Time series in *M* dimensions: The power spectrum

by Roger H. Yetzer

The approach presented here extends the modeling of *M*-dimensional (spatial) time series from the time domain into the frequency domain. The autocovariance function for an *M*-dimensional time series is transformed to obtain the power spectrum in *M* dimensions. The latter describes the variance within the series and can be used to identify dependencies and/or test the adequacy of a fitted model. An example is provided.

Introduction

The analysis of M-dimensional time series is a methodology that can be used to form a composite model of available information across spatial regions over time to identify underlying structure (dependencies). One of its objectives is to identify, independent of a specific site or time, a single model with the least number of parameters that provides an adequate approximation to a stochastic event or process. Some examples of processes that generate time series data are climatic changes, sales volumes, process control, highway deterioration, and pattern recognition.

Spatial time series data or M-dimensional time series data are data collected at a number of different spatial locations over a number of different times. Alternatively, this can be regarded as a set of time series collected simultaneously at a number of different spatial locations. A conceptual example is given by the following: Assume we wanted to understand how the amount of dissolved oxygen evolves over time in a flowing stream. The approach for doing so would be to establish P sampling sites along the stream at different spatial locations. Each location would then generate a time series or a realization of the underlying stochastic process. It would be assumed that the amount of dissolved oxygen present at each site depends upon the amounts present at locations upstream, with a time lag. The objective would then be to use the composite data to determine a single model that is invariant from site to site to describe the spatial and temporal dependencies. This differs from regression analysis, in which one site would be selected as the dependent site.

The primary tool used to determine spatial and temporal dependencies is the autocorrelation function. However, this function is sometimes difficult to interpret and may be misleading. An alternative and perhaps preferred approach is to transform from the time/space domain into the equivalent frequency domain via a Fourier transform of the autocorrelation function.

The concepts of M-dimensional time series have been developed and explained in [1–3]. Described here is the development of an M-dimensional power spectrum that can be used as a tool to analyze the variance. The area under its curve is equal to the total variance of the data. A method for estimating the associated spectrum is presented, along with procedures for using the spectrum for model identification and diagnostic checking of model adequacy. A brief example is provided to illustrate the concepts.

Definitions and terminology

Let the set $\{Z(u, t)\}$ denote the realization of a stochastic process in the M-dimensional space denoted by the spatial vector $u = \langle u_1, u_2, \cdots, u_M \rangle$ for $1 \le u_i \le P_i$ and its temporal index $t = 1, 2, \cdots, N$, where t is a scalar. Each P_i represents the number of sampling locations in dimension i for $1 \le i \le M$. Thus, for each dimension u_i there are P_i sampling locations, where N observations are made over a period of time. In this context, time is viewed as an "index" into the spatial data and not as a dimension. When $M = 0$, $Z(u, t)$ reduces to $Z(t)$, which yields a traditional "Box–Jenkins" time series [4].

As a minimum assumption it is assumed that weak stationarity exists in space and time. This assumption is ensured by the existence of the following:

- The mean $\mu_Z = E[Z(u, t)] = 0$; all trends have been removed.
- The variance $\sigma_Z^2 = E[Z - \mu]^2 < \infty$ is a constant.
- The autocovariance $\gamma(m, n) = \mathrm{COV}\,[Z(u, t), Z(u - m, t - n)]$ depends only on the lags m and n, or on the separation of the data points by a constant interval; it does not depend on the choice of origin.
- The autocorrelation function is given by $\rho(m, n) = \gamma(m, n)/\sigma_Z^2$.
- The random error $a(u, t)$ is an i.i.d. variable with mean zero and $\sigma_a^2 > 0$.

Nonstationary behavior may be rendered stationary by suitable differencing of the series, as given by Perry [5].

The autocovariance function $\gamma(m, n)$ for lags m and n, and its estimate $c(m, n)$ are given by

$$\gamma(m, n) = E[Z(u, t) * Z(u - m, t - n)], \tag{1}$$

$$c(m, n) = \frac{1}{NP} \sum_u \sum_t Z(u, t) * Z(u - m, t - n), \tag{2}$$

where $u = \langle m_i + 1, m_i + 2, \cdots, P_i \rangle$ for positive or negative values of m, for positive n, and for $t = n + 1, n + 2, \cdots, N$. The notation $\sum_u$ is defined as $\sum_{u_M} \cdots \sum_{u_2} \sum_{u_1}$.

