (1)$$

where

Table 1. Summary of drainage-basin and water and sediment discharge characteristics

[mi² = square miles; tons/d = tons per day; ft³/s = cubic feet per second; σ_s^2 = variance of daily mean sediment discharge; σ_w^2 = variance of daily mean water discharge]

River	Drainage area (mi ²)	Station location	Period of record used	Average sediment discharge (tons/d)	σ_s^2 (tons/d) ²	Average water discharge (ft ³ /s)	σ_w^2 (ft ³ /s) ²
Delaware	6,780	Trenton, N.J.	10/01/56–03/06/79	1,173	1.0×10^8	11,295	1.2×10^8
Juniata	3,354	Newport, Pa.	10/01/56–03/06/79	731	3.1×10^7	4,137	3.5×10^7
Monocacy	817	Frederick, Md.	07/14/68–09/30/79	667	1.5×10^7	1,172	6.6×10^6
Potomac	9,651	Point of Rocks, Md.	07/14/68–09/30/79	3,745	4.9×10^8	10,891	2.3×10^8
Rappahannock	620	Remington, Va.	10/01/56–03/06/79	272	2.9×10^6	667	1.3×10^6
Tar	2,140	Tarboro, N.C.	01/09/58–08/18/63	427	4.2×10^5	2,850	8.9×10^6
Yadkin	2,280	Yadkin College, N.C.	10/01/56–03/06/79	2,698	6.8×10^7	3,193	1.0×10^7

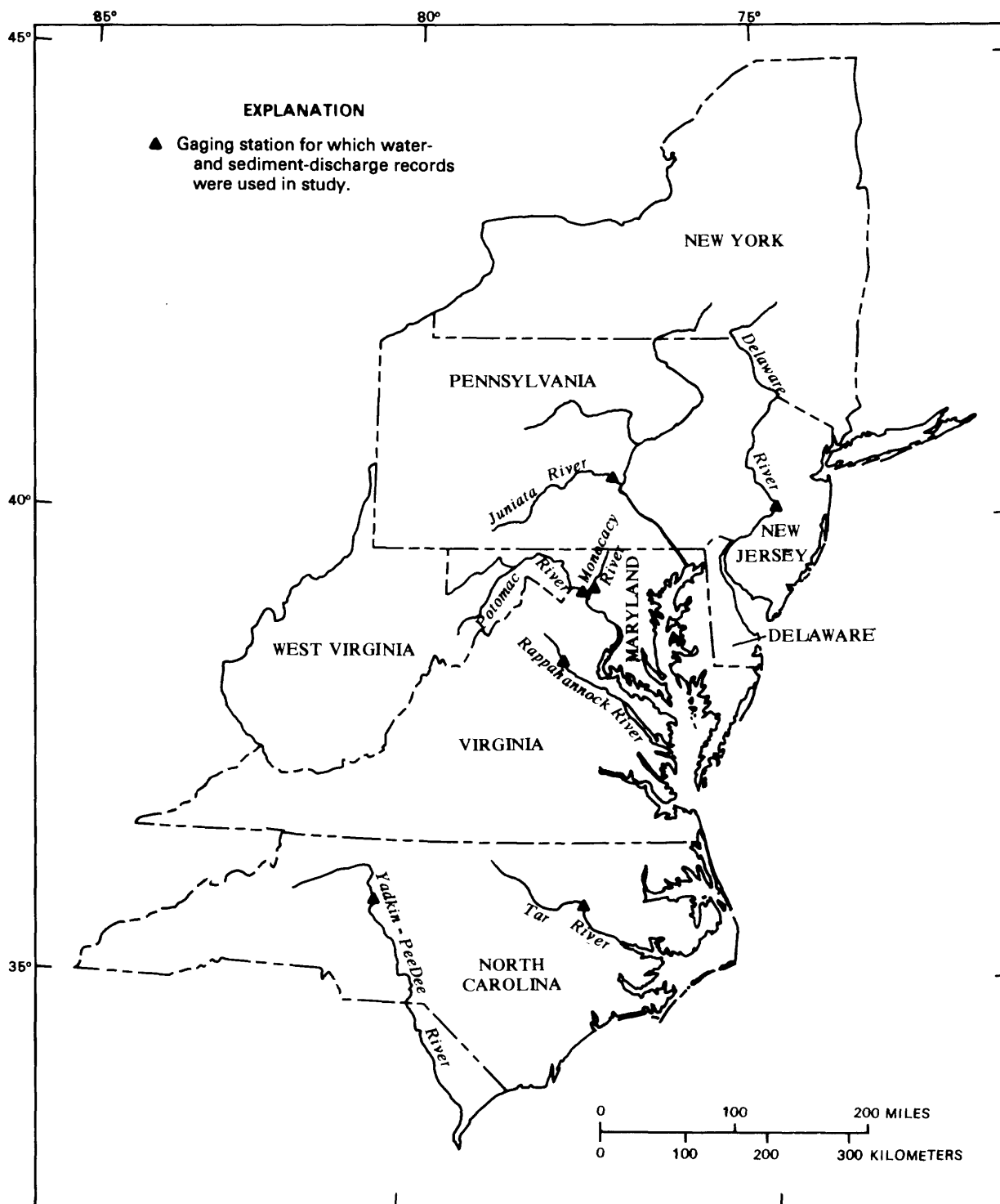


Figure 1. Location of rivers included in study.

$c_{xx}(s)$ is the autocovariance function;

\bar{x} is the sample mean, an estimator of μ ;

x_t is a member of the series;

and

s is the lag.

The estimated autocorrelation at lag s then is defined:

$$r_{xx}(s) = c_{xx}(s) / c_{xx}(0). \quad (2)$$

A plot of the autocorrelation function as a function of lag s indicates the extent to which one value of the process is correlated with previous values. Thus, the autocorrelation function can be used to measure the length and strength of the "memory" of the process, that is, the extent to which the value at time t depends on that of time $t-s$. A "white noise" sequence is one in which $r_{xx}(0) = 1$ and $r_{xx}(s) = 0$ for $s \neq 0$.

The Fourier transform of the autocorrelation function produces its frequency domain equivalent, the periodogram. The periodogram provides information on the oscillations of the time series by showing how the variance of a series, consisting of mixtures of sines and cosines, is distributed among distinct harmonic frequencies. The periodogram is defined as

$$(1/N)[X(\omega)X^*(\omega)] = (1/N)|X(\omega)|^2, \quad (3)$$

where

ω is angular frequency;

$X(\omega)$ is the Fourier transform of $x(t)$,

that is,

$$X(\omega) = \int_{-\infty}^{\infty} x_t e^{-i\omega t} dt;$$

and

$X^*(\omega)$ is the conjugate of $X(\omega)$.

Computation of the Fourier transform is accomplished through the Fast Fourier Transform algorithm (Bloomfield, 1976).

The periodogram is an erratic function of frequency; some method of averaging the available information needs to be used to obtain a consistent spectral estimate, $G_{xx}(\omega)$. One way of doing the averaging is to compute the periodogram and then perform a smoothing opera-

tion on the periodogram to obtain the desired estimate. In this study, smoothing is accomplished by averaging over three frequencies, using a weighting function with coefficients 0.25, 0.50, and 0.25.

This treatment can be extended to two stochastic processes. The cross-covariance function between the series x_t and y_t can be estimated according to the formula

$$c_{xy}(s) = (1/N) \sum_{t=1}^{N-s} (x_t - \bar{x})(y_{t+s} - \bar{y}) \quad s \geq 0 \quad (4)$$

$$= (1/N) \sum_{t=1}^{N-s} (x_{t-s} - \bar{x})(y_t - \bar{y}) \quad s \leq 0. \quad (5)$$

Similarly, the cross-correlation at lag s is defined as

$$r_{xy}(s) = c_{xy}(s) / \sqrt{c_{xx}(0)} \sqrt{c_{yy}(0)}. \quad (6)$$

The cross-spectrum, G_{yx} , defined as the Fourier transform of the cross-covariance function, can be estimated by smoothing

$$(1/N)[Y(\omega)X^*(\omega)], \quad (7)$$

where

ω is angular frequency;

$Y(\omega)$ is the Fast Fourier Transform of y_t ;

and

$X^*(\omega)$ is the conjugate of the Fast Fourier Transform of x_t .

The complex-valued cross-spectrum can be represented by

$$G_{yx}(\omega) = C(\omega) + iQ(\omega), \quad (8)$$

where

$C(\omega)$ is the co-spectrum, a measure of the in-phase covariance;

and

$Q(\omega)$ is the quadrature spectrum, a measure of the out-of-phase covariance.

$C(\omega)$ measures the contribution of oscillations of different frequencies to the total cross-variance at lag zero between two time series. Conversely, $Q(\omega)$ shows the

contribution of the different harmonics to the total cross-variance, when all harmonics of the series x_t are delayed by a quarter-period and y_t remains unchanged (Rodríguez-Iturbe and Nordin, 1968).

The real quantity defined as coherence, $\gamma^2(\omega)$, is a measure of the correlation between frequency components of the two processes:

$$\gamma^2(\omega) = \frac{|G_{yx}(\omega)|^2}{G_{xx}(\omega)G_{yy}(\omega)}, \quad 0 \leq \gamma(\omega) \leq 1, \quad (9)$$

where

$G_{xx}(\omega)$ is the spectrum of x_t ,

and

$G_{yy}(\omega)$ is the spectrum of y_t .

The frequency-response function, $H(\omega)$, describes the input-output relations between series. For example, if $x_t(\omega)$ and $y_t(\omega)$ are the components of x_t and y_t at frequency ω , models can be constructed of the form $y_t(\omega) = H(\omega)x_t(\omega) + \epsilon_t(\omega)$, where $\epsilon_t(\omega)$ is the noise that results from nonlinear effects. If $\epsilon_t(\omega)$ is negligible, the frequency response of the system can be calculated from

$$H(\omega) = \frac{G_{yx}(\omega)}{G_{xx}(\omega)} = A(\omega)e^{i\theta(\omega)}, \quad (10)$$

where $A(\omega)$ is the amplitude gain of the system (the ratio of the amplitudes of the frequency components) and $\theta(\omega)$ is the phase difference between frequency components of the two processes, given by

$$\theta(\omega) = \arctan [Q(\omega)/C(\omega)]. \quad (11)$$

The inverse Fourier transform of the frequency-response function is the impulse-response function.

Time-Series Models

While spectral-analysis techniques provide good exploratory tools for examining the properties within, and relationships between, series of data, they are somewhat deficient for developing stochastic models. These models need to include an error structure that can be statistically analyzed to make predictions with confidence limits or to synthesize stochastic data streams, while preserving the statistics of the observed series. Time-series techniques permit models that provide these capabilities, and these techniques require only a small number of model parameters.

The estimation of model parameters requires that a time series satisfy the assumptions of ergodicity and weak stationarity. Stationarity dictates that the mean and variance of the elements of the series do not change with time and that the covariance between any two points of the series is a function only of the time difference between points. Ergodicity requires that values far apart in the series are virtually uncorrelated (Granger and Newbold, 1977).

Univariate Models

Univariate models of stationary time series generally are one of two types: autoregressive or moving-average models. If a variable, z_t , is a function only of its past values and a current random shock, then it is said to follow an autoregressive process. The general form of an autoregressive process is

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t, \quad (12)$$

where the ϕ_i are the model parameters and a_t is a random variable from a zero mean white-noise process. The order of the autoregressive process is denoted as p . If we define an autoregressive operator of order p by

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (13)$$

where

$$B^m z_t = z_{t-m},$$

then the autoregressive model may be written economically as

$$\phi(B)z_t = a_t. \quad (14)$$

If, however, the variable z_t is a function of only current and lagged values of a_t , then it is said to be a moving-average process of the following form:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \quad (15)$$

The θ_i are the parameters of the moving-average model, and the order of the process is denoted as q . If we define a moving-average operator of order q by

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q, \quad (16)$$

then the moving-average model may be written as

$$z_t = \theta(B)a_t. \quad (17)$$

These models can be combined into an autoregressive moving-average model:

$$z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \dots - \phi_p z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (18)$$

This model can be written economically as

$$\phi(B)z_t = \theta(B)a_t,$$

and it is referred to as a mixed model of order p and q .

