EFFECTS OF TRUNCATION ON DEPENDENCE IN HYDROLOGICAL TIME SERIES

by

Rezaul Karim Bhuiya and Vujica Yevjevich

November 1968

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

31

EFFECTS OF TRUNCATION ON DEPENDENCE IN HYDROLOGICAL TIME SERIES

by

Rezaul Karim Bhuiya

and

Vujica Yevjevich

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80521
U. S. A.

November 1968

No. 31

ACKNOWLEDGMENTS

The writers acknowledge the financial assistance of the U. S. National Science Foundation for the research project support leading to this paper. It is also acknowledged that the National Center for Atmospheric Research at Boulder, Colorado, allowed free computer time on the CDC-6600 for computations of this study. Appreciation is extended to Dr. M. M. Siddiqui, Professor of Mathematics and Statistics at Colorado State University for his valuable suggestions in conducting the research whose results are presented in this paper.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ABSTRACT

The effects of constant value truncation on the dependence property of idealized hydrological time series are investigated. This analysis was carried out for dependent stochastic components, for periodic component and for the sum of stochastic dependent and periodic deterministic components of hydrological time series. The random elements of the process were considered for both normal and lognormal distributions. For all cases, the truncation is considered at a constant level. Analytical equations were developed in approximate forms for correlograms which express the dependence of the truncated time series. Computations of correlograms from these equations were made on a digital computer. Two series of daily river flows were used with the truncation at the mean flow to compare their correlograms with those of the original daily flow time series.

EFFECTS OF TRUNCATION ON DEPENDENCE IN HYDROLOGICAL TIME SERIES

by

Rezaul Karim Bhuiya*

and

Vujica Yevjevich**

Chapter I

INTRODUCTION

1. Character of hydrological time series. The hydrological time series are usually positive valued variables which include also the zero value. In many cases, the observed time series never show a zero value, or the minimal observed values are much greater than the zero. The probability of the lowest discharge of the Mississippi River being zero is practically zero. Also, the probability of monthly precipitation of a wet region being zero is very small. However, many hydrological time series include values which are either zeros or very close to zero.

Hydrological time series are either continuous (especially if rates are observed as the precipitation intensity, flow discharge, sediment discharge, and similar) or discrete (if the values are observed at discrete times or are derived from continuous time series in the form of daily, monthly or annual values of precipitation, river flow, sediment transport, etc.). When a continuous or discrete hydrological time series has zero values from time to time and for some period, the process is called intermittent. The general pattern for discrete time series is that they are less likely to have zero values when the cumulative or average amounts of a hydrological magnitude are obtained over longer time intervals. Every series of daily precipitation has zero values. However, the series of monthly values may or may not have zero values. It is rare that the series of annual precipitation contains zeros. If it does the series usually belongs to a very arid region or to a desert.

Therefore, the continuous series of rates of several hydrological magnitudes is expected to have zero values very frequently and with the longer total duration than the series of its cumulative discrete values of a variable. The continuous series of precipitation intensity has zero values more often and of a longer total duration than the series of daily precipitation. In turn, the daily precipitation series has zero values more often and of a longer duration than the series of monthly precipitation. The same patterns occur for river flows in intermittent rivers. The process of river bed load transport is intermittent, either because of the intermittent flow of the river or because the low flows cannot transport the bed load.

Two questions arise in the treatment of hydrological time series with zero values: (1) How to interpret these values in the sense of probability distribution of the variable; and (2) What is the influence of zero values on properties of hydrological time series.

The total probability of zero values may be interpreted as the probability of all negative values of a distribution function. Therefore, the probability distribution is composed of two parts: (1) Probability mass for the discrete value of zero; and (2) Probability densities for all positive values, zero included. Sometimes, a physical interpretation is attractive for these values of zero. In the case of the series of precipitation intensity, the zero values represent the times when the opposite process to precipitation or the evaporation occurs. In the case of a liquid-gaseous interface at the ground surface, the flux of moisture through it is a continuous process in the form of both precipitation and evaporation (intermittently) or with positive and negative variable values of a continuous probability density.

In an arid region where the rate of evaporation in the air is very high, a considerable portion of precipitation evaporates before it reaches the ground. Sometimes the total amount of precipitation evaporates, and this fact represents an increase in the length of zero values of the precipitation time series.

The second question is the effect of zero values on properties of hydrological time series, and this topic is one of the two subjects of this paper. It is logical to expect that the series with zero values would have some particular properties, at least as it concerns the time dependence and its other parameters, if not its general mathematical models.

2. Truncated series by man-made processes. It often occurs that man-made factors either create hydrological time series with zero values or decrease some values relative to others. The diversion of water from one river basin to another is sufficient to alter or truncate the distribution of the flow of the first river or to alter the distribution of the flow of the second river. In this case, the first river shows a reduction in its flow and also increased length of zero values if the river is dried up due to diversion, while the other river experiences an improved sustained flow and increased low flows. It is expected that the

*Ph.D. Graduate Student of Colorado State University, Civil Engineering Department, Fort Collins, Colorado

**Professor-in-Charge of Hydrology Program, Civil Engineering Department, Colorado State University, Fort Collins, Colorado, USA

simple permanent transfer of water from one river to another decreases time dependence for the former and increases it for the latter. The analytical or experimental studies of these man-made effects are of significant practical importance. This is the second subject treated in this paper.

The various practical manipulations of lowest river flows (withdrawal for water supply, irrigation, recharge of groundwater, flow regulations, etc.) make the time dependence of the resulting new regime a very complex phenomenon. It is treated, however, in this paper in a simplified schematic way because the varieties of resulting effects are not an easy subject for systematization.

3. The approach of studying the effect of zero values and man-made withdrawals on the dependence of time series. No suitable technique has yet been devised to analyze hydrological time series which have truncation either by the natural phenomenon or by man-made withdrawals or additions. Such transformations distort the structure of the continuous time series and thereby change its dependence properties. The present day analysis of time series emphasizes the study of their structure by their decomposition into different components, such as jump or trend, periodic component, stochastic component, etc. But time series which are truncated either by natural or man-made processes should first be studied for the effect of truncation before they are studied for their general structure.

The approach of this investigation is to determine the effects of truncation on various components of time series, such as pure stochastic or pure deterministic components, or their combination.

Chapter II

GENERAL ANALYSIS

1. Definition of the problem. A time series with significant length of zero value for the case of a cumulative variable may be considered as truncation by a physical process. This truncation can be obtained by a deduction from a continuous or discrete time series of a constant, of a linear or non-linear deterministic model, of a dependent or independent stochastic model, or a combination of any or all of these truncations. The physical process of truncation or addition may be very complex. As the simplest step in the investigation of characteristics of a time series under various types of truncation models, the truncation by deduction of a constant value (c) is first considered in this study. The truncation has several effects on a variable and its time series, but in this paper, the effect of such a constant truncation on autocorrelation coefficients of time series only is analyzed.

It is logical first to study the truncation of independent time series. However, this case is usually best shown when the dependence parameters are equated with their values of independent case, in the developed expressions for the case of dependent time series.

2. Effect of truncation on autocorrelation coefficients for the first order Markov linear process. Let x_t be a discrete time, stationary first order Markov linear process. It is truncated at any real constant value c such that

$$y_t = x_t - c \text{ if } x_t > c, \text{ with } -\infty < c < \infty \text{, and}$$
$$y_t = 0 \text{ if } x_t < c . \tag{1}$$

The autocorrelation function for the truncated process is derived for the cases when x_t at any time t is (a) normally distributed; and (b) lognormally distributed variable.

3. Normal dependent process. The first order Markov linear model is defined here in the form

$$x_t = \mu + \rho\,(x_{t-1} - \mu) + \sqrt{1 - \rho^2}\;\varepsilon_t \tag{2}$$

where ε_t is an independent normal variable (independent stochastic component) with the mean zero ($\mu_\varepsilon = 0$), and the variance σ^2 (σ_ε^2) ; x_t is dependent normal variable with the mean μ and the variance σ^2 (same as ε_t) and the population autocorrelation coefficients are: $\rho_\tau = \rho^\tau$.

The process of eq. (2) is truncated by the model of eq. (1) so that

$$E\,(y_t) = \int_c^\infty \frac{x-c}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

$$= \frac{\sigma}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{c-\mu}{\sigma}\right)^2\right\} - (c-\mu)\left[1 - \Phi\left(\frac{c-\mu}{\sigma}\right)\right] \tag{3}$$

$$E\,(y_t^2) = \int_c^\infty \frac{(x-c)^2}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx$$

$$= \left[(c-\mu)^2 + \sigma^2\right]\left[1-\Phi\left(\frac{c-\mu}{\sigma}\right)\right] - \frac{\sigma}{\sqrt{2\pi}}(c-\mu)\exp\left\{-\frac{1}{2}\left(\frac{c-\mu}{\sigma}\right)^2\right\} \tag{4}$$

where $\Phi\left(\frac{c-\mu}{\sigma}\right)$ represents the cumulative value of the standard normal density function from $-\infty$ to $\frac{c-\mu}{\sigma}$.

$$\operatorname{Var} y_t = (c-\mu)^2\left[1-\Phi\left(\frac{c-\mu}{\sigma}\right)\right]\Phi\left(\frac{c-\mu}{\sigma}\right) - \frac{\sigma}{\sqrt{2\pi}}(c-\mu)\left[\Phi\left(\frac{c-\mu}{\sigma}\right)-1\right]$$
$$\exp\left\{-\frac{1}{2}\left(\frac{c-\mu}{\sigma}\right)^2\right\} +$$
$$+\,\sigma^2\left[1 - \Phi\left(\frac{c-\mu}{\sigma}\right)\right] - \frac{\sigma^2}{2\pi}\exp\left\{-\left(\frac{c-\mu}{\sigma}\right)^2\right\} \tag{5}$$

To find the autocorrelation function, first the expected value of the cross-product function is obtained as

$$E\,(y_t y_{t+\tau}) = \int_c^\infty\int_c^\infty \frac{(x_1-c)(x_2-c)}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp - \frac{1}{2(1-\rho^2)}$$
$$\left[\left(\frac{x_1-\mu}{\sigma}\right)^2 + \left(\frac{x_2-\mu}{\sigma}\right)^2 - 2\rho_\tau\left(\frac{x_1-\mu}{\sigma}\right)\left(\frac{x_2-\mu}{\sigma}\right)\right] dx_1\, dx_2 . \tag{6}$$

Standardizing the variables and using the inversion theorem for the characteristic function of bivariate normal distribution and expanding the integrand in ascending power of ρ_τ , the above equation becomes

$$E\,(y_t y_{t+\tau}) = \sum_{j=0}^\infty \frac{(-\rho_\tau)^j}{j!}\left\{\int_{\frac{c-\mu}{\sigma}}^\infty (\sigma x_1 + \mu - c)\left[\frac{1}{2\pi}\int_{-\infty}^\infty u^j\right.\right.$$
$$\left.\left.\exp\left(-\frac{1}{2}u^2 - iux_1\right) du\right] dx_1\right\}\cdot\left\{\int_{\frac{c-\mu}{\sigma}}^\infty (\sigma x_2 + \mu - c)\right.$$
$$\left.\left[\frac{1}{2\pi}\int_{-\infty}^\infty v^j \exp\left(-\frac{1}{2}v^2 - i\,vx_2\right) dv\right] dx_2\right\} . \tag{7}$$

The coefficient of $\frac{(-\rho_\tau)^j}{j!}$ is the product of two integrals of which the first is

$$I(\frac{c-\mu}{\sigma} x_1, u) = \int_{\frac{c-\mu}{\sigma}}^{\infty} (\sigma x_1 + \mu - c)$$

$$[\frac{1}{2\pi} \int_{-\infty}^{\infty} u^j \exp(-\frac{1}{2} u^2 - iux_1) du] dx_1, \qquad (8)$$

and the second is $I(\frac{c-\mu}{\sigma}, x_2, v)$ similar to eq. (8). Both the integrals are essentially the same because x_t is assumed to be a stationary Markov process.

The integral in the square bracket of eq. (8) is

$$(-i)^j H_j(x_1) \alpha(x_1)$$

where $\alpha(x_1) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} x_1^2)$ and $H_j(x_t)$ are Chebyshev-Hermite polynomials [see reference (1)].