An important property of γ is the form of symmetry given by

$$\gamma(-m, -n) = \gamma(m, n),$$

$$\gamma(-m, n) \;\; = \gamma(m, -n); \tag{3}$$

however, $\gamma(m, n) \ne \gamma(-m, n)$ or $\gamma(0, n) \ne \gamma(m, 0)$.

General models

The two basic models used are the autoregressive (AR) model and the moving average (MA) model. The AR model is a form of multiple regression on past values of itself, $Z(u, t)$, and has the form

$$Z(u, t) = \sum_p \sum_q \phi_{pq} Z(u, t) Z(u + p, t - q) + a(u, t). \tag{4}$$

For example,

$$Z(u, t) = \phi_{01} Z(u, t - 1) + \phi_{11} Z(u - 1, t - 1) + a(u, t).$$

The MA model is similar to a linear filter where the inputs are the past random shocks of a stochastic process; it has the form

$$Z(u, t) = \sum_p \sum_q \theta_{pq} Z(u, t) Z(u + p, t - q) + a(u, t). \tag{5}$$

For example,

$$Z(u, t) = \theta_{01} a(u, t - 1) + \theta_{11} a(u - 1, t - 1) + a(u, t).$$

There also exists a duality relationship between the two models: Every stationary MA process can be converted to an AR process.

Power spectrum

The literature does not provide a universal definition for a multidimensional Fourier transform. The definitions cited by [2, 3, 6, 7] all differ from Equation (6) below by a constant multiple and by the range of the definition for the transformation. It should be kept in mind that both $Z(u, t)$ and the error term $a(u, t)$ are regarded as time series. The concepts developed here can be applied to the autocovariance function of either without any loss of generality.

The Fourier transform $f(\omega, \lambda)$ is expressed by

$$f(\omega, \lambda) = \frac{1}{(2\pi)^M} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \gamma(m, n) e^{-i(\omega \cdot m + \lambda n)}. \tag{6}$$

The transform is the amplitude at the frequencies ω in the spatial dimension and λ in the time domain for $-\pi < \omega < \pi$ and $-\pi < \lambda < \pi$. The term ω is the vector given by $\omega = \langle \omega_1, \omega_2, \cdots, \omega_M \rangle$, λ a scalar, and $\omega \cdot m = \sum \omega_i m_i$.

The inverse of the transform is given by

$$\gamma(m, n) = \int_\omega \int_\lambda f(\omega, \lambda) e^{i(\omega \cdot m + \lambda n)} \, d\omega d\lambda. \tag{7}$$

For spatial lags m and temporal lags n, $(-P_i < m_i < P_i, -N < n < N)$, the power spectrum $f(\omega, \lambda)$ can be used to identify variation within the sum of squares. The variance of the error term from an autoregressive model $a(u, t)$ is the total area under the power spectrum. This property is demonstrated as follows.

Let $m = 0$ and $n = 0$ in the inverse transform (7) to obtain

$$\gamma(0, 0) = \int \int f(\omega, \lambda) \, d\omega d\lambda,$$

Table 1 Decomposition of the power spectrum into regions along the axes for $M = 1$ time series.

Regions for m	n	*Pairs of terms*	$= G(m, n)$	*Regions for* $G(m, n)$
0	0	$g(0, 0)$	$= G(0, 0)$	$m = 0, n = 0$
$\pm$	$\pm$	$g(m, n) + g(-m, -n)$	$= G(m, n)$	$m > 0, n > 0$
$\pm$	0	$g(m, 0) + g(-m, 0)$	$= G(m, 0)$	$m > 0, n = 0$
0	$\pm$	$g(0, n) + g(0, -n)$	$= G(0, n)$	$m = 0, n > 0$
$\pm$	$\mp$	$g(m, -n) + g(-m, n)$	$= G(-m, n)$	$m > 0, n > 0$

but $\gamma(0, 0) = E[a(u, t) * a(u - 0, t - 0) = E[a^2(u, t)]$ is the total variance of $a(u, t)$. Hence,

$$E[a^2] = \int_\omega \int_\lambda f(\omega, \lambda) = \int_\omega \int_\lambda \text{power spectrum.} \tag{8}$$

It includes all frequencies in the range of $(-\pi, \pi)$ which might contribute to the variation of the process. Any peak in the spectrum would indicate an important contribution to the variance at frequencies in the appropriate region. This property provides the benefits of using the spectral function over the autocorrelation function when analyzing data. It may be more appropriate to rename the power spectrum as the variance spectrum when dealing with spatial data.