Often time-series sequences that have nonstationary characteristics can be differenced to produce a stationary series. New variables are defined by

$$z'_t = z_t - z_{t-d} \quad (19)$$

where z'_t is a d th-order difference. A mixed model with a d th-order differenced variable is an autoregressive integrated moving-average model and is of the form

$$z'_t - \phi_1 z'_{t-1} - \dots - \phi_p z'_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (20)$$

Seasonality within a time series, like nonstationary characteristics, can sometimes be removed by differencing. Seasonal time series with a period s have similarities in observations that are s time-units apart. Therefore, a seasonal nonstationarity may be eliminated by computing a differenced variable:

$$z_t^s = z_t - z_{t-s} \quad (21)$$

There are three steps to constructing successfully a time-series model from stationary data: (1) model identification, (2) parameter estimation, and (3) diagnostic checking. Although these are listed as three separate steps, there is necessarily some overlap between steps. For a more complete explanation, see Box and Jenkins (1976) and Granger and Newbold (1977).

Identification of the proper model is based on the form of the autocorrelation and partial autocorrelation functions of the series. The partial autocorrelation at lag p is defined as the p th coefficient of an autoregressive model of order p fit to the data (see eq. 12). The partial autocorrelation function then is represented by a plot of these coefficients as a function of lag p (Box and Jenkins, 1976).

The autocorrelation function or partial autocorrelation function can appear to be a damped exponential, a sinusoidal curve, or both, or these functions can appear to "cut off" after k lags. If the autocorrelation function decreases to zero or is damped while the partial autocorrelation function cuts off, the model is identified as autoregressive. If the autocorrelation function cuts off and the partial autocorrelation function decreases to

zero, then the model is moving average. If both functions occur similarly, then a mixed model may be in order. In addition, a very slow decrease to zero of the autocorrelation function may indicate that a differencing operation is needed. The identification process in time-series model building is thus somewhat subjective.

Initial estimation of model parameters is based on initial values of the sample autocorrelation function. These estimates then are refined and optimized based on the value of the residual sum of the squares, which is obtained by checking actual data values with model-predicted values.

The final step in the time-series model building is diagnostic checking. After a set of parameters is optimized, the structure of the model residuals is checked. These checks are made to insure that the residuals, \hat{a}_t (that is, actual value minus estimated value), have an approximately normal distribution and that the autocorrelation function of the residuals indicates a white-noise process; that is, that no autocorrelation exists at lag greater than zero. An approximate test of model adequacy may be made by computing a value of Q , the Portmanteau autocorrelation statistic, defined as

$$Q = N \sum_{k=1}^K r_{\hat{a}\hat{a}}^2(k) \quad (22)$$

and distributed as χ^2 (Box and Jenkins, 1976). In this equation, N is unadjusted degrees of freedom and $r_{\hat{a}\hat{a}}(k)$ is the autocorrelation value at lag k of the residuals \hat{a}_t ; K is assumed to be sufficiently long, so that the weights Ψ in the model, written in the form $\tilde{\omega}_t = \phi^{-1}(B)\theta(B)a_t = \Psi(B)a_t$, will be negligibly small for $j > k$ (Box and Jenkins, 1976).

Transfer-Function Models

Bivariate modeling is an extension of univariate time-series model building. A special case of a bivariate time-series model with broad applications in hydrology involves one stochastic variable as input into a dynamic (unidirectional) input-output system, and a second variable as the resulting output. In such a case, there is no feedback from the output variable to the input variable, and the process can be modeled by using a transfer function. The general form of the model is

$$y_t = v(B)x_t + \eta_t, \quad (23)$$

where

y_t is the output process;

x_t is the input process;

η_t is a white-noise process independent of $v(B)x_t$;

and

$v(B)$ is the transfer function (Box and Jenkins, 1976).

The lag operator, B , indicates that the output y_t can be a function of previous as well as present values of x_t [that is, $v(B)x_t = v_0x_t + v_1x_{t-1} + v_2x_{t-2} + \dots$]. Because $v(B)$ can be written as a ratio of two polynomials, y_t can be a function of previous values of y_t as well as x_t . When this ratio is expressed as

$$v(B) = \frac{\omega(B)}{\delta(B)}, \quad (24)$$

the transfer-function model can be rewritten as

$$y_t = \frac{\omega(B)}{\delta(B)} x_t + \eta_t \quad (25)$$

because of the independence between η_t and $v(B)$. There may be a delay parameter (b) for x_t , in which case x_{t-b} is substituted for x_t . Hence, $\delta(B)$ is an operator of order r in the output series, and $\omega(B)$ is an operator of order s in the input series. The three stages for transfer-function model building, as for univariate model building, are identification, estimation, and diagnostic checking.

Unidirectional causality between two processes can be confirmed through examination of the lagged cross-correlation between the processes. Although the cross-correlation between the actual series could give some information as to causality, a more appropriate signature of causality is the lagged cross-correlation between residuals of the univariately modeled processes. This procedure is called prewhitening the input. If there is unidirectional causality, there would be significant cross-correlation values only at either positive or negative lags, and the sign of the lag would indicate which process is input and which is output.

Initial estimation of transfer-function parameters is based on values of the lagged cross-correlations. These initial parameters are then optimized, until a minimum sum of the squares of residuals from the model is achieved.

As with univariate models, diagnostic checking is based on statistical properties of model residuals. However, the statistical tests not only ascertain the serial independence of the residuals from the transfer-function model, but they also investigate the crosswise lagged independence between the residuals of the transfer-function model, $\hat{\eta}_t$, and the residuals, \hat{a}_t , of the univariate model of the input process. The relevant cross-correlation

statistic is the Portmanteau cross-correlation statistic, S , defined similarly to the Q statistic:

$$S = M \sum_{k=1}^K r_{\hat{\eta}\hat{a}}^2(k), \quad (26)$$

where

M is the adjusted degrees of freedom;

$r_{\hat{\eta}\hat{a}}^2(k)$ is the cross-correlation between the above-mentioned residuals;

and

K is chosen sufficiently large (Box and Jenkins, 1976).

This statistic needs to be distributed as a χ^2 variable. The relevant serial correlation-test statistic of the transfer-function model residuals is the same as that described for the univariate model residuals.

A large value of the Portmanteau cross-correlation statistic indicates that the fitted transfer function, $v(B)$, is inadequate to represent the relationship between input and output processes. Possible remedies range from improving the univariate models to simply increasing the number of lagged parameters in the transfer function.

Failure of the transfer-function residuals to pass the χ^2 test can best be remedied by modeling the residuals series, $\hat{\eta}_t$, with a univariate model. Identification and estimation of these residual noise-model parameters are performed by using the autocorrelation function and partial autocorrelation function of the residuals series, a process identical to the previously discussed univariate modeling.

RESULTS

Data Compilation

Daily water-discharge rates, in cubic feet per second, and sediment-discharge rates, in tons per day, were obtained from the U.S. Geological Survey WATSTORE data files (U.S. Geol. Survey, 1956-70; 1971-79). Sediment loads considered here are suspended loads calculated from daily sediment samples that were used to define a continuous concentration-time relation for each station. The unmeasured sediment load that moves below the lowest sampling point in the cross section (including bedload) was not considered. However, the measured load should serve as a good index of the total sediment discharge in these rivers (Simmons, 1976).

The time-series modeling approach requires constant variance. Increased variance with increased discharge, therefore, dictates that data need to be transformed to satisfy the assumptions. Log-transformation (\log_{10}) of hydrologic data is common and is used here; reasons for this transformation are discussed in a later section, "Time-Series Models." In addition, log-transformed data were standardized by subtracting the mean and dividing by the standard deviation for spectral and cross-spectral analyses. Such standardization rendered the sequences dimensionless and offered convenience in the interpretation of spectral results. Data sets were restricted to 2^N data points (where N is an integer) to speed computation (the Fast Fourier Transform algorithm is much faster if 2^N points are used).

Spectral Results

For the rivers examined, all log-transformed water and sediment-discharge spectra display prominent peaks at the annual frequency, 0.0027 cycle per day (c/d) or 1 cycle per year, and smaller semiannual oscillations (figs. 2 and 3). Annual and semiannual cycles explain a substantially larger percentage of total variance for the water-discharge series than they do for sediment records (table 2). If cyclic models were constructed for each series, these models could explain between 17 and 40 percent of the variance for the water series, but only between 9 and 31 percent for the sediment series. (The small value for the Yadkin River water discharge may be a result of flood-control dams constructed along the river.) The relatively small variance explained by these harmonic components indicates that random fluctuations are the more important component. Furthermore,

frequencies less than 0.024 c/d, or 1 cycle per 40 days, contain as much as 85 percent of the variance for water discharge, but only as much as 55 percent for sediment (70 percent in the Juniata River)—indicating that greater-frequency storms exert considerable effect on the sediment discharge in these rivers (figs. 2 and 3).

Cross-spectral analysis examines the runoff-sediment discharge system. The cross correlogram for the log transforms of the data for each river peaks at lag zero with a large cross-correlation coefficient (fig. 4), indicating that the delay parameter, b , in the transfer-function model is zero. If memory terms (lag greater than zero) are disregarded in the model, more than 67 percent of the variability in daily sediment discharge can be explained by variability in daily water discharge (table 3).

A typical example of the cospectrum, estimated in the analysis for the Delaware River, is shown in figure 5. Low-frequency (less than 0.024 c/d) harmonics appear to account for most of the covariance and, thus, contribute greatly to the large cross-correlation at lag zero. A peak in the cospectrum at the annual cycle (0.0027 c/d) shows that although the annual cycle does not explain much of the variance it is an important factor in the cross-correlation at lag zero. In Piedmont rivers, this annual cycle most likely is related to the springtime snowmelt.

A plot of the coherency spectra (correlations between frequency components of the water and sediment series) shows large values throughout most of the frequency range (fig. 6), particularly at the lower frequencies. These large values of coherency indicate that a linear model is a reasonable representation of the daily runoff-sediment yield process in terms of log-transformed variates. The decrease in coherency at high frequencies may

Table 2. Summary of harmonic analysis

River	Series	Percent variance of log transforms explained by annual cycle	Percent variance of log transforms explained by semiannual cycle
Delaware	Water	27	3
	Sediment	9	2
Juniata	Water	40	2
	Sediment	30	1
Monocacy	Water	29	1
	Sediment	8	1
Potomac	Water	31	1
	Sediment	9	2
Rappahannock	Water	35	2
	Sediment	20	1
Tar	Water	38	7
	Sediment	23	10
Yadkin	Water	15	2
	Sediment	8	2

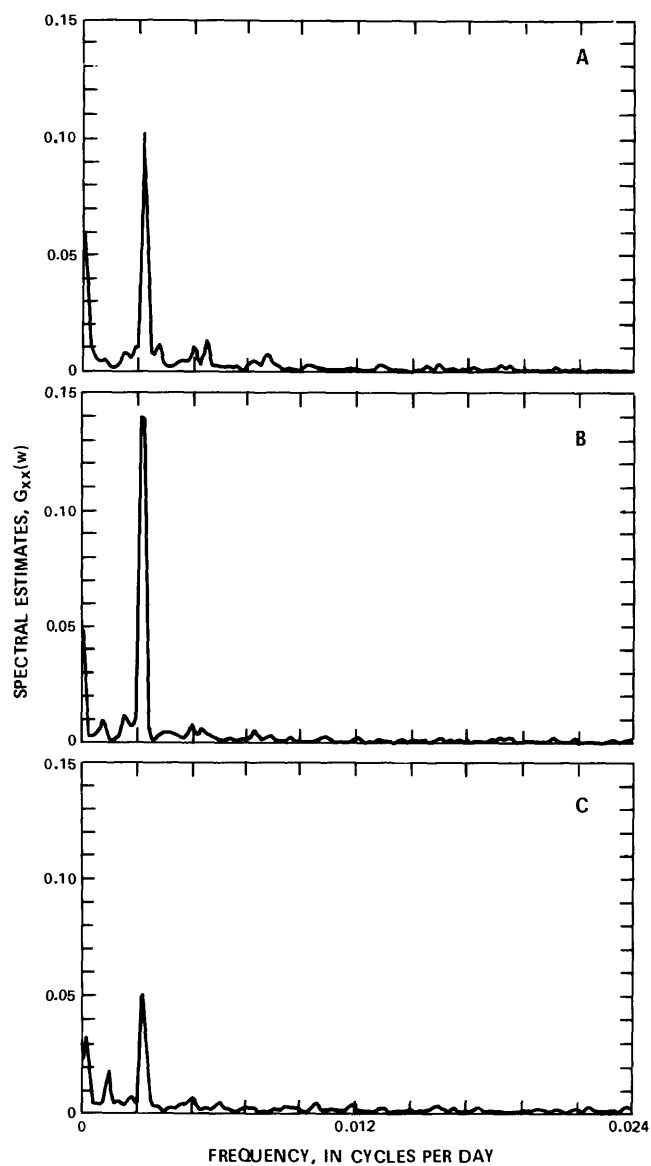


Figure 2. Water-discharge (log transformed) spectra for the (A) Delaware River at Trenton, N.J., (B) Juniata River at Newport, Pa., and (C) Yadkin River at Yadkin College, N.C.