For $j = 0$, eq. (8) becomes

$$I(\frac{c-\mu}{\sigma}, x_1, u) = \int_{\frac{c-\mu}{\sigma}}^{\infty} (\sigma x_1 + \mu - c)$$

$$H_o(x_1) \alpha(x_1) dx_1 = E(y_t). \qquad (9)$$

For $j = 1$, eq. (8) gives

$$I(\frac{c-\mu}{\sigma}, x_1, u) = (-i) \int_{\frac{c-\mu}{\sigma}}^{\infty} (\sigma x_1 + \mu - c)$$

$$H_1(x_1) \alpha(x_1) dx_1 = (-i) \sigma [1 - \Phi(\frac{c-\mu}{\sigma})]. \qquad (10)$$

For $j \geq 2$, eq. (8) becomes

$$I(\frac{c-\mu}{\sigma}, x_1, u) = (-i)^j \int_{\frac{c-\mu}{\sigma}}^{\infty} (\sigma x_1 + \mu - c)$$

$$H_j(x_1) \alpha(x_1) dx_1 \qquad . \qquad (11)$$

Using the identity

$$(-D)^r \alpha(x) = H_r(x) \alpha(x) \qquad (12)$$

with D^r as the symbol for r-th derivative, and integrating by parts, eq. (11) reduces to

$$I(\frac{c-\mu}{\sigma}, x_1, u) = (-i)^j \sigma H_{j-2}(\frac{c-\mu}{\sigma}) \alpha(\frac{c-\mu}{\sigma}). \qquad (13)$$

Substituting eqs. (9), (10) and (13) into eq. (7), then

$$\text{cov}(y_t y_{t+\tau}) = \rho_\tau \sigma^2 [1+\Phi(\frac{c-\mu}{\sigma})]^2 + \frac{\sigma^2}{2\pi} \exp\left\{-(\frac{c-\mu}{\sigma})^2\right\}$$

$$\sum_{j=2}^{\infty} \frac{\rho_\tau^{\,j}}{j!} H_{j-2}^2 (\frac{c-\mu}{\sigma}) \quad . \qquad (14)$$

The autocorrelation function is then given by

$$\rho_\tau(c) = \rho_\tau \sigma^2 [1 - \Phi(\frac{c-\mu}{\sigma})]^2 + \frac{\sigma^2}{2\pi} \exp\left\{-(\frac{c-\mu}{\sigma})^2\right\}$$

$$\sum_{j=2}^{\infty} \frac{\rho_\tau^{\,j}}{j!} H_{j-2}^2 (\frac{c-\mu}{\sigma}) \Big/ (c-\mu)^2 [1-\phi(\frac{c-\mu}{\sigma})] \quad \phi(\frac{c-\mu}{\sigma}) -$$

$$-\frac{\sigma}{\sqrt{2\pi}} (c-\mu)[\Phi(\frac{c-\mu}{\sigma}) - 1] \exp\left\{-\frac{1}{2}(\frac{c-\mu}{\sigma})^2\right\} +$$

$$+ \sigma^2 [1 - \Phi(\frac{c-\mu}{\sigma})] - \frac{\sigma^2}{2\pi} \exp\left\{-(\frac{c-\mu}{\sigma})^2\right\} \quad . \qquad (15)$$

Putting $c \to -\infty$ in eq. (15), $\rho_\tau(c) = \rho_\tau$ as for eq. (2).

For the case of an independent normal random variable $\rho_\tau = 0$, and hence for independent normal random variable eq. (15) reduces to $\rho_\tau(c) = 0$. This shows that, an independent process remains independent after any constant truncation. For the case of the first order Markov linear chain, the autocorrelation coefficients for any constant truncation c are given by eq. (15).

4. Lognormal dependent process. The lognormal variable does not have additive but rather multiplicative reproducing properties. In order to keep the variable lognormal after the application of the first order Markov linear process, the process should be

$$\ln x_t - \mu_n = \rho (\ln x_{t-1} - \mu_n) + \sqrt{1-\rho^2} \; \varepsilon_t \qquad (16)$$

where ε_t is an independent normal variable with mean zero and variance $\sigma_n^{\,2}$, and x_t is a dependent lognormal variable with mean $\mu = e^{\mu_n + \frac{1}{2}\sigma_n^{\,2}}$ and variance $\sigma^2 = e^{\mu_n + \sigma_n^{\,2}} (e^{\sigma_n^{\,2}} - 1)$, where μ_n is the mean of logarithms of the random variable x_t. Equation (16) can also be expressed as

$$x_t = e^{\mu_n + \rho (\ln x_{t-1} - \mu_n)} \; e^{\sqrt{1-\rho^2} \; \varepsilon_t} \quad . \qquad (17)$$

The autocorrelation coefficients of the dependence model of eq. (17) are

$$\rho'_\tau = \frac{e^{\rho_\tau \sigma_n^{\,2}} - 1}{e^{\sigma_n^{\,2}} - 1} \quad . \qquad (18)$$

The process of eq. (17) is truncated according to eq. (1) for $0 < c < \infty$, giving

$$E(y_t) = \int_c^{\infty} \left(\frac{x-c}{\sqrt{2\pi}}\right) \exp\left\{-\frac{1}{2}\left(\frac{\ln x - \mu_n}{\sigma_n}\right)^2\right\} d \ln x$$

$$= \exp\left(\frac{1}{2}\sigma_n^{\,2} + \mu_n\right)\left[1 - \Phi\left(\frac{\ln c - \mu_n}{\sigma_n} - \sigma_n\right)\right] - c\left[1 - \Phi\left(\frac{\ln c - \mu_n}{\sigma_n}\right)\right]. \qquad (19)$$

$$E(y_t^2) = \int_c^\infty \frac{(x-c)^2}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x-\mu_n}{\sigma_n}\right)^2\right\} d \ln x$$

$$= \exp(2\sigma_n^2 + 2\mu_n)\left[1 - \Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - 2\sigma_n\right)\right] -$$

$$- 2c \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right] +$$

$$+ c^2\left[1 - \Phi\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\right] \quad . \qquad (20)$$

From eqs. (20) and (19), the variance of y_t is obtained as var $y_t = E(y_t^2) - [E(y_t)]^2$. Similarly, as in eq. (6) but with the use of eq. (17) for $y_t = x_t - c$, if $x_t > c$ or $y_t = 0$ for $x_t < c$, the expected value of $E(y_t y_{t+\tau})$ is obtained with the replacements $\ln x_t = u$ and $\ln x_{t+\tau} = v$:

$$E(y_t y_{t+\tau}) = \int_{\ln c}^\infty \int_{\ln c}^\infty (e^u-c)(e^v-c) \frac{1}{2\pi\sigma_n^2\sqrt{1-\rho^2}}$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{u-\mu_n}{\sigma_n}\right)^2 - \right.\right.$$

$$\left.\left. - 2\rho_\tau\left(\frac{u-\mu_n}{\sigma_n}\right)\left(\frac{v-\mu_n}{\sigma_n}\right) + \left(\frac{v-\mu_n}{\sigma_n}\right)^2\right]\right\} du\, dv \quad . \qquad (21)$$

Integrating eq. (21) in the same way as eq. (6), the covariance of y_t and $y_{t+\tau}$ can be expressed as

$$\text{cov}\ (y_t y_{t+\tau}) = \rho_\tau\sigma_n^2 \exp(\sigma_n+2\mu_n)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]^2$$

$$+ \sum_{j=2}^\infty \frac{\rho_\tau^{\ j}}{j!}\left\{c\sum_{k=2}^j H_{j-k}\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\alpha\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\sigma_n^{k-1} + \sigma_n^j\right.$$

$$\left.\exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]\right\}^2 \quad . \qquad (22)$$

The autocorrelation function is given by

$$\rho_\tau(c) = \rho_\tau\sigma_n^2 \exp\ (\sigma_n^2 + 2\mu_n)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]^2$$

$$+ \sum_{j=2}^\infty \frac{(\rho_\tau)^j}{j!}\left\{c\sum_{k=2}^j H_{j-k}\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\alpha\left(\frac{\ln c-\mu_n}{.\ \sigma_n}\right)\sigma_n^{k-1} +\right.$$

$$\left. + \sigma_n^j \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]\right\}^2 \Bigg/$$

$$\left\{\exp\ (2\sigma_n^2 + 2\mu_n)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - 2\sigma_n\right)\right]\right.$$

$$- \exp(\sigma_n^2 + 2\mu_n)\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]^2 -$$

$$- 2c \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right)$$

$$\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n} - \sigma_n\right)\right]\Phi\left(\frac{\ln c-\mu_n}{\sigma_n}\right) +$$

$$\left. + c^2\left[1-\Phi\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\right]\Phi\left(\frac{\ln c-\mu_n}{\sigma_n}\right)\right\} \quad . \qquad (23)$$

For the case of an independent lognormal process, $\rho_\tau = 0$ for all τ , and eq. (23) reduces to $\rho_\tau(c) = 0$. So an independent lognormal process remains independent after any constant truncation.

Equation (23) is valid for any positive truncation. For negative truncation, it needs some modification. Since the legnormal variables have only positive values, for any negative truncation the mathematical expectation of any function of x_t is extended over the whole range of the random variable x_t . Therefore setting $\ln c = -\infty$ for $c \leq 0$, as if $c = 0$, eq. (23) can be simplified to

$$\rho_\tau(c) = \frac{e^{\rho_\tau\sigma_n^2} - 1}{e^{\sigma_n^2} - 1} \quad ,$$

in the same way as eq. (18). Therefore, $\rho_\tau(c)$ is independent of c for $c \leq 0$.

$$E(y_t) = \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right) - c \quad , \qquad (24)$$

$$\text{var}\ y_t = \exp\ (\sigma_n^2 + 2\mu_n)\left(e^{\sigma_n^2} - 1\right) \quad . \qquad (25)$$

Therefore, from eqs. (23) through (25), it follows that for any negative truncation $(c < 0)$, only the mean is changed.

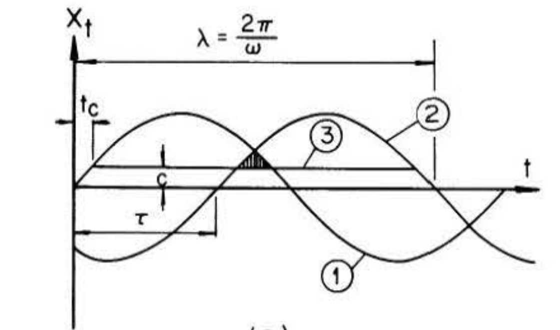
The case of a value of c being negative can be easily conceived as the constant water discharge diverted into a river with relatively low flows.

As hydrologic time series often have a deterministic component (monthly values have the yearly cycle, hourly values have the daily cycle), first a simple sine function is investigated in the following text, and then a combination of sine function and a dependent random variable is studied. However, it is difficult with this general analysis to come very close to the real structure of hydrological time series or to the various complex truncation models which occur in practice.

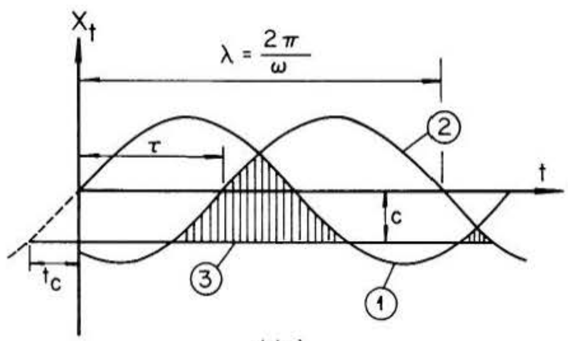
5. Effect of truncation on autocorrelation coefficients for a sinusoidal series. Let $x_t = A \sin \omega t$, a continuous deterministic time series be truncated as follows:

$$\left.\begin{aligned} y_t &= x_t - c \quad \text{if } x_t \geq c \\ y_t &= 0 \quad \text{if } x_t < c \end{aligned}\right\} - A \leq c \leq A \tag{26}$$

Figure 1 shows: (a) the truncation at $0 \leq c \leq A$ and (b) the truncation at $-A \leq c \leq 0$.