The spectrum $f(\omega, \lambda)$ can be normalized by dividing by the variance $\gamma(0, 0)$ to obtain the spectral density function with the property

$$\int\int g(\omega, \lambda)\, d\omega d\lambda = 1.$$

In order to reduce Equation (6) from a transform defined in terms of complex variables to a cosine transform, the following is developed next.

Using the symmetry of the autocovariance function given by property (3) and the identity $e^{-ia} + e^{ia} = 2\cos(a)$, and letting $g(m, n) = \gamma(m, n)e^{-i(\omega \cdot m + \lambda n)}$, the following two lemmas will aid in simplification.

Lemma 1

$$g(-m, -n) + g(m, n) = 2\gamma(m, n)\cos(\omega \cdot m + \lambda n). \tag{9}$$

Proof

$$\begin{aligned} &g(-m, -n) + g(m, n) \\ &= \gamma(-m, -n)e^{-i(\omega \cdot m + \lambda n)} + \gamma(m, n)e^{i(\omega \cdot m + \lambda n)} \\ &= \gamma(m, n)e^{-i(\omega \cdot m + \lambda n)} + \gamma(m, n)e^{-i(\omega \cdot m + \lambda n)} \\ &= \gamma(m, n)[e^{i(\omega \cdot m + \lambda n)} + e^{-i(\omega \cdot m + \lambda n)}] \\ &= \gamma(m, n)[2\cos(\omega \cdot m + \lambda n)]. \end{aligned}$$

Lemma 2

$$g(-m, n) + g(m, -n) = 2\gamma(m, n)\cos(\omega \cdot m + \lambda n). \tag{10}$$

Proof

$$\begin{aligned} &g(-m, n) + g(m, -n) \\ &= \gamma(-m, n)e^{-i(\omega \cdot m + \lambda n)} + \gamma(m, -n)e^{-i(\omega \cdot m + \lambda n)} \\ &= \gamma(-m, n)e^{-i(\omega \cdot m + \lambda n)} + \gamma(-m, n)e^{-i(-\omega \cdot m + \lambda n)} \\ &= \gamma(-m, n)[e^{-i(-\omega \cdot m + \lambda n)} + e^{-i(-\omega \cdot m + \lambda n)}] \\ &= \gamma(-m, n)[2\cos(-\omega \cdot m + \lambda n)]. \end{aligned}$$

Letting $G(m, n) = 2\gamma(m, n)\cos(\omega \cdot m + \lambda n)$, Lemmas 1 and 2 can be summarized as follows:

$$G(\pm m, \pm n) = g(\pm m, \pm n) + g(\mp m, \mp n). \tag{11}$$

M-dimensional power spectrum

The power spectrum $f(\omega, \lambda)$ given by definition (6) can be reduced to a more desirable form, given by expression (12), as discussed below. This is achieved by decomposing each sum along the axes and then recombining in pairs of terms such that all sums will be over their positive indices. The pairing is done on regions symmetrical to the origin, with care taken to avoid double counting. This procedure is illustrated for an $M = 1$-dimensional spatial time series using Equations (9), (10), and (11) to simplify the notation.

As the lags m and n increase to infinity, the autocovariance function $\gamma(m, n)$ converges to zero. A truncated autocovariance function with finite limits can thus be used without any loss of generality for $-\pi < \omega < \pi$, $\pi < \lambda < \pi$. The M-dimensional power spectrum for time series data is then given by

$$f(\omega, \lambda) = \frac{1}{(2\pi)^{M+1}} \left[\gamma(0, 0) + 2 \sum_{\substack{m=-P+1 \\ (m,n)\neq 0 \\ (+m_j, 0)}}^{P-1} \sum_{n=0}^{N} \gamma(m, n)\cos(\omega \cdot m + \gamma n) \right], \tag{12}$$

where $(+m_j, 0)$ is defined, when $n = 0$, as $m_j > 0$, and all succeeding values of m are zero: i.e., $m_{j+1} = m_{j+2} = \cdots m_M = 0$.