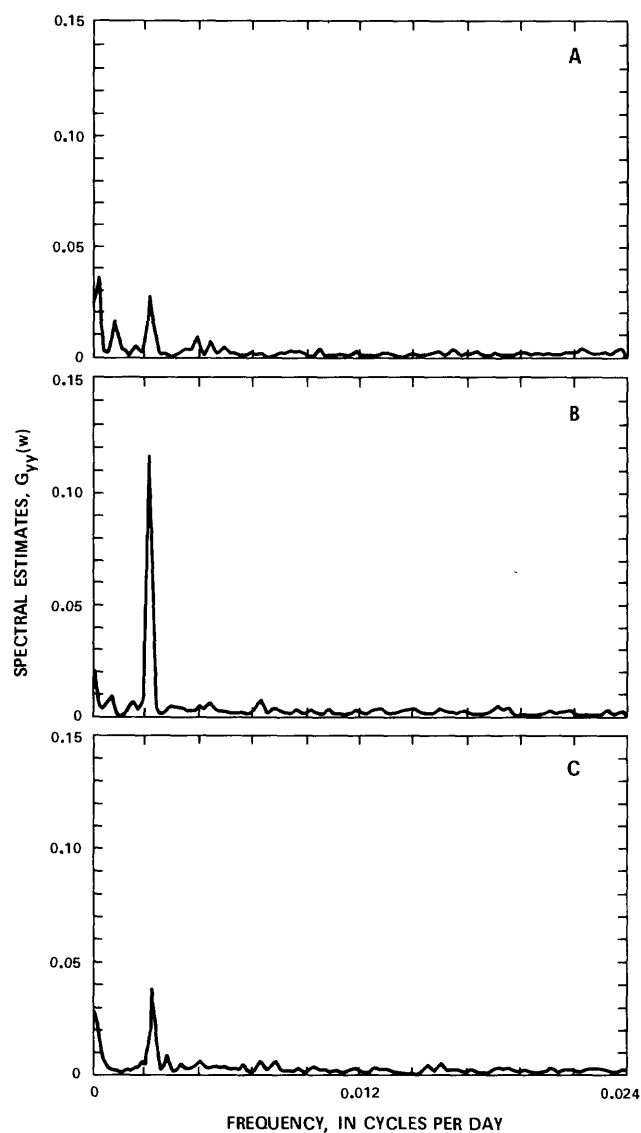


Figure 3. Sediment-discharge (log transformed) spectra for the (A) Delaware River at Trenton, N.J., (B) Juniata River at Newport, Pa., and (C) Yadkin River at Yadkin College, N.C.

Table 3. Cross-spectral characteristics between log-transformed daily discharge of water and sediment

[c/yr = cycle per year]

River	Cross correlation at lag 0	Variance of log transforms explained, with lag 0 model (percent)	Coherence (1 c/yr)	Gain (1 c/yr)	Phase (1 c/yr)
Delaware	0.820	67.2	0.949	0.609	333.7
Juniata	.894	79.9	.926	.878	349.9
Monocacy	.840	70.6	.926	.921	359.2
Potomac	.815	66.4	.985	.526	323.4
Rappahannock	.869	75.5	.990	.734	354.8
Tar	.918	84.3	.974	.777	342.0
Yadkin	.882	77.8	.986	.734	307.6

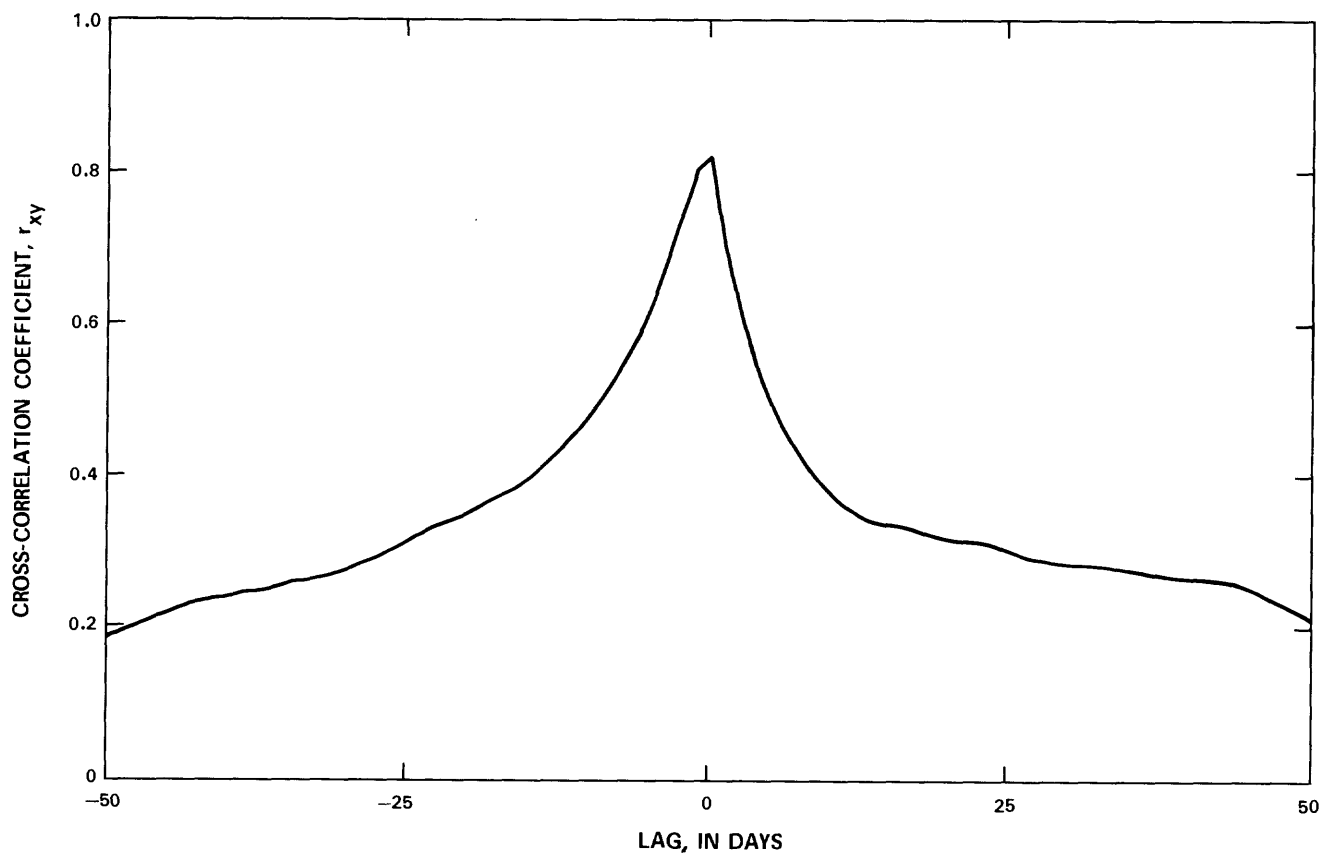


Figure 4. Cross-correlation function for Delaware River at Trenton, N.J., log-transformed water and sediment discharge.

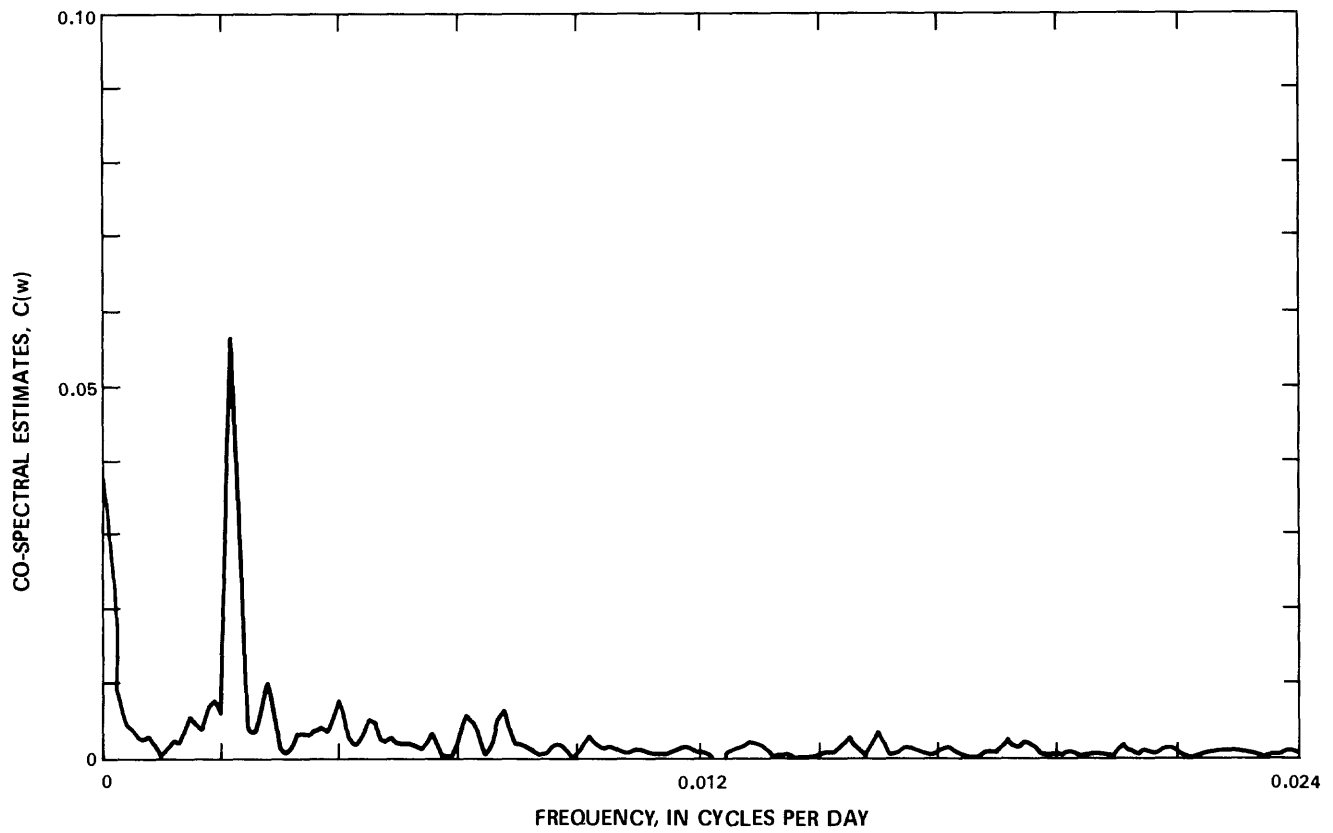


Figure 5. Cospectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J.