(a)



(b)

Fig. 1. Schematic diagram for the lagged cross product of a sinusoidal series: (1) sine series; (2) shifted series for τ ; (3) level of truncation

The level of truncation can be expressed as $A \sin \omega t_c = c$, where

$$t_c = \frac{1}{\omega} \sin^{-1} \frac{c}{A} \tag{27}$$

with

$$-\frac{\pi}{2\omega} \leq t_c \leq \frac{\pi}{2\omega} \tag{28}$$

then the expected values of y_t are

$$E(y_t) = A\left[\frac{\cos \omega t_c}{\pi} - \frac{\sin \omega t_c}{2} + \omega t_c \frac{\sin \omega t_c}{\pi}\right] \tag{29}$$

$$E(y_t^2) = A^2 \left[\frac{2 \sin^2 \omega t_c + 1}{4} - \frac{(2 \sin^2 \omega t_c + 1)\, \omega t_c}{2\pi} - \frac{3 \sin 2 \omega t_c}{4\pi}\right] . \tag{30}$$

The transformed variable y_t is not a continuous function of time. Therefore, to find the cross product of the function with respect to any lag, it is necessary to look into the length of the series where $x_t \leq c$ or $y_t = 0$. When the series is truncated at its negative values, i.e., t_c is negative and $\pi/\omega + 2t_c \leq \tau \leq \pi/\omega - 2t_c$, the cross product with any lag τ is

$$E(y_t y_{t+\tau}) = \frac{1}{\lambda}\left[\int_{(\tau+t_c)}^{\frac{\pi}{\omega} - t_c} (A \sin\omega t - c)(A \sin\omega(t+\tau) - c)\, dt\right.$$

$$\left. + \int_{\frac{2\pi}{\omega} + t_c}^{\frac{\pi}{\omega} + \tau - t_c} (A \sin \omega t - c)(A \sin\omega(t+\tau) - c)\, dt\right] \tag{31}$$

which upon integration becomes

$$\begin{aligned} E(y_t y_{t+\tau}) = \frac{A^2}{2\pi}\Bigg\{&\left[\frac{\pi - \omega\tau}{2} - \omega t_c + \frac{\sin 2 \omega t_c}{4} + \frac{\sin 2\omega(\tau + t_c)}{4}\right] \cos \tau - \\ &- \left[\sin^2 \omega t_c - \sin^2 \omega(\tau + t_c)\right] \frac{\sin \omega\tau}{2} - \\ &- 2\left[\cos \omega t_c + \cos \omega(\tau + t_c)\right] \sin \omega t_c + \\ &+ (\pi - \omega\tau - 2\omega t_c) \sin^2 \omega t_c\Bigg\} + \\ + \frac{A^2}{2\pi}\Bigg\{&\left[\frac{\omega\tau - \pi}{2} - \omega t_c - \frac{\sin 2\omega(\tau - t_c)}{4} + \frac{\sin 2 \omega t_c}{4}\right] \cos \tau - \\ &- \left[\sin^2 \omega(\tau - t_c) - \sin^2 \omega t_c\right] \frac{\sin \omega\tau}{2} - \\ &- 2\left[\cos \omega(\tau - t_c) + \cos \omega t_c\right] \sin \omega t_c + \\ &+ (\omega\tau - 2\omega t_c - \pi) \sin^2 \omega t_c\Bigg\} . \end{aligned} \tag{32}$$

For all truncation when $0 \leq \tau \leq \pi/\omega - 2t_c$, the second integral in eq. (31) does not exist and, therefore,

$$E(y_t y_{t+\tau}) = \frac{A^2}{2\pi}\Bigg\{\left[\frac{\pi - \omega\tau}{2} - \omega t_c + \frac{\sin 2 \omega t_c}{4} + \frac{\sin 2\omega(\tau + t_c)}{4}\right] \cos \omega\tau -$$

$$- \left[\sin^2 \omega t_c - \sin^2 \omega(\tau+t_c)\right] \frac{\sin \omega\tau}{2} -$$

$$- 2\left[\cos \omega t_c + \cos \omega(\tau+t_c)\right] \sin \omega t_c +$$

$$\left. + (\pi - \omega\tau - 2\omega t_c) \sin^2 \omega t_c \right\} . \quad (33)$$

Similarly, when $\pi/\omega + 2t_c \leq \tau \leq 2\pi/\omega$, the first integral of eq. (31) does not exist for which case

$$E(y_t y_{t+\tau}) = \frac{A^2}{2\pi}\left\{\left[\frac{\omega\tau-\pi}{2} - \omega t_c - \frac{\sin 2\omega(\tau-t_c)}{4} + \frac{\sin 2\omega t_c}{4}\right] \cos \tau - \left[\sin^2 \omega(\tau-t_c) - \sin^2 \omega t_c\right] \frac{\sin \omega\tau}{2} - 2\left[\cos \omega(\tau-t_c) + \cos \omega t_c\right] \sin \omega t_c + (\omega\tau-2\omega t_c-\pi) \sin^2 \omega t_c\right\} . \quad (34)$$

When t_c is positive and $\pi/\omega - 2t_c \leq \tau \leq \pi/\omega + 2t_c$, none of the integrals in eq. (31) exists for which

$$E(y_t y_{t+\tau}) = 0 . \quad (35)$$

The evaluation of $E(y_t y_{t+\tau})$ according to eqs. (31) through (35) depends on τ and the level of truncation c. By using $E(y_t)$ from eq. (29) and $E({y_t}^2)$ from eq. (30), the autocorrelation coefficients can be computed by

$$\rho_\tau(c) = \frac{E(y_t y_{t+\tau}) - \left[E(y_t)\right]^2}{E({y_t}^2) - \left[E(y_t)\right]^2} . \quad (36)$$

6. Effect of truncation on the autocorrelation coefficients for a time series composed of a periodic component and a stationary stochastic component. Let the discrete process x_t be composed of a deterministic periodic component $P_t = A \sin \omega t$ plus a stochastic component η_t, where A is the amplitude of the periodic component, ω is the angular frequency and $\lambda = 2\pi/\omega$ becomes the period. The process x_t is analyzed for truncation when η_t at any particular time follows (a) a normal distribution and (b) a lognormal distribution.

Let the process $x_t = \eta_t + A \sin \omega t$ be truncated by any real constant c such that

$$\left.\begin{aligned} y_t &= x_t - c \quad \text{when} \quad x_t \geq c \text{ or } \eta_t \geq c - A \sin \omega t \\ y_t &= 0 \quad \text{when} \quad x_t < c \text{ or } \eta_t < c - A \sin \omega t \end{aligned}\right\} . \quad (37)$$

(a) Normal distribution. Let η_t be normally distributed with mean μ and variance σ^2. It is further assumed that η_t follows the first order Markov linear model according to eq. (2).

Following the truncation of eq. (37) and letting

$$h_t = c - A \sin \omega t \text{ and } h'_{t+\tau} = c - A \sin \omega(t+\tau) ,$$

it can be written that

$$Ey(t) = \frac{1}{\lambda}\sum_{t=1}^{\lambda}\int_{h_t}^{\infty} \frac{(\eta-h_t)}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\eta-\mu}{\sigma}\right)^2\right\} d\eta$$

$$= \frac{1}{\lambda}\sum_{t=1}^{\lambda}\left[\frac{\sigma}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{h_t-\mu}{\sigma}\right)^2\right\} - (h_t-\mu)\left\{1 - \Phi\left(\frac{h_t-\mu}{\sigma}\right)\right\}\right] \quad (38)$$

$$Ey^2(t) = \frac{1}{\lambda}\sum^{\lambda}_{h_t} \frac{(\eta-h_t)^2}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\eta-\mu}{\sigma}\right)^2\right\} d\eta$$

$$= \frac{1}{\lambda}\sum_{t=1}^{\lambda}\left\{\left[(h_t-\mu)^2 + \sigma^2\right]\left[1-\Phi\left(\frac{h_t-\mu}{\sigma}\right)\right] - \sigma\frac{(h_t-\mu)}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{h_t-\mu}{\sigma}\right)^2\right]\right\} . \quad (39)$$

To find the autocorrelation function,

$$Ey_t y_{t+\tau} = \frac{1}{\lambda}\sum_{t=1}^{\lambda}\int_{h_t}^{\infty}\int_{h_{t+\tau}}^{\infty}$$

$$\frac{(\eta_1-h_t)(\eta_2-h_{t+\tau})}{2\pi\sigma^2 \ (1-\rho_\tau^2)} \exp\left\{-\frac{1}{2(1-\rho_\tau^2)}\left[\left(\frac{\eta_1-\mu}{\sigma}\right)^2 + \left(\frac{\eta_2-\mu}{\sigma}\right)^2 - 2\rho_\tau\left(\frac{\eta_1-\mu}{\sigma}\right)\left(\frac{\eta_2-\mu}{\sigma}\right)\right]\right\} d\eta_1 \, d\eta_2 . \quad (40)$$

Following the integration of eq. (6), it can be obtained that

$$Ey_t y_{t+\tau} = \frac{1}{\lambda}\sum_{t=1}^{\lambda}\left[\left\{\frac{\sigma}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{h_t-\mu}{\sigma}\right)^2\right] - (h_t-\mu)\left[1-\Phi\left(\frac{h_t-\mu}{\sigma}\right)\right]\right\}\left\{\frac{\sigma}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{h_{t+\tau}-\mu}{\sigma}\right)^2 - (h_{t+\tau}-\mu)\left[1-\Phi\left(\frac{h_{t+\tau}-\mu}{\sigma}\right)\right]\right\} +\right.$$

$$+ \rho_\tau \sigma^2 \left[1-\Phi\left(\frac{h_t - \mu}{\sigma}\right)\right] \left[1-\Phi\left(\frac{h_{t+\tau} - \mu}{\sigma}\right)\right] +$$

$$+ \frac{\sigma^2}{2\pi} \sum_{j=2}^{\infty} \frac{\rho_\tau^{\,j}}{j!} H_{j-2}\left(\frac{h_t-\mu}{\sigma}\right) H_{j-2}\left(\frac{h_{t+\tau}-\mu}{\sigma}\right)$$

$$\left.\exp\left\{-\frac{1}{2}\left[\left(\frac{h_t-\mu}{\sigma}\right)^2 + \left(\frac{h_{t+\tau}-\mu}{\sigma}\right)^2\right]\right\}\right] \qquad (41)$$

The autocorrelation coefficients can be computed with $Ey_t y_{t+\tau}$, $Ey_t^{\,2}$, Ey_t given by eqs. (38), (39) and (41), respectively.

(b) <u>Lognormal distribution</u>. Let η_t be lognormally distributed with μ_n and σ_n^2 as the mean and variance of the logarithms of the variable. Also, $\ln \eta_t$ follows the first order Markov linear process, so that the dependence model of η_t is given by eq. (17).

Let the process $x_t = \eta_t + A \sin \omega t$ be truncated at any level c according to eq. (37), such that $0 \le c - A \sin \omega t < \infty$. Then substituting $c - A \sin \omega t = h_t$ and $c - A \sin \omega(t+\tau) = h_{t+\tau}$, it can be obtained that

$$E(y_t) = \frac{1}{\lambda} \sum_{t=1}^{\lambda} \left\{\exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right) \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - \sigma_n\right)\right] - h_t \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n}\right)\right]\right\} . \qquad (42)$$

$$E(y_t^2) = \frac{1}{\lambda} \sum_{t=1}^{\lambda} \left\{\exp(2\sigma_n^2 + 2\mu_n)\left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - 2\sigma_n\right)\right] - 2h_t \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right) \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - \sigma_n\right)\right] + h_t^{\,2} \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n}\right)\right]\right\} \qquad (43)$$

and

$$Ey_t y_{t+\tau} = \frac{1}{\lambda} \sum_{t=1}^{\lambda} \left[\left\{\exp\left(\frac{1}{2}\sigma_n^2+\mu_n\right) \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - \sigma_n\right)\right] - h_t\left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n}\right)\right]\right\}\left\{\exp\left(\frac{1}{2}\sigma_n^2+\mu_n\right)\left[1-\Phi\left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n} - \sigma_n\right)\right]\right.\right.$$

$$\left. - h_{t+\tau}\left[1-\Phi\left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n}\right)\right]\right\} + \rho_\tau \sigma_n^2 \exp(\sigma_n^2 + 2\mu_n)$$

$$\left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - \sigma_n\right)\right] \left[1-\Phi\left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n} - \sigma_n\right)\right] +$$

$$+ \sum_{j=2}^{\infty} \frac{\rho_\tau^{\,j}}{j!} \left\{h_t \sum_{k=2}^{j} H_{j-k}\left(\frac{\ln h_t-\mu_n}{\sigma_n}\right) \alpha\left(\frac{\ln h_t-\mu_n}{\sigma_n}\right)\sigma_n^{k-1} + \sigma_n^j\right.$$

$$\left.\exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right) \left[1-\Phi\left(\frac{\ln h_t-\mu_n}{\sigma_n} - \sigma_n\right)\right]\right\} \left\{h_{t+\tau} \sum_{k=2}^{j}\right.$$

$$H_{j-k}\left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n}\right) \alpha \left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n}\right)$$

$$\left.\left.\sigma_n^{k-1} + \sigma_n^{\,j} \exp\left(\frac{1}{2}\sigma_n^2 + \mu_n\right)\left[1-\Phi\left(\frac{\ln h_{t+\tau}-\mu_n}{\sigma_n} - \sigma_n\right)\right]\right\}\right] . \qquad (44)$$

The autocorrelation coefficients can be computed from the values of Ey_t , Ey_t^2 and $Ey_t y_{t+\tau}$ as given by eqs. (42), (43), and (44), respectively.

