THE APPLICATION OF CROSS-SPECTRAL ANALYSIS TO HYDROLOGIC TIME SERIES

By

Ignacio Rodríguez - Iturbe

September 1967

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

24

THE APPLICATION OF CROSS-SPECTRAL ANALYSIS TO HYDROLOGIC TIME SERIES

by

Ignacio Rodríguez - Iturbe

HYDROLOGY PAPERS

COLORADO STATE UNIVERSITY

FORT COLLINS, COLORADO

September 1967

No. 24

ACKNOWLEDGMENTS

This paper is primarily based on research performed by the author during studies toward a Doctor of Philosophy degree at Colorado State University. The author wishes to extend his sincere appreciation to his major professor and adviser, Dr. Vujica Yevjevich, Professor of Civil Engineering, for his counsel and encouragement. Special thanks are given to Dr. Mohammed M. Siddiqui, Professor in the Department of Mathematics and Statistics, whose guidance has been invaluable in the author's work.

The writer gratefully acknowledges the U. S. National Science Foundation (grant GK-1661) under whose sponsorship this research is supported with V. Yevjevich as the principal investigator and from which the writer's graduate research assistantship was awarded. The National Center for Atmospheric Research, Boulder, Colorado, has also substantially helped the study be allowing free computer time to the project.

Financial assistance for the author's academic training was also received from La Universidad del Zulia, Maracaibo, Venezuela, which kindly allowed the author a leave of absence in order to finish his graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES - Continued

LIST OF TABLES

ABSTRACT

The main objective of this paper is to study the potentials of cross-spectrum and multiple cross-spectrum for the analysis of hydrological data.

Groups of precipitation and runoff stations were selected in different climatic environment and complete cross-spectral analyses were performed between those stations. The coherence and partial coherence functions were used for the study of frequency correlations between the series and they show that there exists a very strong correlation between the annual cycles of the stations. Along the Pacific Coast of the United States the annual cycle in precipitation appears to be basically the same up to distances of 1000 Km.

Cyclic regression analysis with the use of the gain and phase functions is shown to work correctly in hydrologic time series. This type of regression may be very useful in regions where frequency components account for a large percentage of the variance of the series.

Cross-spectral characteristics of the moving average and autoregressive processes are shown to be a powerful tool in testing and analyzing these types of generating processes in hydrology. Special significance has the coherence between two 1st order autoregressive processes which is shown to be equal to a constant independent of frequency.

The effects of smoothing or pre-filtering in the cross-spectral properties of two series are studied and recommendations made when working with this practice which is frequently used in hydrology.

THE APPLICATION OF CROSS-SPECTRAL ANALYSIS TO HYDROLOGIC TIME SERIES

by Ignacio Rodríguez - Iturbe*

CHAPTER I

INTRODUCTION

1. Significance of the Study. Many studies have been performed with regard to the spectral characteristics of hydrologic data, but problems involving the simultaneous behavior of two or more series have not been worked on in a wide variety of fields of application, although enough has been done to point the way and suggest the possibilities of hydrologic spectral analyses.

There is an increasing number of problems in the geophysical sciences, which can be approached and solved by multiple regression analysis. They can also be effectively studied by multiple spectral techniques which are precise analogs of multiple regression in spirit and, if care is taken in choice, in the algebraic form of their basic equations (Tukey, 1961). The differences which arise in the development stem from:

(a) the fact that regression goes on separately at each frequency, and

(b) the fact that regression coefficients take complex values rather than real values, which enables one to learn more about the underlying relationship.

In studying time series, as in its more classical situations, regression analysis is a more sensitive and powerful form of analysis than variance component analysis whenever there is a suitable regression variable. As a consequence, one major reason for learning about spectrum analysis is a foundation for learning about cross-spectrum analysis (Tukey, 1961).

2. Main Objectives of This Study. The main objective of this study is to look into the possibilities of the still young techniques of cross-spectrum and multiple cross-spectrum for the analysis of hydrologic data.

Two general approaches may be taken when analyzing interrelations between hydrologic data. The first approach is to represent a given process by a multidimensional model and then study the characteristics of this model, specifically the covariance and spectral matrices. The second approach is to consider each series as a realization of the process and study the interrelations between these realizations. Both approaches are considered here.

Five specific aspects are specially stressed:

(1) The use of the coherence and the partial coherence (defined in Chapter II-3) as correlation measures between the different frequency components of hydrologic time series. Also, how these measures compare with the classical time-domain methods.

(2) Frequency regression analysis of hydrologic time series by using the gain function (defined in Chapter III-1) as a regression coefficient at each frequency.

(3) Potentials in using cross-spectral analysis to study input-output relationships in hydrologic systems.

(4) Cross-spectral characteristics of generating processes in common usage in hydrology.

(5) Effects of smoothing of time series on the coherence and phase functions between two series.

* Assistant Professor of Civil Engineering, Universidad del Zulia, Maracaibo, Venezuela.

CHAPTER II

GENERAL MATHEMATICAL TECHNIQUES FOR CROSS-SPECTRAL ANALYSIS

1. Stationary Random Processes. A random process $[x_k(t)]$, $-\infty < t < \infty$, is an ensemble of real valued or complex valued functions which can be characterized through its probability structure. The variable t can represent any characteristic of a process, although for convenience it will be interpreted as time in the following discussion. Each particular function $x_k(t)$, where t is a variable and k is fixed, is called a sample function. In practice a sample function may be thought of as the observed result of a single experiment.

A particular sample function $x_k(t)$ is, in general, not suitable for representing the entire random process. It is one of the main goals of statistics to estimate the properties of the entire process on the basis of particular sample functions.

Consider two arbitrary random processes $[x_k(t)]$ and $[y_k(t)]$ with mean values

$$\mu_x(t) = E\left[x_k(t)\right] \tag{2-1}$$

$$\mu_y(t) = E\left[y_k(t)\right] \tag{2-2}$$

Their autocovariance functions are defined at arbitrary fixed values of t and $t-\tau$, by

$$\alpha_x(t,\, t-\tau) = E\left[\left(x_k(t) - \mu_x(t)\right)\left(x_k(t-\tau) - \mu_x(t-\tau)\right)\right] \tag{2-3}$$

$$\alpha_y(t,\, t-\tau) = E\left[\left(y_k(t) - \mu_y(t)\right)\left(y_k(t-\tau) - \mu_y(t-\tau)\right)\right] \tag{2-4}$$

Similarly, the cross-covariance function is defined by

$$\alpha_{xy}(t, t-\tau) = E\left[\left(x_k(t)-\mu_x(t)\right)\left(y_k(t-\tau) - \mu_y(t-\tau)\right)\right] \tag{2-5}$$

In general, all the preceding quantities vary for different values of t and τ

Other statistical quantities can be defined over the ensemble by fixing three or more times. The probability structure is thus described in finer and finer detail by increasing the number of fixed times. If all possible probability distributions involving $x_k(t)$ are independent of the absolute times t_1 , t_2 , ---- , t_n , ---- and are only function of the intervals τ_1 , τ_2 , --- , τ_n , --- , then the process is said to be strongly stationary. If only the first "n" probability distributions are independent of the absolute times, the process is called n^{th}-order stationary. In order to prove n^{th}-order stationarity it is only necessary to prove that the n^{th} probability density is independent of absolute times because the first (n-1) probability densities are obtained from the n^{th} density by successive integrations.

In the special case of a Gaussian independent process, the mean value and the covariance function provide a complete description of the underlying probability structure. In this case, second order stationarity or weak stationarity is equivalent to strong stationarity because the former implies the mean and covariance function are independent of absolute times and this inturn implies all the possible probability distributions are independent of absolute times because all of them may be derived from the mean value and the covariance function.

2. Spectral Density Functions. It will now be assumed that for the two stationary processes $[x_k(t)]$ and $[y_k(t)]$, the functions $\alpha_x(\tau)$, $\alpha_y(\tau)$ and $\alpha_{xy}(\tau)$ exist and have Fourier transforms $S_x(f)$, $S_y(f)$ and $S_{xy}(f)$ given by:

$$S_x(f) = \int_{-\infty}^{\infty} \alpha_x(\tau)\, e^{-2\pi f \tau i}\, d\tau \tag{2-6}$$

$$S_y(f) = \int_{-\infty}^{\infty} \alpha_y(\tau)\, e^{-2\pi f \tau i}\, d\tau \tag{2-7}$$

$$S_{xy}(f) = \int_{-\infty}^{\infty} \alpha_{xy}(\tau)\, e^{-2\pi f \tau i}\, d\tau \tag{2-8}$$

$S_x(f)$ and $S_y(f)$ are defined as the power spectra of the stochastic processes $[x_k(t)]$ and $[y_k(t)]$. $S_{xy}(f)$ is defined as the cross-spectrum function between these processes.

It is convenient to define the so-called physically realizable one-sided power spectra and cross-spectrum functions. These functions given by

$$G_x(f) = 2\, S_x(f),\ 0 \le f < \infty,\ \text{otherwise zero} \tag{2-9}$$

$$G_y(f) = 2\, S_y(f),\ 0 \le f < \infty,\ \text{otherwise zero} \tag{2-10}$$

$$G_{xy}(f) = 2\, S_{xy}(f),\ 0 \le f < \infty,\ \text{otherwise zero} \tag{2-11}$$

are the quantities measured by direct procedures in practice.

In the case of real valued process all the previous equations may be simplified. The real valued two-sided power spectrum is obtained from equation 2-6 by making the imaginary part equal to zero:

$$S_x(f) = \int_{-\infty}^{\infty} \alpha_x(\tau) \cos 2\pi f\tau \, d\tau \qquad (2\text{-}12)$$

Due to the fact that the covariance is an even function,

$$S_x(f) = 2\int_{0}^{\infty} \alpha_x(\tau) \cos 2\pi f\tau \, d\tau \qquad (2\text{-}13)$$

and

$$G_x(f) = 4 \int_{0}^{\infty} \alpha_x(\tau) \cos 2\pi f\tau \, d\tau \qquad (2\text{-}14)$$

for $0 \leq f < \infty$, otherwise zero.

The physically realizable one-sided cross-spectrum function can also be expressed as

$$G_{xy}(f) = 2 \int_{0}^{\infty} \alpha_{xy}(\tau) \, e^{-2\pi f\tau i} \, d\tau \qquad (2\text{-}15)$$

and being a complex number, it can be written as:

$$G_{xy}(f) = C_{xy}(f) - i\, Q_{xy}(f) \qquad (2\text{-}16)$$

where $C_{xy}(f)$ and $Q_{xy}(f)$ are the co-spectrum and quadrature spectrum, respectively. Following Bendat and Piersol (1966) the co-spectrum can be thought of as the average product of $x(t)$ and $y(t)$ within a narrow frequency interval, between f and $f + \Delta f$, divided by the frequency interval, Δf. The quadrature spectrum is the same except that either $x(t)$ or $y(t)$, not both, is shifted in time sufficiently to produce a 90-degree phase shift at the frequency f. In this manner, $C_{xy}(f)$ is a measure of the in-phase-covariance, and $Q_{xy}(f)$ is a measure of the out-of-phase covariance. In more practical words, the co-spectrum measures the contribution of oscillations of different frequencies to the total cross-covariance at the lag zero between two time series. The quadrature spectrum measures the contribution of the different harmonics to the total cross-covariance between the series when all the harmonics of the series $x(t)$ are delayed by a quarter period but the series $y(t)$ remains unchanged (Panofsky and Brier, 1958).

Equation 2-16 may be inverted to give the cross covariance function

$$\alpha_{xy}(\tau) = \int_{0}^{\infty} \left[C_{xy}(f) \cos 2\pi f\tau + Q_{xy}(f) \sin 2\pi f\tau \right] df \qquad (2\text{-}17)$$

Because $\alpha_{xy}(\tau)$ satisfies the relation

$$\alpha_{xy}(-\tau) = \alpha_{yx}(\tau)$$

the co-spectrum may be expressed as:

$$C_{xy}(f) = \int_{0}^{\infty} \left[\alpha_{xy}(\tau) + \alpha_{yx}(\tau) \right] \cos 2\pi f\tau \, d\tau = C_{xy}(-f) \qquad (2\text{-}18)$$

meaning that $C_{xy}(f)$ displays symmetry about the ordinate.

Similarly the quadrature spectrum may be expressed as:

$$Q_{xy}(f) = \int_{0}^{\infty} \left[\alpha_{xy}(\tau) - \alpha_{yx}(\tau) \right] \sin 2\pi f\tau \, d\tau = - Q_{xy}(-f) \qquad (2\text{-}19)$$

meaning that $Q_{xy}(f)$ is an odd function. From equations 2-18 and 2-19 and from the definitions

$$G_{yx}(f) = C_{yx}(f) - iQ_{yx}(f)$$

$$G_{xy}(f) = C_{xy}(f) - iQ_{xy}(f)$$

one obtains,

$$C_{xy}(f) = \frac{1}{2} \left[G_{xy}(f) + G_{yx}(f) \right] \qquad (2\text{-}20)$$

$$Q_{xy}(f) = \frac{1}{2i} \left[G_{xy}(f) - G_{yx}(f) \right] \qquad (2\text{-}21)$$

An alternative way to describe $G_{xy}(f)$ is by the complex polar form

$$G_{xy}(f) = \left| G_{xy}(f) \right| \, e^{-i\,\theta_{xy}(f)} \qquad 0 \leq f < \infty$$

where

$$\left| G_{xy}(f) \right| = \sqrt{C_{xy}^2(f) + Q_{xy}^2(f)} \qquad (2\text{-}22)$$

and

$$\theta_{xy}(f) = \mathrm{Tan}^{-1} \left[\frac{Q_{xy}(f)}{C_{xy}(f)} \right] \qquad (2\text{-}23)$$

which is called the phase function.

The physical meaning of $G_{xy}(f)$ and $\theta_{xy}(f)$ and the role they play in a linear system will be explained in Chapter III-2. By interchanging $x(t)$ and $y(t)$ one finds that $C_{yx}(f) = C_{xy}(f)$ and $Q_{yx}(f) = -Q_{xy}(f)$. Therefore, one can write:

$$G_{yx}(f) = G^*_{xy}(f) \qquad (2\text{-}24)$$

and

$$G_{yx}(f) = G_{xy}(-f) \qquad (2\text{-}25)$$

where $G^*_{xy}(f)$ denotes the complex conjugate of $G_{xy}(f)$.

3. Coherence Functions. The coherence function is a real valued quantity $\gamma^2_{xy}(f)$ defined as

$$\gamma^2_{xy}(f) = \frac{\left| G_{xy}(f) \right|^2}{G_x(f)\, G_y(f)} = \frac{\left| S_{xy}(f) \right|^2}{S_x(f)\, S_y(f)} \qquad (2\text{-}26)$$

For a better understanding of the coherence function it is useful to make an analogy with the classical results of correlation and regression analysis.

In statistical analysis of real variables, the correlation coefficient between two variables x and y with mean values of zero is defined as

$$\rho_{xy} = \frac{E[xy]}{(E[x^2]\ E[y^2])^{1/2}} = \frac{\text{cov}[x,y]}{\sigma_x \sigma_y} = \frac{\alpha_{xy}}{\sigma_x \sigma_y} \tag{2-27}$$

where σ_x^2 and σ_y^2 represent the variances of x and y, respectively.

Similarly, if complex numbers X and Y are being considered the square of the correlation coefficient becomes:

$$\rho_{XY}^2 = \frac{E[XY^*]\ E[XY^*]^*}{E[XX^*]\ E[YY^*]} \tag{2-28}$$

where the (*) symbol represents the complex conjugate of the term in question.

Rewriting equation 2-28 one gets:

$$\rho_{XY}^2 = \frac{E[XY^*]\ E[X^*Y]}{E[XX^*]\ E[YY^*]} = \frac{|E[XY^*]|^2}{\sigma_X^2 \sigma_Y^2} = \frac{|\alpha_{XY}|^2}{\sigma_X^2 \sigma_Y^2} \tag{2-29}$$

From equation 2-29 it is seen that the coherence function may be thought of as a correlation coefficient squared if we replace α_{XY} with $S_{xy}(f)$, σ_X^2 with $S_x(f)$ and σ_Y^2 with $S_y(f)$. We will proceed to show the meaning of these changes.

Cramer's representation of a stationary process with zero mean gives:

$$x(t) = \int_{-\infty}^{\infty} e^{2\pi itf}\, dz_x(f) \tag{2-30}$$

where we have written $dz_x(f)$ for $z_x(df)$ and $z_x(f)$ is an orthogonal set function with

$$E\left|dz_x(f)\right|^2 = dS_x(f) = S_x(f)\,df \tag{2-31}$$

$S_x(f)$ being the power spectrum of the process $\{x(t)\}$. Similarly we have:

$$E\left|dz_y(f)\right|^2 = dS_y(f) = S_y(f)\,df \tag{2-32}$$

and

$$E\left[dz_x(f) \cdot dz_y^*(f)\right] = dS_{xy}(f) = S_{xy}(f)\,df \tag{2-33}$$

Comparing equations 2-31, 2-32 and 2-33 with the expressions

$$\sigma_X^2 = E\left[XX^*\right] = E\left|X\right|^2 \tag{2-34}$$

$$\sigma_Y^2 = E\left[YY^*\right] = E\left|Y\right|^2 \tag{2-35}$$

$$\alpha_{XY} = E\left[XY^*\right] \tag{2-36}$$

it may be seen that the coherence function can be interpreted as a correlation coefficient squared between the spectral variables $z_x(f)$ and $z_y(f)$ calculated at each frequency f.

It should now be clear that it is often advantageous to study correlation problems in the frequency domain rather than in the time domain. Working in the frequency domain, any stationary series can be considered as a sum of components or frequency bands, each component being statistically independent of the others. One of the important things that the theory of stationary processes tells us is that not only is the component with center f_j independent of all the other components of the process, but it is also independent of all components of another process except for the component centered on f_j. In this manner when the coherence between two time series is calculated one looks for correlations among them in a very small range of frequencies. On the other hand, with the cross covariance function one is looking for correlations between the two processes considering each one as a whole.

4. Partial Coherence Functions. Consider two real-valued stationary processes $[x(t)]$ and $[y(t)]$ and assume that the mean values are zero in order to simplify the notation. The residual random variable $\Delta y(t)$ of $y(t)$ from $x(t)$ is defined by:

$$\Delta y(t) = y(t) - \hat{y}(t) \tag{2-37}$$

where $\hat{y}(t)$ is the least squares prediction of $y(t)$ from $x(t)$,

$$\hat{y}(t) = \frac{\alpha_{xy}}{\alpha_{xx}}\, x(t) \tag{2-38}$$

Consider now three real-valued stationary random processes $[x_1(t)]$, $[x_2(t)]$ and $[y(t)]$ where the mean values are assumed to be zero. One can define the partial correlation coefficient $\rho_{1y\cdot 2}$ by

$$\rho_{1y\,2}^2 = \frac{\alpha^2_{\Delta x_1\, \Delta y}}{\alpha_{\Delta x_1\, \Delta x_1}\, \alpha_{\Delta y \Delta y}} = \frac{\alpha^2_{1y\cdot 2}}{\alpha_{11\cdot 2}\, \alpha_{yy\cdot 2}} \tag{2-39}$$

where

$$\alpha_{\Delta x_1\, \Delta x_1} = \alpha_{1y\cdot 2} = \alpha_{11}(1 - \rho_{21}^2) \tag{2-40}$$

$$\alpha_{\Delta y\, \Delta y} = \alpha_{yy\cdot 2} = \alpha_{yy}(1 - \rho_{2y}^2) \tag{2-41}$$

$$\alpha_{\Delta x_1 y} = \alpha_{1y.2} = \alpha_{1y}\left(1 - \frac{\alpha_{12}\alpha_{2y}}{\alpha_{22}\alpha_{1y}}\right) \quad (2\text{-}42)$$

Similar to the partial correlation coefficient in the time domain, it is possible to define in the frequency domain a partial coherence function between $x_1(t)$ and $y(t)$ with $x_2(t)$ removed at every t by least squares prediction from $x_1(t)$ and $y(t)$:

$$\gamma^2_{1y\cdot 2}(f) = \frac{\left|S_{1y\cdot 2}(f)\right|^2}{S_{11\cdot 2}(f)\, S_{yy\cdot 2}(f)} = \frac{\left|G_{1y\cdot 2}(f)\right|^2}{G_{11\cdot 2}(f)\, G_{yy\cdot 2}(f)} \quad (2\text{-}43)$$

The terms in equation 2-43 are called residual or partial spectra and are defined by:

$$S_{1y\cdot 2}(f) = S_{1y}(f)\left[1 - \frac{S_{12}(f)\, S_{2y}(f)}{S_{22}(f)\, S_{1y}(f)}\right] \quad (2\text{-}44)$$

$$S_{11\cdot 2}(f) = S_{11}(f)\,(1 - \gamma^2_{12}(f)) \quad (2\text{-}45)$$

$$S_{yy\cdot 2}(f) = S_{yy}(f)\,(1 - \gamma^2_{2y}(f)) \quad (2\text{-}46)$$

The proof that the partial coherence is nothing else but an analog of the partial correlation coefficient between the spectral variables, calculated at each frequency f, can be carried out by following the same procedure used for the normal coherence.

The case of multiple processes is only a generalization of the three variable case explained before. The partial coherence function bewteen $x_1(t)$ and $y(t)$ with $x_2(t)$, $x_3(t)$, ---, $x_n(t)$ removed at every t by least squares prediction from $x_1(t)$ and $y(t)$, is defined by

$$\gamma^2_{1y\cdot 23.....n}(f) = \frac{\left|S_{1y\cdot 23.....n}(f)\right|^2}{S_{11\cdot 23....n}(f)\cdot S_{yy\cdot 23....n}(f)} \quad (2\text{-}47)$$

The definition and calculation of the partial spectra of formula 2-47 has been done in matrix form by Goodman (1965) in a very suitable form for the use of high speed digital computers. Their meaning is essentially the same as those of formula 2-43.

Similarly to the development made for the partial coherence function it is possible to define the partial phase function between $x_1(t)$ and $y(t)$ with $x_2(t)$, $x_3(t)$, ----, $x_n(t)$ removed at every t by lease squares prediction from $x_1(t)$ and $y(t)$,

$$\theta_{1y\cdot 23.....n} = \mathrm{Tan}^{-1}\frac{\text{Imag. part of } S_{1y\cdot 23....n}}{\text{Real part of } S_{1y\cdot 23.....n}} \quad (2\text{-}48)$$

5. Application of Partial Coherence Functions. When more than two variables are being considered, the partial coherence function, rather than the ordinary coherence, gives a quantitative indication of the degree of linear dependence between the variables. An example of erroneous high coherence is shown in Figure 2.1.

Assume that a coherence function value near unity is computed between the variables $x_1(t)$ and $y(t)$. One would be inclined to believe that there is a linear system relating these two variables.

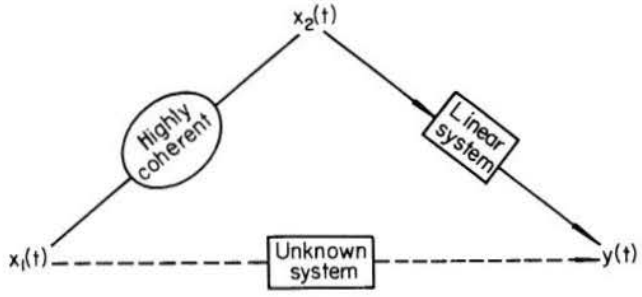

Figure 2.1 Example of erroneous high coherence (Bendat and Piersol, 1966).

Suppose there is a third variable $x_2(t)$ which is highly coherent with $x_1(t)$ and also passes through a linear system to make up $y(t)$. In this type of situation, the high coherence computed between $x_1(t)$ and $y(t)$ might only be due to the fact that $x_2(t)$ is highly coherent with $x_1(t)$. If this is in fact the situation, the partial coherence between $x_1(t)$ and $y(t)$ will be very low.

On the other hand, the opposite situation can exist. If two uncorrelated inputs $x_1(t)$ and $x_2(t)$ pass through existing linear systems to make up the output $y(t)$, the coherence functions $\gamma^2_{1y}(f)$ and $\gamma^2_{2y}(f)$ will appear less than unity since there will exist a contribution due to the other input which will appear as noise. If the partial coherences are computed, the effects of the other input will be subtracted out and the true coherence will be obtained.