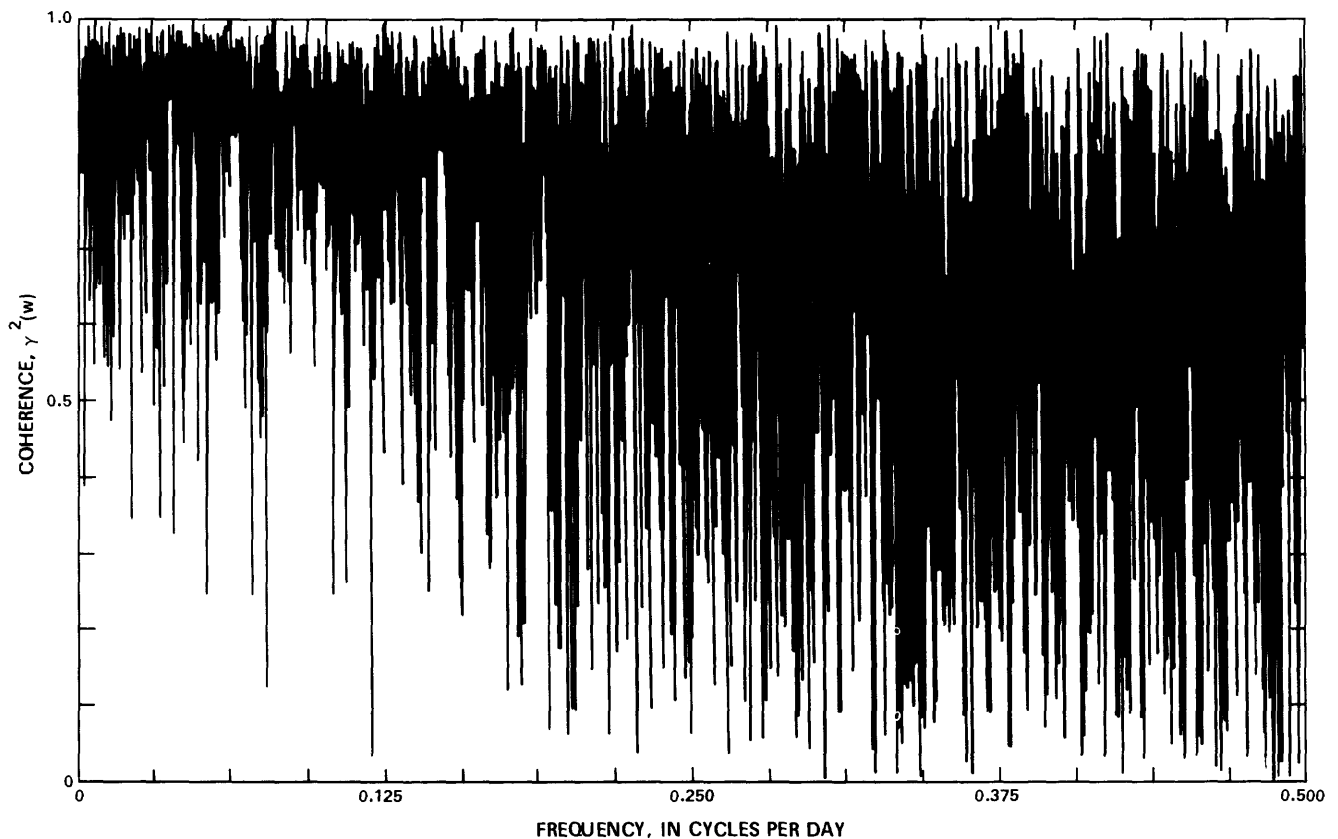


Figure 6. Coherency spectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J.

result from the disturbing effect of storm runoff on sediment yield.

The cross-correlation function shows that water and sediment discharge are in phase when considered as a whole. However, most high frequency components of the sediment series are leading the corresponding frequency components of the water series (fig. 7). This indicates that the sediment system responds quickly to frequent storms and peaks sooner than does the water system. One possible explanation is that the rainfall intensity associated with frequent storms removes the erodible sediment, producing a rapid increase in sediment yield that decreases as the storm progresses.

The gain function consists of the ratios of amplitudes of frequency components. In general, for the rivers examined, gain fluctuates about unity (fig. 8). However, gain at frequencies less than 0.024 c/d does show some evidence of being less than 1, while at frequencies greater than 0.024 c/d, gain is often greater than 1. These results indicate that regularly occurring storms are important to sediment discharge. Guy (1964) asserts that, for most streams, the bulk of suspended-sediment transport occurs during relatively brief storm periods when the drainage basin collects and routes excess precipitation.

Time-Series Models

Initial attempts at univariate modeling of water and sediment discharges were made with the assumption that no data transformations were necessary. However, the models' residuals for each data set did not approximate a normal distribution, nor was the variance of the residuals sufficiently constant with respect to the magnitudes of the discharges. These two departures from the assumptions necessary to statistically analyze the residuals of time-series models indicated that a transformation of the data was needed, and the log transformation was used to alleviate these problems. Sharma and Dickinson (1980) indicate that the log transformation is adequate in transfer-function modeling of a water-sediment discharge system.

The autocorrelation plot of each data set, as indicated by figure 9 for the Tar River, does not decrease quickly. This plot of the autocorrelation function indicated that a differencing operation was necessary. Therefore, a first difference was applied to the logarithms. Autocorrelations of the resulting data sets had a rapid exponential or sinusoidal decrease (fig. 10).

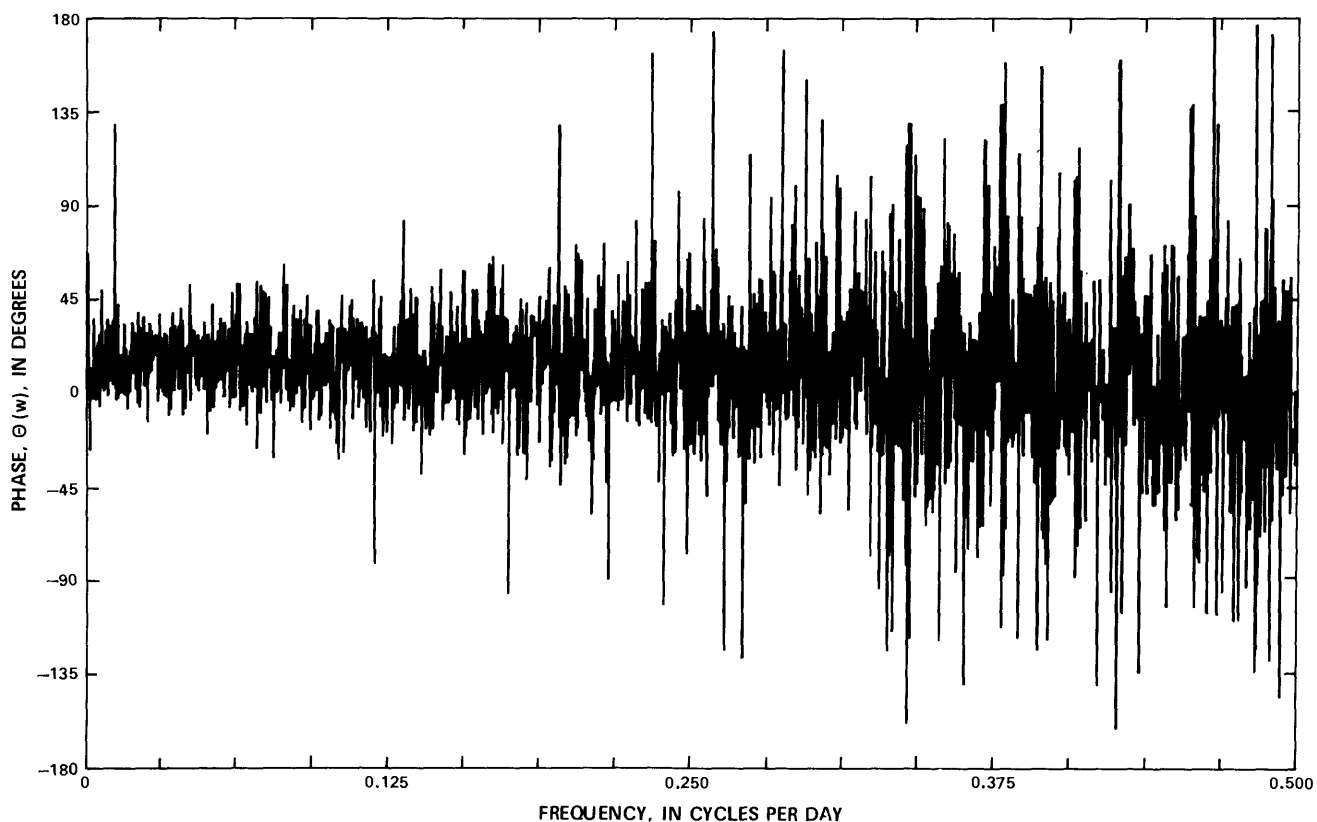


Figure 7. Phase spectrum between water and sediment discharges (log transformed) for the Delaware River at Trenton, N.J.

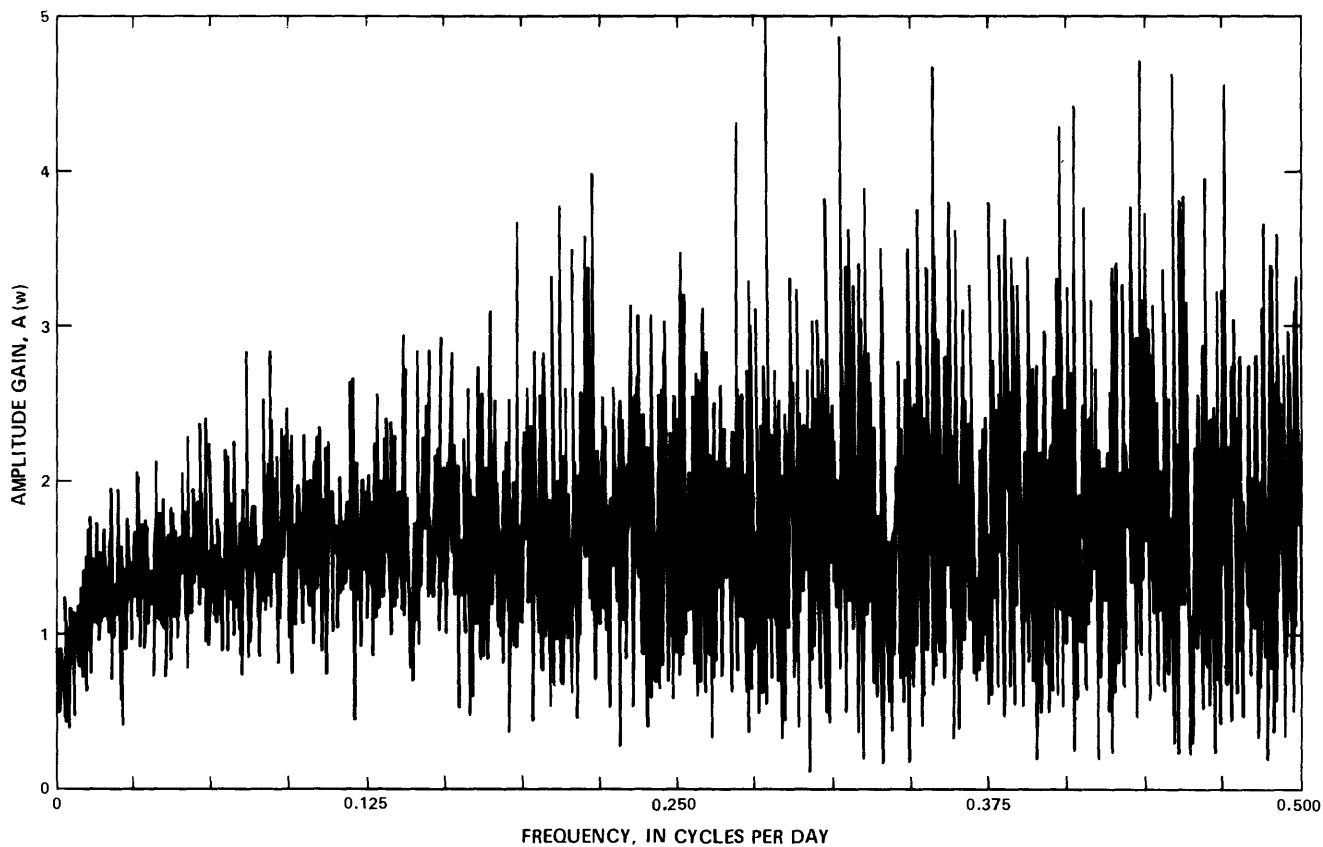


Figure 8. Gain function between water and sediment discharges (log transformed) for the Delaware River at Trenton, N. J.