When $h_t \le 0$ or $h_{t+\tau} \le 0$, i.e., $c < A \sin 2\omega t$ or $A \sin 2\omega(t+\tau)$, the range of expectation of $y(t)$ extends over the entire region of the lognormal variable, i.e., from 0 to ∞ . Therefore, in eqs. (42), (43) and (44), $\ln h_t$ or $\ln h_{t+\tau}$ is to be replaced by $-\infty$ where $h_t \le 0$ and $h_{t+\tau} \le 0$. When $h_t \le 0$ and $h_{t+\tau} \le 0$ for all t , then the autocorrelation coefficients for the truncated series can be obtained as

$$\rho_\tau(c) = \frac{\exp(\sigma_n^2+2\mu_n)(e^{\rho_\tau \sigma_n^2}-1) + \frac{A^2}{\lambda} \sum_{t=1}^{\lambda} \sin\omega t \sin\omega(t+\tau)}{\exp(\sigma_n^2+2\mu_n)(e^{\sigma_n^2}-1) + \frac{A^2}{\lambda} \sum_{t=1}^{\lambda} \sin^2 \omega t} . \qquad (45)$$

The correlogram of the untruncated series can also be obtained from eq. (45) since it is independent of truncation under the conditions of its derivation.

Chapter III

PRESENTATION AND DISCUSSION OF COMPUTATIONAL RESULTS

1. Computation of correlograms. The effects of truncations by constant values of c on the autocorrelation coefficients of time series can be determined by two methods as follows: (a) numerical solutions of equations developed in Chapter II, by approximations and (b) generating large samples of time series by Monte Carlo method (data generation method), truncating them with various constant values and computing correlograms of the truncated time series. The first approach is selected here for simple cases as investigated in Chapter II. However, the second approach is always available for complex cases. It is applied here only in two cases of daily river flows and only for a given case of constant truncation at the mean value.

All equations theoretically derived in Chapter II for the autocorrelation coefficients of stochastic processes involve the Chebyshev-Hermite polynomials [1]. It was found that for both truncated normal and truncated lognormal random variables the equations which involve sums of polynomials converge for all values of ρ and for all finite levels of truncation to the true finite values as j increases. The convergence is faster for lower values of ρ than for higher values, or fewer polynomials are necessary for the same accuracy for small ρ than for large ρ values. In all computations 14 polynomials were used. Tables 1 and 2 show the comparison of the first serial correlation coefficients computed with 9 polynomials with those computed with 14 polynomials for

Table 1. Comparative study of the convergence of eq. (15) by using 9 and 14 Chebyshev-Hermite polynomials in the computation of $\rho_1(c)$ for different values of ρ and different levels of truncation

Level of truncation	$\rho = 0.2$		$\rho = 0.4$		$\rho = 0.6$		$\rho = 0.8$	
	9 polynomials	14 polynomials	9 polynomials	14 polynomials	9 polynomials	14 polynomials	9 polynomials	14 polynomials
$-\infty$	.2	.2	.4	.4	.6	.6	.8	.8
-1	.18886	.18886	.38149	.38149	.57868	.57868	.78159	.78162
-0.5	.17606	.17606	.36172	.36172	.55797	.55797	.76671	.76674
0	.15587	.15587	.33086	.33086	.52633	.52634	.74538	.74548
0.5	.12498	.12498	.28002	.28002	.46827	.46827	.69565	.69575
1.0	.08696	.08696	.21215	.21215	.38355	.38356	.61314	.61341
2.0	.02140	.02140	.07022	.07022	.16617	.16618	.33951	.34003
3.0	.00155	.00155	.00863	.00863	.03232	.03234	.09780	.09835

Table 2. Comparative study of the convergence of eq. (23) by using 9 and 14 Chebyshev-Hermite polynomials in the computation of $\rho_1(c)$ for different values of ρ and different levels of truncation

Level of truncation	$\rho = 0.2$		$\rho = 0.4$		$\rho = 0.6$		$\rho = 0.8$	
	9 polynomials	14 polynomials	9 polynomials	14 polynomials	9 polynomials	14 polynomials	9 polynomials	14 polynomials
0	.12885	.12885	.28623	.28623	.47845	.47845	.71324	.71324
1	.10801	.10801	.25304	.25304	.44256	.44256	.68533	.68533
2	.08114	.08114	.20653	.20653	.38901	.38901	.64335	.64337
4	.05022	.05022	.14809	.14809	.31864	.31864	.59379	.59385
6	.03253	.03253	.10908	.10908	.26335	.26335	.54336	.54355
8	.02218	.02218	.08276	.08276	.21996	.21997	.49402	.49247
10	.01569	.01569	.06496	.06496	.18895	.18896	.45860	.45886
14	.00889	.00889	.04352	.04352	.14663	.14664	.40382	.40426
18	.00566	.00566	.03240	.03240	.12445	.12447	.38181	.38265

both normal and lognormal processes. The difference is not significant, and, hence, 14 polynomials are considered to be adequate for all computations. All computations are carried out on a high speed digital computer (CDC 6600) by using equations theoretically derived in Chapter II.

2. Normal dependent process. For a normal dependent stochastic process of the first order Markov linear model the autocorrelation coefficients of the truncated series were found to be dependent on four parameters: (1) ρ, the first autocorrelation coefficient of the Markov model; (2) c, truncation constant; (3) μ, mean of the non-truncated series; and (4) σ^2, variance of the non-truncated series, as it is shown by eq. (15).

As it is complicated to study the dependence of $\rho_\tau(c)$ of eq. (15) in function of all four parameters, the case is made simpler by investigating $\rho_\tau(c)$ only for a standardized normal variable with $\mu = 0$ and $\sigma^2 = 1$. Also, the truncation constant c is expressed in terms of the standard deviation σ of the original series. Four values of ρ in the equation $\rho_\tau = \rho^\tau$ are used: $\rho = 0.2$, $\rho = 0.4$, $\rho = 0.6$ and $\rho = 0.8$.

Figure 2 gives four graphs of autocorrelation coefficients $\rho_\tau(c)$, each for a given value of ρ. The first two graphs show three curves (for $c = -\infty$, $c = 0$, and $c = \sigma$), and the last two graphs give four curves (for $c = -\infty$, $c = 0$, $c = \sigma$, and $c = 2\sigma$). The case $c = -\infty$ corresponds to the first order Markov linear model, $\rho_\tau = \rho^\tau$, without truncation.

For $\rho = 0$, it follows that $\rho_\tau(c) = 0$ for any c. The independent normal process remains, therefore, independent after any constant truncation. The effect of the constant truncation c of the standard normal first order Markov linear model on

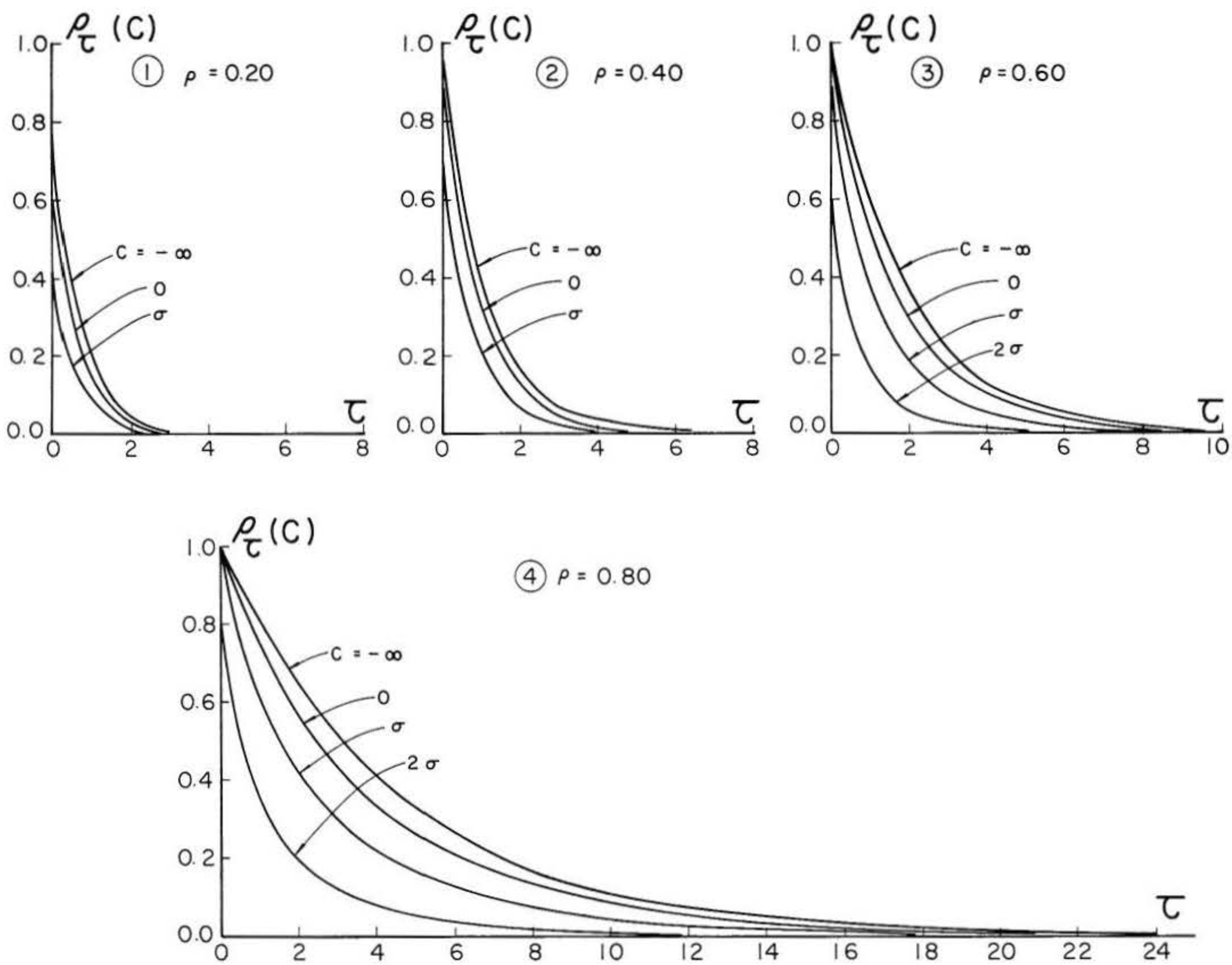


Fig. 2. Correlograms of the truncated normal process of the first order Markov linear model for: (1) $\rho = 0.20$; (2) $\rho = 0.40$; (3) $\rho = 0.60$; and (4) $\rho = 0.80$. The truncation constants c are either $-\infty$, 0 and σ; or $-\infty$, 0, σ and 2σ

the first autocorrelation coefficient $\rho_1(c)$ is shown in Fig. 3.

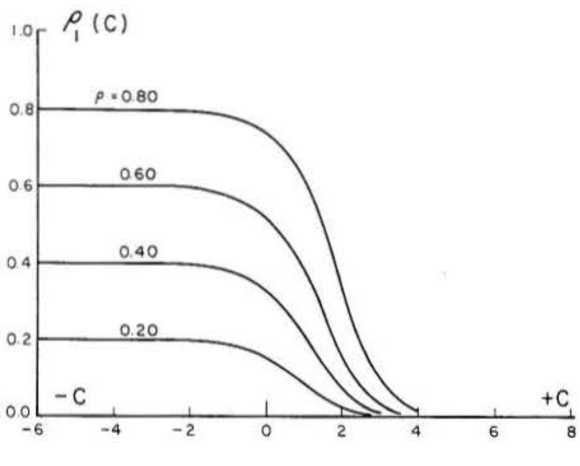


Fig. 3. First autocorrelation coefficient as a function of the level of truncation c for a normal process of the first order Markov linear model, for four values of $\rho_1 = \rho$: 0.20 ; 0.40 ; 0.60 and 0.80

The means and variances which correspond to the analytically derived eqs. (3) and (5), respectively, are computed for various values of truncation constant c , for $\mu = 0$ and $\sigma^2 = 1$, and are presented in Table 3.