Equation (12) is derived by first rewriting Equation (6) in the form $f(\omega, \lambda) = k \sum \sum g(m, n)$ and then decomposing each sum along the axes, as illustrated in **Table 1**. Combining the regions shown in Table 1 gives

$$g(0,0) + \underbrace{\sum_{m=1}^{P-1}\sum_{n=1}^{N-1} G(m,n) + \sum_{m=1}^{P-1} G(m,0) + \sum_{n=1}^{N-1} G(0,n)} + \sum_{m=1}^{P-1}\sum_{n=1}^{N-1} G(-m,n)$$

$$= g(0,0) + \sum_{\substack{m=0 \\ (+m,0) \\ (m,n)\neq 0}}^{P-1}\sum_{n=0}^{N-1} G(m,n) + \sum_{m=-P+1}^{P-1}\sum_{n=1}^{N-1} G(m,n)$$

$$= g(0,0) + \sum_{\substack{m=-P+1 \\ (+m,0) \\ (m,n)\neq 0}}^{P-1}\sum_{n=0}^{N-1} G(m,n),$$

where $(+m, 0)$ corresponds to $m > 0$ when $n = 0$. This yields Equation (12) for $M = 1$, viz.,

$$f(\omega,\lambda) = \frac{1}{(2\pi)^2}\left[\gamma(0,0) + 2\sum_{\substack{m=-P+1 \\ (m,n)\neq 0 \\ (+m,0)}}\sum_{n=0}^{N} \gamma(m,n)\cos(\omega \cdot m + \gamma n)\right]. \quad (13)$$

Estimation of the power spectrum

The estimate of the spectrum should not be obtained by replacing the theoretical autocovariance $\gamma(m, n)$ with the estimates $c(m, n)$. The sample spectrum of a stationary time series fluctuates about the theoretical spectrum. This is analogous to using too small a group interval for the histogram when estimating an ordinary probability distribution.

Since the cosine function is periodic, variations at frequencies higher than π cannot be distinguished, and appear "aliased" with lower frequencies in $(-\pi, \pi)$. The phenomenon of aliasing depends on the sampling rate. One should try to choose an appropriate sampling interval so that the maximum frequency that can be detected from the data exceeds the maximum frequency present and is at least twice the frequency of interest. These concepts are further discussed in [4, 7, 8].

By using a modified or tapered estimate of the sample autocovariance $c(m, n)$, it is possible to increase the "bandwidth" and obtain a smoother estimate for the spectrum. A popular method of smoothing is achieved by multiplying the autocovariance function by a weighting or tapering function that gives less weight to values of $c(m, n)$ as $|m|$ and n increase [6, 8, 9]. One such weighting function, or "lag window," is given by

$$h(m,n) = \frac{1}{2^{P+1}}\,[1 + \cos(n\pi/R)]\,\prod\,[1 + \cos(m_i\pi/S)], \quad (14)$$

with truncation points R and S. The associated estimate of $f(\omega, \lambda)$ is given by

$$\hat{f}(\omega,\lambda) = \frac{1}{(2\pi)^2}\left[\gamma(0,0) + 2\sum_m\sum_n h(m,n)\gamma(m,n)\cos(\omega \cdot m + \gamma n)\right]. \quad (15)$$

The proper choice of R and S may be difficult to make in attempting to determine the overall nature of the true spectral density. Low values will give an indication of where the large peaks exist, but the curve is likely to be too smooth. Large values are likely to produce a curve showing a large number of peaks, some of which may be spurious. A compromise can be achieved through the use of intermediate values.

Use of spectral analysis

The modeling of time series data follows a three-stage iterative approach: model identification, parameter estimation, and testing of adequacy. The goal of the modeling is to identify a model with the smallest number of parameters that provides an adequate approximation for the process of interest.

The spectral function applied to a spatial time series provides insight into spatial and temporal dependencies and thus aids in model selection and identification. This provides an alternative approach to viewing the visual appearance of the autocorrelation function to analyze the total variance (mean squared value) of the data. From peaks in the spectral function it can be inferred which terms should be

Table 2 Site locations.