6. Procedure of Computation. All the computations were carried out in a 6600 CDC digital computer.

The first step in computing the spectrum is the calculation of the autocovariance function of the series according to the formula,

$$\hat{\alpha}_{xx}(i) = \frac{1}{N-i}\left[\sum_{t=1}^{N-i} x_t x_{t+i} - \frac{1}{N-i}\left(\sum_{t=i+1}^{N} x_t\right)\left(\sum_{t=1}^{N-i} x_t\right)\right] \quad (2\text{-}49)$$

for $i = 0, 1, 2, \ldots, m$, where m is the maximum number of lags and N the number of observations.

Next, the finite cosine series transform function of the autocovariances is calculated according to the formula (Blackman and Tukey, 1958),

$$\tilde{G}_x(i) = \sum_{j=0}^{m} \hat{\alpha}'_{xx}(j)\cos\frac{ij\pi}{m} \quad (2\text{-}50)$$

for $i = 0, 1, 2, \ldots, m$. Where,

$$\left.\begin{aligned} \hat{\alpha}'_{xx}(0) &= \hat{\alpha}_{xx}(0) \\ \hat{\alpha}'_{xx}(i) &= 2\hat{\alpha}_{xx}(i) \quad \text{for } 1 \le i \le m-1 \end{aligned}\right\} \quad (2\text{-}51)$$

and

$$\hat{\alpha}'_{xx}(m) = \hat{\alpha}_{xx}(m)$$

The spectrum is calculated according to the spectral window formulas (Blackman and Tukey, 1958):

$$\left.\begin{aligned} \hat{G}_x(0) &= 0.5\,\tilde{G}_x(0) + 0.5\,\tilde{G}_x(1) \\ \hat{G}_x(i) &= 0.25\,\tilde{G}_x(i-1) + 0.5\,\tilde{G}_x(i) \\ &\quad + 0.25\,\tilde{G}_x(i+1) \\ &\quad \text{for } i = 1, 2, 3, \ldots, m-1 \end{aligned}\right\} \quad (2\text{-}52)$$

and

$$\hat{G}_x(m) = 0.5\,\tilde{G}_x(m-1) + 0.5\,\tilde{G}_x(m)$$

In computing the cross spectra, the first step is the calculation of the cross-covariance functions of series x and y :

$$\hat{\alpha}_{xy}(-i) = \frac{1}{N-i}\left[\sum_{t=1}^{N-i} x_t y_{t+i} - \frac{1}{N-i}\left(\sum_{t=i+1}^{N} y_t\right)\left(\sum_{t=1}^{N-i} x_t\right)\right] \quad (2\text{-}53)$$

$$\hat{\alpha}_{yx}(-i) = \frac{1}{N-i}\left[\sum_{t=1}^{N-i} y_t x_{t+i} - \frac{1}{N-i}\left(\sum_{t=i+1}^{N} x_t\right)\left(\sum_{t=1}^{N-i} y_t\right)\right] \quad (2\text{-}54)$$

Next, the cross-covariance transform functions are computed according to the formulas (Granger and Hatanaka, 1964):

$$\tilde{C}_{xy}(i) = \frac{1}{2}\sum_{j=0}^{m}\left[\hat{\alpha}'_{xy}(j) + \hat{\alpha}'_{yx}(j)\right]\cos\frac{ji\pi}{m} \quad (2\text{-}55)$$

and

$$\tilde{Q}_{xy}(i) = \frac{1}{2}\sum_{j=0}^{m}\left[\hat{\alpha}'_{xy}(j) - \hat{\alpha}'_{yx}(j)\right]\sin\frac{ji\pi}{m} \quad (2\text{-}56)$$

for $i = 0, 1, 2, \ldots, m$

where

$$\left.\begin{aligned} \hat{\alpha}'_{xy}(0) &= \hat{\alpha}_{xy}(0) \\ \hat{\alpha}'_{xy}(i) &= 2\hat{\alpha}_{xy}(i) \quad \text{for } 1 \le i \le m-1 \\ \hat{\alpha}'_{xy}(m) &= \hat{\alpha}_{xy}(m) \end{aligned}\right\} \quad (2\text{-}57)$$

and similarly for $\hat{\alpha}'_{yx}$.

In order to obtain the real and imaginary part of the cross spectrum, the cross-covariance transform functions are weighted according to the spectral window formulas:

$$\left.\begin{aligned} \hat{C}_{xy}(0) &= 0.5\,\tilde{C}_{xy}(0) + 0.5\,\tilde{C}_{xy}(1) \\ \hat{C}_{xy}(i) &= 0.25\,\tilde{C}_{xy}(i-1) + 0.50\,\tilde{C}_{xy}(i) \\ &\quad + 0.25\,\tilde{C}_{xy}(i+1) \\ &\quad \text{for } i = 1, 2, 3, \ldots, m-1 \\ \hat{C}_{xy}(m) &= 0.5\,\tilde{C}_{xy}(m-1) + 0.5\,\tilde{C}_{xy}(m) \end{aligned}\right\} \quad (2\text{-}58)$$

and

$$\left.\begin{aligned} \hat{Q}_{xy}(0) &= 0.5\,\tilde{Q}_{xy}(0) \\ \hat{Q}_{xy}(i) &= 0.25\,\tilde{Q}_{xy}(i-1) + 0.50\,\tilde{Q}_{xy}(i) \\ &\quad + 0.25\,\tilde{Q}_{xy}(i+1) \\ &\quad \text{for } i = 1, 2, 3, \ldots, m-1 \\ \hat{Q}_{xy}(m) &= 0.5\,\tilde{Q}_{xy}(m) \end{aligned}\right\} \quad (2\text{-}59)$$

The resulting $\hat{C}_{xy}(i)$ and $\hat{Q}_{xy}(i)$ are the estimated co-spectrum and quadrature spectrum, respectively.

The gain being similar to a regression coefficient, of series y on series x at each frequency is computed by

$$\left|\hat{H}(i)\right| = \frac{\sqrt{\hat{C}^2_{xy}(i) + \hat{Q}^2_{xy}(i)}}{\hat{G}_x(i)} \quad (2\text{-}60)$$

for $i = 0, 1, 2, \ldots, m$

The phase is estimated by

$$\hat{\theta}(i) = \tan^{-1}\left|\frac{\hat{Q}_{xy}(i)}{\hat{C}_{xy}(i)}\right| \quad (2\text{-}61)$$

for $i = 0, 1, 2, \ldots, m$

and the coherence becomes,

$$\hat{\gamma}^2(i) = \frac{\hat{C}^2_{xy}(i) + \hat{Q}^2_{xy}(i)}{\hat{G}_x(i)\ \hat{G}_y(i)} \qquad (2\text{-}62)$$

for $i = 0, 1, 2, \ldots, m$

Frequently there are obtained unstable values of the gain in the higher frequencies of the analysis performed. These values reflect no more than the rounding-off error in the division of two small quantities (Jenkins, 1963). This unrealiability carries over into the coherence function and the phase angle and it is common to obtain nonsense values for these function in the higher frequencies of the analysis. Because of this, the gain, the coherence and the phase should always be interpreted in the light of the information given by the cross-amplitude curve defined as

$$A(i) = \sqrt{\hat{C}^2_{xy}(i) + \hat{Q}^2_{xy}(i)} \qquad (2\text{-}63)$$

for $i = 0, 1, 2, \ldots, m$

7. Confidence Limits. The distribution of the cross-spectral estimates has been studied by Goodman (1957) with the main assumption being the process (x_t, y_t) has a bivariate normal distribution. If the sample size is N and the cross-spectrum is estimated over m frequency bands, the distribution of the estimated coherence, $\hat{\gamma}^2(f)$, when the true coherence is zero at a given frequency, is given by

$$F(u) = 1 - (1 - u^2)^{\frac{N}{m-1}} \qquad (2\text{-}64)$$

Equation 2-63 enables one to fix confidence limits for the coherence. Tables which give these limits have been presented by Granger and Hatanaka (1964).

Goodman's work also provides a frequency function for the estimated phase angle, $\hat{\theta}(f)$. This frequency function is extremely complicated but two important simplifications are noted by Granger and Hatanaka (1964):

i) When the true coherence is zero, $\hat{\theta}(f)$ is rectangularly distributed over the entire admissible range of values.

ii) When the true coherence is one, the variance of $\hat{\theta}(f)$ is zero.

Jenkins (1962) deduces $\hat{\theta}(f)$ approximately normally distributed with mean $\theta(f)$ and variance given by,

$$\mathrm{Var}\,(\hat{\theta}(f)) \sim \frac{1}{2}\,\frac{km}{N}\left[\frac{1}{\gamma^2(f)} - 1\right] \qquad (2\text{-}65)$$

where m and N are the same as in equation 2-63 and k is a constant associated with the particular spectral window used. The values of k are described by Parzen (1961).

Jenkins' approach has been used here to fix confidence limits for $\hat{\theta}(f)$.

CHAPTER III

MATHEMATICAL TECHNIQUES OF SPECTRAL ANALYSIS FOR LINEAR SYSTEMS

1. Frequency Response Functions. A physically realizable, constant parameter linear system is defined by the convolution integral

$$y(t) = \int_0^\infty h(\tau)\, x(t-\tau)\, d\tau \qquad (3\text{-}1)$$

The value of the output $y(t)$ is given as a weighted linear sum over the entire history of the input $x(t)$. The weighting function $h(\tau)$ associated with the system is defined as the output or response of the system to a unit impulse function, and is measured as a function of time, τ, from the moment of occurrence of the impulse input.

The dynamic characteristics of this type of system can be represented by the Fourier transform of $h(\tau)$

$$H(f) = \int_0^\infty h(\tau)\, e^{-i2\pi f\tau}\, d\tau \qquad (3\text{-}2)$$

The frequency response function is of great interest since it contains both amplitude magnification and phase shift information. Since $H(f)$ is complex valued, it can be expressed as

$$H(f) = \left| H(f) \right| e^{-i\phi(f)} \qquad (3\text{-}3)$$

The absolute value $|H(f)|$ is called the system gain factor and the angle $\phi(f)$ is called the system phase factor.

From equations 3-1 and 3-3 it is easily shown that the response of the system to a sinusoidal input of the type

$$x(t) = a \sin(2\pi f t + \zeta) \qquad (3\text{-}4)$$

can be expresses as

$$y(t) = a \left| H(f) \right| \sin\left[2\pi f t + \zeta + \phi(f)\right] \qquad (3\text{-}5)$$

Therefore, the gain $|H(f)|$ measures the amplitude magnification at frequency f when the input is a sinusoid of frequency f, while $\phi(f)$ gives the corresponding phase shift.

2. Single Input Linear Systems. From the start, it is good to notice that the largest part of the responses in geophysical systems are nonlinear. When the deviations from the linear case are not too large, the output can be written in the form

$$y(t) = \int_0^\infty h(\tau)\, x(t-\tau)\, d\tau + \eta(t) \qquad (3\text{-}6)$$

where $\eta(t)$ is a noise term which arises because the input and output variables may not be well controlled. $\eta(t)$ may also include quadratic and higher terms omitted in the linear approximation.

If $x(t)$ and $y(t)$ may be regarded as stationary time series, and $\eta(t)$ can be neglected, it can be shown the following relations hold for the system represented by equation 3-1 (Enochson, 1964),

$$G_y(f) = \left| H(f) \right|^2 G_x(f) \qquad (3\text{-}7)$$

$$G_{xy}(f) = H(f)\, G_x(f) \quad . \qquad (3\text{-}8)$$

From equation 3-8 we get:

$$\left| H(f) \right| = \frac{\left| G_{xy}(f) \right|}{G_x(f)} \qquad (3\text{-}9)$$

and

$$\theta_{xy}(f) = \phi(f) \qquad (3\text{-}10)$$

Equation 3-7 contains only the gain factor and in this manner it only gives amplitude information. Equation 3-8 is actually a pair of equations containing both the gain and the phase factor. By means of equation 3-8, if the input and corresponding output of a system are known, we can estimate $H(f)$ which will be of great importance in predicting future responses of the system. If the input $x(t)$ in equation 3-6 is of the type $x(t) = a \cos(2\pi f t + \zeta)$, the ouput of the system is:

$$y(t) = a \cdot \frac{\left| G_{xy}(f) \right|}{G_x(f)} \cos\left(2\pi f t + \zeta + \phi(f)\right) + \eta(t) \qquad (3\text{-}11)$$

where the spectrum of the residuals term $\eta(t)$ is given by Jenkins (1963) as

$$G_{\eta\eta}(f) = G_{yy}(f)\left(1 - \gamma_{xy}^2(f)\right) \qquad (3\text{-}12)$$

$G_{\eta\eta}(f)$ will give an idea of possible other periodicities in the series $y(t)$ which are not shared by $x(t)$.

It is important to notice that the frequency

response function for a constant parameter linear system is a function of frequency only. If the system were nonlinear the weighting function, $h(\tau)$, would be a function of the applied input, $h_x(\tau)$, and then the frequency response function would be a function of both, frequency and applied input. If the parameters of the system were not constant, the dynamic properties would have to be described by a time-varying weighting function, $h(\tau, t)$, which is defined as the output of the system at any time t to a unit impulse input at time $t - \tau$. In this case the frequency response function would be a function of both, frequency and time.

For a linear system, equations 3-7 and 3-8 may be substituted into the definition of coherence (equation 2-26) giving

$$\gamma^2_{xy}(f) = \frac{G_x^2(f)\,|H(f)|^2}{G_x(f)\,|H(f)|^2\,G_x(f)} = 1 \qquad (3\text{-}13)$$

Thus, the coherence function may be thought of as a measure of linear relationship in the sense that it attains a theoretical maximum of one for all f in a single input linear system.

Goodman et al. (1961) examined a single input linear system with the assumption there was noise in the measurement of the output.

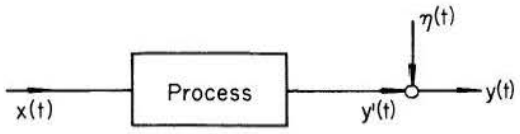

Figure 3.1 Linear system with noise in the measurement of the output (Goodman et al. 1961)

Assuming $n(t)$ and $x(t)$ statistically uncorrelated and all three processes $x(t)$, $y(t)$ and $n(t)$ stationary Gaussian noises, the effect of the disturbance $n(t)$ appears only in the coherence $\gamma^2_{xy}(f)$ which is now in the form

$$\gamma^2_{xy}(f) = \frac{1}{1 + \dfrac{G_\eta(f)}{G_{y'}(f)}} \qquad (3\text{-}14)$$

It is seen from equation 3-14 that the coherence decreases as the size of the disturbance increases.

Enochson (1964) considered a general case of noise in both input and output measuring devices. Assuming that a measured input $x(t)$ and a measured output $y(t)$ are composed of true signals $u(t)$ and $v(t)$ and uncorrelated noise components $n(t)$ and $m(t)$ respectively as shown below,

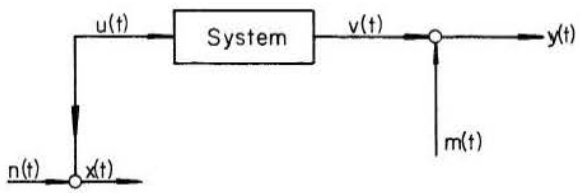

Figure 3.2 Linear system with noise in the measurement of the input and output (Enochson, 1964).

then the "desired" coherence function is

$$\gamma^2_{uv}(f) = \frac{|G_{uv}(f)|^2}{G_u(f)\,G_v(f)} \qquad (3\text{-}15)$$

but the measured coherence function will be

$$\gamma^2_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_x(f)\,G_y(f)}$$

$$= \frac{|G_{uv}(f)|^2}{\left[G_u(f) + G_n(f)\right]\left[G_v(f) + G_m(f)\right]} \qquad (3\text{-}16)$$

Thus, theoretically, the measured coherence function will always be less than the desired coherence function.

The concepts outlined above are most important in the analysis of multiple hydrologic time series and in the design of adequate hydrologic instrumentation (Eagleson and Shack, 1966).

3. Multiple Input Linear Systems. Constant parameter linear systems responding to multiple inputs from stationary random processes will now be considered. It will be assumed that N inputs are occurring with a single output being measured. The output may be considered as the sum of the

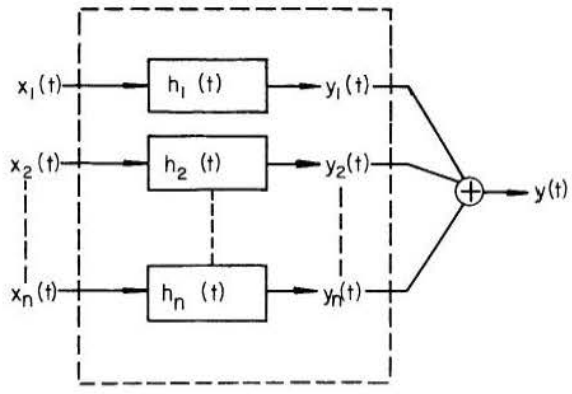

Figure 3.3 Multiple input linear system.

N partial-non-measured-outputs $y_i(t)$, i = 1, 2, ... , N . That is,

$$y(t) = \sum_{i=1}^{N} y_i(t) \qquad (3\text{-}17)$$

where $y_i(t)$ is defined as that part of the output which is produced by the i^{th} input when all the other inputs are zero.

The cross-spectral relations between the inputs and the output can be expressed concisely with matrix notation. The following formulation of results is contained in Enochson (1964) and in Bendat and Piersol (1966).

First define a N-dimensional input vector

$$\left[x(t)\right] = \left[x_1(t) , x_2(t) , ... , x_N(t)\right] \qquad (3\text{-}18)$$

Let H(f) be a N-dimensional frequency response function vector

$$\left[H(f)\right] = \left[H_1(f) , H_2(f) , ... , H_N(f)\right] \qquad (3\text{-}19)$$

Next, define a N-dimensional cross-spectrum vector of the output y(t) with the inputs $x_i(t)$,

$$\left[S_{xy}(f)\right] = \left[S_{1y}(f) , S_{2y}(f) , ... , S_{Ny}(f)\right] \qquad (3\text{-}20)$$

where

$$S_{iu}(f) = S_{x_i y}(f) , \quad i = 1, 2, ... , N \qquad (3\text{-}21)$$

Finally, define the N x N cross-spectral matrix of the inputs $x_i(t)$:

$$\left[S_{xx}(f)\right] = \begin{bmatrix} S_{11}(f) & S_{12}(f) & . . . . . & S_{1N}(f) \\ S_{21}(f) & S_{22}(f) & . . . . . & S_{2N}(f) \\ \vdots & & & \vdots \\ S_{N1}(f) & S_{N2}(f) & . . . . . & S_{NN}(f) \end{bmatrix} \qquad (3\text{-}22)$$

where

$$S_{ij}(f) = S_{x_i x_j}(f) , \quad i , j = 1 , 2 , ... , N \qquad (3\text{-}23)$$

The fundamental equation for multiple input, constant parameter linear systems can be written as:

$$S_{yy}(f) = \left[H(f)\right] \left[S_{xx}(f)\right] \left[H^{*\prime}(f)\right] \qquad (3\text{-}24)$$

where $[H^{*\prime}(f)]$ denotes the complex conjugate transpose vector of $[H(f)]$.

The basic equation which gives the transfer functions $H_j(f)$ in the case of multiple correlated inputs is

$$\left[S'_{xy}(f)\right] = \left[S_{xx}(f)\right] \left[H'(f)\right] \qquad (3\text{-}25)$$

Equation 3-25 may be rewritten as the system of equations

$$S_{iy}(f) = \sum_{j=1}^{N} H_j(f)\, S_{ij}(f) \qquad (3\text{-}26)$$

Solving equation 3-25 for the transposed row vector $[H'(f)]$ we get

$$\left[H'(f)\right] = \left[S_{xx}(f)\right]^{-1} \left[S'_{xy}(f)\right] \qquad (3\text{-}27)$$

Equation 3-27 gives each $H_i(f)$ as a function of the input-output cross spectrum and holds whether or not the inputs are correlated

The solution of equation 3-27 has been presented by Goodman(1965) in the form,

$$\left.\begin{aligned} H_1(f) &= \frac{S_{1y \cdot 234 \ldots . N}(f)}{S_{11 \cdot 234 \ldots . N}(f)} \\ &\vdots \\ H_N(f) &= \frac{S_{Ny \cdot 1234 \ldots . N-1}(f)}{S_{NN \cdot 1234 \ldots . N-1}(f)} \end{aligned}\right\} \qquad (3\text{-}28)$$

Equation 3-28 will be used in Chapter VI, where monthly rainfall in different parts of a watershed is considered as the multiple input vector which gives as total output the monthly runoff at the outlet.

Chapter IV

THEORY OF CROSS-SPECTRAL ANALYSIS OF LINEARLY DEPENDENT STOCHASTIC PROCESSES

1. Moving Average Process. The moving average process of m^{th} order is defined by

$$y(t) = \sum_{j=0}^{m} \alpha_j x(t-j) \qquad (4\text{-}1)$$

where $x(t)$ is a random process uncorrelated with $X(t-j)$ for all $j > 0$, and the α's are weights assigned to each past value of $x(t)$.

This process may be used, at least as a first approximation, as the generating scheme for certain hydrologic phenomena. For example, consider that runoff for a given interval of time is a function of all climatic factors, present and past, since the beginning of time. The dominant factor is effective precipitation which is defined as total precipitation less all losses. Because the effect of effective precipitation in the present runoff decreases with an increase in antecedency, each value of effective precipitation must be given a weight whose value decreases with an increase in antecedency. If the present runoff is essentially independent of the effective precipitation beyond the m^{th} antecedent interval of time, then runoff may be represented as being generated by a moving average of extent m of effective precipitation (Matalas, 1966).

From equations 3-7 and 3-8 we can write the spectrum of the process $y(t)$ as

$$G_y(f) = \left| \sum_{t=o}^{m} \alpha_t e^{-2\pi i f} \right|^2 G_x(f) \qquad (4\text{-}2)$$

and the cross-spectral density function $G_{xy}(f)$ as:

$$G_{xy}(f) = \left(\sum_{t=o}^{m} \alpha_t e^{-2\pi i f} \right) G_x(f) \qquad (4\text{-}3)$$

where the form of $G_y(f)$ has been studied by Siddiqui (1962). Making

$$\sum_{t=o}^{m} \alpha_t e^{-2\pi i f} = H(f) \qquad (4\text{-}4)$$

the spectral matrix of the process $y(t)$ can be written as:

$$\left[G(f)\right] = \begin{bmatrix} G_{xx}(f) & G_{xy}(f) \\ G_{yx}(f) & G_{yy}(f) \end{bmatrix} =$$

$$G_{xx}(f) \begin{bmatrix} 1 & H(f) \\ H^*(f) & |H(f)|^2 \end{bmatrix} \qquad (4\text{-}5)$$

where the use has been made of equation 2-24.

From equation 4-5 one gets,

$$\gamma^2_{xy}(f) = \frac{H(f) \cdot H^*(f)}{|H(f)|^2} = 1 \qquad (4\text{-}6)$$

The coherence function is one for all f essentially because the process $y(t)$ is a deterministic linear function of the process $x(t)$. Equation 4-6 provides a simple means of testing the validity to assume a moving average process.

2. Autoregressive Processes. In a wide variety of geophysical problems, multidimensional types of autoregressive schemes are of common application. The n^{th} dimensional autoregressive process of m^{th} order is defined by

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \sum_{j=1}^{m} \begin{bmatrix} a_{11}^{(j)} & a_{12}^{(j)} & \cdots & a_{1n}^{(j)} \\ a_{21}^{(j)} & a_{22}^{(j)} & \cdots & a_{2n}^{(j)} \\ \vdots & \vdots & & \vdots \\ a_{n1}^{(j)} & a_{n2}^{(j)} & \cdots & a_{nn}^{(j)} \end{bmatrix} \begin{bmatrix} x_1(t-j) \\ x_2(t-j) \\ \vdots \\ x_n(t-j) \end{bmatrix} + \begin{bmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_n(t) \end{bmatrix} \qquad (4\text{-}7)$$

where $[z_j(t)]$ is a random component uncorrelated with $[x_j(t-k)]$ and $[z_j(t-k)]$ for all $k > 0$. In practical cases, the random component, $[z_j(t)]$, may be interpreted as the residuals which represent the action of events, other than $[x_j(t-k)]$, affecting $[x_j(t)]$.