Table 3. Means and variances of truncated standard normal dependent process ($\mu = 0$, $\sigma^2 = 1$, ρ)

Level of truncation c	Mean	Variance
$-\infty$	∞	1
-1	1.08054	.75186
-0.5	.69554	.55189
0	.39840	.34127
0.5	.19554	.17609
1	.08054	.07406
2	.00731	.00861
3	.00022	.00072

From Figs. 2 and 3 it is seen that all the autocorrelation coefficients decrease steadily with an increase in the level of truncation. For the normal dependent process the dependence model is almost unaffected when the relative level $(c-\mu)/\sigma$ of truncation is ≤ -2 . The rate of decrease is relatively rapid in the range $-2 \leq c-\mu/\sigma \leq 3$. The process becomes almost independent when $(c-\mu)/\sigma \geq 4$. It can be observed that the autocorrelation coefficients of untruncated series are independent of μ and σ^2 , whereas for the truncated series the opposite is true.

The closest practical hydrological case to this theoretical dependent time series of normal variable may be found in the time series of annual river flows with distributions close to normal, having drainage basins with large water storage capacities (or relatively large changes of water carryover in basins from year to year), so that the first order Markov linear models are approximately applicable. The truncations by a constant value would correspond to the diversion of a constant annual amount of water into other river basins, or the consumption of a constant water quantity every year by evaporation and evapotranspiration.

The annual river flows of the Saint Lawrence River below the Great Lakes are approximately normally distributed with the first serial correlation coefficient of about $\rho = 0.71$. A constant annual withdrawal from the Great Lakes (and confined or diverted in other river basins) would be equivalent to a constant truncation c . Therefore, the annual flows remaining in the river below the lakes would be less dependent in sequence after those constant diversions out of the basin.

3. Lognormal dependent process. As for the normal dependent process, the lognormal dependent stochastic process of eq. (17) of series truncated by a constant value has autocorrelation coefficients $\rho_\tau(c)$ which are dependent on the following four parameters: (1) the first autocorrelation coefficient ρ of the correlogram of eq. (18); (2) the truncation constant c ; (3) the mean μ_n of logarithms; and (4) the variance σ_n^2 of the logarithms of the lognormal dependent random variable; this is shown by eq. (23).

As for the normal dependent process, the parameters μ_n and σ_n^2 are selected as particular values, namely $\mu_n = 0$ and $\sigma_n^2 = 1$, so that only the relationship of $\rho_\tau(c) = f(\rho , c , \tau)$ is investigated by detailed computations. Also, the truncation constant c is expressed in terms of the standard deviation σ_n of the logarithms of the original time series. Four values of ρ in eq. (18) are used: 0.2 , 0.4 , 0.6 and 0.8.

Figure 4 gives four graphs of autocorrelation coefficients $\rho_\tau(c)$, each for a given value of ρ . The first two graphs are shown for three values of c ($c = 0$, $c = 0.5\sigma_n$, $c = \sigma_n$), the third graph for four values of c ($c = 0$, $c = 0.5\sigma_n$, $c = \sigma_n$, $c = 2\sigma_n$), and the fourth also for four values of c ($c = 0$, $c = \sigma_n$, $c = 2\sigma_n$ and $c = 3\sigma_n$). The case $c = 0$ corresponds to the first order Markov model of eq. (18).

For $\rho = 0$, it follows $\rho_\tau(c) = 0$ for any c . Therefore, an independent lognormal process remains independent after any constant truncation. The effect of the constant truncation c of the lognormal dependent process of eqs. (17) and (18) on the first autocorrelation coefficient $\rho_1(c)$ is shown in Fig. 5.

The mean of truncated lognormal dependent time series is given by eq. (19), and the variance [var $y_t = E(y_t^2) - E^2(y_t)$] are given by eqs. (20) and (19). They are computed for various values of truncation constant c , for $\mu_n = 0$ and $\sigma_n^2 = 1$, and are presented in Table 4.

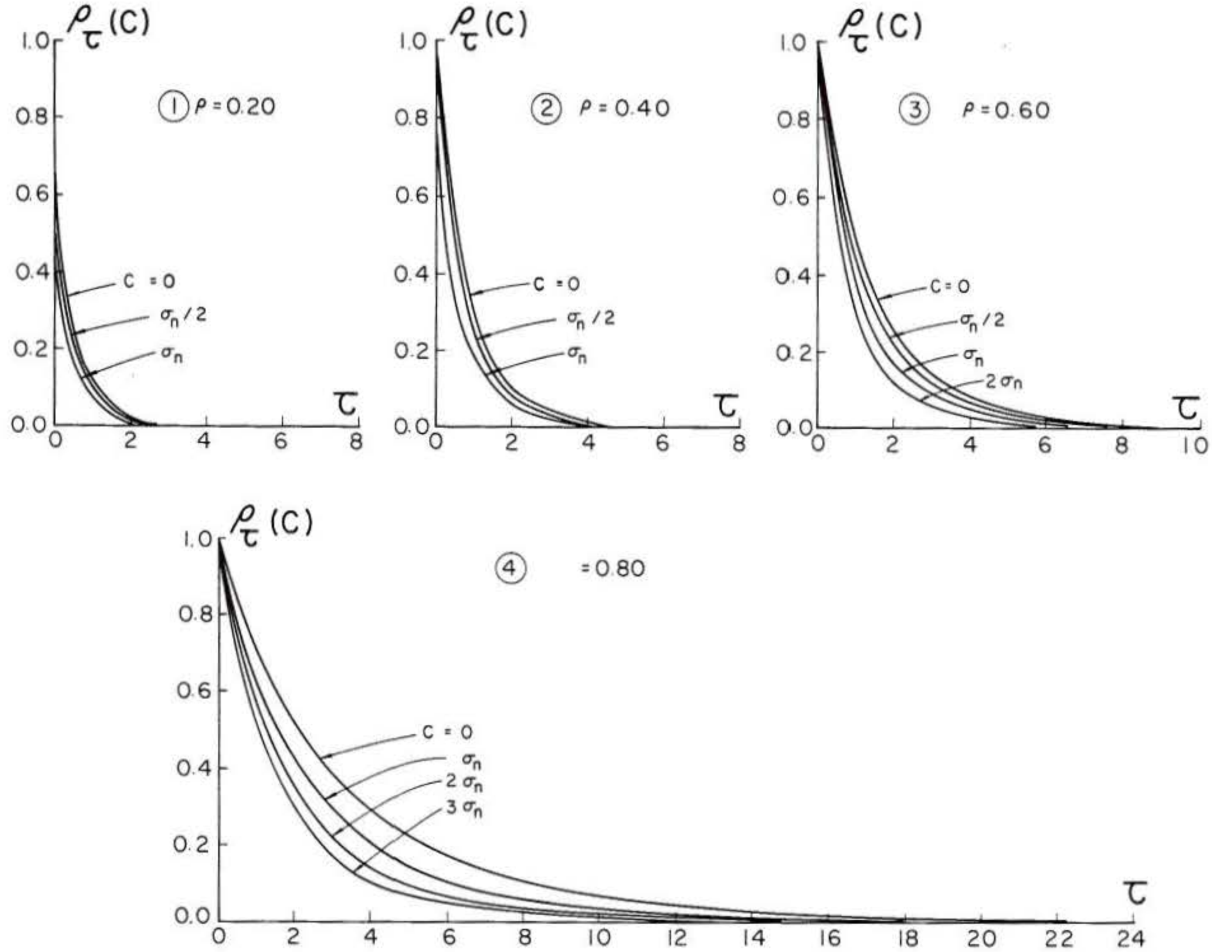


Fig. 4. Correlograms of the truncated lognormal process of the first order Markov linear model for: (1) $\rho = 0.20$; (2) $\rho = 0.40$; (3) $\rho = 0.60$ and (4) $\rho = 0.80$. The truncation constants c are either 0, $\sigma_n/2$, σ_n (first two graphs), or 0, $\sigma_n/2$, σ_n, $2\sigma_n$ (third graph), or 0, σ_n, $2\sigma_n$, $3\sigma_n$ (fourth graph)

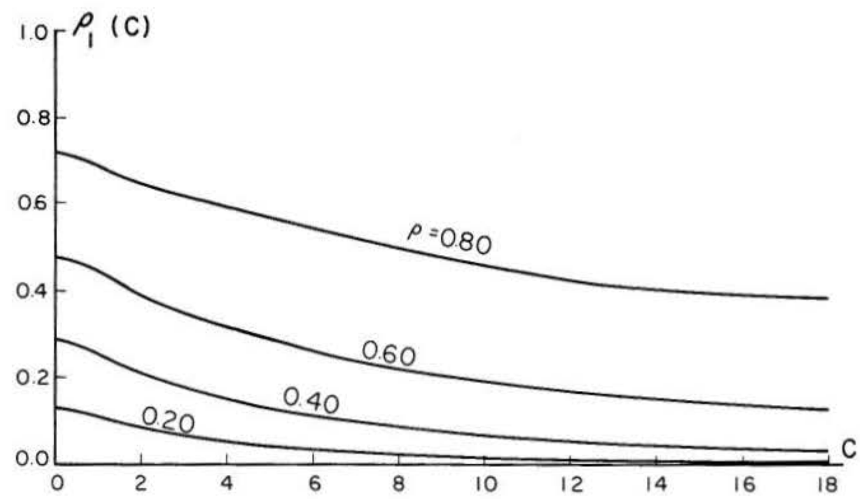


Fig. 5. First autocorrelation coefficient as a function of the level of truncation c for lognormal process of the first order Markov linear model of eqs. (17) and (18), for four values of $\rho_1 = \rho$: 0.20 , 0.40 , 0.60 and 0.80

Table 4. Means and variances of truncated lognormal dependent process ($\mu_n = 0$, $\sigma_n^2 = 1$, ρ)

Level of truncation c	Mean	Variance
0	1.64862	4.67077
1	.88311	4.17074
2	.52854	3.29986
4	.24515	2.02526
6	.13395	1.33399
8	.08101	.95680
10	.05260	.69867
12	.03522	.53504
14	.02487	.41307
16	.01851	.32053
18	.01427	.24993

As for the normal dependent process, the autocorrelation coefficients of a truncated lognormal dependent process steadily decrease with an increase of truncation level c . Figure 5 shows that $\rho_\tau(c)$ continuously decreases with an increase of c from zero to greater positive values. It should be noted that ρ_τ of untruncated series are dependent on σ_n^2 of logarithms of a lognormal variable, as shown by eq. (18), while for the truncated series the values $\rho_\tau(c)$ are also dependent on μ_n or the mean of logarithms of a lognormal variable.

As annual flows of many rivers are asymmetrically distributed (usually well fitted by lognormal distribution), time dependent because of changes in water storage in river basins and often well fitted by the first order Markov linear model of eq. (17) or similar. The above case of truncated lognormal dependent random variables is applicable whenever a constant annual amount of water is assigned to a diversion or unreturnable consumptive use. The above results can be applied to many rivers under these approximate conditions.

4. Truncation of a sinusoidal series by a constant value. The autocorrelation function of a truncated sinusoidal series depends on the amplitude of this series and the level of truncation. The autocorrelogram of a sinusoidal series is a cosine function. The period of the autocorrelogram is the same as the period of the original series, and its amplitude is unity. In eqs. (32), (33), (34) and (35), a complete cycle of the process was considered. The correlogram of the sinusoidal series at different levels of truncation, with both amplitude and angular frequency taken to be equal to unity, is plotted in Fig. 6. The total number of τ-values in a full period $\lambda = 2\pi/\omega$ is taken as 20 for plotting the graph in Fig. 6.

Figure 6 shows that the average absolute autocorrelation coefficient $E\{|\rho_\tau(c)|\}$ decreases with an increase of c from c = -1 to c = +0.707. or, in other words, the dependence decreases with an increase of c .