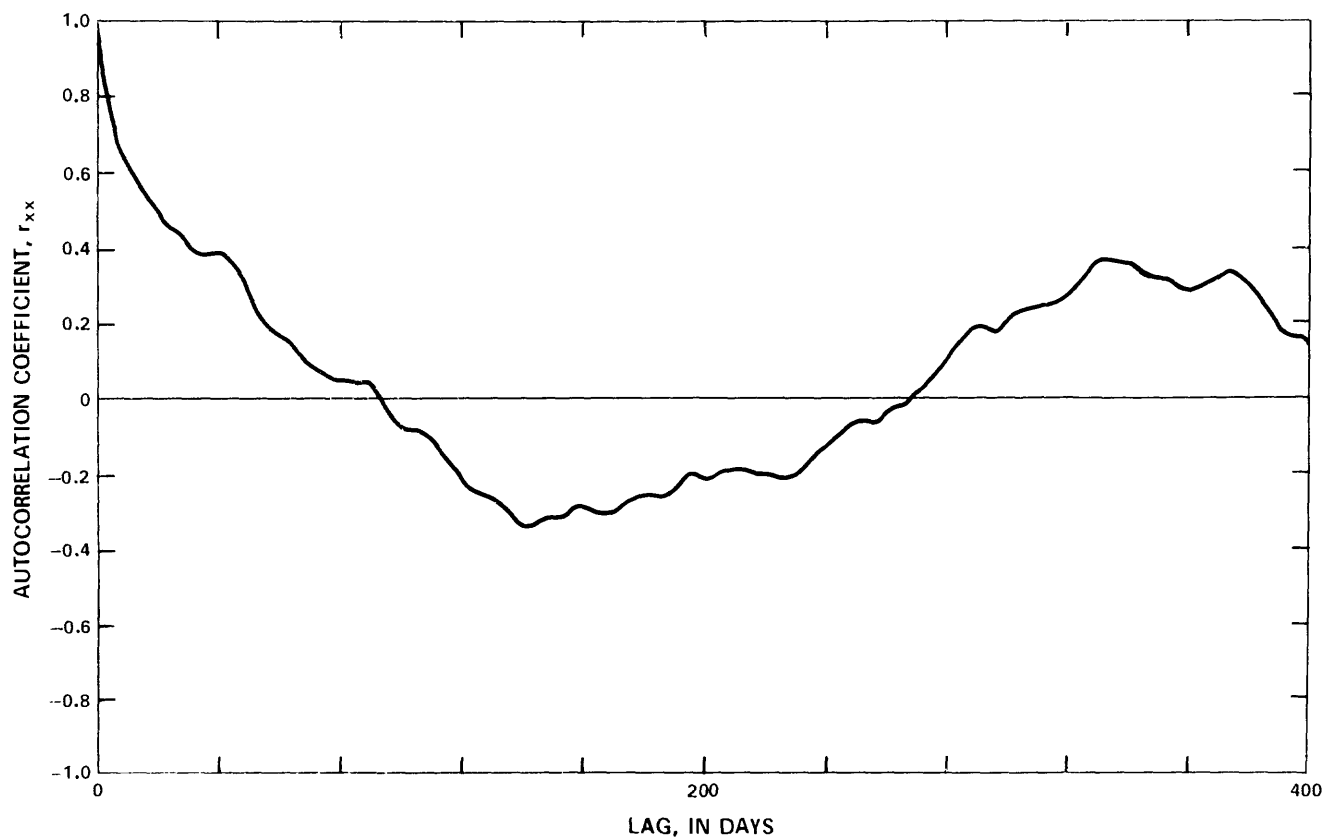


Figure 9. Autocorrelation function for water discharge, Tar River at Tarboro, N.C., log transformed and undifferenced.

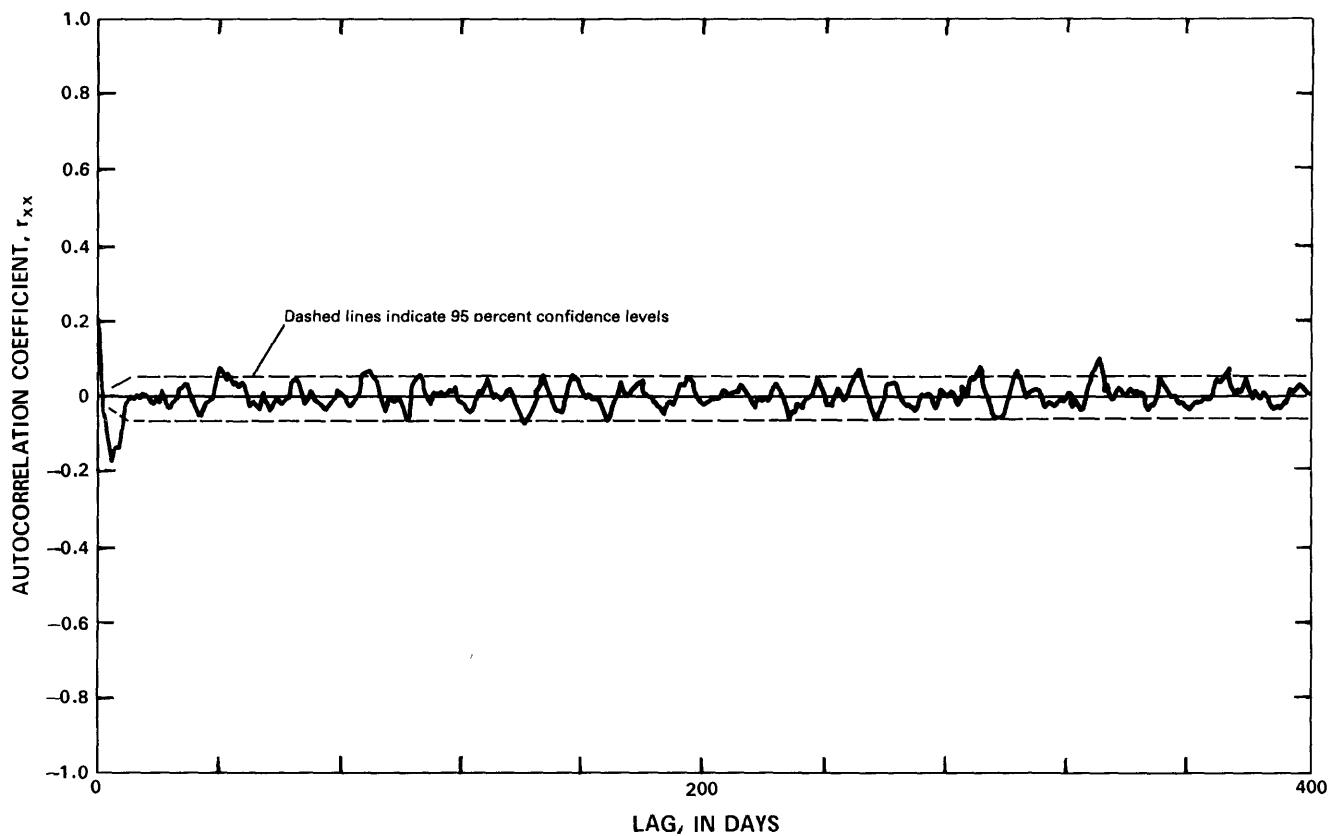


Figure 10. Autocorrelation function for water discharge, Tar River at Tarboro, N.C., log transformed and differencing of order 1.

A final consideration for model adequacy was an inspection of the autocorrelation function for possible seasonalities. Spectral analysis showed that daily water and sediment discharge have a periodic component with a 1-year cycle. The autocorrelation function of the undifferenced data for each data set indicated that an annual cycle exists. However, Kavvas and Delleur (1975) show that first differencing can decrease significantly a seasonality, although not completely remove it. An inspection of the autocorrelation function of the differenced logarithms (fig. 10) showed that the first differencing operation had produced virtually the necessary condition of second-order stationarity for the data sets. As a consequence of the above considerations, the basis of each univariate and transfer-function model for all rivers is a set of first-differenced logarithms of the original data.

Univariate Models

The tendency of the autocorrelation functions to decrease exponentially for most of the data sets indicated that autoregressive models are potential choices for the respective sets (fig. 11A for the Tar River). However, response of the partial-autocorrelation functions for these same data sets did not give a clear distinction between an exponential decrease and a cutoff after three or four lags (fig. 11B for the Tar River). As a result, for these data sets, mixed models were attempted, as well as high-order autoregressive models. The only discrepant series were those of the Monocacy River (fig. 12), which were identified as a moving-average process. For the remaining rivers, mixed models yielded best results in terms of a minimum residual mean-square error criterion. The type and order of the univariate models for each of the river's water and sediment discharge are listed in table 4. The percentage of the variance explained by the univariate models, calculated from the residual mean square and standard deviation of the logarithms, ranged from 81 to 97 percent for water and from 70 to 90 percent for sediment. Residuals for all of the final models passed the Portmanteau autocorrelation test, based on the statistic S , at the 10-percent level, indicating a white-noise structure.

The types of models adopted for each of the rivers can be analyzed qualitatively in terms of drainage-basin features. Delleur and others (1976) suggest that autoregressive coefficients are related to the longer-term storage characteristics of the watershed, whereas moving-average coefficients are related to the rainfall or storm inputs. Steep slopes, small drainage areas, and nonvegetated catchment basins, which respond quickly to storm rainfall, could translate into moving-average coefficients; slight slopes and grassy timbered areas could be associated with autoregressive parameters.

Of the seven rivers analyzed, the Monocacy and the

Rappahannock seem to diverge from the rest in terms of model type. The pure moving-average process of both water and sediment discharge for the Monocacy flow system may be a primary result of drainage area size. The Rappahannock River, with a smaller drainage area than the Monocacy, has a limited autoregressive contribution. This may be caused by more effective storage overriding contributions to a pure moving-average process.

Transfer-Function Models

Sharma and Dickinson (1980) demonstrate that a water-sediment discharge system can be regarded as a lumped linear system. In this system, the log-transformed water discharge is input, and the log-transformed sediment discharge is output. The noise of the system would encompass such errors as model error, measurement error, or time-discretization error.

The first step in building a transfer function for the linear system is determination of causality. Lagged cross-correlation values between the residuals of each set of univariate models were obtained. This prewhitening process insured the signature of the direction and amount of causality. Although the lag-zero cross-correlation value was by far the most significant (indicating a zero-delay parameter), prior physical considerations eliminated the prospect of sediment leading water discharge.

Initial estimates of the transfer functions for the logarithms of the original data are based on transfer functions relating residuals of the univariate models of water and sediment discharges. The transfer function model for the residuals is

$$\hat{a}_{ts} = \hat{v}(B)\hat{a}_{tw} + e_t, \quad (27)$$

where

\hat{a}_{ts} are residuals from univariate model for sediment discharge,

\hat{a}_{tw} are residuals from univariate model for water discharge,

$\hat{v}(B)$ is transfer function relating sediment and water discharge residuals,

e_t is a white-noise process.