Examples of this type of process occur frequently in hydrology. In the case of representing runoff as a moving average process of effective precipitation, m may happen to be very large. It would be a mistake to try to reduce the order of the moving average because it is convenient to consider the effective precipitation for all antecedent intervals of time, however small their contribution may be to the present runoff. In this case, the generating scheme for the runoff can be represented by an autoregressive process which involves far fewer coefficients than the moving average process (Matalas, 1966).

River flows can be represented in many instances by autoregressive processes (Yevjevich 1964, Roesner 1965, Quimpo 1966).

In the analysis of the cross-spectral characteristics of autoregressive processes, the mathematical complications increase very rapidly with the order of the process. Fortunately, most of the autoregressive processes used in hydrology are the first or second order Markov linear processes. The analysis and results obtained for the cross-spectral characteristics of Markov linear processes is believed to be new by the author.

First order Markov linear process. Let us assume that a certain process, such as annual runoff, can be represented by a first order model,

$$\left[x_j(t)\right] = \left[a_{ij}\right]\left[x_j(t-1)\right] + \left[z_j(t)\right] \qquad (4\text{-}8)$$

where the terms are the matrices of equation 4-7 when $m = 1$. Two realizations of this process can be expressed by

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} a_{11}a_{12} \\ a_{21}a_{22} \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix} + \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad (4\text{-}9)$$

which results from making $m = 1$, $n = 2$ in equation 4-7.

Equation 4-9 is equivalent to the system:

$$x_1(t) = a_{11}x_1(t-1) + a_{12}x_2(t-1) + z_1(t) \qquad (4\text{-}10)$$

$$x_2(t) = a_{21}x_1(t-1) + a_{22}x_2(t-1) + z_2(t) \qquad (4\text{-}11)$$

Quenouille (1957) shows the coefficient matrix of a first order autoregressive scheme to be equal to:

$$\left[a_{ij}\right] = \left[\alpha_{ij}(1)\right] \cdot \left[\alpha_{ij}(0)\right]^{-1} \qquad (4\text{-}12)$$

where $[\alpha_{ij}(1)]$ is the covariance matrix for the lag one,

$$\left[\alpha_{ij}(1)\right] = \begin{bmatrix} \alpha_{11}(1) & \alpha_{12}(1) \\ \alpha_{21}(1) & \alpha_{22}(1) \end{bmatrix} \qquad (4\text{-}13)$$

and similarly

$$\left[\alpha_{ij}(0)\right] = \begin{bmatrix} \alpha_{11}(0) & \alpha_{12}(0) \\ \alpha_{21}(0) & \alpha_{22}(0) \end{bmatrix} \qquad (4\text{-}14)$$

the inverse of $[\alpha(0)]$ denoted by $[\alpha(0)]^{-1}$.

From equation 4-12 it is directly obtained:

$$a_{11} = \frac{\alpha_{11}(1)\,\alpha_{22}(0) - \alpha_{12}(1)\,\alpha_{21}(0)}{\alpha_{11}(0)\,\alpha_{22}(0) - \alpha_{21}(0)\,\alpha_{12}(0)} = \frac{\rho_{11}(1) - \rho_{12}(1)\,\rho_{12}(0)}{1 - \rho_{21}^2(0)} \qquad (4\text{-}15)$$

where

$$\rho_{12}(s) = \rho_{21}(-s) = \frac{\alpha_{12}(s)}{\left[\alpha_{11}(0) \cdot \alpha_{22}(0)\right]^{1/2}} \qquad (4\text{-}16)$$

The terms $\alpha_{11}(0)$ and $\alpha_{22}(0)$ represent the variances of $x_1(t)$ and $x_2(t)$, respectively:

$$\alpha_{11}(0) = \sigma_1^2 \qquad (4\text{-}17)$$

$$\alpha_{22}(0) = \sigma_2^2 \qquad (4\text{-}18)$$

Similarly we can obtain:

$$a_{22} = \frac{-\alpha_{21}(1)\,\alpha_{12}(0) + \alpha_{22}(1)\,\alpha_{11}(0)}{\alpha_{11}(0)\,\alpha_{22}(0) - \alpha_{21}(0)\,\alpha_{12}(0)} = \frac{\rho_{22}(1) - \rho_{12}(0)\,\rho_{21}(1)}{1 - \rho_{12}^2(0)} \qquad (4\text{-}19)$$

$$a_{12} = \frac{-\alpha_{11}(1)\,\alpha_{12}(0) + \alpha_{12}(1)\,\alpha_{11}(0)}{\alpha_{11}(0)\,\alpha_{22}(0) - \alpha_{21}(0)\,\alpha_{12}(0)} =$$

$$\frac{\sigma_1 \left[\rho_{12}(1) - \rho_{11}(1)\,\rho_{12}(0)\right]}{\sigma_2 \left[1 - \rho_{12}^2(0)\right]} \qquad (4\text{-}20)$$

$$a_{21} = \frac{\alpha_{21}(1)\,\alpha_{22}(0) - \alpha_{22}(1)\,\alpha_{21}(0)}{\alpha_{11}(0)\,\alpha_{22}(0) - \alpha_{21}(0)\,\alpha_{12}(0)} =$$

$$\frac{\sigma_2 \left[\rho_{21}(1) - \rho_{22}(1)\,\rho_{12}(0)\right]}{\sigma_1 \left[1 - \rho_{12}^2(0)\right]} \qquad (4\text{-}21)$$

Assuming $z_1(t)$ uncorrelated with $x_1(t-s)$ and multiplying equation 4-10 by $x_1(t-s)$, one gets after taking expected values:

$$\sigma_1^2\,\rho_{11}(s) = a_{11}\sigma_1^2\,\rho_{11}(s-1) + a_{12}\sigma_1\sigma_2\rho_{21}(s-1) \qquad (4\text{-}22)$$

Assuming now that one has two first order Markov linear processes, or in other words that the coefficients a_{12} and a_{21} are identically equal to zero, further simplifications are possible. Substituting in equation 4-22 the well known result (Kendall, 1966):

$$\rho_{11}(s) = \rho_{11}^{|s|}(1) \qquad (4\text{-}23)$$

one gets:

$$\sigma_1^2\,\rho_{11}^{|s|}(1) = a_{11}\sigma_1^2\,\rho_{11}^{|s-1|}(1) + a_{12}\sigma_1\sigma_2\,\rho_{21}(s-1) \qquad (4\text{-}24)$$

Using for a_{11} and a_{12} the values given by equations 4-15 and 4-20 and performing some simplifications, equation 4-24 may be reduced to:

$$\rho_{21}(s-1) = \rho_{11}^{|s-1|}(1)\,\rho_{12}(0) \qquad (4\text{-}25)$$

From equation 4-25 it becomes:

$$\rho_{21}(s) = \rho_{11}^{|s|}(1)\,\rho_{12}(0) \qquad (4\text{-}26)$$

and

$$\alpha_{21}(s) = \rho_{11}^{|s|}(1)\,\rho_{12}(0)\,\sigma_1\sigma_2 \; . \qquad (4\text{-}27)$$

Taking the Fourier transform of $\alpha_{21}(s)$ we get the cross-spectrum between the series $x_1(t)$ and $x_2(t)$:

$$S_{21}(f) = \sum_{s=-\infty}^{\infty} \rho_{11}^{|s|}(1)\,\rho_{12}(0)\,\sigma_1\sigma_2\,e^{-2\pi i f s} \qquad (4\text{-}28)$$

or

$$S_{21}(f) = \sum_{s=0}^{\infty} \rho_{11}^{s}(1)\,\rho_{12}(0)\,\sigma_1\sigma_2 e^{2\pi i f s} +$$

$$\sum_{s=0}^{\infty} \rho_{11}^{s}(1)\,\rho_{12}(0)\,\sigma_1\sigma_2 e^{-2\pi i f s} - \rho_{12}(0)\sigma_1\sigma_2 =$$

$$\frac{\rho_{12}(0)\sigma_1\sigma_2}{1-\rho_{22}(1)e^{-2\pi i f}} + \frac{\rho_{12}(0)\sigma_1\sigma_2}{1-\rho_{11}(1)e^{2\pi i f}} - \rho_{12}(0)\,\sigma_1\sigma_2 \; . \qquad (4\text{-}29)$$

Equation 4-29 gives the cross-spectrum between two first order Markov linear processes.

Siddiqui (1962) shows that the spectrum of a first order Markov linear process is equal to:

$$S_{11}(f) = \frac{\sigma_1^2 \left[1 - \rho_{11}^2(1)\right]}{\rho_{11}^2(1) + 1 - 2\rho_{11}(1)\cos 2\pi f} \qquad (4\text{-}30)$$

which can be written as,

$$S_{11}(f) = \frac{\sigma_1^2 \left[1 - \rho_{11}^2(1)\right]}{\left[1 - \rho_{11}(1)\,e^{2\pi i f}\right]\left[1 - \rho_{11}(1)\,e^{-2\pi i f}\right]} \qquad (4\text{-}31)$$

Similarly,

$$S_{22}(f) = \frac{\sigma_2^2 \left[1 - \rho_{22}^2(1)\right]}{\left[1 - \rho_{22}(1)\,e^{2\pi i f}\right]\left[1 - \rho_{22}(1)\,e^{-2\pi i f}\right]} \qquad (4\text{-}32)$$

Using equations 4-29, 4-31 and 4-32 and making some simplifications, the coherence between the two processes is found to be equal to:

$$\gamma_{12}^2(f) = \frac{\left|S_{12}(f)\right|^2}{S_{11}(f)\,S_{22}(f)} =$$

$$\frac{\left(1 - \rho_{11}(1)\,\rho_{22}(1)\right)^2 \rho_{12}^2(0)}{\left(1 - \rho_{11}^2(1)\right)\left(1 - \rho_{22}^2(1)\right)} \qquad (4\text{-}33)$$

Equation 4-33 shows the coherence function between two first order autoregressive processes equal to a constant independent of frequency*. This result should provide a valuable tool in the analysis of this type of processes.

Second order Markov linear process. The method that will be presented now can be used to study the spectral matrix of any autoregressive process regardless of its order or dimension.

Quimpo and Yevjevich (1967) have shown that 2^{nd} order autoregressive schemes may be used to fit the patterns in the sequence of daily river flows after the periodic component has been removed from the series.

The equation for a two-dimensional 2^{nd} order autoregressive process is obtained by making $n = 2$ and $m = 2$ in equation 4-7:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} a_{12} \\ a_{21} a_{22} \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix} + \begin{bmatrix} b_{11} b_{12} \\ b_{21} b_{22} \end{bmatrix} \begin{bmatrix} x_1(t-2) \\ x_2(t-2) \end{bmatrix} + \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad (4\text{-}34)$$

which is equivalent to the system:

$$x_1(t) = a_{11}x_1(t-1) + a_{12}x_2(t-1) + b_{11}x_1(t-2) + b_{12}x_2(t-2) + z_1(t) \qquad (4\text{-}35)$$

$$x_2(t) = a_{21}x_1(t-1) + a_{22}x_2(t-1) + b_{21}x_1(t-2) + b_{22}x_2(t-2) + z_2(t) \qquad (4\text{-}36)$$

In order to obtain the coefficient matrix $[a_{ij}]$ as was done for the 1^{st} order Markov process, it will be necessary to transform the 2^{nd} order process to a 1^{st} order process. This, in turn, can be accomplished by doing

$$x_1(t-1) = x_3(t) \qquad (4\text{-}37)$$

and

$$x_2(t-1) = x_4(t) \qquad (4\text{-}38)$$

Equations 4-34, 4-37, and 4-38 can be written as:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} a_{11} a_{12} & 0 & 0 \\ a_{21} a_{22} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \\ x_3(t-1) \\ x_4(t-1) \end{bmatrix} + \begin{bmatrix} 0 & 0 & b_{11} b_{12} \\ 0 & 0 & b_{21} b_{22} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \\ x_3(t-1) \\ x_4(t-1) \end{bmatrix} + \begin{bmatrix} z_1(t) \\ z_2(t) \\ 0 \\ 0 \end{bmatrix} \qquad (4\text{-}39)$$

which can be expressed:

$$\left[x_j(t)\right] = \left[a_{ij}\right]\left[x_j(t-1)\right] + \left[b_{ij}\right]\left[x_j(t-1)\right] + \left[z_j(t)\right] \qquad (4\text{-}40)$$

or

$$\left[x_j(t)\right] = \left[u_{ij}\right]\left[x_j(t-1)\right] + \left[z_j(t)\right] \qquad (4\text{-}41)$$

where

$$\left[u_{ij}\right] = \begin{bmatrix} a_{11} & a_{12} & b_{11} & b_{12} \\ a_{21} & a_{22} & b_{21} & b_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad (4\text{-}42)$$

and $[x_j(t)]$ now represents a four-dimensional 1^{st} order process for which we can obtain $[u_{ij}]$:

---

* This result was first pointed out to the author by Dr. M. M. Siddiqui of Colorado State University

$$\left[u_{ij}\right] = \left[\alpha_{ij}(1)\right] \cdot \left[\alpha_{ij}(0)\right]^{-1} \qquad (4\text{-}43)$$

$[\alpha_{ij}(1)]$ and $[\alpha_{ij}(0)]$ represent now four by four covariance matrices.

Using straightforward relations like:

$$\alpha_{13}(1) = \text{Cov}\left[x_1(t)\, x_3(t-1)\right] =$$

$$\text{Cov}\left[x_1(t)\, x_1(t-2)\right] = \alpha_{11}(2)$$

we can express $[\alpha_{ij}(1)]$ by:

$$\left[\alpha_{ij}(1)\right] = \begin{bmatrix} \alpha_{11}(1) & \alpha_{12}(1) & \alpha_{11}(2) & \alpha_{12}(2) \\ \alpha_{21}(1) & \alpha_{22}(1) & \alpha_{21}(2) & \alpha_{22}(2) \\ \alpha_{11}(0) & \alpha_{12}(0) & \alpha_{11}(1) & \alpha_{12}(1) \\ \alpha_{21}(0) & \alpha_{22}(0) & \alpha_{21}(1) & \alpha_{22}(1) \end{bmatrix} \qquad (4\text{-}44)$$

and similarly,

$$\left[\alpha_{ij}(0)\right] = \begin{bmatrix} \alpha_{11}(0) & \alpha_{12}(0) & \alpha_{11}(1) & \alpha_{12}(1) \\ \alpha_{21}(0) & \alpha_{22}(0) & \alpha_{21}(1) & \alpha_{22}(1) \\ \alpha_{11}(1) & \alpha_{21}(1) & \alpha_{11}(0) & \alpha_{12}(0) \\ \alpha_{12}(1) & \alpha_{22}(1) & \alpha_{21}(0) & \alpha_{22}(0) \end{bmatrix} \qquad (4\text{-}45)$$

After the calculation of $\left[u_{ij}\right]$, we may return to the original model of equation 4-34 which can be written:

$$\begin{bmatrix} 1 - a_{11}D - b_{11}D^2 & -a_{12}D - b_{12}D^2 \\ -a_{21}D - b_{21}D^2 & 1 - a_{22}D - b_{22}D^2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad (4\text{-}46)$$

or in shorter notation:

$$\left[p_{ij}(D)\right] \cdot \left[x_j(t)\right] = \left[q_{ij}\right]\left[z_j(t)\right] \qquad (4\text{-}47)$$

where the operator D is defined by:

$$D^n x(t) = x(t-n) \qquad (4\text{-}48)$$

The spectral matrix of the process is now obtained using a formula given by Whittle (1954):

$$\left[S_{ij}(f)\right] = \left[v_{ij}(z)\right] \cdot \left[v^*_{ij}{}'(z)\right] \qquad (4\text{-}49)$$

where

$$z = e^{2\pi f i}$$

and

$$\left[v_{ij}(D)\right] = \left[p_{ij}(D)\right]^{-1}\left[q_{ij}\right] \qquad (4\text{-}50)$$

with $[v^*_{ij}{}'(z)]$ standing for the complex conjugate transpose matrix of $[v_{ij}(z)]$.

Equation 4-49 is valid provided the series $x_1(t)$ and $x_2(t)$ have unit variance.

In the case studied here $[q_{ij}]$ is the unit matrix, therefore

$$\left[v_{ij}(D)\right] = \left[p_{ij}(D)\right]^{-1} \qquad (4\text{-}51)$$

and

$$\left[p_{ij}(D)\right]^{-1} = \frac{1}{\det\left[p_{ij}(D)\right]} \cdot \begin{bmatrix} 1-a_{22}D-b_{22}D^2 & a_{12}D+b_{12}D^2 \\ a_{21}D+b_{21}D^2 & 1-a_{11}D-b_{11}D^2 \end{bmatrix} \qquad (4\text{-}52)$$

The spectral matrix of the process can be written:

$$\left[S_{ij}(f)\right] = \frac{1}{\left|\det\left[p_{ij}(z)\right]\right|^2} \cdot \begin{bmatrix} F_{11}(z) & F_{12}(z) \\ F_{21}(z) & F_{22}(z) \end{bmatrix} \qquad (4\text{-}53)$$

where

$$F_{11}(z) = 1 + a_{12}^2 + b_{22}^2 + a_{22}^2 + b_{12}^2 + (a_{22}b_{22} + a_{12}b_{12} - a_{22})(z^{-1} + z) - b_{22}(z^{-2} + z^2) \qquad (4\text{-}54)$$

$$F_{22}(z) = 1 + a_{21}^2 + b_{11}^2 + a_{11}^2 + b_{21}^2 + (a_{21}\, b_{21} - a_{11} + a_{11}\, b_{11})\,(z + z^{-1}) - b_{11}\,(z^{-2} + z^2) \qquad (4\text{-}55)$$

$$F_{12}(z) = F^*_{21}(z) = (1 - a_{22}\, z - b_{22}\, z^2) \cdot (a_{21}\, z^{-1} + b_{21}\, z^{-2}) + (a_{12} z + b_{12}\, z^2) \cdot (1 - a_{11}\, z^{-1} + b_{11}\, z^{-2}) \qquad (4\text{-}56)$$

The coherence function of the process can then be expressed as:

$$\gamma^2(f) = \frac{\left| F_{12}(f) \right|^2}{F_{11}(f)\, F_{22}(f)} \qquad (4\text{-}57)$$

After going through some algebra, $\gamma^2(f)$ can be written:

$$\gamma^2(f) = \frac{2\,\theta \cos 2\pi f + 2\,\psi \cos 4\pi f + 2\nu \cos 6\,\pi f + 2\,\zeta \cos 8\pi f + \rho}{\left[\alpha + 2\beta \cos 2\pi f - 2b_{22} \cos 4\pi f\right]\left[k + 2\delta \cos 2\pi f - 2b_{11} \cos 4\pi f\right]} \qquad (4\text{-}58)$$

where

$$\alpha = 1 + a_{12}^2 + b_{22}^2 + a_{22}^2 + b_{12}^2$$

$$\beta = a_{22}\, b_{22} + a_{12}\, b_{12} - a_{22}$$

$$k = 1 + a_{21}^2 + b_{11}^2 + a_{11}^2 + b_{21}^2$$

$$\delta = a_{21}\, b_{21} + a_{11}\, b_{11} - a_{11}$$

$$\epsilon = a_{21} - a_{22}\, b_{21} + a_{12}\, b_{11}$$

$$\lambda = -b_{22}\, a_{21} + a_{12} - b_{12}\, a_{11}$$

$$\eta = b_{12}\, b_{11} - a_{22} a_{21} - b_{22} b_{21} - a_{12} a_{11}$$

$$\rho = \epsilon^2 + \eta^2 + b_{21}^2 + b_{12}^2 + \lambda^2$$

$$\theta = \epsilon b_{21} + \epsilon\eta + b_{12}\,\lambda + \lambda\eta$$

$$\psi = \epsilon\lambda + b_{21}\eta + b_{12}\,\eta$$

$$\nu = \epsilon b_{12} + b_{21}\,\lambda$$

$$\zeta = b_{21}\, b_{12}$$

3. Mathematical Development of the Cross-Spectral Characteristics of Filtered Series. The smoothing of time series by moving average schemes and other types of filters is a practice sometimes applied in hydrology and other geophysical sciences. So it is of great practical importance to understand clearly the different effects that filters can have in the cross-spectral characteristics of time series. In this chapter, we will study the behavior of the cross-spectral density function, the coherence function and the phase function when one or both of the series in which the analysis will be performed have been pre-filtered.

Let us have two random input functions $x_1(t)$ and $x_2(t)$ related to two output functions $y_1(t)$ and $y_2(t)$ through a linear filter function $h(t)$ by means of a simple convolution,

$$y_1(t) = \int_{-\infty}^{\infty} x_1(t - u)\; h_1(u)\, du \qquad (4\text{-}59)$$

$$y_2(t) = \int_{-\infty}^{\infty} x_2(t - s)\; h_2(s)\; ds \quad . \qquad (4\text{-}60)$$

The cross-covariance function between the filtered output functions is:

$$\alpha_{y_1 y_2} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y_1(t)\; y_2(t + \tau)\, dt \qquad (4\text{-}61)$$

which can be written:

$$\alpha_{y_1 y_2}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt \int_{-\infty}^{\infty} x_1(t-u) h_1(u) du \cdot \int_{-\infty}^{\infty} x_2(t+\tau-s) h_2(s) ds = \int_{-\infty}^{\infty} h_1(u)\, du \int_{-\infty}^{\infty} h_2(s)\, ds \cdot \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x_1(t-u)\, x_2(t+\tau-s) dt = \int_{-\infty}^{\infty} h_1(u)\, du \cdot \int_{-\infty}^{\infty} h_2(s)\, ds \cdot \alpha_{x_1 x_2}(\tau + u - s) \qquad (4\text{-}62)$$

In order to go from the time domain to the frequency domain, we take Fourier transforms at both sides of equation 4-62,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \alpha_{y_1 y_2}(\tau)\; e^{-i\omega\tau} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} d\tau \cdot \int_{-\infty}^{\infty} h_1(u)\, du \int_{-\infty}^{\infty} h_2(s)\, ds \cdot \alpha_{x_1 x_2}(\tau + u - s) \qquad (4\text{-}63)$$

where ω represents the angular frequency (radians per unit of time).

The left hand side of equation 4-63 is the cross-spectrum between $y_1(t)$ and $y_2(t)$: $G_{y_1y_2}(\omega)$

Doing $\ell = \tau + u - s$ in equation 4-63 we get:

$$G_{y_1y_2}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega(\ell + s - u)} d\ell \int_{-\infty}^{\infty} h_1(u)\, du \cdot$$

$$\int_{-\infty}^{\infty} h_2(s)\, ds\, \alpha_{x_1x_2}(\ell) = \int_{-\infty}^{\infty} h_1(u)\, e^{i\omega u}\, du \cdot$$

$$\int_{-\infty}^{\infty} h_2(s)\, e^{-i\omega s} ds \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} \alpha_{x_1x_2}(\ell)\, e^{-i\omega\ell}\, d\ell$$

Therefore,

$$G_{y_1y_2}(\omega) = G_{x_1x_2}(\omega) \cdot \int_{-\infty}^{\infty} h_1(u)\, e^{i\omega u}\, du \cdot$$

$$\cdot \int_{-\infty}^{\infty} h_2(s)\, e^{-i\omega s}\, ds \qquad (4\text{-}64)$$

Equation 4-64 gives the relation between the corss-spectrum of the filtered series as a function of the cross-spectrum of the original series and of the linear operators $h_1(u)$ and $h_2(s)$

The transfer function of a filter $h(t)$ is defined as the Laplace transform of $h(t)$:

$$R(\omega) = \int_{-\infty}^{\infty} h(t)\, e^{-i\omega t}\, dt = \left| R(\omega) \right| e^{-i\phi(\omega)}$$

$$= \int_{-\infty}^{\infty} h(t) \cos \omega t\, dt - i \int_{-\infty}^{\infty} h(t) \sin \omega t\, dt$$

$$= \mathrm{Re} \left\{ R(\omega) \right\} + i\, \mathrm{Im} \left\{ R(\omega) \right\} \qquad (4\text{-}65)$$

where

$$\left| R(\omega) \right| = \left[(\mathrm{Re} \left\{ R(\omega) \right\})^2 + (\mathrm{Im} \left\{ R(\omega) \right\})^2 \right]^{\frac{1}{2}} \qquad (4\text{-}66)$$

and

$$\phi(\omega) = \tan^{-1} \left[\mathrm{Im} \left\{ R(\omega) \right\} \Big/ \mathrm{Re} \left\{ R(\omega) \right\} \right] \qquad (4\text{-}67)$$

The angle $\phi(\omega)$ represents the phase shift which the filtering function $h(t)$ produces at the frequency ω .