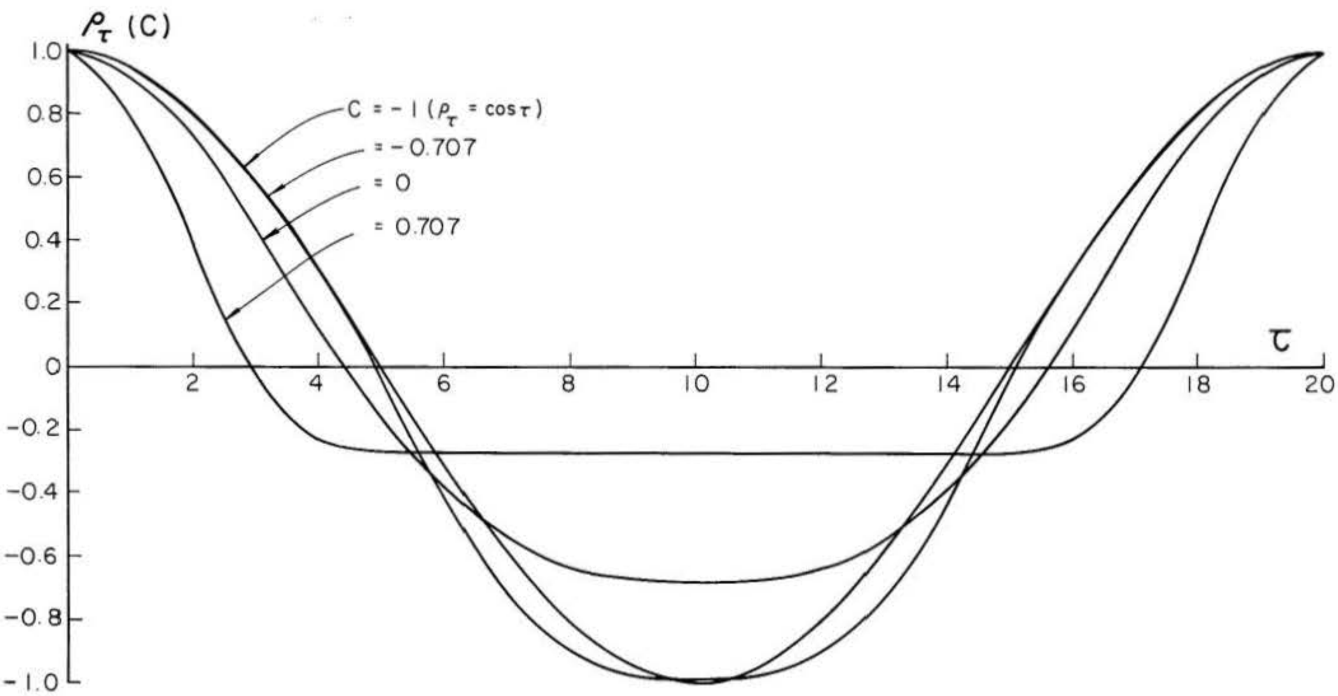


Fig. 6. Autocorrelation coefficients of a sinusoidal series $x_t = A \sin \omega t$, (with both the amplitude A and angular frequency ω being unities) truncated at four constant values: c = -1 ; c = - .707 ; c = 0 ; and c = 0.707

5. Truncation of composite series of a dependent normal random component and a deterministic sinusoidal component. For a series having a periodic and a dependent random component, the correlogram depends on the variance ratio of the two components. Considering the time average over a complete cycle, the autocorrelation coefficients of the series composed of a periodic and a normal dependent random component truncated at constant levels are computed from the values of Ey_t , Ey_t^2 and $Ey_t y_{t+\tau}$ obtained by eqs. (38), (39) and (41). The mean and variance of the random component are taken to be zero and unity, respectively. Then for different ratios of variances of the random component and the periodic component in Fig. 7, Appendix 1, are plotted for $\rho = 0$, 0.2 , 0.4 , 0.6 , and 0.8 with the truncation of series at different levels. The levels of truncation are expressed in terms of the total standard deviation of the process.

Figure 7 is presented in Appendix 1 by the following scheme of variance ratio and ρ (the first autocorrelation coefficient of untruncated dependent normal variable) given in Table 5.

In the case of a dependent normal random variable (first order Markov linear dependence) plus the sinusoidal component, when truncated by a constant c (expressed in ratio to σ of x_t), the first few autocorrelation coefficients show the similar pattern in $\rho_\tau(c)$ and decrease with an increase of c as with a pure dependent normal random variable (with no periodic component). However, for larger values of τ , the autocorrelation coefficients $\rho_\tau(c)$ begin to oscillate with the same period as the periodic component and with the amplitude which depends on the variance ratio and the truncation constant as shown in Figs. 7.1 through 7.15 in Appendix 1.

Table 5. Scheme of the arrangement of Figures in Appendix 1

First auto-correlation coefficient ρ	Variance Ratio (v.r.)		
	1.0	2.0	3.0
0.0	Fig. 7.1	7.6	7.11
0.2	Fig. 7.2	7.7	7.12
0.4	Fig. 7.3	7.8	7.13
0.6	Fig. 7.4	7.9	7.14
0.8	Fig. 7.5	7.10	7.15

This case of a normally distributed dependent random variable superposed on a periodic component may roughly approximate the case of monthly and daily river flows, which have a clear 12-month or 365-days cycle in the form of a sine-function, plus a normal random variable of the first order Markov linear dependence. However, this case departs from the reality in three ways: (1) it treats the population autocorrelation function (no noise in the correlogram); (2) the sine-cycle means that only the cyclicity exists in the mean monthly or mean daily flows, while it is known that a corresponding cycle exists in the standard deviations of monthly or daily flows; and (3) stochastic components of monthly or daily flows are rarely normally distributed once the periodic component (in both mean and standard deviation of flow) is removed. Regardless of these three limitations, Figs. 7.1 through 7.15 give insight into how the truncation works on a composite time series.

6. Truncation of composite series of a dependent lognormal component and a deterministic sinusoidal component. Similarly, for the series with a periodic deterministic and a lognormal random component, the autocorrelation coefficients are computed by the eqs. (42), (43) and (44) or (45). Figures 8.1 through 8.15, Appendix 2, give correlograms of truncated series composed of a sinusoidal series superposed on a dependent lognormal process according to the model of eq. (17) for $\rho = 0$, 0.2 , 0.4 , 0.6 and 0.8. The levels of truncation are expressed in terms of the total standard deviation of the series, and the variance ratios are 1.0, 2.0 and 3.0.

Figure 8 is sorted in Appendix 2 as Figs. 8.1 through 8.15 following the same scheme as shown for Figs. 7.1 through 7.15 in Table 5.

Similar results are shown in Fig. 8 as in Fig. 7, namely, the first few values of $\rho_\tau(c)$ decrease both with a decrease of ρ and an increase of c . For higher values of $\rho_\tau(c)$, the correlogram fluctuates with the same period as the periodic component and with the amplitude which depends on both the variance ratio and the truncation constant c

Because of lognormal distribution of the stochastic component, this analysis approximates better the case of monthly or daily river flows than the previous case of normal distribution. However, this lognormal case still departs significantly from the real time series of monthly or daily hydrological time series. Though it is only a rough approximation, it gives a clear picture of what occurs to correlograms of time series when a given constant water quantity is either diverted out of a river basin or is consumed in it without a return flow.

7. Two examples with hydrological time series of daily river flows. In order to show what happens to series of daily river flows when a large truncation is made, the truncation constant is taken to be the mean discharge. Water withdrawal from a river is a physical analogy of truncation and to illustrate its effect on the dependence properties the daily flows of the Batten Kill River at Battenville, New York, and the Madison River at West Yellowstone, Montana, have been taken as examples. The time series of both rivers are truncated at the mean flow. Both correlograms, with and without truncation, are shown in Fig. 9 for the Batten Kill River and in Fig. 10 for the Madison River.

The correlograms of the truncated series of daily flows of the Batten Kill River and the Madison River when truncated shows a reduction of the average absolute autocorrelation coefficient. Since both series of river flows are periodic and the stochastic component follows the first order Markov linear model [2], the correlograms of truncated series show the same patterns as the correlograms obtained by theoretical analysis and shown in Figs. 7 and 8.

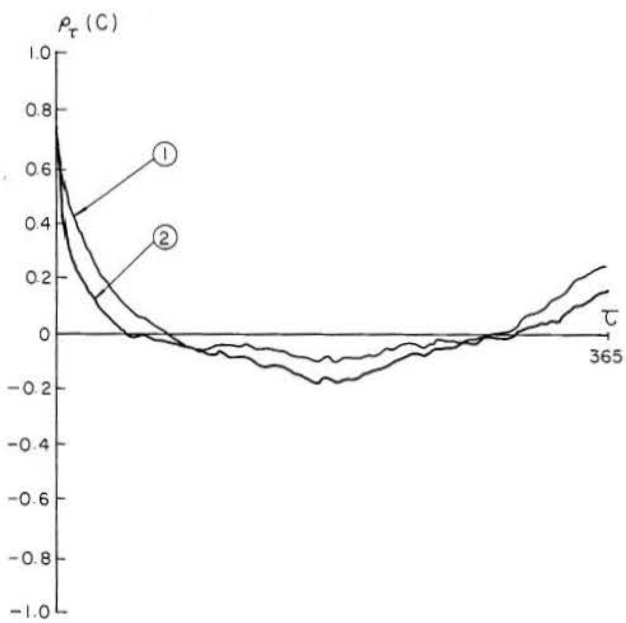


Fig. 9. Correlogram of truncated and untruncated series of daily flows of the Batten Kill River at Battenville, N. Y.: (1) correlogram of untruncated series; and (2) correlogram of truncated series; with $c = \overline{x}$

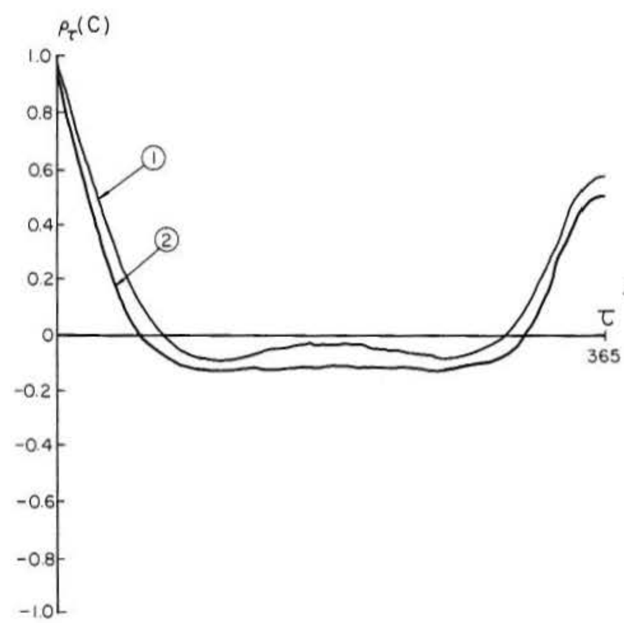


Fig. 10. Correlogram of truncated and untruncated series of daily flows of the Madison River at W. Yellowstone, Montana: (1) correlogram of untruncated series; and (2) correlogram of truncated series; with $c = \overline{x}$

Chapter IV

CONCLUSIONS

The following conclusions may be drawn from the previous analytical study of the effect of truncation on a dependent random process with or without a periodic component:

1. Truncation at any constant level of a dependent random process reduces the dependence in sequence of a time series. The magnitude of the reduction depends on the constant of truncation. The general pattern of correlograms for truncated series in the case of the first order Markov linear process remains similar as for the untruncated case. Therefore, when analyzing a physical process with zero values, a reasonable assumption is to consider it as a truncated process.

2. First autocorrelation coefficients of precipitation series in an arid region should be, in general and neglecting other factors which affect the dependence, less on the average than in a wet region, if the annual evaporation is assumed to be a constant. The mathematical dependence model of precipitation series for the two regions should however be the same. Because the evaporation in a dry region is a more complex process than in a humid region, the above effect may not prevail in the final result of all factors affecting the dependence.

3. The truncation of a periodic process at a constant level reduces the amplitude of the correlogram while the period remains the same. It can also be said that the periodic truncation of a purely stochastic process makes the correlogram periodic.

4. Water withdrawal from a river will decrease the dependence when this process dries up the river for some time, thus creating zero values in the hydrological time series.