Site	Station	Location	U.T.M. coordinates	
H	3101–16	Niagara & Hawley, Lockport	199.2E	4786.1N
D	1451–03	Audubon Golf Course, Amherst	193.0E	4766.5N
B	1401–18	Dingens & Weiss, Buffalo	188.8E	4754.0N
A	1402-01	Lehigh St., Lackawanna	186.1E	4747.7N

U.T.M. = Universal Transverse Mercator

Table 3 Average concentration of sulfur dioxide emissions (micrograms per cubic meter).

Month (1982)	Sites				
	A	B	D	H	Mean
Dec	0.010	0.020	0.014	0.012	0.014
Nov	0.008	0.011	0.013	0.013	0.011
Oct	0.007	0.013	0.012	0.009	0.010
Sept	0.007	0.011	0.006	0.009	0.008
Aug	0.006	0.012	0.007	0.017	0.011
Jul	0.005	0.011	0.009	0.012	0.009
Jun	0.007	0.010	0.007	0.012	0.009
May	0.008	0.015	0.009	0.021	0.013
Apr	0.013	0.016	0.012	0.018	0.015
Mar	0.016	0.016	0.015	0.022	0.017
Feb	0.019	0.024	0.021	0.022	0.022
Jan	0.017	0.014	0.020	0.024	0.019
Mean (1980–82)	0.013	0.016	0.011	0.013	0.013

Partial data set.
Source: New York State Department of Environmental Conservation, 50 Wolf Road, Albany, New York 12233

considered. Any peaks in the spectrum would indicate an important contribution to the variance at frequencies in the appropriate region.

After fitting an AR or MA model to a set of spatial data, a test of model adequacy is performed. The spatial function can then be applied to the corresponding residuals $a(u, t)$ as an effective tool to test model adequacy.

If $\{a(u, t)\}$ were a purely random process,

$$\gamma(m, n) = \begin{cases} \sigma^2(a) & m = 0,\ n = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

causing the power spectral function

$$f(\omega, \lambda) = \sigma^2(a)/\pi^2 \tag{17}$$

to be a constant.

If $f(\omega, \lambda)$ were plotted for $\{\omega, \lambda\}$ in $(-\pi, \pi)$, a purely random process would yield a plane in the $M + 2$ dimensional space $\langle \omega, \lambda \rangle$. The power (variance) spectral distribution function describes how the variability (power) of the process is distributed over a continuous range of frequencies. A large spectrum at low frequencies indicates the possibility of nonstationarity. Peaks in the spectrum indicate a contribution to the variance at frequencies in the appropriate regions, or a lack of fit. A topic for future work would be to extend the Kolmogorov–Smirnov goodness-of-fit test [10] to several variables to provide a statistical test for deviations from a constant.

Example: Sulfur dioxide emissions

The purpose of the following example is primarily to illustrate how the power spectrum can be used as a tool for identifying an underlying mechanism and ascertaining model adequacy.

The New York State Department of Environmental Conservation maintains continuous air monitors throughout the State of New York. Four sites from the Buffalo area (Niagara Frontier) were selected for analysis. The characteristic analyzed was the monthly average concentrations in micrograms per cubic meter of sulfur dioxide contained in air samples. The sites used, their relative locations, and the monthly average concentrations of sulfur dioxide for a representative set of data are presented in **Tables 2** and **3**. The sites were located along a N–NE path, while the prevailing winds were in the NE, E–NE direction. The data were obtained over the 36-month period from January 1980 to December 1982. The means shown are for the total period.

The data were initially modeled with an autoregressive model of the form (4) without any seasonal terms, viz.,

$$z(u, t) = 0.27z(u, t - 1) + 0.53z(u - 1, t - 1) + a(u, t). \tag{18}$$

The indicated coefficients were estimated using an ordinary least-squares estimate of properly ordered data.

Figure 1 depicts the power spectrum for time lags of $n = 18$ and spatial lags between sites of $m = 3$, plotted for intervals of 0.5 units. The peaks in the time direction occur at a lag of 3 or $18/3 = 6$-month period. The peaks in the spatial direction occur at lags of 0 and 3, thus showing a dependency of a site with itself and with the site at $3/3 = 1$ location away. Note that the time lags are positive, whereas the spatial lags are both positive and negative.

An analysis of the power spectrum reveals that in addition to a six-month seasonal "characteristic" of sulfur dioxide emissions, there was also a strong dependency on neighboring sites. On the basis of the spectrum, it might be useful to entertain a model containing a six-month seasonal term and additional site dependencies.