For smoothing and filtering functions having $(n + m + 1)$ discrete weights, the transfer function is computed by the following form of equation 4-65:

$$R(\omega) = \sum_{k=-n}^{m} h_k \cos \omega k - i \sum_{k=-n}^{m} h_k \sin \omega k \qquad (4\text{-}68)$$

In the spectral analysis of a time series, the effect of applying a filter to the series is to multiply the power spectrum by $|R(\omega)|^2$, (Siddiqui, 1962). When analyzing a single time series, the phase shifts will have no effect on the spectral analysis since the spectrum suppresses all phase information. This is not the case in cross-spectral analysis where the phase diagram is a very useful one. So, it is highly desirable that smoothing and filtering functions do not shift the phase of waves of any frequency. The shift angle can be made equal to zero by requiring that the imaginary part of $R(\omega)$ be zero. This, in turn, can be accomplished by requiring the filter function $h(t)$ to be even, for if $h(t)$ is even, the terms containing the sines in equations 4-65 and 4-68 are zero, and $R(\omega)$ is a pure real quantity computed by

$$R(\omega) = 2 \int_{o}^{\infty} h(t) \cos \omega t\, dt \qquad (4\text{-}69)$$

for continuous $h(t)$ functions, or by

$$R(\omega) = \sum_{k=-n}^{n} h_k \cos \omega k = h_o + 2 \sum_{k=1}^{n} h_k \cos \omega k \qquad (4\text{-}70)$$

for smoothing and filtering functions having $(2n + 1)$ discrete weights.

Using the definition of $R(\omega)$, equation 4-65, we can write equation 4-64 as

$$G_{y_1y_2}(\omega) = G_{x_1x_2}(\omega) \cdot R^*_{h_1}(\omega) \cdot R_{h_2}(\omega) \qquad (4\text{-}71)$$

From equation 4-71 it is seen that the cross-spectrum of the filtered series will be different of the cross-spectrum of the original series.

It is of fundamental interest to know if linear filters like those of equations 4-59 and 4-60 will change the coherence between the series. It is known (Siddiqui, 1962) that the individual spectra of the filtered series are equal to

$$G_{y_1}(\omega) = \left| R_{h_1}(\omega) \right|^2 G_{x_1}(\omega) \qquad (4\text{-}72)$$

and

$$G_{y_2}(\omega) = \left| R_{h_2}(\omega) \right|^2 G_{x_2}(\omega) \quad . \qquad (4\text{-}73)$$

In this manner the coherence $\gamma^2_{y_1y_2}(\omega)$ can be written as:

$$\gamma^2_{y_1 y_2}(\omega) = \frac{\left|G_{y_1 y_2}(\omega)\right|^2}{G_{y_1}(\omega)\, G_{y_2}(\omega)} =$$

$$= \frac{\left|R^*_{h_1}(\omega)\right|^2 \left|R_{h_2}(\omega)\right|^2 \left|G_{x_1 x_2}(\omega)\right|^2}{\left|R_{h_1}(\omega)\right|^2 G_{x_1}(\omega) \left|R_{h_2}(\omega)\right|^2 G_{x_2}(\omega)} =$$

$$= \frac{\left|G_{x_1 x_2}(\omega)\right|^2}{G_{x_1}(\omega)\, G_{x_2}(\omega)} = \gamma^2_{x_1 x_2}(\omega) \qquad (4\text{-}74)$$

Equation 4-74 shows that the coherence function of the filtered series is the same as the coherence function between the original series.

The main conclusions of this chapter with regard to the use of linear filters before any cross-spectral analysis is performed may be summarized as follows:

a. The phase function of the filtered series is different from the phase function of the original series except if both filters are even, in which case, the phase function will remain unchanged.

b. The cross-spectrum of the filtered series is different from the cross-spectrum of the original series, their relation being given by equation 4-71.

c. The coherence function will remain unchanged after the use of any kind of linear filter.

CHAPTER V

DATA ASSEMBLY AND PROCEDURE FOR THE ANALYSIS OF HYDROLOGIC SERIES BY CROSS-SPECTRAL TECHNIQUES

1. Data Selection. One of the aspects this dissertation was directly concerned with was to study the frequency correlations between hydrologic time series and the characteristics of the gain and phase functions between them. To attempt this, several stations were chosen and complete cross-spectral analyses were performed between them and groups of other stations in the same or different environment.

Precipitation data consisted of annual and monthly series. Only monthly runoff data were analyzed. Figure 5.1 shows the location of the stations used in the analysis. A detailed description of these stations is done in the appendix.

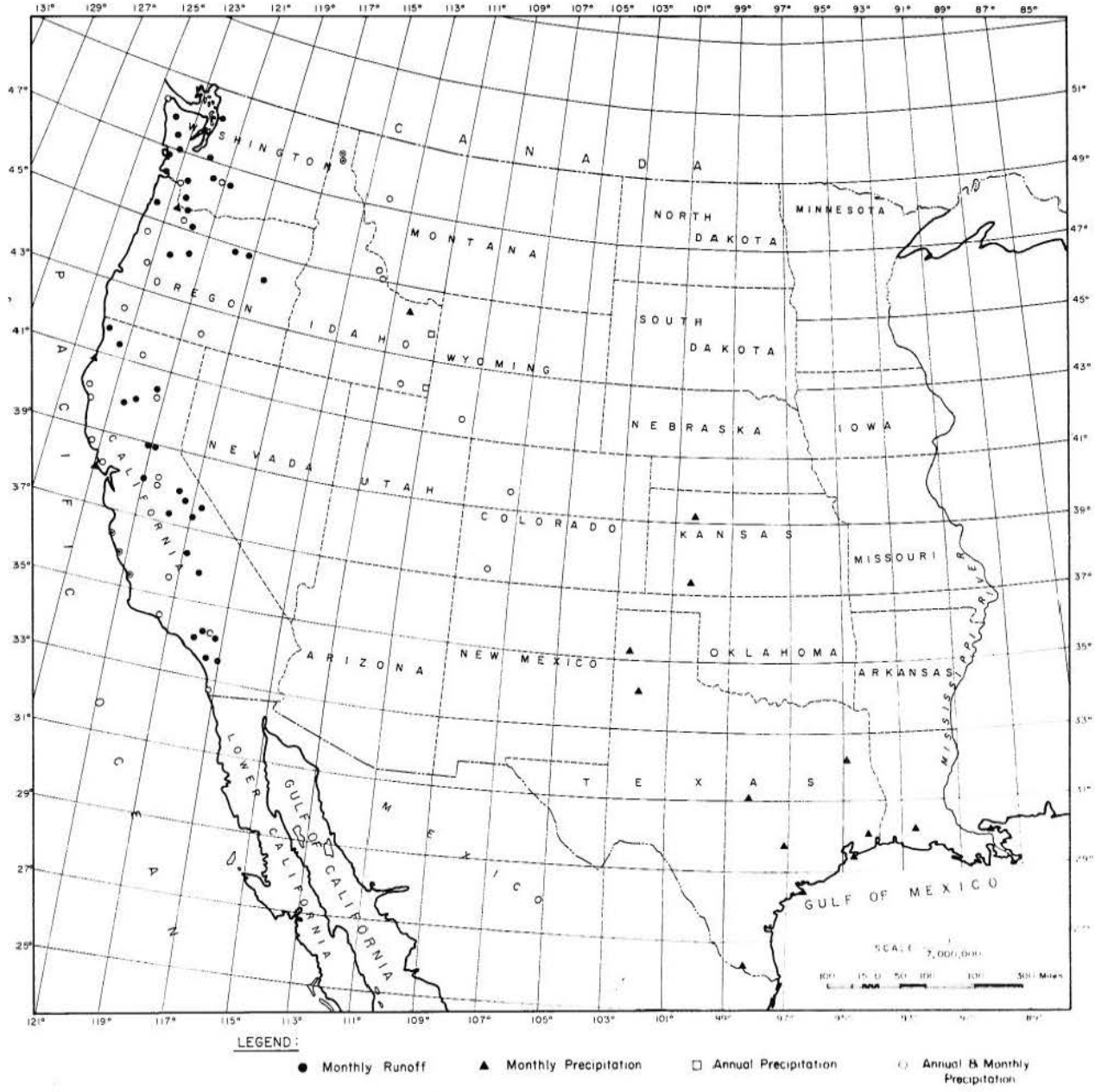

Fig. 5.1 Geographic distribution of stations used in the analysis.

In the analysis of annual precipitation, there were 27 stations with an average length of data of 62 years per station. There were 41 stations with monthly precipitation data averaging 57 years per station. The rainfall stations were divided into five regions in each of which one base station was fixed. Complete cross-spectral analyses were then made between the base stations and all the stations in the region. The characteristics of the base stations are as follows:

Region No. 1 Pacific Coast (California, Oregon, Washington) with 14 stations and the base station: San Diego Lat: 32.733 Long: 117.167 Period of records: 1850-1960.

Region No. 2 Valley Environment (California, Oregon) with 8 stations and the base station: Sonora Lat: 37.983 Long: 120.383 Period of records: 1888-1960.

Region No. 3 Mountain Environment (Colorado, Wyoming, Idaho, Montana) with 6 stations and the base station: Durango Lat: 37.280 Long: 107.880 Period of records: 1895-1960.

Region No. 4 Gulf of Mexico (Texas, Louisiana) with 7 stations and the base station: New Orleans Lat: 29.95 Long: 90.07 Period of records: 1870-1960.

Region No. 5 Plain Regions (Texas, Oklahoma, Kansas) with 6 stations and the base station: Lampasas Lat: 31.05 Long: 98.18 Period of records: 1895-1960.

There were 37 runoff stations, all of them located west of the 117° meridian. The average length of series among these stations was 36 years. The base station characteristics are:

Middle Fork American River near Auburn, California. Lat: 38.92 Long: 121.00 Period of records: 1912-1960.

2. Method of Analysis. All the data used were previously standardized by the transformation:

$$x^{*}(t) = \frac{x(t) - \bar{x}}{s(x)} \qquad (5\text{-}1)$$

where $x(t)$ is the series of values of each station, $\bar{x}$ is the mean of $x(t)$, and $s(x)$ is the standard deviation of the series. Because of the previous standardization, the spectral ordinates are all in $[\text{cycles per unit of time}]^{-1}$.

With the annual data there were used 15 lags with a resolution of 0.067 cycles per year. In the monthly analysis there were 24 lags with a resolution of 0.042 cycles per month.

CHAPTER VI

APPLICATION OF CROSS-SPECTRAL TECHNIQUES TO HYDROLOGIC TIME SERIES

1. Analysis of Monthly Precipitation Data. Monthly hydrologic data can be regarded as consisting of two parts: periodic and stochastic (Roesner and Yevjevich, 1966)

$$x(t) = m(t) + z(t) \tag{6-1}$$

where

$$m(t) = \bar{x} + \sum_j \left(a_j \cos \frac{2\pi jt}{12} + b_j \sin \frac{2\pi jt}{12}\right) \tag{6-2}$$

$$\bar{x} = N^{-1} \sum_t x(t) \tag{6-3}$$

and the summation over j can vary from $j = 1$ (common case for rainfall stations), or $j = 1, 2$ (frequently found in runoff series) to $j = 1, 2, 3, 4, 5$ in some extreme cases. $z(t)$ represents the "noise" or random component which for monthly precipitation series follows a random independent model and for the monthly runoff series follows a first order autoregressive process (Roesner, 1966).

In the monthly precipitation data of regions 1 and 2, the variance explained by the annual oscillation appears to depend upon the climatic and thermal conditions of the station, being large where the annual temperature range is small and less pronounced where the latter is large. Peak changes in the spectrum can be considerable even over short distances: at Lakeview, Oregon, where the mean annual temperature range is about 40°F, the peak in the spectrum at the annual cycle of precipitation is 3.098; at Auburn, California, 490 Km. away from Lakeview, where the mean annual temperature range is about 33°F, the peak in the spectrum is 6.007; in Tatoosh Island, Washington, where the temperature range is only 13.7°F, the peak is 7.073.

In order to investigate more fully the frequency correlations between temperature, atmospheric pressure and precipitation, the cross-spectra and partial cross-spectra of these three series were calculated for five of the stations located in region 1. Figures 6.1 through 6.10 show the coherence and partial coherence functions obtained in this analysis. It appears from these figures that, for the stations considered, there exists a real relation between the annual cycle in temperature and the annual cycle in precipitation, but only an apparent correlation exists between the annual cycles in pressure and precipitation. Notice that the high coherence between the annual cycles of pressure and precipitation is in all cases non-significantly different from zero when the effect of temperature is subtracted from the analysis. On the other hand, the partial coherence between the annual cycles of temperature and pressure is high in all but one case.

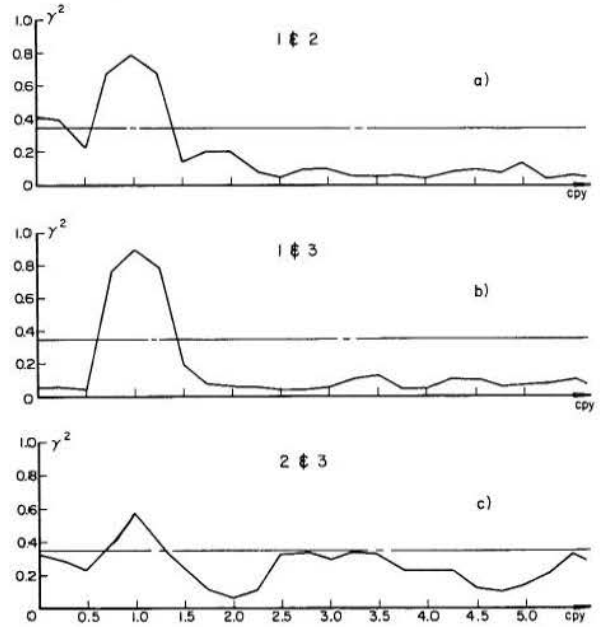

Fig. 6.1 Coherence functions for monthly data or temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Eureka (California) with their corresponding 95% significance levels.

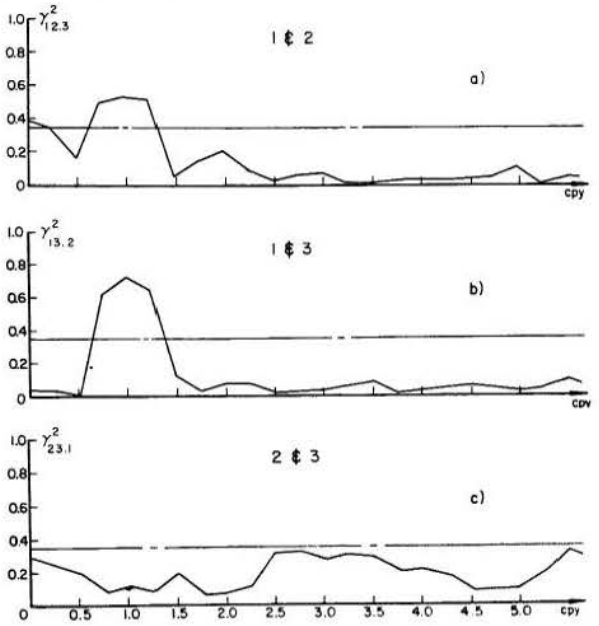

Fig. 6.2 Partial coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Eureka (California) with their corresponding 95% significance levels.

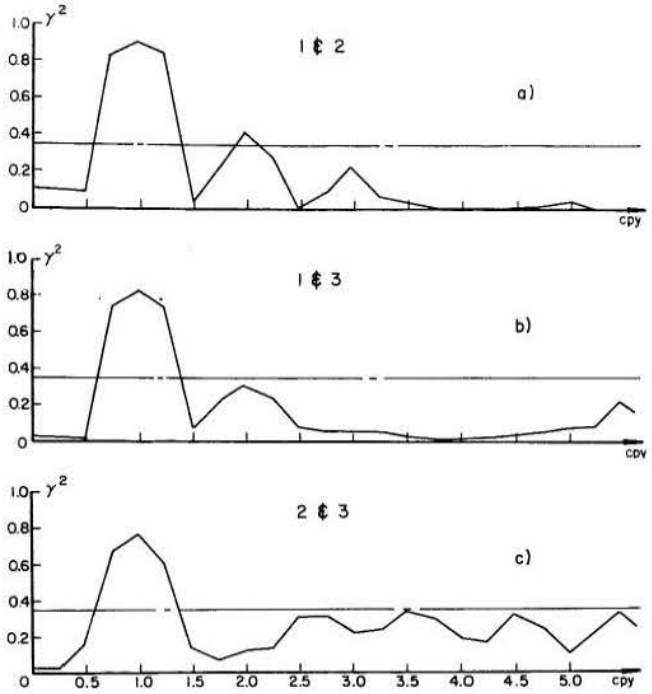

Fig. 6.3 Coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at San Francisco (California) with their corresponding 95% significance levels.

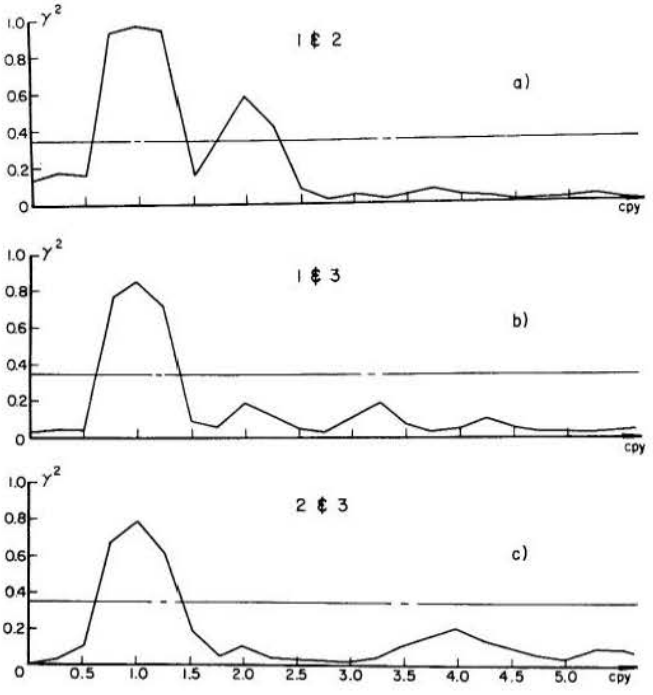

Fig. 6.5 Coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at San Diego (California) with their corresponding 95% significance levels.

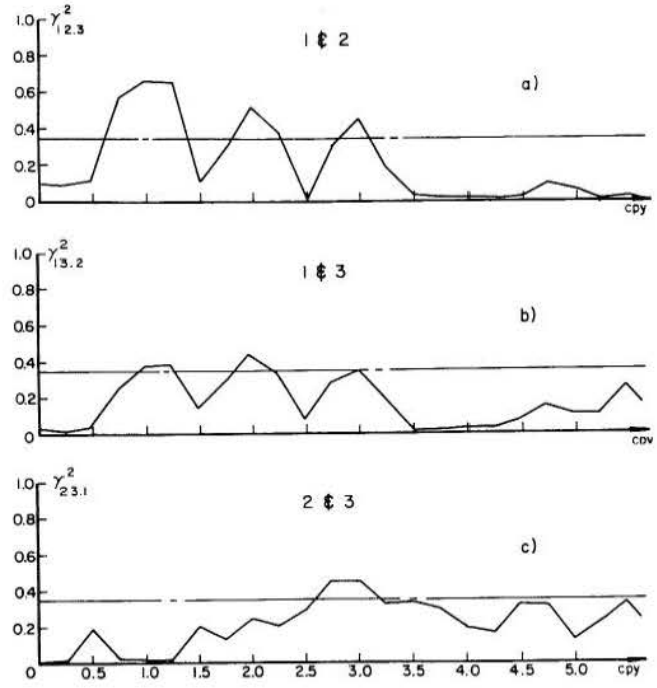

Fig. 6.4 Partial coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at San Francisco (California) with their corresponding 95% significance levels.

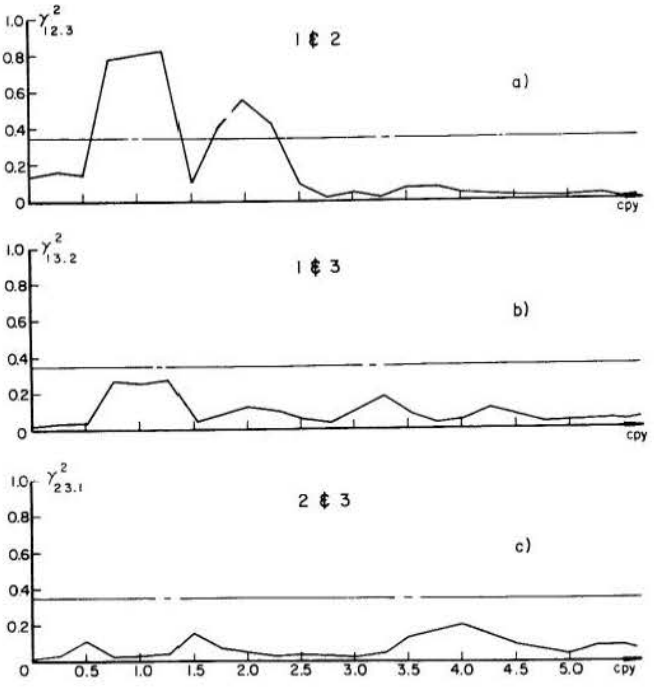

Fig. 6.6 Partial coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at San Diego (California) with their corresponding 95% significance levels.

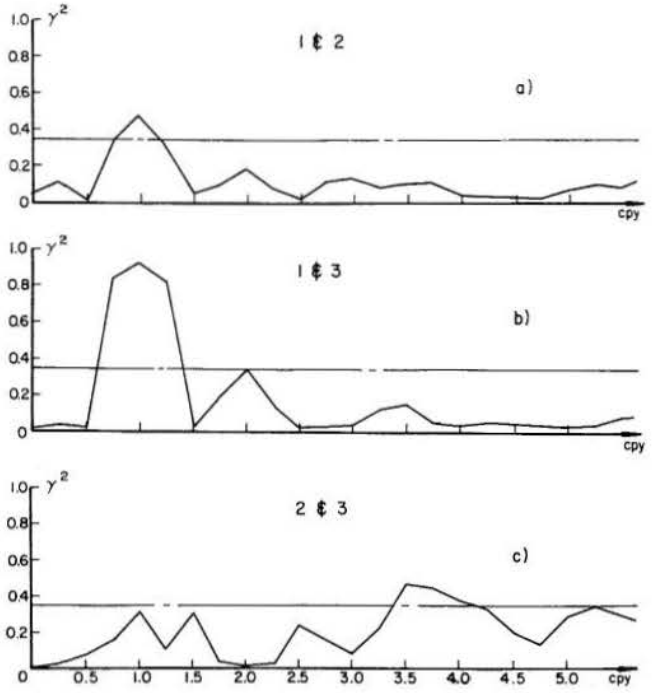

Fig. 6.7 Coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Portland (Oregon) with their corresponding 95% significance levels.

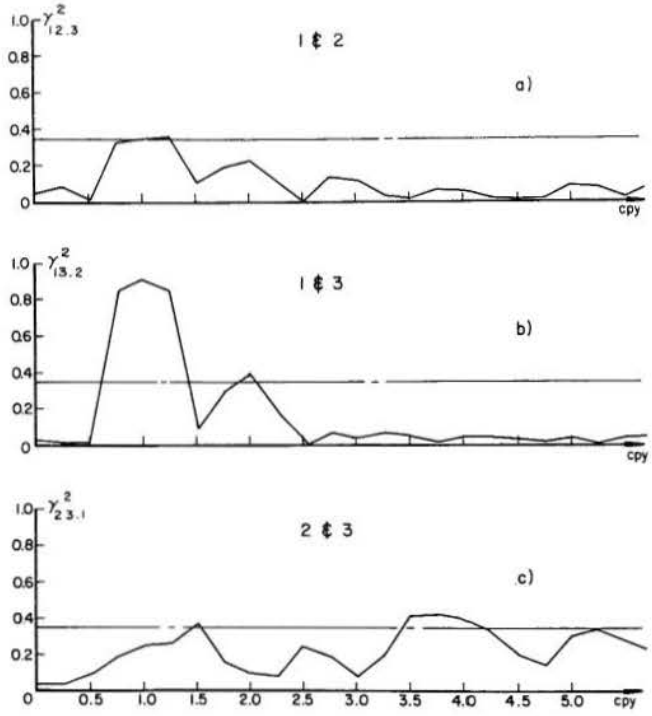

Fig. 6.8 Partial coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Portland (Oregon) with their corresponding 95% significance levels.