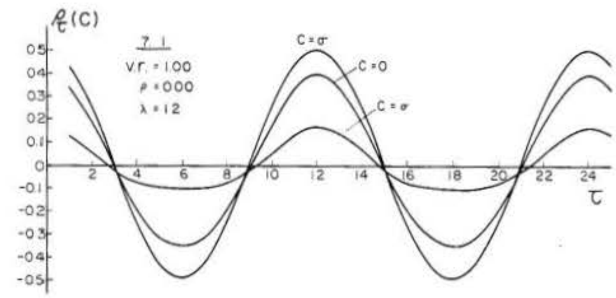
BIBLIOGRAPHY

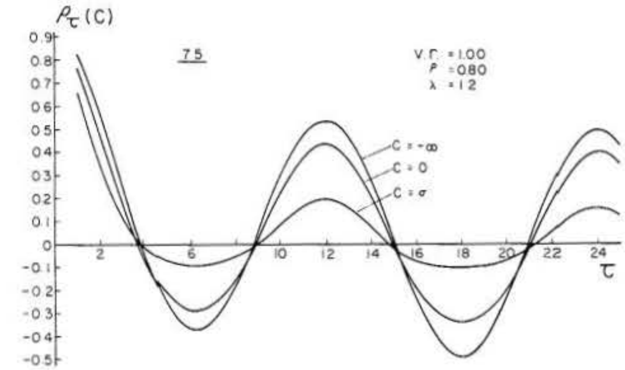
1. Kendall, M. G. and A. Stuart, The Advanced Theory of Statistics, Volume 1. Charles Griffin and Company Limited, London 1958, pp. 135, 155-156

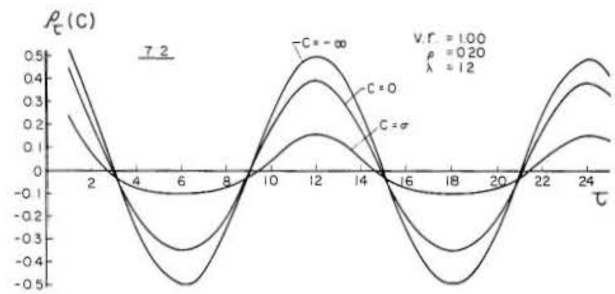
2. Quimpo, R. G., Stocastic Model of Daily River Flow Sequences. Colorado State University Hydrology Paper No. 18, February 1967.

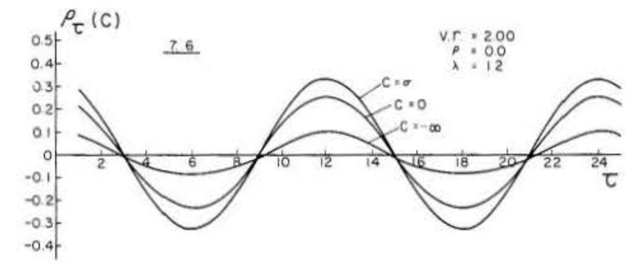
APPENDIX 1

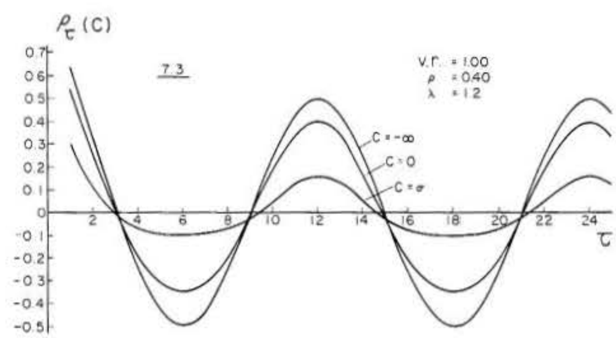
Figs. 7.1 through 7.15 Correlogram of $x_t = A \sin \omega t + \eta_t$, truncated at constant level for $\lambda = 12$ and different values of ρ and variance ratios, with η_t normally distributed.

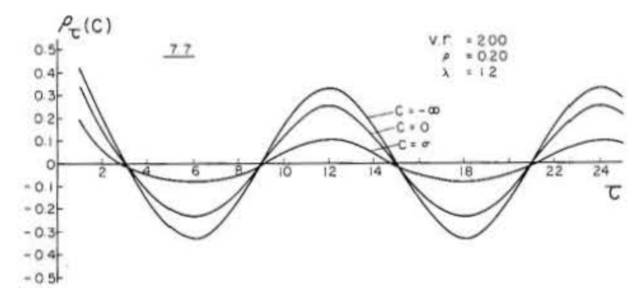

$\rho_\tau(C)$
7.1
V.r. = 1.00
ρ = 0.00
λ = 12
C = σ
C = 0
C = σ
τ

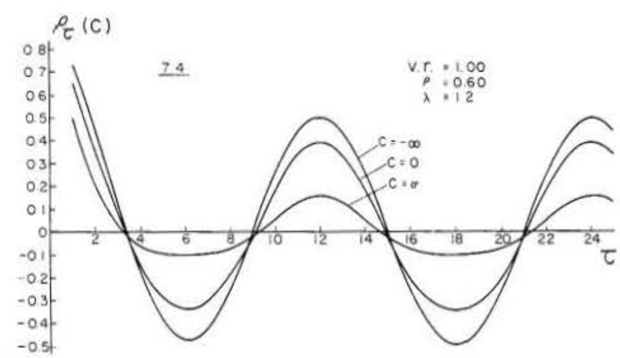

$\rho_\tau(C)$
7.5
V.r. = 1.00
ρ = 0.80
λ = 12
C = -∞
C = 0
C = σ
τ

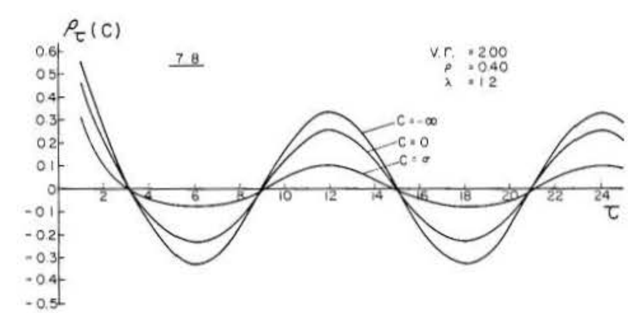

$\rho_\tau(C)$
7.2
V.r. = 1.00
ρ = 0.20
λ = 12
C = -∞
C = 0
C = σ
τ


$\rho_\tau(C)$
7.6
V.r. = 2.00
ρ = 0.0
λ = 12
C = σ
C = 0
C = -∞
τ


$\rho_\tau(C)$
7.3
V.r. = 1.00
ρ = 0.40
λ = 12
C = -∞
C = 0
C = σ
τ

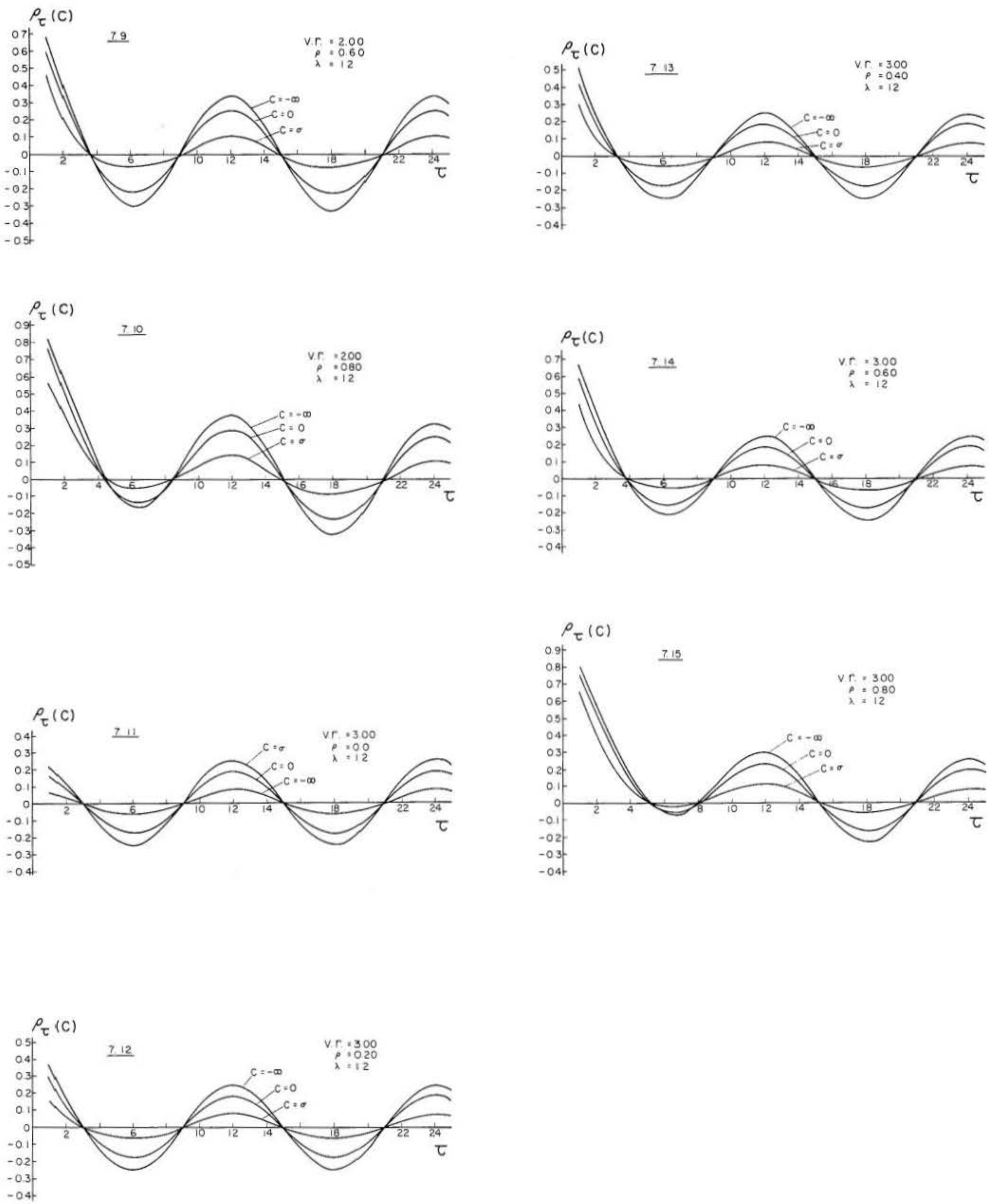

$\rho_\tau(C)$
7.7
V.r. = 2.00
ρ = 0.20
λ = 12
C = -∞
C = 0
C = σ
τ


$\rho_\tau(C)$
7.4
V.r. = 1.00
ρ = 0.60
λ = 12
C = -∞
C = 0
C = σ
τ


$\rho_\tau(C)$
7.8
V.r. = 2.00
ρ = 0.40
λ = 12
C = -∞
C = 0
C = σ
τ

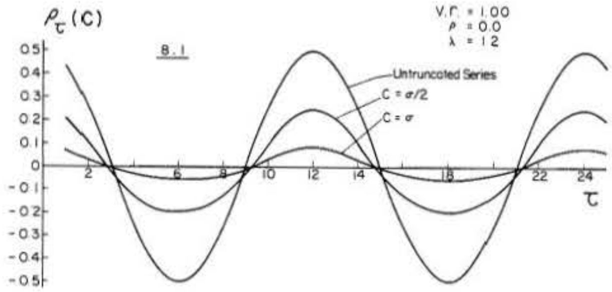
Continued

Figure 7- continued:

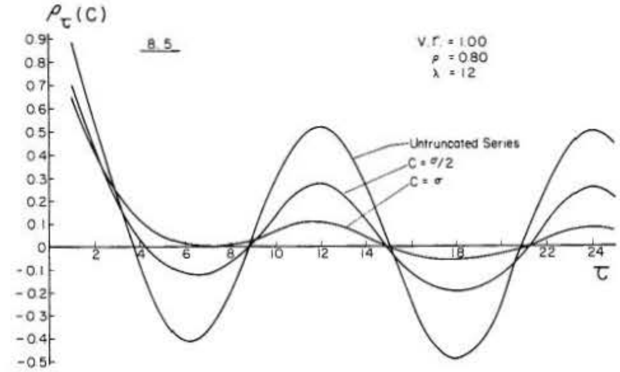


APPENDIX 2

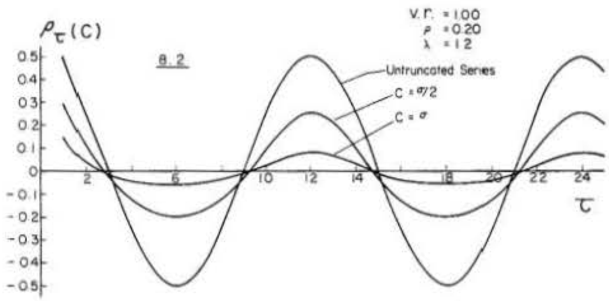
Figs. 8.1 through 8.15 Correlogram of $x_t = A \sin \omega t + \eta_t$, truncated at constant level for $\lambda = 12$ and different values of ρ and variance ratios, with η_t lognormally distributed