Cross-correlation values between residuals at nonnegative lags determine the transfer function for the residuals $\hat{v}(B)$, that is,

$$v_k = r_{\hat{a}_{tw}\hat{a}_{ts}}(k) \frac{\sigma_{\hat{a}_{ts}}}{\sigma_{\hat{a}_{tw}}} \quad (28)$$

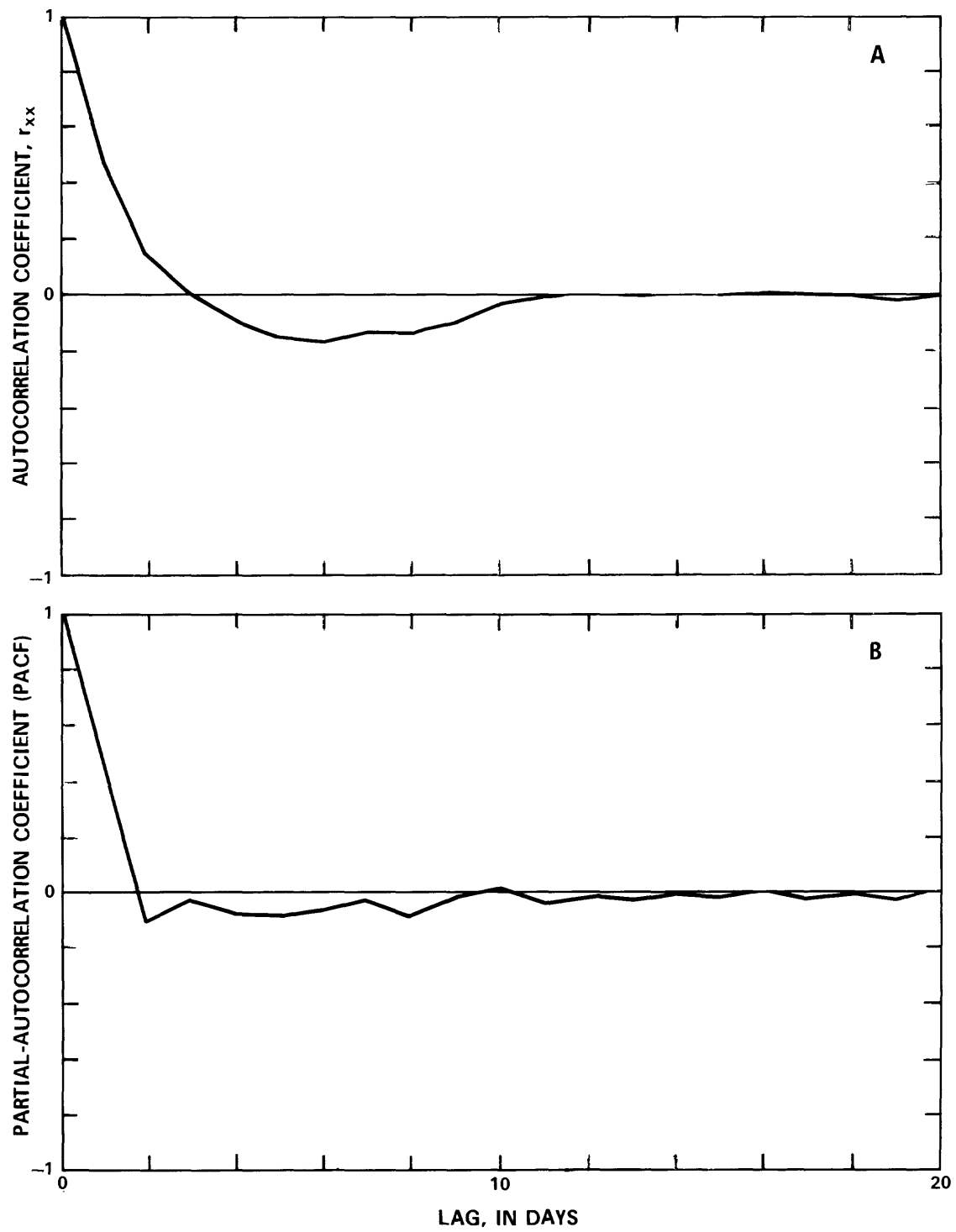


Figure 11. Autocorrelation (A) and partial-autocorrelation (B) functions for water discharge (log transformed and first differenced), Tar River at Tarboro, N.C.

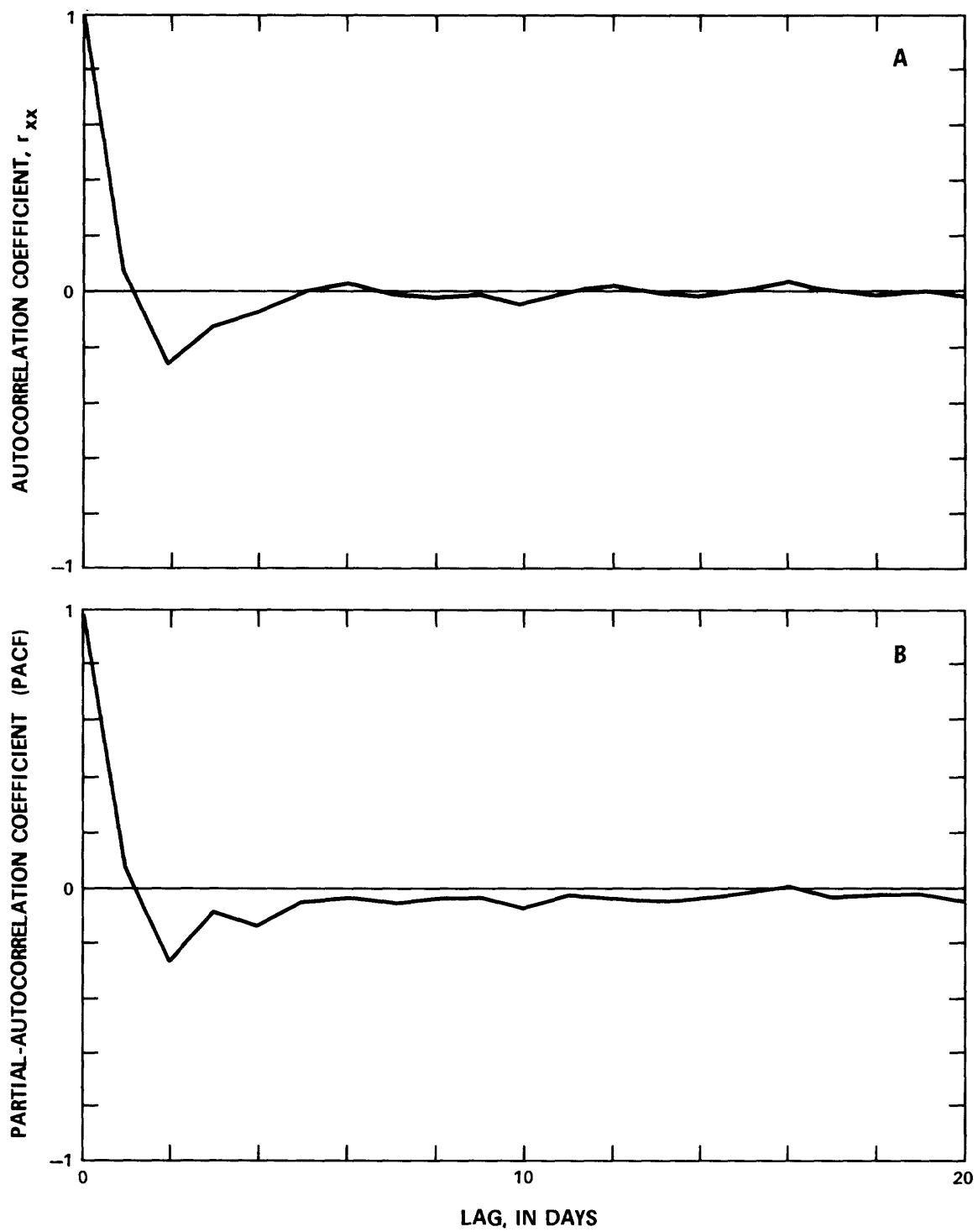


Figure 12. Autocorrelation (A) and partial-autocorrelation (B) functions for water discharge (log transformed and first differenced), Monocacy River at Frederick, Md.

Table 4. Estimates of parameters for univariate models*

[p=order of autoregressive model; d=order of differencing; q=order of moving-average model; ϕ_n =autoregressive parameter of order n; θ_n =moving-average parameter of order n; σ_a^2 =variance of residuals; R^2 =variance explained; 95-percent confidence limits in parentheses]

River	Series	Order of model (p,d,q)	ϕ_1	ϕ_2	θ_1	θ_2	θ_3	θ_4	σ_a^2	R^2 (percent)
Delaware	Water	(2,1,2)	0.982 ($\pm .049$)	-0.153 ($\pm .050$)	0.591 ($\pm .047$)	0.311 ($\pm .049$)	—	—	0.006	94.9
			.924 ($\pm .033$)	-.115 ($\pm .038$)	.765 ($\pm .028$)	.214 ($\pm .032$)	—	—	.097	81.8
	Sediment	(2,1,2)	.958 ($\pm .055$)	.110 ($\pm .056$)	.619 ($\pm .052$)	.295 ($\pm .055$)	—	—	.009	95.5
Juniata	Water	(2,1,2)	.919 ($\pm .066$)	-.093 ($\pm .068$)	.699 ($\pm .064$)	.247 ($\pm .069$)	—	—	.083	89.8
			—	—	-.033 ($\pm .031$)	.325 ($\pm .031$)	0.167 ($\pm .031$)	0.103 ($\pm .031$)	.033	83.8
	Sediment	(2,1,2)	—	—	-.035 ($\pm .031$)	.391 ($\pm .030$)	.223 ($\pm .030$)	.128 ($\pm .031$)	.190	70.0
Monocacy	Water	(0,1,4)	1.119 ($\pm .053$)	-.282 ($\pm .054$)	.574 ($\pm .052$)	.328 ($\pm .053$)	—	—	.007	95.8
			1.054 ($\pm .063$)	-.233 ($\pm .066$)	.711 ($\pm .063$)	.256 ($\pm .068$)	—	—	.062	88.2
	Sediment	(2,1,2)	.533 ($\pm .087$)	.063 ($\pm .079$)	.495 ($\pm .083$)	.341 ($\pm .081$)	—	—	.027	89.8
Rappahannock	Water	(2,1,2)	.630 ($\pm .032$)	— ($\pm .036$)	.754 ($\pm .028$)	.156 ($\pm .028$)	—	—	.272	77.1
			1.367 ($\pm .050$)	-.465 ($\pm .048$)	.875 ($\pm .035$)	.063 ($\pm .027$)	—	—	.007	96.5
	Sediment	(1,1,2)	1.214 ($\pm .078$)	-.323 ($\pm .096$)	1.005 ($\pm .073$)	-.037 ($\pm .096$)	—	—	.042	90.3
Tar	Water	(2,1,2)	.634 ($\pm .053$)	-.010 ($\pm .049$)	.479 ($\pm .048$)	.441 ($\pm .048$)	—	—	.012	81.2
			.748 ($\pm .073$)	-.053 ($\pm .069$)	.677 ($\pm .072$)	.273 ($\pm .075$)	—	—	.105	76.4
	Sediment	(2,1,2)	—	—	—	—	—	—	—	—

*Note: All series are subjected to differencing of order 1.

and

$$\hat{v}(B) = \hat{v}_0 + \hat{v}_1(B) + \hat{v}_2(B^2) + \dots + \hat{v}_k(B^k), \quad (29)$$

where

$\sigma_{\hat{a}_{tw}}$ is standard error of univariate model for water discharge;

$\sigma_{\hat{a}_{ts}}$ is standard error of univariate model for sediment discharge;

\hat{v}_k is impulse response weight at lag k ;

and

$r_{\hat{a}_{tw}\hat{a}_{ts}}$ is cross correlation between residuals of water and sediment discharge univariate models.

Substituting, in equation 27, the expression for \hat{a}_{tw} and \hat{a}_{ts} in terms of their univariate models yields an initial estimate for the transfer function for the lags of the original data. For example, if the models for sediment and water discharge are

$$\phi_1(B)y_t = \theta_1(B)\hat{a}_{ts} \quad (30)$$

and

$$\phi_2(B)x_t = \theta_2(B)\hat{a}_{tw}, \quad (31)$$

then

$$\hat{a}_{ts} = \frac{\phi_1(B)}{\theta_1(B)} y_t \quad (32)$$

and

$$\hat{a}_{tw} = \frac{\phi_2(B)}{\theta_2(B)} x_t, \quad (33)$$

where

y_t is sediment output,

and

x_t is water output.

These expressions for \hat{a}_{ts} and \hat{a}_{tw} yield initial parameter estimates for a transfer-function model in terms of x_t and y_t . The assumed model is

$$y_t = v(B)x_t + \eta_t \quad (34)$$

with the transfer function $v(B)$ written as

$$v(B) = \frac{\omega(B)}{\delta(B)}$$

where

y_t is the sediment output;

x_t is the water input;

$\omega(B)$ is the input lag function;

$\delta(B)$ is the output lag function;

and

η_t is the noise contribution.

Written in this form, sediment discharge is not only a function of lagged values of water discharge, but also a function of lagged values of sediment discharge.

None of the transfer-function models originally passed the Portmanteau autocorrelation test. Therefore, it was necessary to model the η_t noise process with autoregressive, moving-average modeling. Since adequate noise models were found, the Portmanteau autocorrelation statistic passed at the 10-percent level for all the transfer-function models. Similarly, cross-correlations between residuals of the transfer-function model and residuals of the input (water) univariate model passed the Portmanteau cross-correlation test at the 10-percent level. Final coefficient values of the ω and δ functions, as well as the magnitudes of the residual mean-square error, are shown in table 5.