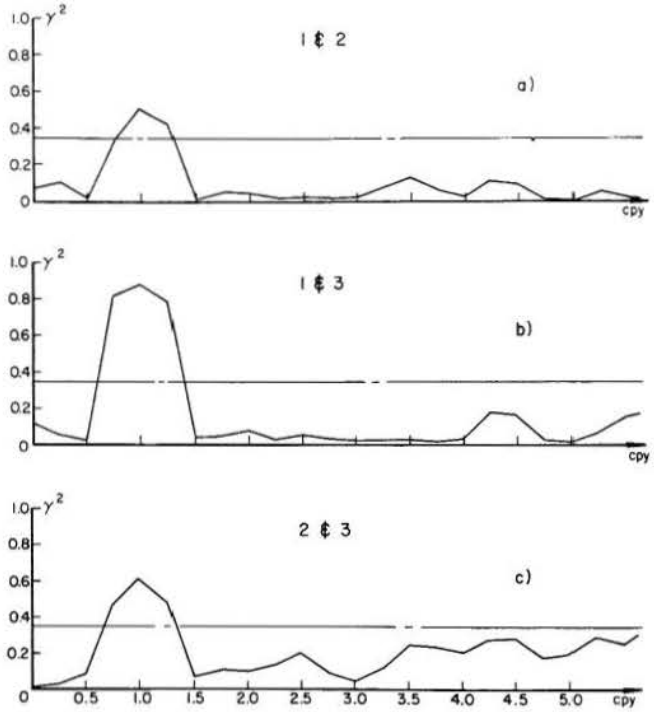

Fig. 6.9 Coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Tatoosh Island (Washington) with their corresponding 95% significance levels.

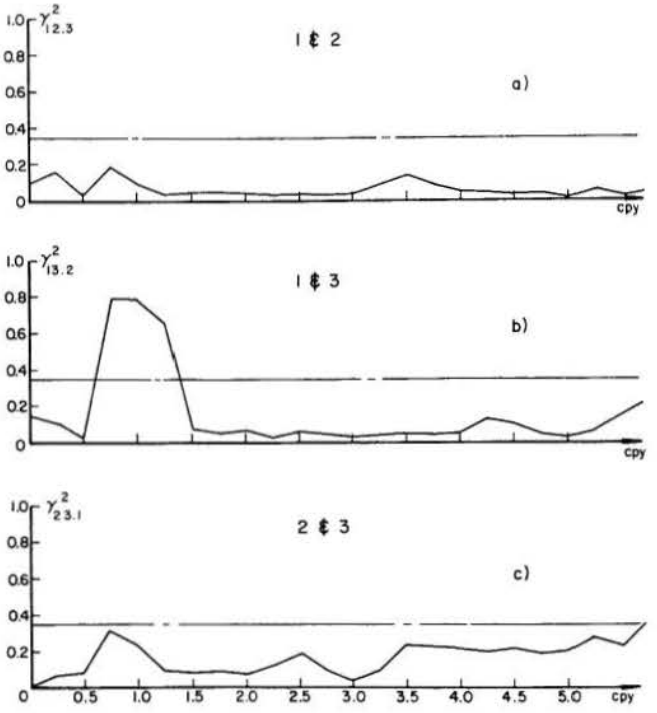

Fig. 6.10 Partial coherence functions for monthly data of temperature (series No. 1), atmospheric pressure (series No. 2) and precipitation (series No. 3) at Tatoosh Island with their corresponding 95% significance levels.

Except for Tatoosh Island, the coherence between the annual cycles of temperature and precipitation remains significantly different from zero when the effect of pressure is subtracted from the analysis. This indicates that for these stations, the annual cycle in pressure is related to the annual cycle in precipitation only through the annual cycle in temperature.

Figures 6.11, 6.12 and 6.13 show some typical phase diagrams obtained in the analysis of the series of temperature, atmospheric pressure and precipitation for the stations previously mentioned. When the phase is about 180 degrees in all the frequency range, it indicates that when a given frequency component of one series increases, the corresponding frequency component of the other series decreases and therefore the relationship between the variables is an inverse one. Both the phase and partial phase between temperature and atmospheric pressure seems to oscillate about 180°. This is also the case for precipitation and atmospheric pressure which should be expected from physical reasons when precipitation is caused by low pressure centers. The phase and partial phase between the series of temperature and precipitation exhibit much larger variations than for the previous series.

Tables 6.1 and 6.2 present the cross-spectral characteristics at the annual frequency for some of the stations considered in regions 1 and 2. Up to distances of 1000 Km in region 1, the gain is very

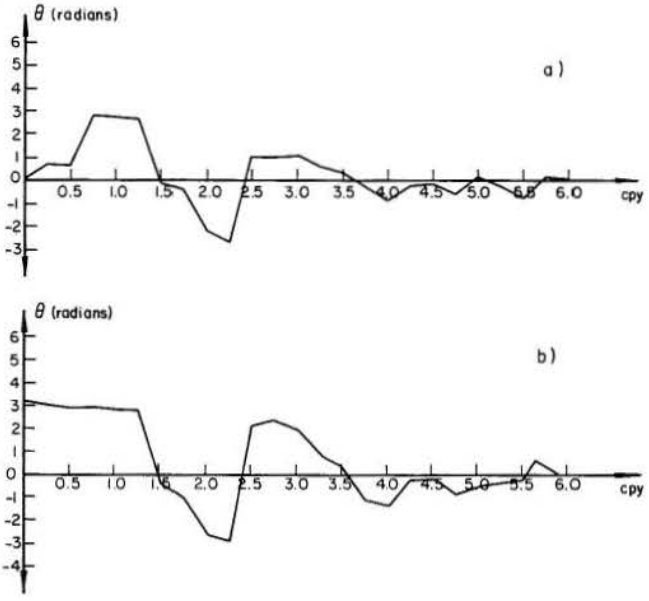

Fig. 6.12 a) Phase diagram for monthly data of temperature and precipitation at Eureka (California).

b) Partial phase diagram for the same series of a) when atmospheric pressure is subtracted from the analysis.

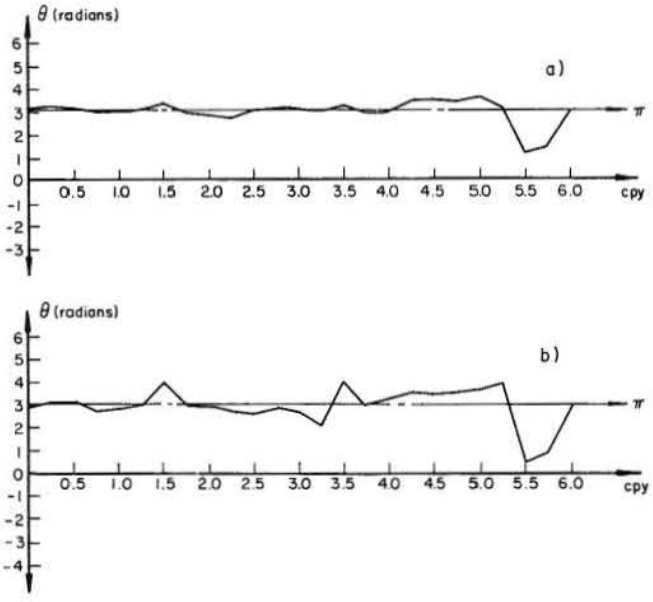

Fig. 6.11 a) Phase diagram for monthly data of temperature and atmospheric pressure at Eureka (California).

b) Partial phase diagram for the same series of a) when precipitation is subtracted from the analysis

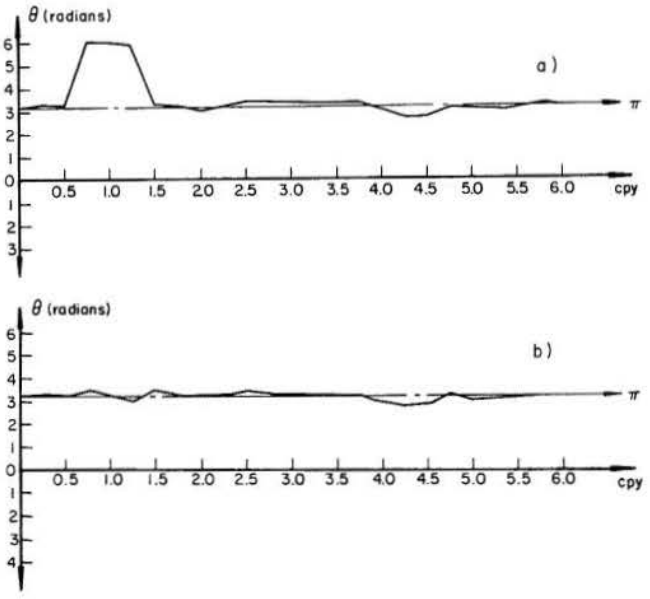

Fig. 6.13 a) Phase diagram for monthly data of atmospheric pressure and precipitation at Eureka (California).

b) Partial phase diagram for the same series of a) when temperature is subtracted from the analysis.

TABLE 6.1 CROSS-SPECTRAL CHARACTERISTICS AS FUNCTIONS OF DISTANCE FOR MONTHLY PRECIPITATION SERIES (REGION NO. 1)

Base station: San Diego WB APT.

Name	Station I. D.	Lat.	Long.	Years of records used	Distance from base station (Km.)	Coherence γ^2 (1 cpy)	95% significance level for γ^2	Gain $\|H(1\ cpy)\|$	Phase θ (1 cpy) (radians)	95% conficence bands for θ
Ojai	4.640	34.450	119.250	56	280 ↑	0.908	0.320	1.000	-0.081	±0.034
San Luis Obispo	4.785	35.300	120.667	91	445 ↑	0.904	0.220	0.907	-0.003	±0.017
Big Sur St. Park	4.079	36.250	121.783	46	581 ↑	0.898	0.355	0.847	-0.017	±0.034
Antioch F. Mills	4.023	38.017	121.767	81	756 ↑	0.867	0.241	0.884	0.056	±0.026
Fort Ross	4.319	38.517	123.250	85	861 ↑	0.805	0.230	0.820	0.074	±0.063
Fort Bragg	4.316	39.950	123.800	61	1120 ↑	0.859	0.300	0.782	0.105	±0.069
Cottage Grove	35.190	43.783	123.067	44	1480 ↑	0.800	0.365	0.755	0.235	±0.071
Longview	45.477	46.167	122.917	36	1624 ↑	0.747	0.401	0.770	0.337	±0.089
Tatoosh Island	45.833	48.383	124.733	77	1897 ↑	0.764	0.248	0.718	0.446	±0.103

↑(northward) ↓(southward)

TABLE 6.2 CROSS-SPECTRAL CHARACTERISTICS AS FUNCTIONS OF DISTANCE FOR MONTHLY PRECIPITATION SERIES (REGION NO. 2)

Base station: Sonora

Name	Station I. D.	Lat.	Long.	Years of records used	Distance from base station (Km.)	Coherence γ^2 (1 cpy)	95% significance level for γ^2	Gain $\|H(1\ cpy)\|$	Phase θ (1 cpy) (radians)	95% confidence bands for θ
Auburn	4.038	38.900	121.067	61	133 ↑	0.975	0.300	0.959	0.041	±0.000
Chester	4.170	40.300	121.217	50	287 ↑	0.949	0.341	0.952	0.098	±0.008
Wasco	4.945	35.600	119.333	61	287 ↓	0.890	0.300	1.016	-0.114	±0.010
McCloud	4.545	41.267	122.133	50	444 ↑	0.924	0.341	0.950	0.114	±0.008
Lakeview	35.467	42.183	120.350	48	490 ↑	0.728	0.345	1.112	-0.056	±0.103
Lytle Creek	4.522	34.200	117.450	55	518 ↓	0.930	0.360	1.082	0.000	±0.017
Estacada	35.269	45.267	122.317	52	854 ↑	0.868	0.331	0.899	0.268	±0.005

↑(northward) ↓(southward)

close to one and the phase practically zero if considered with its confidence bands. This indicates that in this range of distances the annual precipitation cycle can be considered the same for stations located in the U. S. Pacific Coast. In region 2, although the gain remains close to one, phase differences among the stations seem to be larger than in region 1. In both regions the coherence at the annual frequency is very high and always significantly different from zero.

In region 1 the amplitude of the annual oscillation decreases when advancing in the northward direction and at Tatoosh Island, 1900 Km away from the station base, the gain is 0.718 and the phase 0.446 radians. In region 2 the same trend is observed in all the stations with the exception of Lakeview.

The slow and progressive variations of the gain and phase make it easier to predict models like equation 3-5 for one station on the basis of another one. Due to the high percentage of variance explained in some regions by the annual cycle and its subharmonics (Roesner and Yevjevich, 1966), this type of regression by frequencies will be very useful in the prediction and simulation of hydrologic time series. One example of this will be given later in this chapter.

Figures 6.14, 6.15, 6.17, and 6.19 show some typical coherence and cross-correlation functions obtained for regions 1 and 2. For these regions, the coherence in the monthly analysis is very high regardless of frequency up to distances of 250 Km. This indicates that one series could be completely predicted on the basis of the other one because all the respective frequency components are very well correlated. For longer distances the coherence remains very high only at the annual frequency indicating that with the exception of the annual oscillation, the other frequency components can be considered as uncorrelated noise. The average coherence and cross-correlation functions for regions 1 and 2 are shown in Figures 6.16, 6.18 and 6.20 with their respective variances. These average functions were obtained by calculating the average value of the function over all stations in the region at each frequency (for the coherence function) or at each lag (for the cross-correlation function).

For the simulation and statistical prediction of hydrologic events at future points in time, covariance analysis might sometimes be more appropriate than spectral analysis since the time domain is the natural domain in which to operate. As an example, Figure 6.17b) shows the cross-correlation function between monthly precipitation at Tatoosh Island and monthly precipitation at San Diego. A difference in phase of about 2 months is apparent. On the other hand, the phase diagram between those stations gives a value of 0.446 radians at 1 cpy, equivalent to

$$12 \text{ months} \times \frac{0.446}{2\pi} = 0.852 \text{ months} = 26 \text{ days}$$

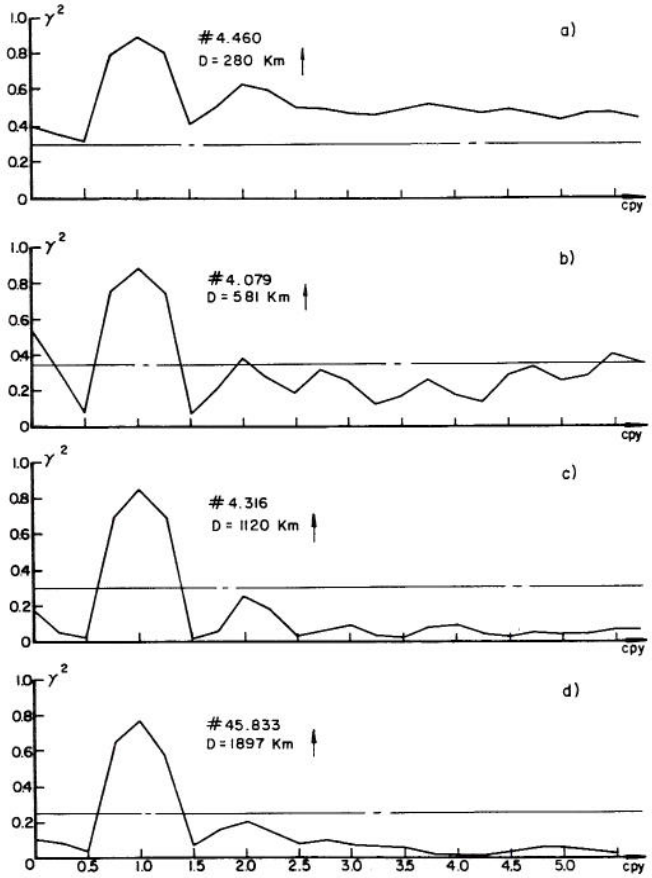

Fig. 6.14 Examples of coherence functions for monthly data of region 1 with their corresponding 95% significance levels. Base station #4.774.

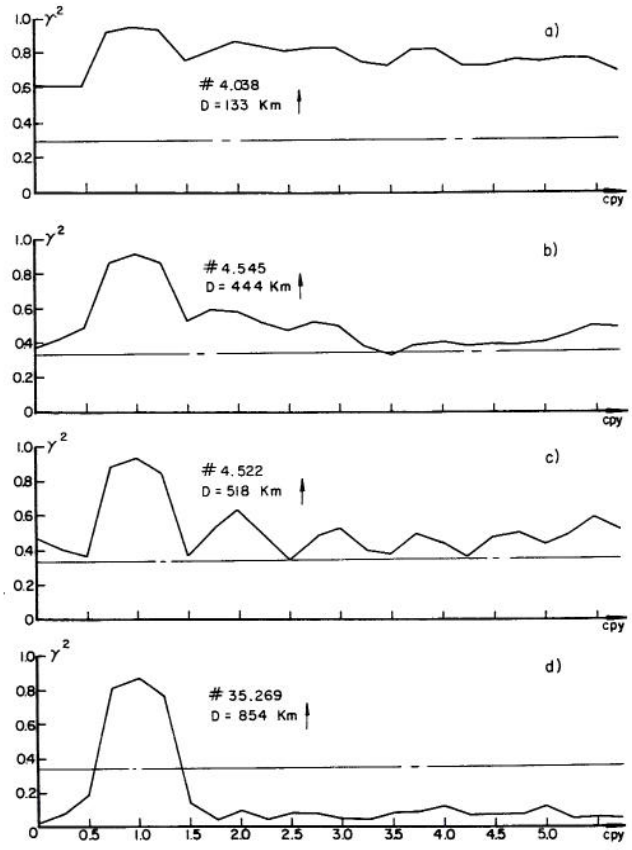

Fig. 6.15 Examples of coherence functions for monthly data of region 2 with their corresponding 95% significance levels. Base station #4.835.

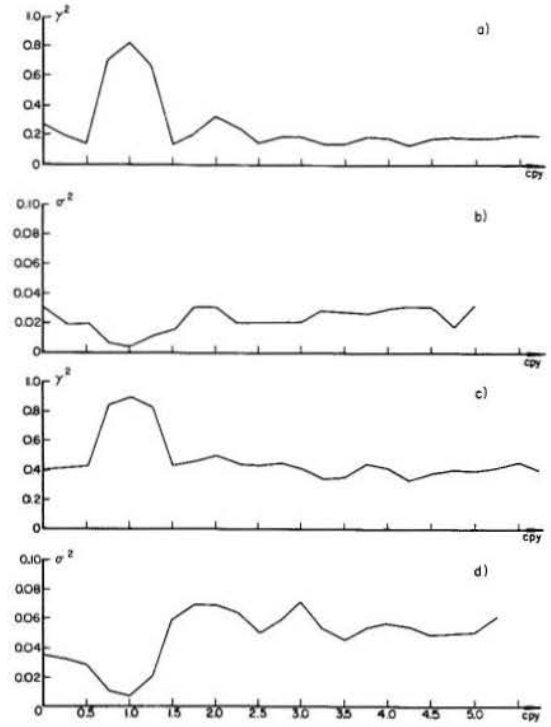

Fig. 6.16 a) Average coherence function for monthly data of region 1
b) Variance of a) c) Average coherence function for monthly data of region 2
d) Variance of c).

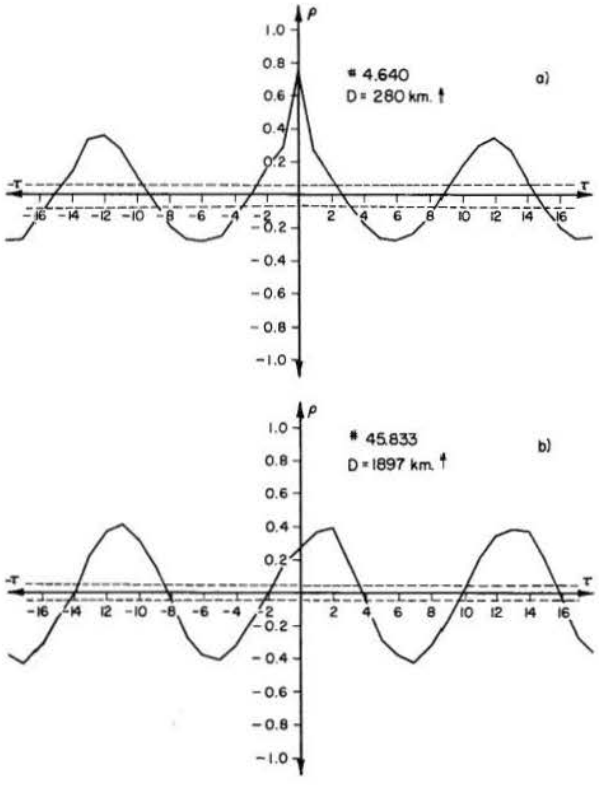

Fig. 6.17 Examples of cross-correlation functions for monthly data of region 1 with their corresponding 95% significance levels. Base station #4.774.

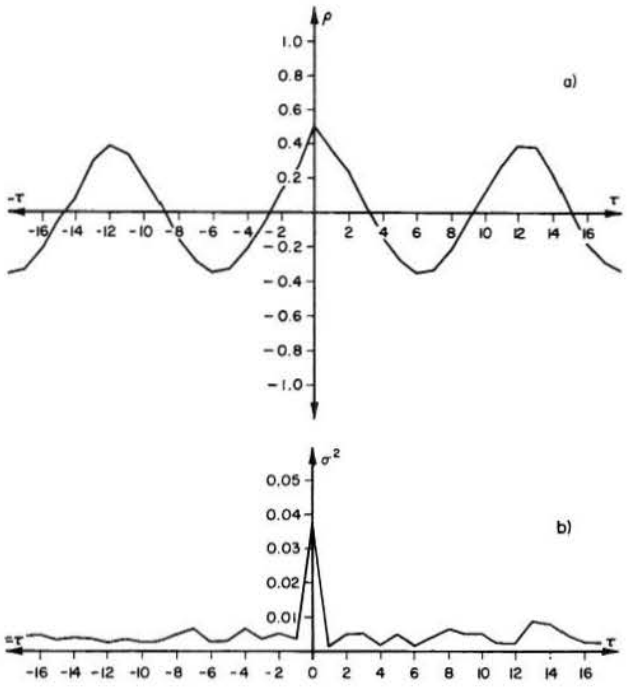

Fig. 6.18 a) Average cross-correlation function for monthly data of region 1
b) Variance of a).

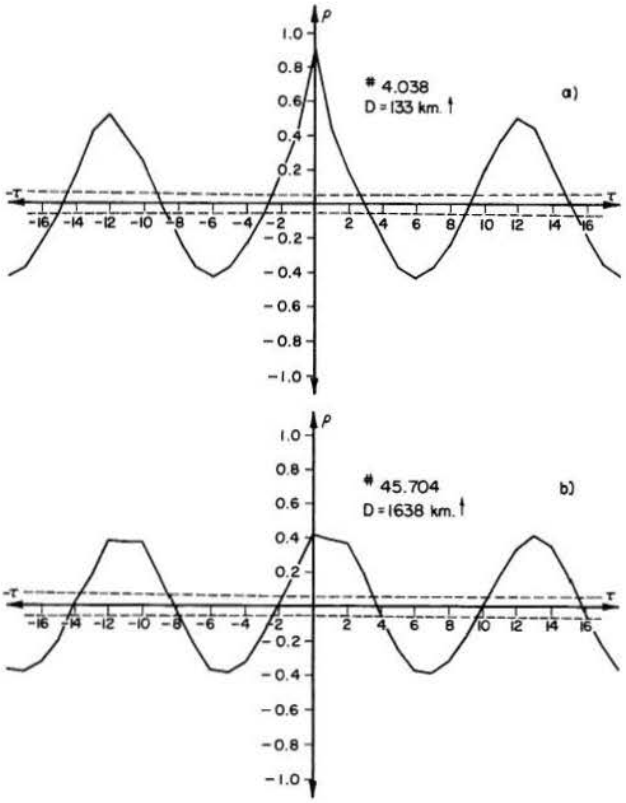

Fig. 6.19 Examples of cross-correlation functions for monthly data of region 2 with their corresponding 95% significance levels. Base station #4.835.