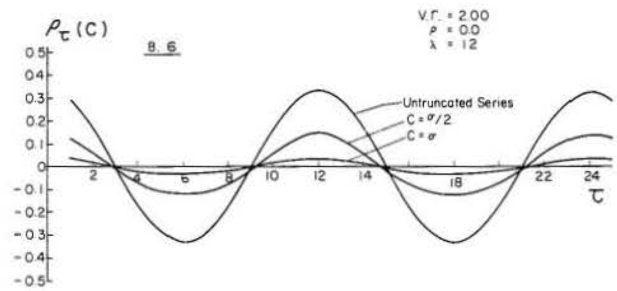
$\rho_\tau(C)$
8.1
V.r. = 1.00
ρ = 0.0
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



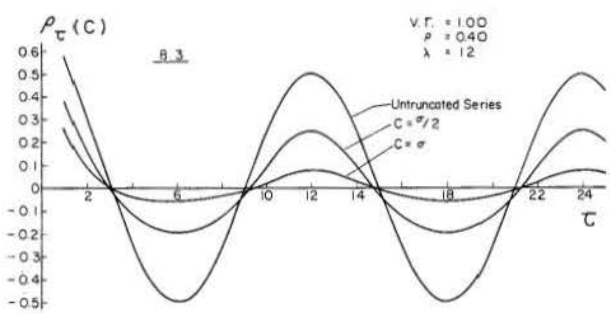
$\rho_\tau(C)$
8.5
V.r. = 1.00
ρ = 0.80
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



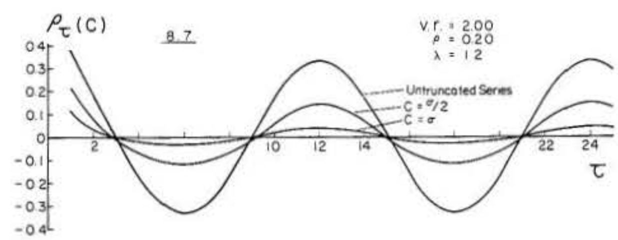
$\rho_\tau(C)$
8.2
V.r. = 1.00
ρ = 0.20
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



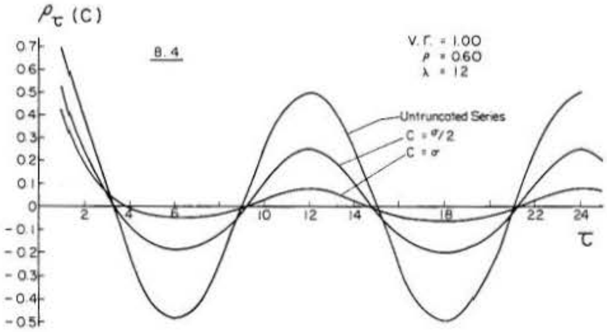
$\rho_\tau(C)$
8.6
V.r. = 2.00
ρ = 0.0
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



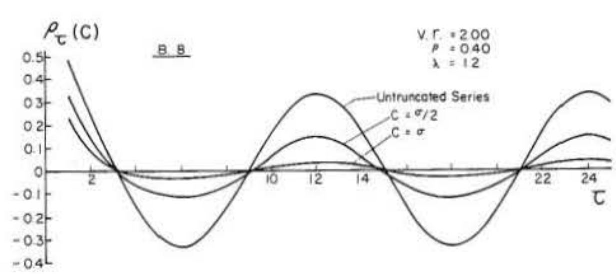
$\rho_\tau(C)$
8.3
V.r. = 1.00
ρ = 0.40
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



$\rho_\tau(C)$
8.7
V.r. = 2.00
ρ = 0.20
λ = 12
Untruncated Series
C = σ/2
C = σ
τ



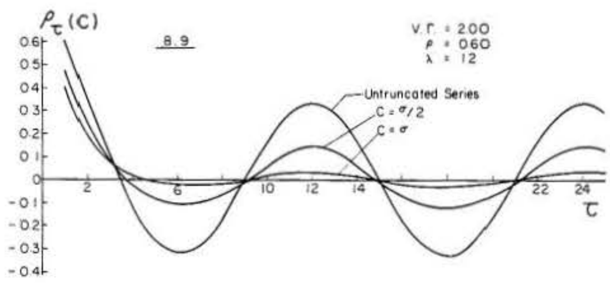
$\rho_\tau(C)$
8.4
V.r. = 1.00
ρ = 0.60
λ = 12
Untruncated Series
C = σ/2
C = σ
τ

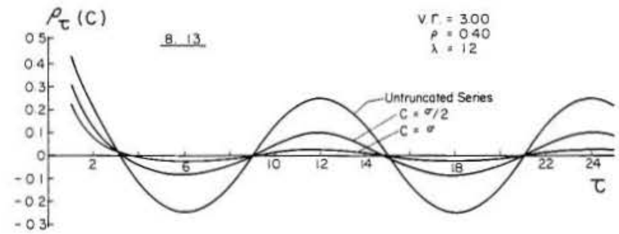


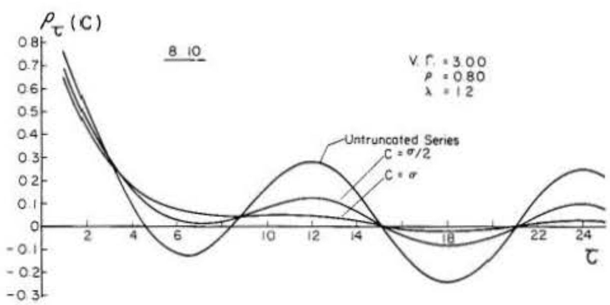
$\rho_\tau(C)$
8.8
V.r. = 2.00
ρ = 0.40
λ = 12
Untruncated Series
C = σ/2
C = σ
τ

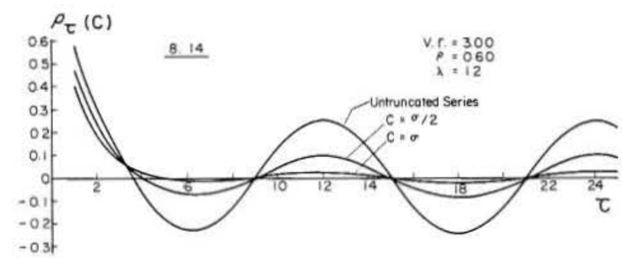
Continued

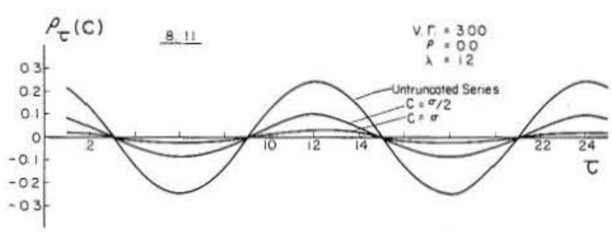
Figure 8 - Continued:

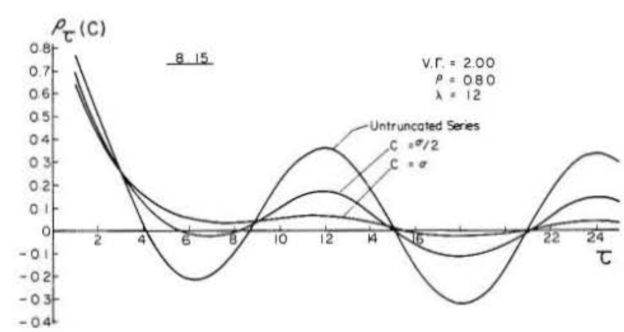


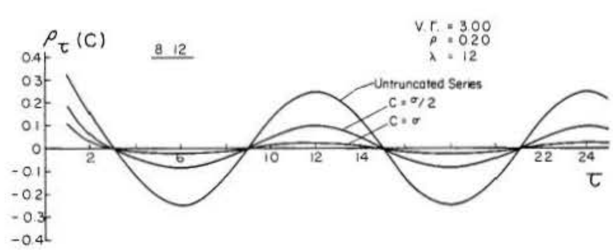












Key Words: Hydrological time series, Truncation of time series, Zero values, Normal process, Lognormal process, First order Markov linear chain

Abstract: The effects of the constant value truncation on the dependence property of idealized hydrological time series are investigated. The analysis was carried out for dependent stochastic components, for pure periodic component and for a sum of stochastic dependent and periodic deterministic components. The random elements of the process were considered for both normal and lognormal distribution. For all cases, the truncation is considered at a constant level. Analytical equations were developed in approximate forms for correlograms which express the dependence of the truncated time series. Computations of correlograms from these equations were made on a digital computer. Two series of daily river flows were used with the truncation at the mean flow to compare their correlograms with those of the original daily flow series.

Key Words: Hydrological time series, Truncation of time series, Zero values, Normal process, Lognormal process, First order Markov linear chain

Abstract: The effects of the constant value truncation on the dependence property of idealized hydrological time series are investigated. The analysis was carried out for dependent stochastic components, for pure periodic component and for a sum of stochastic dependent and periodic deterministic components. The random elements of the process were considered for both normal and lognormal distribution. For all cases, the truncation is considered at a constant level. Analytical equations were developed in approximate forms for correlograms which express the dependence of the truncated time series. Computations of correlograms from these equations were made on a digital computer. Two series of daily river flows were used with the truncation at the mean flow to compare their correlograms with those of the original daily flow series.

Key Words: Hydrological time series, Truncation of time series, Zero values, Normal process, Lognormal process, First order Markov linear chain

Abstract: The effects of the constant value truncation on the dependence property of idealized hydrological time series are investigated. The analysis was carried out for dependent stochastic components, for pure periodic component and for a sum of stochastic dependent and periodic deterministic components. The random elements of the process were considered for both normal and lognormal distribution. For all cases, the truncation is considered at a constant level. Analytical equations were developed in approximate forms for correlograms which express the dependence of the truncated time series. Computations of correlograms from these equations were made on a digital computer. Two series of daily river flows were used with the truncation at the mean flow to compare their correlograms with those of the original daily flow series.

Key Words: Hydrological time series, Truncation of time series, Zero values, Normal process, Lognormal process, First order Markov linear chain

Abstract: The effects of the constant value truncation on the dependence property of idealized hydrological time series are investigated. The analysis was carried out for dependent stochastic components, for pure periodic component and for a sum of stochastic dependent and periodic deterministic components. The random elements of the process were considered for both normal and lognormal distribution. For all cases, the truncation is considered at a constant level. Analytical equations were developed in approximate forms for correlograms which express the dependence of the truncated time series. Computations of correlograms from these equations were made on a digital computer. Two series of daily river flows were used with the truncation at the mean flow to compare their correlograms with those of the original daily flow series.

PREVIOUSLY PUBLISHED PAPERS

Colorado State University Hydrology Papers

No. 23 "An Objective Approach to Definitions and Investigations of Continental Hydrologic Droughts," by Vujica Yevjevich, August 1967

No. 24 "Application of Cross-Spectral Analysis to Hydrologic Time Series," by Ignacio Rodriguez-Iturbe, September 1967.

No. 25 "An Experimental Rainfall-Runoff Facility," by W. T. Dickinson, M. E. Holland and G. L. Smith, September 1967.

No. 26 "The Investigation of Relationship Between Hydrologic Time Series and Sun Spot Numbers," by Ignacio Rodriguez-Iturbe, and Vujica Yevjevich, April 1968.

No. 27 "Diffusion of Entrapped Gas from Porous Media," by Kenneth M. Adam, and Arthur T. Corey, April 1968.

No. 28 "Sampling Bacteria in a Mountain Stream," by Samuel H. Kunkle and James R. Meiman, March 1968.

No. 29 "Estimating Design Floods from Extreme Rainfall," by Frederick C. Bell, July 1968.

No. 30 "Conservation of Ground Water by Gravel Mulches," by A. T. Corey and W. D. Kemper, May 1968.

Colorado State University Fluid Mechanics Papers

No. 4 "Experiment on Wind Generated Waves on the Water Surface of a Laboratory Channel," by E. J. Plate and C. S. Yang, February 1966.

No. 5 "Investigations of the Thermally Stratified Boundary Layer," by E. J. Plate and C. W. Lin, February 1966.

No. 6 "Atmospheric Diffusion in the Earth's Boundary Layer--Diffusion in the Vertical Direction and Effects of the Thermal Stratification," by Shozo Ito, February 1966.

Colorado State University Hydraulics Papers

No. 1 "Design of Conveyance Channels in Alluvial Materials," by D. B. Simons, March 1966.

No. 2 "Diffusion of Slot Jets with Finite Orifice Length-Width Ratios," by V. Yevjevich, March 1966.

No. 3 "Dispersion of Mass in Open-Channel Flow," by William W. Sayre, February 1968.