The reason for using a transfer-function approach in modeling the water-sediment discharge system is the search for a better predictor of sediment yield than either a univariate model or a simple regression of sediment and water discharge (memoryless model) can provide. In all cases, the residual sum of the squares is

Table 5. Estimates of parameters for transfer-function models*

[δ_n = output operator of order n; ω_n = input operator of order n; σ_η^2 = variance of residuals; R^2 = variance explained; 95-percent confidence limits for parameters in parentheses]

River	δ_1	δ_2	ω_0	ω_1	ω_2	σ_η^2	R^2 (percent)
Delaware	0.887 ($\pm .018$)	—	3.275 ($\pm .056$)	3.403 ($\pm .119$)	-0.296 ($\pm .065$)	0.036 —	— 93.3
Juniata	0.798 ($\pm .048$)	—	2.477 ($\pm .038$)	2.097 ($\pm .143$)	.021 ($\pm .046$)	.027 —	— 96.7
Monocacy	—	0.745 ($\pm .054$)	2.206 ($\pm .035$)	1.449 ($\pm .028$)	1.695 ($\pm .115$)	.039 —	— 93.9
Potomac	1.433 ($\pm .030$)	-.469 ($\pm .010$)	2.346 ($\pm .054$)	3.501 ($\pm .127$)	-1.203 ($\pm .055$)	.023 —	— 95.6
Rappahannock	.866 ($\pm .042$)	—	2.450 ($\pm .044$)	2.311 ($\pm .133$)	-.080 ($\pm .052$)	.107 —	— 91.0
Tar	.338 ($\pm .114$)	—	2.039 ($\pm .062$)	1.057 ($\pm .192$)	0.000 —	.013 —	— 97.0
Yadkin	-.588 ($\pm .117$)	—	2.555 ($\pm .034$)	-1.310 ($\pm .305$)	0.000 —	.029 —	— 93.7

*Note: All input and output series were subjected to log transforms and differences (∇) of order 1.

much smaller and R^2 greater for transfer-function models of sediment yield than for univariate models (tables 4 and 5).

Accuracy of Model Prediction

Predicting system response to severe flood or drought is a major objective of statistical investigation. Although transfer-function models better account for variance in sediment yield than univariate models do, each series represents only one realization of the sediment-yield process. Notwithstanding that several of the data sets went back longer than 22 years, the question is how well do corresponding transfer-function models describe critical periods outside that time frame.

Hurricane Connie, during 1955, represents the storm of record for the Delaware River, but is not included in the data set used for model construction. To evaluate the model, an attempt to predict the sediment discharge associated with this storm was made, using the final model with existing discharge measurements as input. The data in table 6 show sequential, updated 5-day predictions for the logarithms of the sediment discharge; thus both rising and falling stages are included. the model is accurate in predicting peak sediment discharge 1 day ahead, but it overestimates the rising stage. However, actual values are within the 95-percent confidence limits; therefore, the model is useful in development design. Although the transfer-function model predicts the storm of record reasonably well, the question remains as to how long a record is necessary to construct a model with sufficient accuracy for practical use in development design and management.

DISCUSSION

While spectral analysis can be used to make inferences about an underlying process, time-domain techniques are used to formulate parsimonious models of the process itself. Both approaches form a basis for inferences drawn about the dynamic components of that process.

A comparison of model statistics between drainage basins could serve as a means of evaluating the general effects of differing land use and drainage-basin management. The unit step-response function provides one way of representing these composite effects (Sharma and Dickinson, 1980). For the runoff-sediment-yield system, the step-response function describes the system response (sediment yield) to a step increase in runoff. Step-response weights V_k are calculated, using impulse-response weights:

$$V_k = v_0 + v_1 + \dots + v_k,$$

where v_k 's are defined by $y = v(B)x$. A constancy in step-response weights indicates a system without memory and produces a constant ratio, V_k/V_0 , of 1. Such a system might indicate a drainage basin with readily erodible,

thick soils; dense stream network; and sparsely vegetated land. A decrease of V_k/V_0 from unity with respect to k would form an index of the retardance capability of a fluvial system in transporting eroded sediment. Differences in the plots of V_k/V_0 versus lag k between rivers would express differences in drainage-basin characteristics and in land management.

Plots of V_k/V_0 versus k (fig. 13) were constructed for each river. The Yadkin River drainage basin showed the most production of sediment while the Delaware and Tar Rivers, the least production. Meade (1982) states that sediment concentrations are consistently large and increase significantly with streamflow in rivers of the Southern Piedmont (for example, the Yadkin River). Because of these large concentrations, sediment yields from the Southern Piedmont are consistently the largest per unit area of any physiographic province on the Atlantic slope (fig. 14). By contrast, in the Coastal Plain (for example, the Tar River), sediment yields per unit area are among the smallest. The minimal sediment production of the Delaware, in turn, may be a consequence of reservoir construction and urbanization, large drainage area, or its glaciated upper reaches.

Storm-period variables have been shown to have considerable effect on sediment production in fluvial systems (Guy, 1964). Spectral and cross-spectral analyses indicate that frequent storms have a significant role in the sediment-yield process; this conclusion is supported by the strong moving-average component in sediment models. In addition, frequency analysis of sediment discharge shows that between 50 and 90 percent of the suspended sediment is discharged in only 10 percent of the time (table 7).

SUMMARY

The theory of time series and spectral analysis is presented in some detail, and the techniques are applied to log-transformed, daily water- and sediment-discharge records from seven rivers in the Eastern United States. The method presented not only provides for the formulation of predictive models, but also gives insight into the hydrologic processes operating within a drainage basin. Through the combination of spectral and time-domain techniques, statistical properties are related to physical characteristics of the drainage basin, and the hydrologic factors affecting interrelations of water and sediment discharges are indicated.

All spectra of water and sediment discharge show significant peaks at the annual frequencies, probably related to the springtime snowmelt. The annual cycle accounts for 17 to 40 percent of the variance for the water series, but for only 9 to 31 percent for the sediment, indicating the importance of random fluctuations in the sediment series. The small annual peak for the Yadkin

Table 6. Predictions of logarithms of storm of record sediment discharge for Delaware River at Trenton, N.J., using transfer-function model

[Prediction origin is final point in data set used in predicting]

Days ahead	Lower confidence limit	Predicted value	Upper confidence limit	Actual value
Prediction origin is 4 days prior to storm peak (with 95-percent confidence limits).				
1	2.240	3.079	3.918	3.238
2	1.956	3.207	4.457	3.754
3	4.896	6.359	7.822	5.928
*4	4.990	6.580	8.171	6.036
5	2.958	4.633	6.308	4.728
Prediction origin is 3 days prior to storm peak (with 95-percent confidence limits).				
1	2.495	3.335	4.174	3.754
2	5.190	6.441	7.691	5.928
*3	5.180	6.643	8.105	6.036
4	3.099	4.690	6.280	4.728
5	2.268	3.943	5.618	4.136
Prediction origin is 2 days prior to storm peak (with 95-percent confidence limits).				
1	5.940	6.779	7.618	5.928
*2	5.608	6.859	8.109	6.036
3	3.391	4.854	6.317	4.728
4	2.502	4.092	5.683	4.136
5	1.997	3.672	5.347	3.801
Prediction origin is 1 day prior to storm peak (with 95-percent confidence limits).				
*1	5.333	6.172	7.011	6.036
2	3.165	4.415	5.666	4.728
3	2.296	3.759	5.222	4.136
4	1.779	3.369	4.960	3.801
5	1.541	3.216	4.891	3.604
Prediction origin is storm peak (with 95-percent confidence limits).				
1	3.466	4.305	5.144	4.728
2	2.438	3.689	4.939	4.136
3	1.853	3.316	4.778	3.801
4	1.577	3.168	4.758	3.604
5	1.119	2.794	4.469	3.448

*Denotes storm peak.

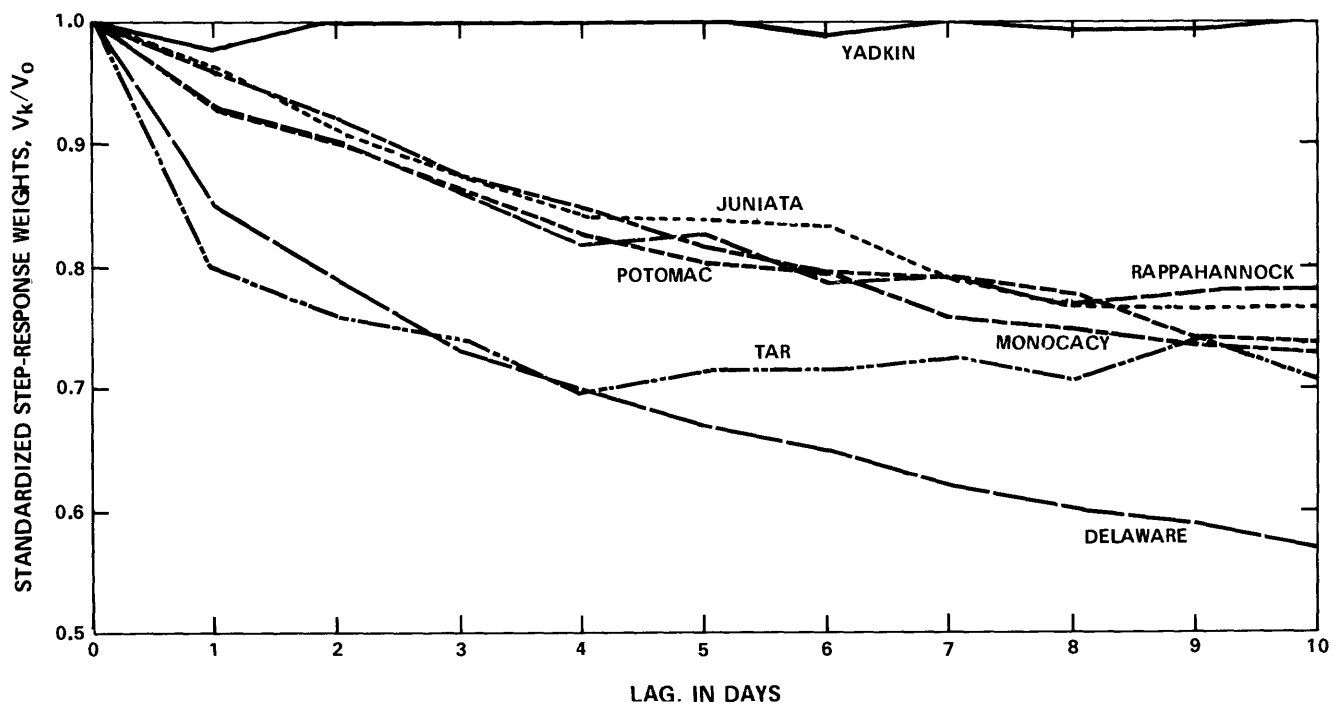


Figure 13. Standardized step-response weights as a function of lag.