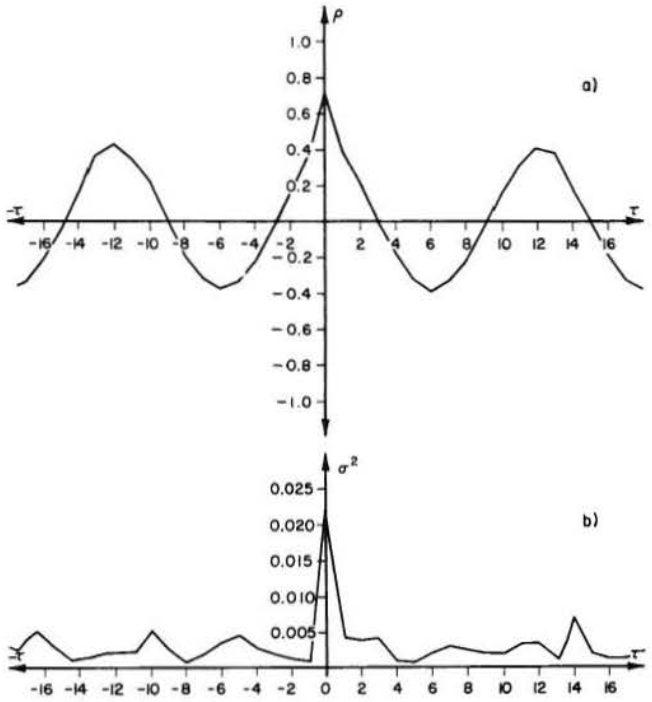

Fig. 6.20 a) Average cross-correlation function for monthly data of region 2
b) Variance of a).

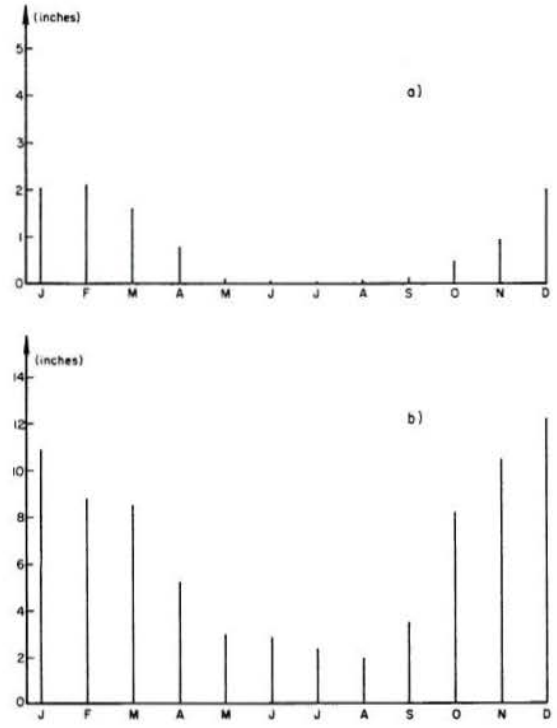

Fig. 6.21 a) Mean monthly precipitation at San Diego (California)
b) Mean monthly precipitation at Tatoosh Island (Washington)

The result from the phase diagram will give much more accurately the difference in phase between the annual cycles in both stations due to the fact that when working in the time domain we consider each process as a whole without any consideration of frequency components. On the other hand, there are many cases in which the engineer prefers to consider each process as a whole and to try the prediction in time of the general relations of rainfall at both stations. Then covariance analysis seems the most appropriate one. The series of rainfall at Tatoosh Island and San Diego are shown in Figure 6.21. The peak in precipitation at San Diego occurs two months later than the peak at Tatoosh Island, as result which is in agreement with the value obtained from the cross-correlation function.

Some examples of the gain and phase functions obtained for regions 1 and 2 are shown in Figures 6.22, 6.23, and 6.24.

In general, for the precipitation stations located in regions 3, 4 and 5, no significant peaks were detected in the spectra. For these stations the coherence was low over the frequency range and only when the stations were very close was the coherence significantly different from zero although without peaks. Figures 6.25 and 6.26 show typical coherence and cross-correlation functions for regions 3, 4 and 5.

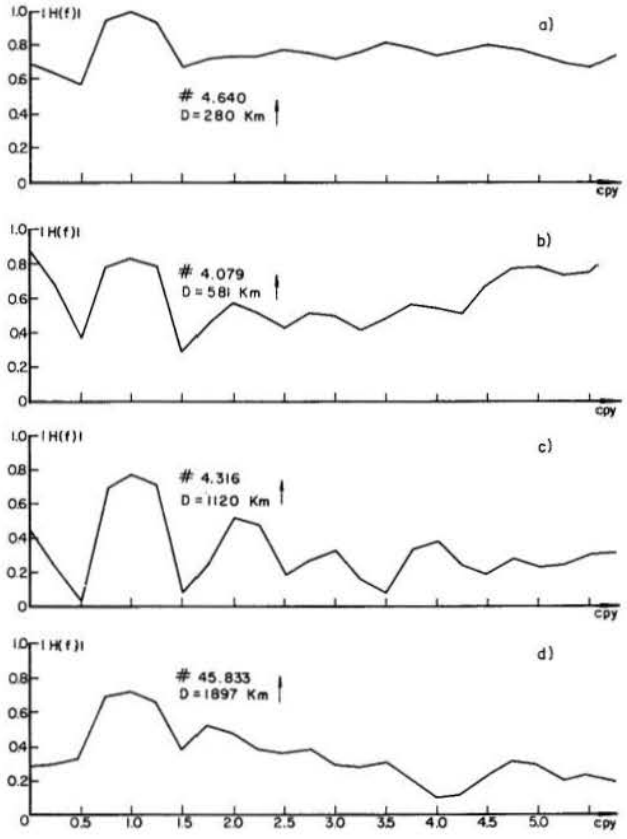

Fig. 6.22 Examples of gain functions for monthly data of region 1 Base station #4.774

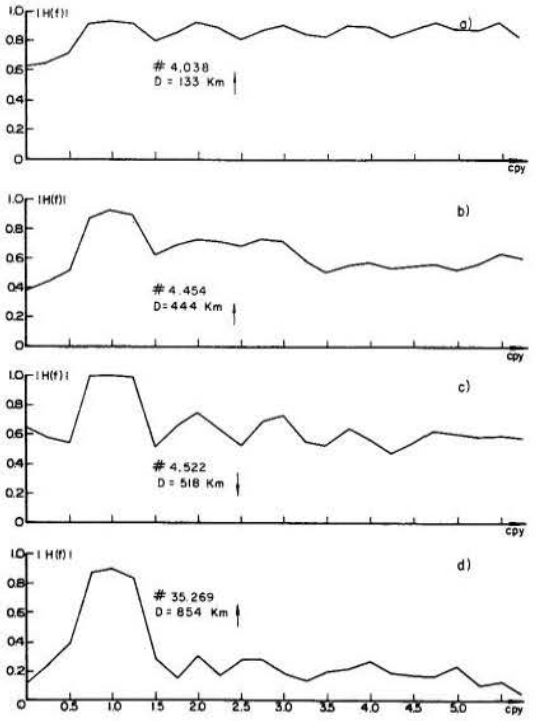

Fig. 6.23 Examples of gain functions for monthly data of region 2 Base station #4.835

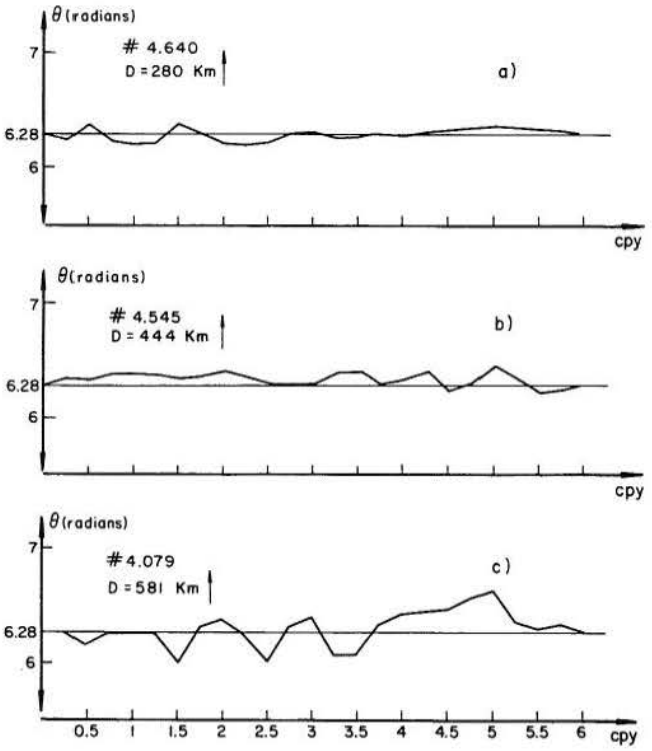

Fig. 6.24 a) and b) Examples of phase functions for monthly data of region 1. Base station #4.774. c) Example of phase function for monthly data of region 2. Base station #4.835

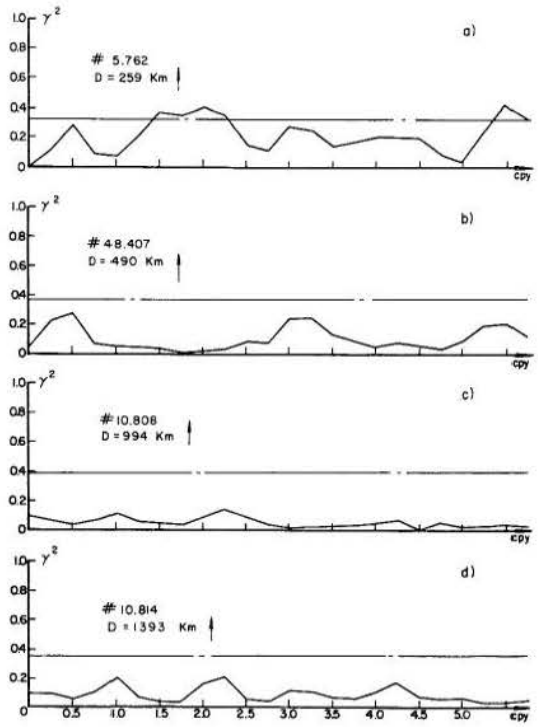

Fig. 6.25 Examples of coherence functions for monthly data of region 3 with their corresponding 95% significance levels. Base station #5.243

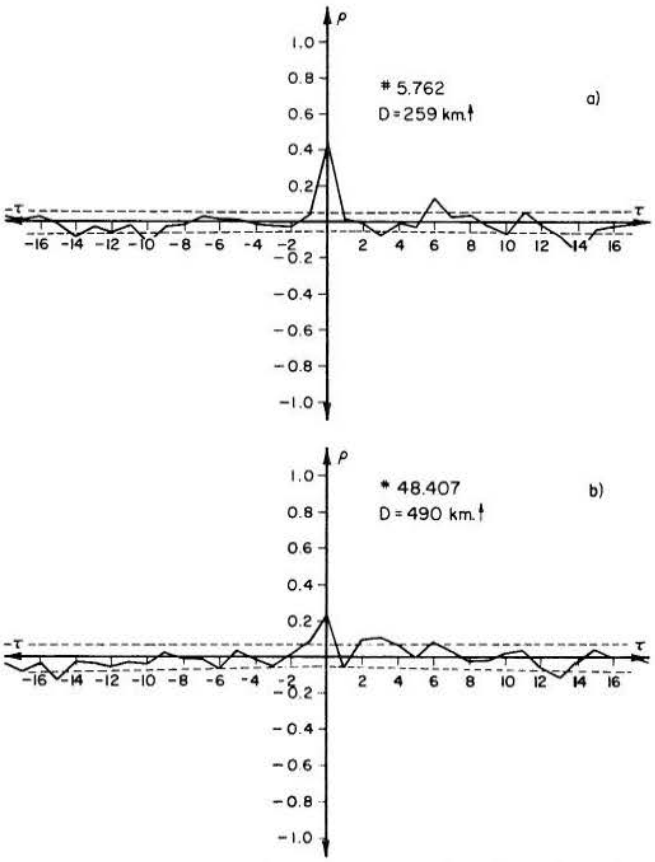

Fig. 6.26 Examples of cross-correlation functions for monthly data of region 3 with their corresponding 95% significance levels. Base station #5.243

2. Analysis of Monthly Runoff Data. Some examples of coherence and cross-correlation functions found in monthly runoff data are shown in Figures 6.27 and 6.28. The average values of these functions and their respective variances are shown in Figures 6.29 and 6.30. Table 6.3 presents the cross-spectral characteristics at the annual frequency for some of the stations considered in the analysis. Contrary to the case of monthly precipitation, there is no definite trend in the amplitude of the annual runoff cycle, a fact that is reflected in the lack of a trend in the gain function. The explanation of this lies in the influence of watershed characteristics on the properties of the runoff time series and also in the influence of the evaporation and storage which vary from region to region.

The coherence at the annual cycle always has a highly significant meaning, and there is a strong correlation between the annual oscillation of these stations. In many cases, the semi-annual cycles were also found to be strongly correlated. For distances up to 370 Km the coherence was significantly different from zero over the frequency range, meaning that one of the series could be completely related to the other one because of the interdependence of all the correspondent frequency components.

Some of the phase and gain diagrams obtained in the study are shown in Figures 6.31 and 6.32. From the phase diagrams it is observed that stations south of the base station have runoff series that lag behind the base series. On the other hand, stations located north of the base station have series whose frequency components precede those of the base station.

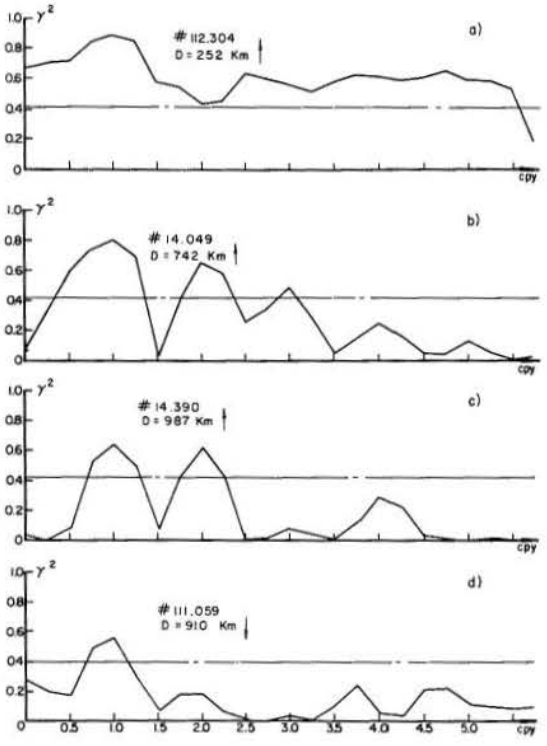

Fig. 6.27 Examples of coherence function for monthly runoff series with their corresponding 95% significance levels. Base station #112.402

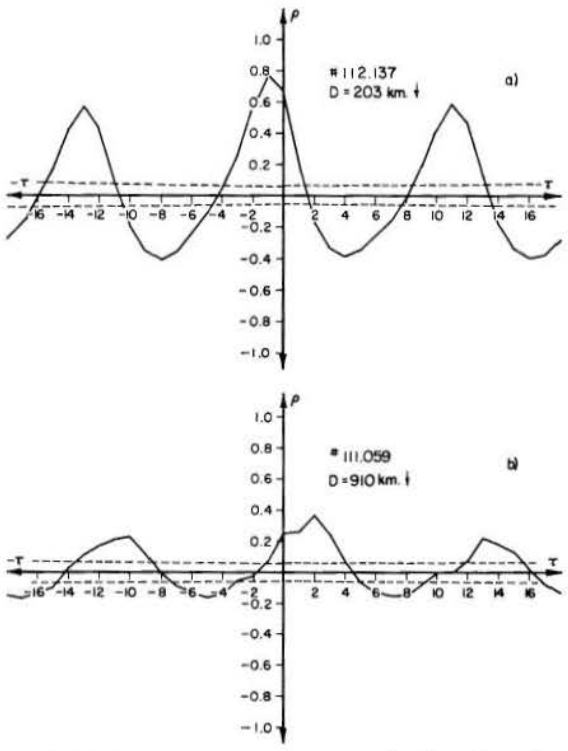

Fig. 6.28 Examples of cross-correlation functions for monthly runoff series with their corresponding 95% significance levels. Base station #112.402

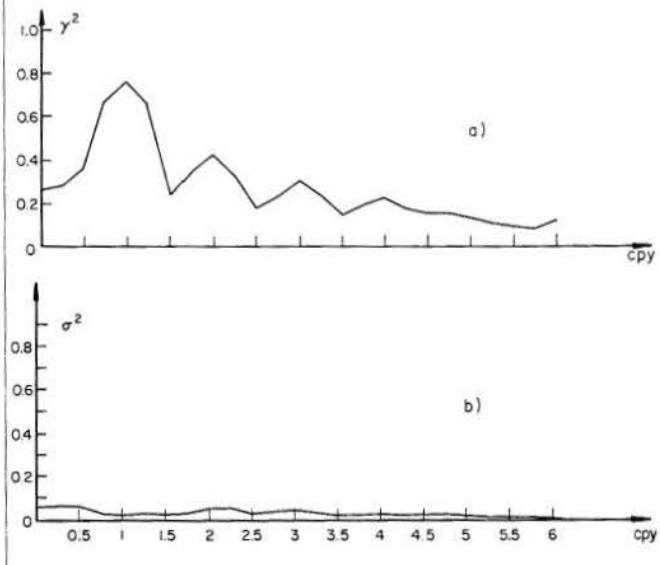

Fig. 6.29 a) Average coherence function for monthly runoff series. Base station #112.402
b) Variance of a)

TABLE 6.3 CROSS-SPECTRAL CHARACTERISTICS AS FUNCTIONS OF DISTANCE FOR MONTHLY RUNOFF SERIES

Base station: Middle Fork American River at Auburn (Calif.)

Name	Station I.D.	Lat.	Long.	Years of records used	Distance from base station (Km.)	Coherence γ^2(1 cpy)	95% significance level for γ^2	Gain $\lvert H(1\ \text{cpy})\rvert$	Phase θ(1 cpy) (radians)	95% confidence bands for θ
Thomes Creek	112.308	39.88	122.55	40	161 ↑	0.887	0.381	0.994	0.593	±0.035
Falls Creek	112.137	37.97	119.77	45	203 ↓	0.942	0.360	0.989	-0.684	±0.017
Salmon River	111.393	41.38	123.47	37	350 ↑	0.914	0.397	0.945	0.296	±0.035
Kern River	112.001	35.93	118.48	48	427 ↓	0.906	0.345	0.986	-0.772	±0.035
McKenzie River	14.278	44.13	122.47	36	616 ↑	0.827	0.400	0.959	0.676	±0.105
Toutle River	14.419	46.33	122.73	31	861 ↑	0.744	0.436	0.891	0.967	±0.157
Murrieta Creek	111.059	33.48	117.15	36	910 ↓	0.562	0.400	1.383	0.695	±0.208
Hoh River	12.050	47.80	124.10	34	1057 ↑	0.624	0.415	1.021	1.401	±0.192

↑(northward) ↓(southward)

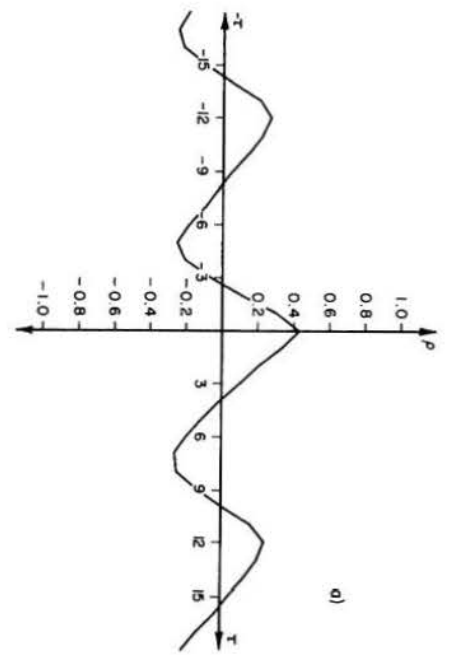

Fig. 6.30 a) Average cross-correlation function for monthly runoff series. Base station #112.402
b) Variance of a)

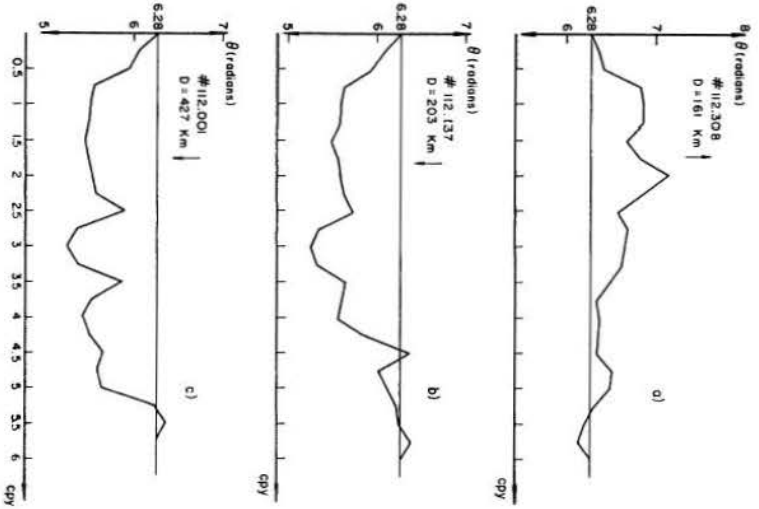

Fig. 6.31 Examples of phase functions for monthly runoff series. Base station #112.402

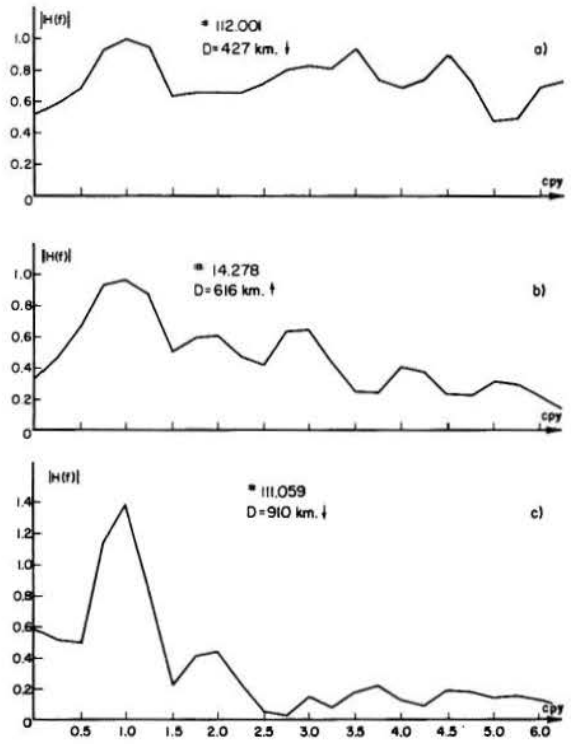

Fig. 6.32 Examples of gain functions for monthly runoff series. Base station #112.402

3. Joint Analysis of Monthly Rainfall and Monthly Runoff in a Watershed. Two cases were considered here:

a) Single input system: when all the precipitation was considered coming from one station.

b) Multiple input system: when precipitation at different stations was considered as different inputs.

Single input system. 35 years of monthly precipitation data at Auburn with the corresponding discharges of the Middle Fork American River at Auburn were used. The analysis was performed with a maximum of 36 lags. Figure 6.33a) shows the coherence function between both series. It indicates a strong correlation between the annual cycles and an almost null correlation between the semi-annual cycles. For frequencies larger than 1 cpy, the coherence function is not significantly different from zero and its shape indicates that in a monthly basis the rainfall-runoff process is far from being linear.