Nevertheless, considering the locations of the sites and the general wind patterns, it would seem feasible that the spatial dependencies observed reflected the effects of wind on the emissions.

Summary

This treatment has extended the analysis of M-dimensional spatial time series from the time/space domain into its equivalent frequency domain. The resulting function, the power spectrum, was derived as a cosine transform of the autocovariance function. The value of analyzing the power spectrum is that it describes how the total variance of the data is distributed within the continuous range of frequencies from $-\pi$ to π. Some of the difficulty associated with the interpretation and use of the autocorrelation function are thereby avoided. The power spectrum provides a means for identifying existing temporal/spatial dependencies and testing the adequacy of a fitted model. An example has been included, with some discussion of how to estimate an associated power spectrum using the sample autocovariance.

References

1. Leo A. Aroian, "Time Series in *M* Dimensions: Past, Present, and Future," *Time Series Analysis: Theory and Practice 6*, O. D. Anderson, J. K. Ord, and E. A. Robinson, Eds., Elsevier Scientific Publishers B.V., Netherlands, 1985, pp. 241–261.
2. Leo A. Aroian and Josef Schmee, "General Results, Time Series in *M*-Dimensions," *Proceedings of the 11th Annual Pittsburgh Conference on Modeling Simulation*, University of Pittsburgh, PA, 1980, pp. 1517–1522.
3. Robert J. Perry and Leo A. Aroian, "Autoregressive Model in *M*-Dimensions, $M = 1$, Theory and Examples," *Time Series Analysis: Theory and Practice* 6, O. D. Anderson, J. K. Ord, and E. A. Robinson, Eds., Elsevier Science Publishers, Netherlands, 1985, pp. 263–271.
4. G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day Publishing Co., San Francisco, 1976.
5. Robert J. Perry, "Time Series in M-Dimensions, Autoregressive Models, $M = 1$," Ph.D. Dissertation, Union College, Schenectady, NY, University Microfilms, Ann Arbor, MI, 1981.
6. C. Chatfield, *The Analysis of Time Series: Theory and Practice*, Chapman & Hall Publishing Co., London, 1975.
7. G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*, Holden-Day Publishing Co., San Francisco, 1968.
8. J. N. Rayner, *An Introduction to Spectral Analysis*, Pion Limited, London, 1972.
9. B. D. Ripley, *Spatial Statistics*, John Wiley & Sons, Inc., New York, 1981.
10. For example, Alexander M. Mood, Franklin A. Graybill, and Duane C. Boes, *Introduction to the Theory of Statistics*, Third Ed., McGraw-Hill Book Co., Inc., New York, 1974.

Received May 12, 1988; accepted for publication November 1, 1988

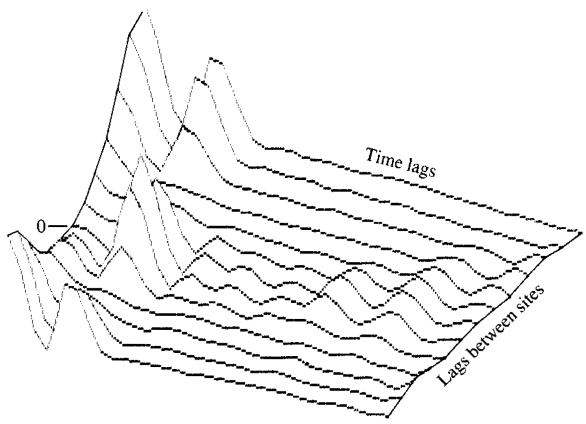

Figure 1

Power spectrum of sulfur dioxide emissions.

Roger H. Yetzer *IBM Corporation, Industrial Sector Organization, Kingston, New York 12401.* Dr. Yetzer received his B.A. degree from Marist College, Poughkeepsie, New York, and his Ph.D. degree in Administrative and Engineering systems from Union College, Schenectady, New York, in 1987. His Ph.D. dissertation pertained to tests of model adequacy for *M*-dimensional time series analysis. Prior to joining IBM, Dr. Yetzer was a college professor at several academic institutions. Since joining IBM in 1983 he has worked in the East Fishkill Development Laboratory; he is currently working in the area of engineering graphics products in the Kingston facility. He is a member of the mid-Hudson chapter of the American Statistical Association.