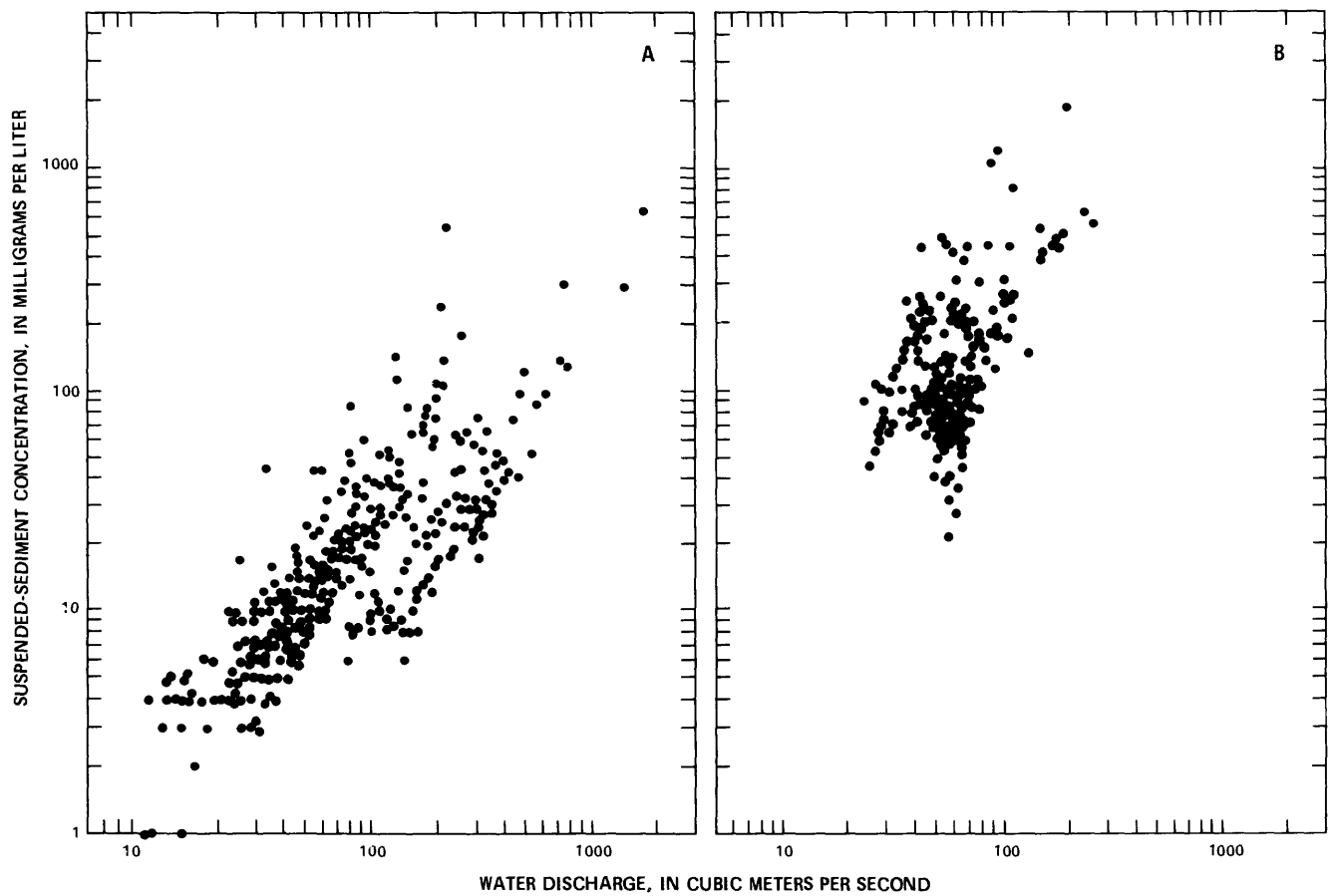


Figure 14. Sediment concentration-water discharge relations for (A) Juniata River at Newport, Pa., and (B) Yadkin River at Yadkin College, N.C. (from Meade, 1982).

Table 7. Frequency of suspended-sediment discharge

[Data from Meade, 1982]

River and gaging station	Years of record	Period of record	Mean annual sediment discharge (10 ³ tons)	Percentage of suspended-sediment discharge in			
				1 percent of time	2 percent of time	5 percent of time	10 percent of time
Delaware River at Trenton, N.J.	30	1950-79	749	49	61	76	85
Delaware River at Trenton, N.J.	29 ^a	{1950-54 1956-79}	696	44	57	74	84
Juniata River at Newport, Pa.	28	1952-79	281	44	59	75	85
Juniata River at Newport, Pa.	27 ^b	{1952-71 1973-79}	252	41	54	72	83
Potomac River at Point of Rocks, Md.	19	1961-79	1,173	49	62	78	87
Monocacy River ¹ near Frederick, Md.	19	1961-79	195	46	62	82	90
Rappahannock River at Remington, Va.	28	1952-79	98	48	63	80	88
Tar River at Tarboro, N.C.	10	1958-67	123	10	17	34	52
Yadkin (Pee Dee) River at Yadkin College, N.C.	28	1952-79	963	25	38	58	72

¹Tributary of Potomac River that enters Potomac downstream of Point of Rocks.^aMinus 1955, the year of Hurricanes Connie and Diane.^bMinus 1972, the year of Hurricane Agnes. These storms produced floods and large sediment discharges whose recurrence intervals may be longer than the period of daily sediment record.

River water discharge may be a result of flood-control projects.

Storms with frequencies of less than 40 days account for as much as 45 percent of the variance in the sediment series. Furthermore, frequency analyses indicate that between 50 and 90 percent of the suspended sediment is discharged in only 10 percent of the time. These data indicate that storms have a significant effect on sediment yield.

Cross-spectral analyses show that both water and sediment discharge are in phase, when considered as a whole, with cross-correlation coefficients at lag zero ranging from 0.82 to 0.92. However, most high-frequency components of the sediment series lead respective frequency components of the water series, indicating that the sediment system responds quickly to storms. One explanation is that rainfall intensity associated with storms quickly removes erodible sediment, producing a rapid peak in sediment discharge that decreases as the storm progresses.

Univariate models constructed for individual records of water and sediment discharge explain between 81 and 97 percent of the variance for the logarithms of water discharge, and between 70 and 90 percent for the logarithms of sediment. Six of the rivers analyzed produce mixed models. If it is accepted that autoregressive coefficients are related to long-term storage characteristics whereas moving-average coefficients are related to storm inputs, then both long-term and storm records are

important. However, the Monocacy River diverges from the rest in terms of model type. The pure moving-average process of both water and sediment discharge for the Monocacy River may result from its small drainage area.

Large values of coherence between water and sediment series indicate that a linear representation of the runoff-sediment yield system is reasonable. At high frequencies, coherence does decrease, a consequence of the effects of storm runoff on sediment yield.

Bivariate models with water discharge as input and sediment discharge as output account for a larger percentage of the variance of the sediment series than univariate models do; thus, they are better predictors. Step-response weights calculated from these transfer-function models describe system response (sediment yield) to a step increase in water discharge, which provides an index of sediment production. The Yadkin River had the most production of sediment, and the Delaware and Tar Rivers, the least production. The Yadkin is located in the Southern Piedmont, where sediment yields are the greatest per unit area of any physiographic province in the Eastern United States. By contrast, the Tar River is located in the Coastal Plain, where sediment yields per unit area are among the smallest. The minimal sediment production of the Delaware may be a consequence of urbanization, drainage area, or its glaciated upper reaches. The plotting of yearly changes in step-response weights may provide possible explanations for variable

sediment production.

The facility of transfer-function models warrants their consideration in examining the economics of data collection for a water-sediment discharge system. The use of models in split-sample analyses (Burkham and Dawdy, 1968) can provide some insight into the aspect of time-sampling error in network-design problems. Cumulative or random subsets of the time series are used to estimate model parameters, and the predicting accuracy of each successive model is the basis of model performance. Marginal improvements in model performance as a function of the data base extension could thus be evaluated. The merit of the transfer-function models also could be assessed in similar network-analysis methods such as that developed by Moss and Gilroy (1980).

Another potential extension of transfer-function modeling is that involving multiple-input series. The multiple input can be similar inputs from different locations, differing inputs applied to the same location, or a combination of each. An example of the first situation is a model of water discharge from two or more upstream stations relating to sediment yield at a downstream location. An example of the second situation is a model of sediment yield with water discharge and rainfall as inputs.

Intervention analysis, a special case of multiple-input transfer-function modeling, could be considered as a procedure for determining effects of abrupt changes in a drainage basin. The intervention series, usually represented by a step function, could address the effects of drainage development, basin development, or climatic change on the sediment yield.

Some applications of the systems approach may focus on a limited frequency range because it is often unreasonable to suppose that the same parametric model is valid throughout a wide range of frequencies. Modeling that describes various frequency ranges of hydrological systems is a potential area of significant return. For example, autoregressive, moving-average models with time-varying autoregressive coefficients can be made to preserve the long-range dependence of hydrologic series (Delleur and others, 1976).

REFERENCES

- Baracos, P. C., Hipel, K. W., and McLeod, A. I., 1981, Modeling hydrologic time series from the Arctic: *Water Resources Research*, v. 17, no. 3, p. 414-422.
- Bloomfield, Peter, 1976, *Fourier analysis of time series—An introduction*: New York, John Wiley and Sons, 258 p.
- Box, G. E. P., and Jenkins, G. M., 1976, *Time series analysis—Forecasting and control* (Revised ed.): San Francisco, Holden-Day, 575 p.
- Box, G. E. P., and Tiao, G. C., 1975, *Intervention analysis with applications to economic and environmental problems*: *Journal of the American Statistical Association*, v. 70, no. 349, p. 70-79.
- Burkham, D. E., and Dawdy, D. R., 1968, Error analysis of streamflow data for an alluvial stream: U.S. Geol. Survey Prof. Paper 655-C, 13 p.
- Delleur, J. W., Tao, P. C., and Kavvas, M. L., 1976, An evaluation of the practicality and complexity of some rainfall and runoff time series models: *Water Resources Research*, v. 12, no. 5, p. 953-970.
- Granger, C. W. J., and Newbold, Paul, 1977, *Forecasting economic time series*: New York, Academic Press, 333 p.
- Guy, H. P., 1964, An analysis of some storm-period variables affecting stream sediment transport: U.S. Geol. Survey Prof. Paper 462-E, 46 p.
- Jenkins, G. M., and Watts, D. G., 1968, *Spectral analysis and its applications* (2d ed.): San Francisco, Holden-Day, 525 p.
- Kavvas, M. L., and Delleur, J. W., 1975, Removal of periodicities by differencing and monthly mean subtraction: *Journal of Hydrology*, v. 26, p. 335-353.
- Matalas, N. C., 1967, Mathematical assessment of synthetic hydrology: *Water Resources Research*, v. 3, no. 4, p. 937-945.
- Meade, R. H., 1982, Sources, sinks, and storage of river sediment in the Atlantic drainage of the United States: *Jour. Geology*, v. 90, no. 3, p. 235-252.
- Moss, M. E., and Gilroy, E. G., 1980, Cost effective stream-gaging strategies for the lower Colorado River basin: The Blythe field office operations: U.S. Geol. Survey open-file report 80-1048, 111 p.
- Otnes, R. K., and Enochson, Loren, 1972, *Digital time series analysis*: New York, John Wiley and Sons, Inc., 552 p.
- Rodriguez-Iturbe, Ignacio, and Nordin, C. F., Jr., 1968, Time series analysis of water and sediment discharges: *Bulletin International Association of Scientific Hydrology*, v. 13, no. 2, p. 69-84.
- _____, 1974, Frequency domain analysis of hydrological systems with monthly data: *Mathematical Models in Hydrology—Symposium*, Warsaw, Poland, 1971, IAHS-AISH Publication no. 100, p. 343-354.
- Sharma, T. C., and Dickinson, W. T., 1980, System model of daily sediment yield: *Water Resources Research*, v. 16, no. 3, p. 501-506.
- Simmons, C. E., 1976, Sediment characteristics of streams in the eastern Piedmont and western Coastal Plain regions of North Carolina: U.S. Geol. Survey Water-Supply Paper 1798-0, 28 p.
- U.S. Geol. Survey, 1956-70, Quality of surface waters of the United States, Parts 1-4: U.S. Geol. Survey Water-Supply Papers, issued annually.
- _____, 1971-79, *Water Resources Data for Maryland and Delaware*, Parts 1 and 2: Towson, Md., issued annually.
- _____, 1971-79, *Water Resources Data for New Jersey*, Parts 1 and 2: Trenton, N.J., issued annually.
- _____, 1971-79, *Water Resources Data for North Carolina*, Parts 1 and 2: Raleigh, N.C., issued annually.
- _____, 1971-79, *Water Resources Data for Pennsylvania*, Parts 1 and 2: Harrisburg, Pa., issued annually.
- _____, 1971-79, *Water Resources Data for Virginia*, Parts 1 and 2: Richmond, Va., issued annually.

Metric Conversion Table

<i>Multiple inch-pound unit</i>	<i>By</i>	<i>To obtain metric unit</i>
mile	1.609	kilometer
cubic feet per second (ft ³ /s)	0.02832	cubic meter per second
ton per day (ton/d)	0.9072	megagram per day
square mile	2.590	square kilometer