The phase diagram is shown in Figure 6.33b), it seems to oscillate around a value of 0.652 radians, or in other words, the precipitation series appears to be, as a whole, 1.24 months ahead of the runoff series. The sign test rejects at the 95% probability level, the hypothesis that the phase function is oscillating around zero.

Figure 6.34 shows the spectra of both series together with the spectrum of the residual terms as defined in Chapter III. From the spectra of the residuals it is seen that the annual cycle of one series could be completely explained by means of a cyclical regression with the other series. This kind of regression was carried out with the help of the phase diagram and the gain function, the later one shown in Figure 6.35. A straightforward harmonic

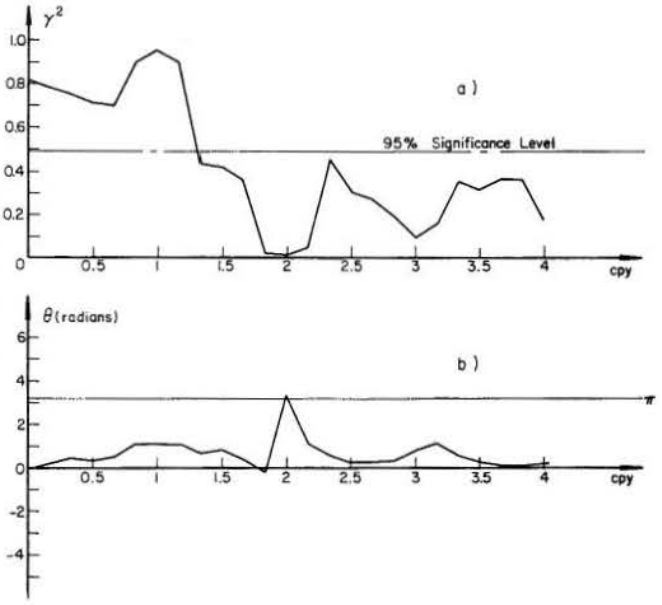

Fig. 6.33 a) Coherence function between monthly precipitation at Auburn and monthly runoff of the Middle Fork American River at Auburn with 95% significance level .
b) Phase function between the same series of part a).

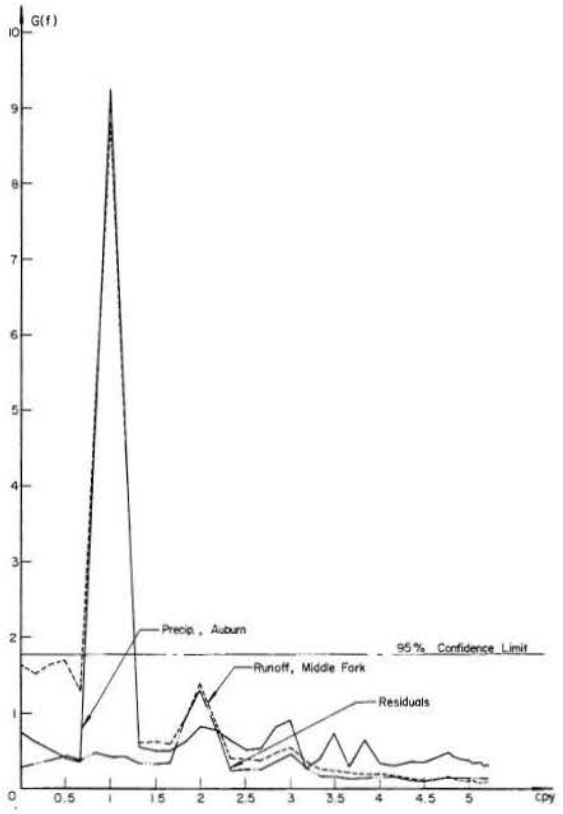

Fig. 6.34 Comparison of spectra obtained in the joint analysis of monthly precipitation at Auburn and monthly runoff of the Middle Fork American River at Auburn.

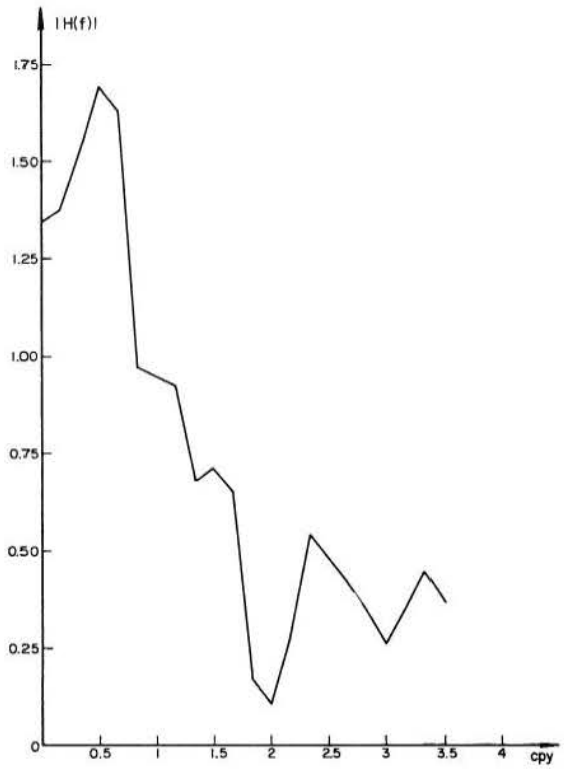

Fig. 6.35 Gain function of monthly runoff of the Middle Fork American River at Auburn based on monthly precipitation at Auburn.

analysis gave for the annual cycle of the Middle Fork American River the expression:

$$y(t) = 0.947 \cos\left(\frac{2\pi}{12} t + 0.354\right) \qquad (6\text{-}4)$$

Similarly, for the precipitation at Auburn, it was obtained:

$$x(t) = 0.969 \cos\left(\frac{2\pi}{12} t - 0.820\right) \qquad (6\text{-}5)$$

The gain and phase at 1 cpy were 0.951 and 1.145 radians, respectively. Using equation 3-5 the predicted annual cycle of the Middle Fork becomes:

$$y_*(t) = 0.969 \times 0.951 \cos\left(\frac{2\pi}{12} t - 0.820 + 1.145\right) \qquad (6\text{-}6)$$

$$= 0.922 \cos\left(\frac{2\pi}{12} t + 0.325\right) \qquad (6\text{-}7)$$

Figure 6.36 shows a comparison of equations 6-4 and 6-7. This type of regression holds theoretically between any corresponding frequency components but it is most useful for those components which explain large percentage of the variance of the series.

Multiple inputs system. The partial cyclical regressions intended here were not successful, nevertheless valuable experience was obtained.

Three watersheds were studied here in which precipitation at different places of the watershed were considered as different inputs producing as sole output the runoff at the outlet.

The practical difficulty in obtaining reasonable results arise from the form of the equation for the partial transfer function (equation 3-23).

$$H_1(f) = \frac{S_{1y\cdot 23\ldots N}(f)}{S_{11.23\ldots N}(f)} \qquad (6\text{-}8)$$

where the partial cross spectrum $S_{11.23\ldots N}(f)$ has for the case of two inputs the form (equation 2-53):

$$S_{11\cdot 2}(f) = S_{11}(f)\left[1 - \gamma^2_{12}(f)\right] \qquad (6\text{-}9)$$

For the multiple inputs case $S_{11.23\ldots N}(f)$ has expressions similar to equation 6-9 and it is observed that we need inputs not very well correlated, $\gamma^2_{12}(f) < 1$, in order to obtain meaningful results for $H_1(f)$. This has proved to be a most difficult

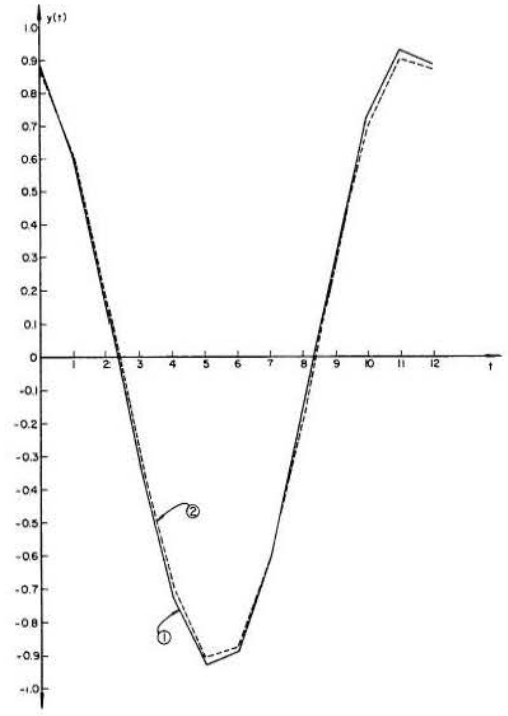

Fig. 6.36 1) Actual annual cycle of Middle Fork American River at Auburn. (Eq. 6-4)
2) Annual cycle of Middle Fork predicted in basis of precipitation at Auburn. (Eq. 6-7)

case for monthly rainfall data coming from stations in the same watershed although it is probable the theory could be applied successfully in other geophysical problems. The fact that $\gamma^2(f)$ is close to one for monthly rainfall data coming from stations in the same watershed shows that one station is representative of the monthly rainfall regime in the watershed and therefore it is unnecessary to analyze the system as one with multiple inputs.

4. Analysis of Annual Precipitation Data. The analysis of annual precipitation data did not show any significant peak in the spectra. The coherence diagrams vary drastically from station to station in all the regions studied and a common peak or feature in them does not appear to exist.

Some examples of the coherence and cross-correlation functions found are shown in Figures 6.37 and 6.38 with the corresponding 95% significance levels.

From the spectra it is concluded these series can be considered as "white noise" or random independent data. When two stations were at short distance one from the other, it was common to find noise correlation at one or several frequencies (Figure 6.37a). When the distance between the stations increases the coherence function is not significantly different from zero in all the frequency range (Figures 6.37b and 6.37c). For nearby stations located in the same environment, it is common to find significant zero-lag correlation (Figure 6.38a) meaning there is a strong relation in the rainfall at the same year in both stations. This significant zero-lag correlation completely disappears when the distance between the stations increases.

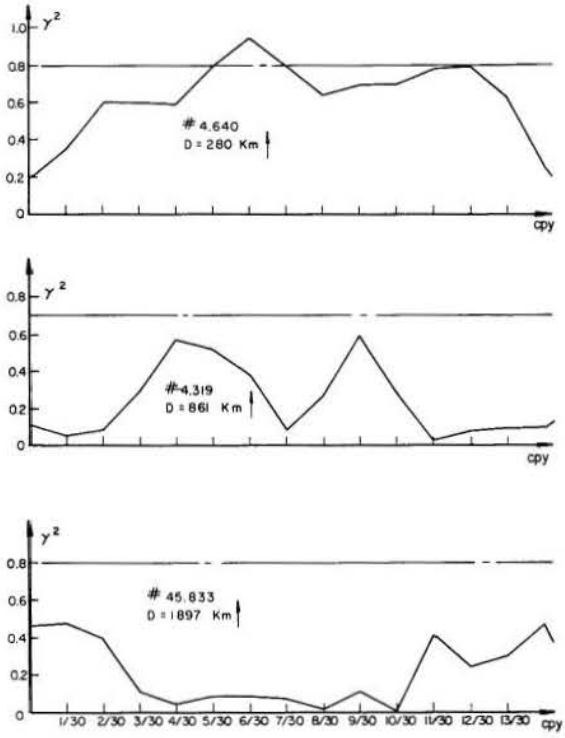

Fig. 6.37 Examples of coherence functions for annual data of region 1 with their corresponding 95% significance levels. Base station #4.774.

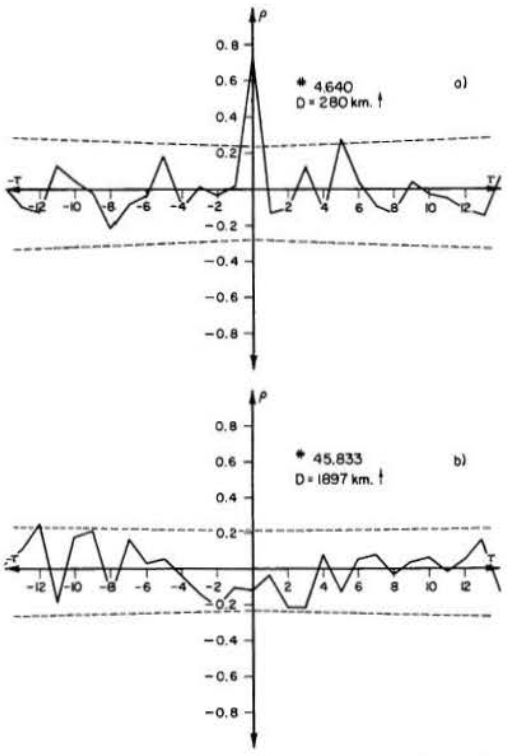

Fig. 6.38 Examples of cross-correlation functions for annual data of region 1 with their corresponding 95% significance levels. Base station #4.774.

5. The Application of Cross-Spectral Analysis in the Study of Hydrologic Stochastic Processes.

First order Markov linear process. Cross-spectral methods are used here to test the assumption of a first order Markov linear generating process in two runoff time series. Use has been made of the theory developed in Chapter IV.

As an example, a complete analysis was performed for annual standardized discharge data of the Wolf River at New London (Wisconsin) and annual standardized discharge data of the Fox River at Berlin (Wisconsin). The data used was from 1898 to 1957 and the analysis was done wtih 12 lags. The reason for using annual data in this example is that Yevjevich (1964) has shown that 1st order autoregressive schemes may in many instances fit sufficiently well the patterns in the sequence of annual river flows.

Figures 6.39 and 6.40 present the estimated power spectra and also the theoretical ones obtained from equations 4-31 and 4-32. The estimated coherence function and the constant value obtained from equation 4-33 are shown in Figure 6.41. Considering that all the data is obtained from a limited sample, the fit can be considered a good one.

Second order Markov linear process. Use has been made of the theory developed in Chapter IV.

As an example we have used two of the rivers analyzed by Quimpo (1967): Boise River near Twin Springs (Idaho) and St. Maries River near Lotus (Idaho).

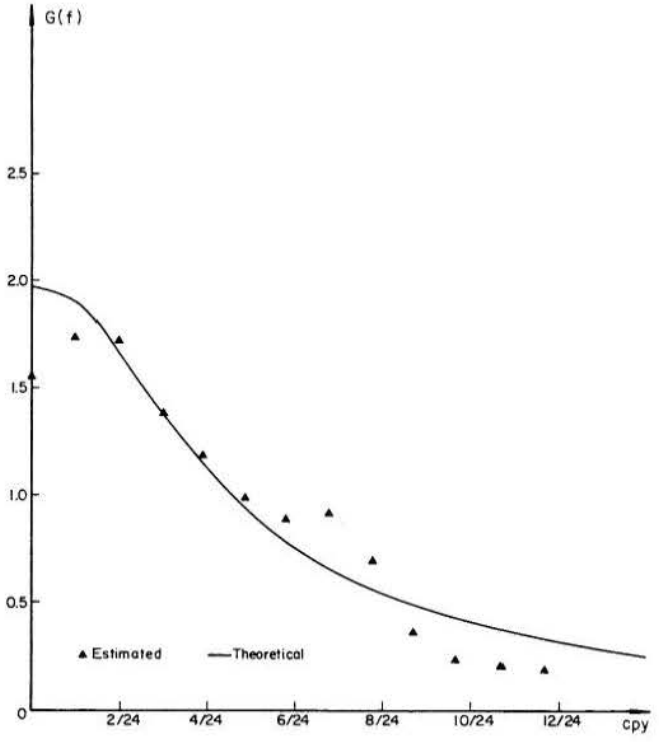

Fig. 6.39 Power spectrum of the Wolf River's annual standardized flows.

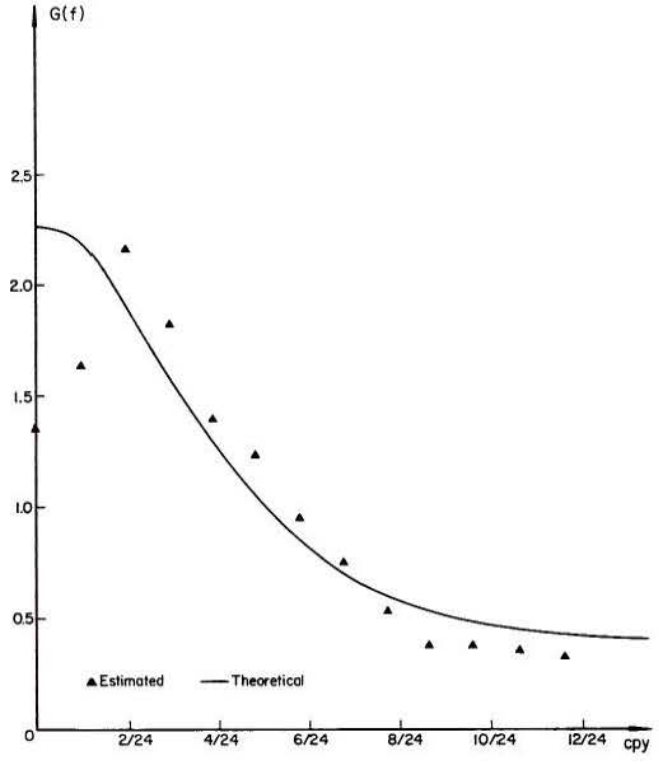

Fig. 6.40 Power spectrum of the Fox River's annual standardized flows.

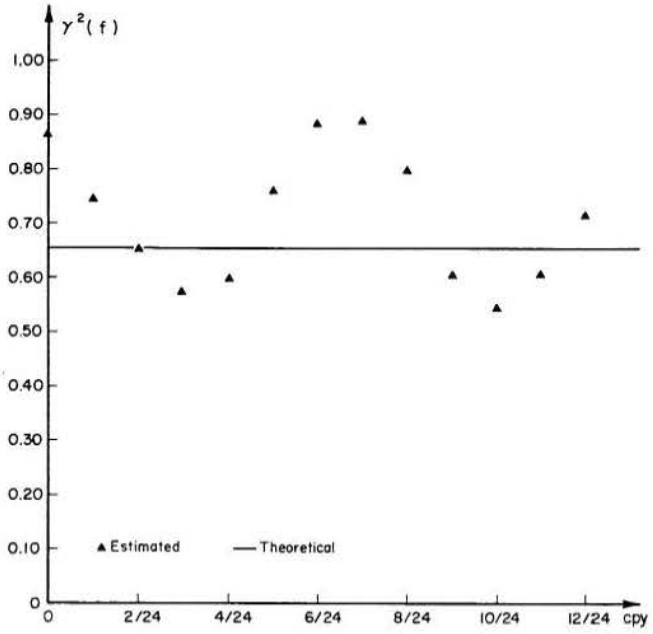

Fig. 6.41 Coherence function between the annual flows of the Wolf and the Fox Rivers.

Figure 6.42 (after Quimpo, 1967) shows the spectral density of the original series of daily river flows of Boise River and the spectral density of the residuals obtained after removal of the annual cycle and its first three sub-harmonics. In the analysis performed in this example the period of record used was from 1950 to 1960 and the residuals were previously standardized. There were 600 lags throughout the analysis.

The covariance matrices were in this case:

$$\left[\alpha_{ij}(1)\right] = \begin{bmatrix} 0.901 & 0.154 & 0.760 & 0.132 \\ 0.169 & 0.109 & 0.148 & 0.074 \\ 1.000 & 0.174 & 0.901 & 0.154 \\ 0.174 & 1.000 & 0.169 & 0.109 \end{bmatrix}$$

and

$$\left[\alpha_{ij}(0)\right] = \begin{bmatrix} 1.000 & 0.174 & 0.901 & 0.154 \\ 0.174 & 1.000 & 0.169 & 0.109 \\ 0.901 & 0.169 & 1.000 & 0.174 \\ 0.154 & 0.109 & 0.174 & 1.000 \end{bmatrix}$$

The matrix of coefficients became equal to:

$$[u_{ij}] = \begin{bmatrix} 1.166 & -0.006 & -0.286 & 0.006 \\ 0.183 & 0.086 & -0.034 & 0.040 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

In this manner, the following two equations were obtained for the standardized stochastic components of the above rivers:

$$x_1(t) = 1.166x_1(t-1) - 0.006x_2(t-1) - 0.286x_1(t-2) + 0.006x_2(t-2) + z_1(t)$$

$$x_2(t) = 0.183x_1(t-1) + 0.086x_2(t-1) - 0.034x_1(t-2) + 0.040x_2(t-2) + z_2(t)$$

where the subscript 1 refers to the Boise River and the subscript 2 refers to the St. Maries River.

Equation 6-58 gave for the coherence function the expression:

$$\gamma^2(f) = \frac{-0.016 \cos 2\pi f - 0.008 \cos 4\pi f + 0.004 \cos 6\pi f + 0.0004 \cos 8\pi f + 0.036}{\left[1.010 - 0.166 \cos 2\pi f - 0.080 \cos 4\pi f\right]\left[2.476 - 3.010 \cos 2\pi f + 0.572 \cos 4\pi f\right]}$$

Figure 6.43 shows a comparison of the obtained and theoretical coherence.

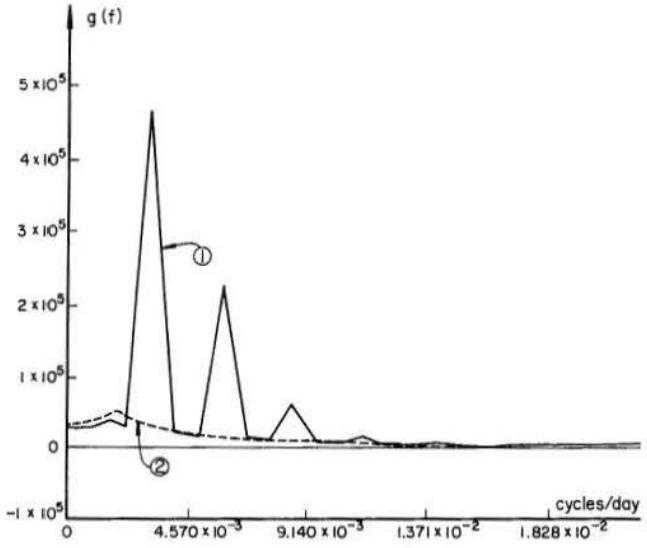

Fig. 6.42 Boise River spectral densities: (1) daily flows, (2) stochastic components of daily flows (after Quimpo, 1967).

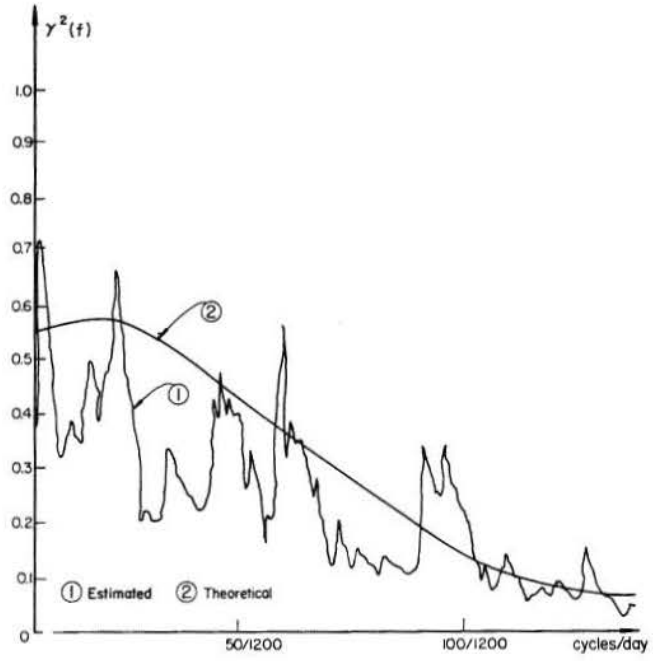

Fig. 6.43 Coherence function between stochastic components of daily flows of the Boise and St. Maries Rivers.

CHAPTER VII

CONCLUSIONS

The investigations performed here gave rise to the following conclusions:

1) In a specific application it seems there there are two main considerations which should enter into the choice between cross-correlation analysis and cross spectrum:

a) the use to be made of the estimated quantities,

b) ease of physical interpretation.

If the ultimate objective of the analysis is the construction of models like equation 3-5 then cross-spectrum certainly provides the answer to the problem. Covariance analysis could be used for frequency response studies, but the calculations would be considerably more difficult. Also as indicated in Section 3-1 cross-spectrum has a direct physical interpretation in this case.

Since prediction is done in time, it is natural to work in the time domain. For the simpler linear prediction models, since the parameters are estimated by functions of the auto- and cross- correlation, it is natural to work with these quantities.

2) A highly significant coherence between the annual cycles of hydrologic time series has been observed, making it possible to predict the annual oscillation in one station on the basis of the annual oscillation at another station.

3) For stations located in the Pacific Coast of the United States, the amplitude of the annual cycle in precipitation appears to decrease when advancing in the northerly direction. Up to distances of 1000 Km the annual cycle can be considered practically the same for all stations in the region. On the other hand, no similar trend was observed in the runoff stations.

4) No significant coherences were observed in the annual series of precipitation and runoff except in very close stations for which the coherence was high in all the frequency range.

5) The coherence function may be used as a measure of linearity of the rainfall-runoff process. In the monthly basis this process appears to be highly nonlinear.

6) By partial cross-spectral analysis it was found that in the Pacific Coast of the United States the annual cycle in temperature is highly correlated with the annual cycle in precipitation, but the annual cycle in atmospheric pressure seems to be related to the cycle in precipitation only throughout the cycle in temperature.

7) Coherence analysis stands as a powerful tool to analyze and test generating processes commonly used in hydrology. Special mention should be given to the coherence function between two 1st order autoregressive processes which was proven to be equal to a constant independent of frequency.

8) The use of prefiltering or smoothing before the series have been analyzed was shown to change the cross-spectrum between the series but the coherence remains the same if the linear filters are used. The phase function will be altered unless even filters are used.

BIBLIOGRAPHY

1. Bendat, J. S., 1958, Principles and applications of random noise theory. Wiley and Sons, New York.

2. Bendat, J. S. and Piersol, A. G., 1966, Measurement and analysis of random data. Wiley and Sons, New York.

3. Blackman, R. B. and Tukey, J. W., 1958, The measurement of power spectra from the point of view of communications engineering. Dover Publications Inc., New York.

4. Bryson. R. A. and Dutton, J. A., 1961, Some aspects of the variance spectra of tree rings and varves. Annals of the New York Academy of Sciences, Vol. 95, Art. 1, pp. 580-604.

5. Coleman, T., Press, H. and Meadows, M., 1958, An evaluation of flexibility on wing strains in rough air for a large swept-wing airplane by means of experimentally determined frequency-response functions. NACA TN-4291.

6. Cox, D. R. and Miller, H. D., 1965, The theory of stochastic processes. Wiley and Sons, New York.

7. Eagleson, P. S. and Shack, W. J., 1966, Some criteria for the measurement of rainfall and runoff. Water Resources Research, Vol. 2, No. 3, pp. 427-436.

8. Enochson, L. D., 1964, Frequency response functions and coherence functions for multiple input linear systems. NASA CR-32.

9. Goodman, N. R., 1957, On the joint estimation of the spectra, cospectrum and quadrature spectrum of a two-dimensional stationary gaussian process. Scientific Paper No. 10, Engineering Statistics Laboratory, New York University, (also Ph. D., Princeton University).

10. Goodman, N. R., Katz, S., Kramer, B. H. and Kuo, M. T., 1961, Frequency response from stationary noise: two case histories, Technometrics, Vol. 3, No. 2, pp. 245-268.

11. Goodman, N. R., 1965, Measurement of matrix frequency response functions and multiple coherence functions, AFFDL TR 65-56, Research and Technology Division, AFSC, Wrigth-Patterson AFB, Ohio.

12. Granger, C. W. J. and Hatanaka, M., 1964, Spectral analysis of economic time series. Princeton University Press, New Jersey.

13. Grenander, U. and Rosenblatt, M., 1957, Statistical analysis of stationary time series. Wiley and Sons, New York.

14. Hannan, E. J., 1960, Time series analysis. Methuen, London.

15. Holloway, J. L., 1958, Smoothing and filtering in time series and space fields. 4 Advances in Geophysics (Ed. Landsberg, H. E.) pp. 351-389. Academic Press, New York.

16. Jenkins, G. M., 1961, General considerations in the analysis of spectra. Technometrics, Vol. 3, No. 2, pp. 133-166.

17. Jenkins, G. M., 1962, Cross-spectral analysis and the estimation of linear open loop transfer functions. Time Series Analysis (Ed. Rosenblatt, M.) pp. 267-276.

18. Jenkins. G. M., 1963 An example of the estimation of a linear open loop transfer function. Technometrics, Vol. 5, No. 2, pp. 227-245.

19. Karreman, H. F., 1963, Computer programs for spectral analysis of economic time series. Econometric Research Program, Research Memorandum No. 59, Princeton Universtiy.

20. Kendall, M. G., 1966, The advanced theory of statistics. Vol. III, Charles Griffin and Co., London.

21. Lee, Y. W., 1963, Statistical theory of communications. Wiley and Sons, New York.

22. Matalas, N. C., 1966, Some aspects of time series analysis in hydrologic studies. Proceedings, 5th Canadian Hydrology Symposium, McGill Univ., Montreal, Canada.

23. Neidell, N. S., 1966, Spectral studies of marine geophysical profiles, Geophysics, Vol, XXXI, No. 1, pp. 122-134.

24. Panofsky, H. A. and Brier, G. W., 1958, Some applications of statistics to meteorology. Mineral Industries Extension Services, Pennsylvania State University, University Park, Pa.

25. Parzen, E., 1961, Mathematical considerations in the estimation of spectra. Technometrics, Vol. 3, No. 2, pp. 167-190.

26. Quimpo, R. G., 1966, Stochastic analysis of daily river flows. Unpublished Ph. D. Dissertation, Colorado State University, Fort Collins, Colorado.

27. Quimpo, R. G., 1967, Stochastic model of daily river flow sequences. Colorado State University Hydrology Paper No. 18, Fort Collins, Colorado.

28. Quimpo, R. G. and Yevjevich, V., 1967, Stochastic description of daily river flows. Proceedings, International Hydrology Symposium, I.A.S.H., Fort Collins, Colorado.

29. Quenouille, M. H., 1957, The analysis of multiple time series. Griffin, London

BIBLIOGRAPHY - continued

30. Roden, G. I., 1964, On the duration of nonseasonal temperature oscillations. Journal of the Atmospheric Sciences, Vol, 22, pp. 520-528.

31. Roden, G. I., 1966, A modern statistical analysis and documentation of historical temperature records in California, Oregon and Washington, 1821-1964, Journal of Applied Meteorology, Vol. 5, No. 1, pp. 3-24.

32. Roesner, L. A., 1965, Analysis of time series of monthly precipitation and monthly river flows. Unpublished Masters Thesis, Colorado State University, Fort Collins, Colorado.

33. Roesner, L. A. and Yevjevich. V., 1966, Mathematical models for time series of monthly precipitation and monthly runoff. Colorado State University Hydrology Paper No. 15, Fort Collins, Colorado.

34. Siddiqui, M. M., 1962, Some statistical theory for the analysis of radio propagation data. Journal of Research of the National Bureau of Standards, 66D, pp. 571-580.

35. Tick, L. J., 1962, Conditional spectra, linear systems, and coherency. Time Series Analysis (Ed. Rosenblatt, M.) pp. 197-203.

36. Tukey, J. W., 1961, Discussion, emphasizing the connection between analysis of variance and spectrum analysis. Technometrics, Vol, 3, No. 2, pp. 191-219.

37. Wainstein, L. A. and Zubakov, V. D., 1962, Extraction of signals from noise. Prentice-Hall Inc., Englewood Cliffs, New Jersey.

38. Whittle, P., 1954, Appendix to second edition of Stationary Time Series, by H. Wold, Almquist and Wiksell, Stockholm.

39. Wiener, N., 1948, Extrapolation, interpolation and smoothing of stationary time series. Wiley and Sons, New York.

40. Wold, H., 1954, A study in the analysis of stationary time series. Uppsala, Almquist and Wiksell.

41. Yaglom, A. M., 1962, An introduction to the theory of stationary random runctions. Prentice-Hall Inc., Englewood Cliffs, New Jersey.

42. Yevjevich, V., 1963, Fluctuations of wet and dry years. Part I, Research data assembly and mathematical models. Colorado State University Hydrology Paper No. 1, Fort Collins, Colorado.

43. Yevjevich, V., 1964, Fluctuations of wet and dry years, Part II, Analysis by serial correlation. Colorado State University Hydrology Paper No. 4, Fort Collins, Colorado.

APPENDIX 1

PRECIPITATION STATIONS USED FOR THE INVESTIGATIONS

O Station used with monthly and annual data
* Station used with monthly data only
** Station used with annual data only

Station Ident.	Name of Station	Lat.	Long.	Years of Record	Type of Data
		REGION NO. 1			
4.023	Antioch F. Mills	38.017	121.767	81	O
4.079	Big Sur State Park	35.250	121.783	46	O
4.291	Eureka	40.800	124.167	45	*
4.316	Fort Bragg	39.950	123.800	61	O
4.319	Fort Ross	38.517	123.250	85	O
4.640	Ojai	34.450	119.250	56	O
4.774	San Diego WB Apt	32.733	117.167	111	O
4.777	San Francisco WB Apt	37.800	123.017	45	*
4.785	San Luis Obispo Poly	35.300	119.250	91	O
35.190	Cottage Grove	43.783	123.067	44	O
35.675	Portland WB	45.517	122.667	45	*
45.477	Longview	46.167	122.917	36	O
45.833	Tatoosh Island WB	48.383	124.733	77	O
		REGION NO. 2			
4.038	Auburn	38.900	121.067	61	O
4.170	Chester	40.300	121.217	50	O
4.522	Lytle Creek PH	34.200	117.450	55	O
4.545	McCloud	41.267	122.133	50	O
4.835	Sonora	37.983	120.383	73	O

APPENDIX 1

PRECIPITATION STATIONS USED FOR THE INVESTIGATIONS-Continued

O Station used with monthly and annual data
* Station used with monthly data only
** Station used with annual data only

Station Ident.	Name of Station	Lat.	Long.	Years of Record	Type of Data
	REGION NO. 2 (Continued)				
4.945	Wasco	35.600	119.333	61	O
35.269	Estucada 2 SE	45.267	122.317	52	O
35.467	Lakeview	42.183	120.350	48	O
45.704	Rimrock Teton Dam	46.650	121.133	51	O
	REGION NO. 3				
5.243	Durango	37.283	107.883	67	O
5.762	Shoshone	39.567	107.233	51	O
10.271	Dubois Exp. Station	44.250	112.200	39	*
10.628	Driggs	43.733	111.117	51	**
10.808	Salmon	45.183	113.883	50	O
10.814	Sand Point Exp. Station	48.283	116.567	50	O
24.279	Saint Ignatius	47.317	114.100	52	**
48.091	Border 3N	42.250	111.033	59	**
48.407	Green River	41.533	109.483	51	O

APPENDIX 1

PRECIPITATION STATIONS USED FOR THE INVESTIGATIONS-Continued

O Station used with monthly and annual data
* Station used with monthly data only
** Station used with annual data only

Station Ident.	Name of Station	Lat.	Long.	Years of Record	Type of Data
		REGION NO. 4			
16.470	Tennings	30.23	92.67	63	*
16.666	New Orleans	29.95	90.07	91	*
41.061	Beaumont	30.08	94.10	68	*
41.318	Flatonia	29.68	97.10	53	*
41.343	Galveston WB City	29.30	94.83	89	*
41.351	George West	28.35	98.12	45	*
41.597	Mission	26.22	98.32	40	*
		REGION NO. 5			
14.643	Plains	37.27	100.58	51	*
14.664	Quinter	39.07	100.23	30	*
41.408	Henderson	32.15	94.80	52	*
41.502	Lampasas	31.05	98.18	66	*
41.721	Post	33.20	101.37	44	*
41.933	Vega	35.25	102.43	30	*

APPENDIX 2

MONTHLY RUNOFF STATIONS USED FOR THE INVESTIGATIONS

Station Ident.	Station Name	Lat.	Long.	Area	No. Years of Record
14. 049	Strawberry Creek AB Slide Creek Nr. Prairie City, Oreg.	44. 33	118. 65	7. 20	30
14. 059	Middle Fork John Day River at Ritter, Oreg.	44. 88	119. 13	515. 00	31
14. 064	John Day River at Service Creek, Oreg.	44. 80	120. 00	509. 00	31
14. 141	Lake Creek Nr. Sisters, Oreg.	44. 43	121. 73	22. 20	45
14. 241	Little Sandy River Nr. Bull Run, Oreg.	45. 42	122. 17	22. 30	41
14. 278	McKenzie River Nr. Vida, Oreg.	44. 13	122. 47	930. 00	36
14. 359	Clackmas River at Big Bottom, Oreg.	45. 02	121. 92	136. 00	40
14. 382	East Fork Lewis River Nr. Weisson, Wash.	45. 83	122. 47	125. 00	31
14. 390	Cowletz River at Packwood, Wash.	46. 62	121. 68	287. 00	31
14. 419	Toutle River Nr. Silver Lake, Wash.	46. 33	122. 73	474. 00	31
14. 438	Wilson River Nr. Tillanook, Oreg.	45. 48	123. 72	159. 00	30
12. 001	Naselle River Nr. Naselle, Wash.	46. 37	123. 75	55. 30	31
12. 006	North River Nr. Raymond, Wash.	46. 82	123. 85	219. 00	33
12. 040	Satsop River Nr. Satsop, Wash.	47. 00	123. 50	290. 00	31
12. 047	Quinault River at Quinault Lake, Wash.	47. 47	123. 90	264. 00	49
12. 050	Hoh River Nr. Spruce, Wash.	47. 80	124. 10	208. 00	34
12. 127	Carbon River Nr. Fairfax, Wash.	47. 03	122. 03	78. 90	31
12. 261	North Fork Stillaguamish Rv. Nr. Arlington, Wash.	48. 27	122. 05	269. 00	32
12. 667	North Fork Ahtanum Creek Nr. Tampico, Wash.	46. 57	120. 92	68. 90	30

APPENDIX 2

MONTHLY RUNOFF STATIONS USED FOR THE INVESTIGATIONS-Continued

Station Ident.	Station Name	Lat.	Long.	Area	No. Years of Record
112.001	Kern River Nr. Kernville,Calif.	35.93	118.48	865.00	48
112.032	North Fork Kaweah River at Kaweah, Calif.	36.48	118.92	128.00	49
112.066	Mono Creek Nr. Vermilion Valley, Calif.	37.37	118.98	92.00	39
112.112	Chowchilla Rv. at Buchanan Dam Site,Calif.	37.22	119.98	238.00	30
112.120	Merced River at Happy Isles Bridge Nr. Yosemite,Calif.	37.73	119.55	181.00	45
112.137	Falls Creek Nr. Hetchy Hetchy,Calif.	37.97	119.77	45.20	45
112.259	Hat Creek Nr. Hat Creek,Calif.	40.68	121.42	122.00	30
112.304	Mill Creek Nr. Los Molinos,Calif.	40.05	122.02	134.00	32
112.308	Thomes Creek at Paskenta,Calif.	39.88	122.55	188.00	40
112.402	Middle Fork American River Nr. Auburn,Calif.	38.92	121.00	616.00	49
111.059	Murrieta Creek at Temecula,Calif.	33.48	117.15	220.00	36
111.066	Arroyo Trabuco Nr. San Juan Capistrano, Calif.	33.53	117.67	36.50	30
111.083	Cajon Creek Nr. Keenbrook,Calif.	34.27	117.47	40.90	40
111.153	Santa Anita Creek Nr. Sierra Madre, Calif.	34.20	118.02	10.50	44
111.393	Salmon River at Somesbar,Calif.	41.38	123.47	746.00	37
111.411	Smith River Nr. Crescent City,Calif.	41.78	124.05	613.00	30
10.275	Big Rock Creek Nr. Valyermo,Calif.	34.42	117.83	23.00	37
10.278	Convict Creek Nr. Mammoth Lakes, Calif.	37.62	118.85	18.70	35

APPENDIX 3

LIST OF SYMBOLS

Symbol	Definition
$A(f)$	Cross-amplitude function
$C_{xy}(f)$, $\hat{C}_{xy}(f)$, $\tilde{C}_{xy}(f)$	True, smoothed estimate, and raw estimate of the co-spectrum
D^n	Mathematical operator
f	Frequency
$G_x(f)$, $\hat{G}_x(f)$, $\tilde{G}_x(f)$	True, smoothed estimate, and raw estimate of the physically realizable one-sided power spectrum
$G_{xy}(f)$, $\hat{G}_{xy}(f)$, $\tilde{G}_{xy}(f)$	True, smoothed estimate, and raw estimate of the physically realizable one-sided cross-spectrum
$g_x(f)$, $\hat{g}_x(f)$	True, and estimated spectral density
$H(f)$	Frequency response function
$\lvert H(f)\rvert$, $\lvert\hat{H(f)}\rvert$	True, and estimated, gain function
$h(\tau)$	Unit impulse response function
$Q_{xy}(f)$, $\hat{Q}_{xy}(f)$, $\tilde{Q}_{xy}(f)$	True, smoothed estimate, and raw estimate of the quadrature spectrum
$R(\omega)$	Transfer function
$S_x(f)$	Two-sided power spectrum
$S_{xy}(f)$	Two-sided cross-spectrum
$S_{1y\cdot 2}(f)$	Residual or partial spectrum
$s(x)$	Standard deviation of the series $x(t)$
X	Complex random variable
X^*	Complex conjugate of X
$x_k(t)$	Sample function of a stationary process
$[x_k(t)]$	Stationary random process
$\hat{y}(t)$	Predicted $y(t)$
$\Delta y(t)$	Residual random variable
$\alpha_x(\tau)$, $\hat{\alpha}_x(\tau)$	True, and estimated, autocovariance function of a stationary process
$\alpha_{xy}(\tau)$, $\hat{\alpha}_{xy}(\tau)$	True, and estimated, cross-covariance function between two stationary processes
$\gamma^2_{xy}(f)$, $\hat{\gamma}^2_{xy}(f)$	True, and estimated, coherence function
$\gamma^2_{1y\cdot 2}(f)$, $\hat{\gamma}^2_{1y\cdot 2}(f)$	True, and estimated, partial coherence function
$\theta_{xy}(f)$, $\hat{\theta}_{xy}(f)$	True, and estimated, phase function
$\theta_{1y\cdot 2}(f)$, $\hat{\theta}_{1y\cdot 2}(f)$	True, and estimated, partial phase function

APPENDIX 3

LIST OF SYMBOLS - Continued

Symbol	Definition
$\mu_x(t)$, $\bar{x}$	True, and estimated, mean value of a stationary process
$\rho_{xy}(\tau)$	Autocorrelation function
$\rho_{1y \cdot 2}$	Partial correlation coefficient
σ_x^2	Variance of the series x(t)
$\phi(f)$	System phase factor
ω	Angular frequency

Key Words: Hydrology, Cross-Spectrum, Coherence, Correlation, Time Series.

Abstract: The main objective of this paper is to study the potentials of cross-spectrum and multiple cross-spectrum for the analysis of hydrological data.

Groups of precipitation and runoff stations were selected in different climatic environment and complete cross-spectral analyses were performed between those stations. The coherence and partial coherence functions were used for the study of frequency correlations between the series and they show that there exists a very strong correlation between the annual cycles of the stations. Along the Pacific Coast of the United States the annual cycle in precipitation appears to be basically the same up to distances of 1000 Km.

Cyclic regression analysis with the use of the gain and phase functions is shown to work correctly in hydrologic time series. This type of regression may be very useful in regions where frequency components account for a large percentage of the variance of the series.

Cross-spectral characteristics of the moving average and autoregressive processes are shown to be a powerful tool in testing and analyzing these types of generating processes in hydrology. Special significance has the coherence between two 1st order autoregressive processes which is shown to be equal to a constant independent of frequency.

The effects of smoothing or pre-filtering in the cross-spectral properties of two series are studied and recommendations made when working with this practice which is frequently used in hydrology.

Reference: Rodríguez - Iturbe, Ignacio, Colorado State University Hydrology Paper No. 24 (September 1967), "The Application of Cross-Spectral Analysis to Hydrologic Time Series."

Key Words: Hydrology, Cross-Spectrum, Coherence, Correlation, Time Series.

Abstract: The main objective of this paper is to study the potentials of cross-spectrum and multiple cross-spectrum for the analysis of hydrological data.

Groups of precipitation and runoff stations were selected in different climatic environment and complete cross-spectral analyses were performed between those stations. The coherence and partial coherence functions were used for the study of frequency correlations between the series and they show that there exists a very strong correlation between the annual cycles of the stations. Along the Pacific Coast of the United States the annual cycle in precipitation appears to be basically the same up to distances of 1000 Km.

Cyclic regression analysis with the use of the gain and phase functions is shown to work correctly in hydrologic time series. This type of regression may be very useful in regions where frequency components account for a large percentage of the variance of the series.

Cross-spectral characteristics of the moving average and autoregressive processes are shown to be a powerful tool in testing and analyzing these types of generating processes in hydrology. Special significance has the coherence between two 1st order autoregressive processes which is shown to be equal to a constant independent of frequency.

The effects of smoothing or pre-filtering in the cross-spectral properties of two series are studied and recommendations made when working with this practice which is frequently used in hydrology.

Reference: Rodríguez - Iturbe, Ignacio, Colorado State University Hydrology Paper No. 24 (September 1967), "The Application of Cross-Spectral Analysis to Hydrologic Time Series."

Key Words: Hydrology, Cross-Spectrum, Coherence, Correlation, Time Series.

Abstract: The main objective of this paper is to study the potentials of cross-spectrum and multiple cross-spectrum for the analysis of hydrological data.

Groups of precipitation and runoff stations were selected in different climatic environment and complete cross-spectral analyses were performed between those stations. The coherence and partial coherence functions were used for the study of frequency correlations between the series and they show that there exists a very strong correlation between the annual cycles of the stations. Along the Pacific Coast of the United States the annual cycle in precipitation appears to be basically the same up to distances of 1000 Km.

Cyclic regression analysis with the use of the gain and phase functions is shown to work correctly in hydrologic time series. This type of regression may be very useful in regions where frequency components account for a large percentage of the variance of the series.

Cross-spectral characteristics of the moving average and autoregressive processes are shown to be a powerful tool in testing and analyzing these types of generating processes in hydrology. Special significance has the coherence between two 1st order autoregressive processes which is shown to be equal to a constant independent of frequency.

The effects of smoothing or pre-filtering in the cross-spectral properties of two series are studied and recommendations made when working with this practice which is frequently used in hydrology.

Reference: Rodríguez - Iturbe, Ignacio, Colorado State University Hydrology Paper No. 24 (September 1967), "The Application of Cross-Spectral Analysis to Hydrologic Time Series."

Key Words: Hydrology, Cross-Spectrum, Coherence, Correlation, Time Series.

Abstract: The main objective of this paper is to study the potentials of cross-spectrum and multiple cross-spectrum for the analysis of hydrological data.

Groups of precipitation and runoff stations were selected in different climatic environment and complete cross-spectral analyses were performed between those stations. The coherence and partial coherence functions were used for the study of frequency correlations between the series and they show that there exists a very strong correlation between the annual cycles of the stations. Along the Pacific Coast of the United States the annual cycle in precipitation appears to be basically the same up to distances of 1000 Km.

Cyclic regression analysis with the use of the gain and phase functions is shown to work correctly in hydrologic time series. This type of regression may be very useful in regions where frequency components account for a large percentage of the variance of the series.

Cross-spectral characteristics of the moving average and autoregressive processes are shown to be a powerful tool in testing and analyzing these types of generating processes in hydrology. Special significance has the coherence between two 1st order autoregressive processes which is shown to be equal to a constant independent of frequency.

The effects of smoothing or pre-filtering in the cross-spectral properties of two series are studied and recommendations made when working with this practice which is frequently used in hydrology.

Reference: Rodríguez - Iturbe, Ignacio, Colorado State University Hydrology Paper No. 24 (September 1967), "The Application of Cross-Spectral Analysis to Hydrologic Time Series."