Commun. Fac. Sci. Univ. Ank. Ser. A_1
V. 37, pp 67–94 (1988)

SPATIAL PROCESSES: MODELLING, ESTIMATION, AND HYPOTHESIS TESTING

ÖMER L. GEBİZLİOĞLU

Department of Statistics, Middle East Technical University, Ankara, Turkey

ABSTRACT

This paper develops purely spatial ARMA models. Suitable models, based on the correlational structure of observed spatial processes, are fitted. The parameters are estimated by minimum variance prediction error, least squares, or maximum likelihood estimation methods. The techniques of hypothesis testing are included.

INTRODUCTION

Spatio–temporal analysis are concerned with the analysis of series of observations of variables measured over a period of time and over a sample region in space. When the history of spatially located observations is considered it is not possible to impose strict ordering on the directions and orders of dependencies in the space domain, although strict ordering in time is preserved. Bennett (1979) presents an overview of such approaches including a detailed list of works. An approach to the analysis of spatio–temporal processes with ARMA family models which emphasize the autocorrelation structure in time and space as the only determinant of an appropriate class of models has been shown by Aroian (1979, 1980, 1981); Oprian, Taneja, Voss, and Aroian (1980); Taneja and Aroian (1980); and Voss, Oprian, and and Aroian (1980).

In many cases the time of observations may be redundant, either because it is not possible to observe the temporal evolution, or because it is of no substantive significance if the system that generates the process under consideration has reached a static equilibrium pattern. These cases constitute purely spatial processes. Geological, agricultural, environmental observations referenced in a spatial coordinate system may be cited as examples of realizations of purely spatial processes.

Wheat yield data of Mercer and Hall(1911), and fruit trees yield data of Batchelor and Reed (1918) are particular examples. An account of several approaches to the modelling of purely spatial interaction may be found in the work of Ripley (1981). Analysis of purely spatial processes with simple autoregressive models were first suggested by Whittle (1954, 1963), and Heine (1955). Thereafter, Besag (1972, 1974), Bartlett (1979), Cliff and Ord (1973, 1975), and Ord (1975) discussed spatial autocorrelation function by emphasizing simple autoregressive schemes. Considering the results of these works and taking the theory of M-dimensional time series analysis developed by Aroian and his co-authors as a base, Aroian and Gebizlioğlu (1980), and Gebizlioğlu (1981, 1982) discussed purely spatial autoregressive, (AR), moving average (MA), and autoregressive–moving average, (ARMA), processes and their properties for the univariate case. Process identification, model estimation, and validation for such processes will be presented in the following sections.

GENERAL MODEL AND PROPERTIES

Let R^m denote an m–dimensional space with a coordinate system whose elements are $X = (x_1, x_2, \ldots, m_m)$. Associated with each point X is a random variable Z_X which represents values of a character of a purely spatial process. If $\mathbb{P} = < \Omega, s, P) >$ denotes a fixed probability space on a non–empty set Ω with sigma algebra ζ of sets in Ω and probability measure $\mathbb{P}$ on ζ, each family of random variables $Z = \{Z_X \mid X \in R^m\}$ on $\mathbb{P}$ designates an m–dimensional random field. Spatial interection on this random field can be expressed by the general probabilistic model

$$\tilde{Z}_X = \sum_{n=-p}^{q} \Phi_n \tilde{Z}_{X+n} - \sum_{n=-u}^{v} \Theta_n a_{X+n} + a_X \tag{1}$$

where $\tilde{Z}_X = Z_X - Z(E_X)$, $\varnothing_0 = \theta_0 = 0$, and a's are random shock variables. $n = (n_1, n_2, \ldots, n_m)$ and $\Phi^1{}_n = (\varnothing_{n_1}, \varnothing_{n_2}, \ldots, \varnothing_{n_m})$, $\Theta^1{}_n = (\theta_{n_1}, \theta_{n_2}, \ldots, \theta_{n_m})$ are parameter vectors of autoregressive and moving average components, respectively. p, q, u, and v are positive integers and $\sum\limits_{n=-p}^{q}$ $\sum\limits_{n=-v}^{v}$ are m–fold sums. Hereafter for convenience $\mu_Z = E(Z_X) = 0$, and the curls $\sim$ on Z_X are omitted.

Assumptions and Stationarity

Assumptions are kept at minimum. Random shock variables are i.i.d. random variables with E $(a_X) = 0$ and finite variance $\sigma^2_a > 0$. Also a_X is independent of Z_X unless they have the same location. Let $l' = (l_1, l_2, \ldots, l_m)$ be a vector of spatial lags, so E $(a_X Z_X) = \sigma^2_a$ if $l = 0$. Weak stationarity is assumed.

The autocovariance and autocorrelation functions are defined as

$$\gamma_l = E(Z_X Z_{X+l})$$

$$p_l = \gamma_l / \gamma_0$$

where $\gamma_0 = \sigma^2_z$, variance of process.

Weak stationarity assumption allows the description of a random field by its mean, variance, and autocorrelation functions. γ_l and p_l are function of l rather than X, and are symmetric to the origin $X = (0, 0, \ldots, 0)$. That is, if $m = 1$, $\rho_{l_1} = \rho_{-l_1}$, and if $m = 2$, $\rho_{l_1, l_2} = \rho_{-l_1, -l_2}$; $\rho_{-l_1, l_2} = \rho_{l_1, -l_2}$; $\rho_{0, l_2} = \rho_{0, -l_2}$; $\rho_{l_1, 0} = \rho_{-l_1, 0}$; $\rho_{0, 0} = 1$

Of course the autocorrelation matrix of a process is positive definite. This and symmetricity to origin determine the range of values of autocorrelation coefficients.

The general model for ARMA processes which is of order (p, q; u, v) may be expressed in terms of backward of forward shift operators B_X; and F_X, such that $B_X Z_{X_i} = Z_{X_i} - \delta_i$, where $\delta_i = (\delta_{i_1}, \delta_{i_2}, \ldots, \delta_{i_m})$ and $\delta_{ij} = 1$ if $i = j$, zero otherwise. Note that $B^{-1} = F$. The model in (1) expressed in terms of F is

$$\left(1 - \sum_{n=-p}^{q} \Phi_n F_X^n\right) Z_X = \left(1 - \sum_{n=-v}^{v} N \quad \Theta_n F_X^n\right) a_X \tag{3}$$

or

$$\varnothing\ (B_X)\ (Z_X) = \theta\ (B_X)\ a_X \tag{4}$$

where $\varnothing\ (B_X)$ and $\theta\ (B_X)$ are characteristic equations for the AR and MA parts, respectively. The model can be expressed as an infinite AR process, as

$$a_X = \sum_{j=o}^{\infty} \left(\sum_{n=-v}^{v} \theta_n F_X^n\right)^j \left(1 - \sum_{n=-p}^{q} \varnothing_n F_X^n\right) Z_X \tag{5}$$

or, as an infinite MA process

$$Z_x = \sum_{j=o}^{\infty} \left(\sum_{n=-p}^{q} \varnothing_n F_x^n \right)^j \left(1 - \sum_{n=-v}^{v} \theta_n F_x^n \right) a_x \tag{6}$$

The autocovariance of Z_x is

$$\gamma(B_x) = \sigma_a^2 \left[\left(\sum_{j=o}^{\infty} \left(\sum_{n=-p}^{q} \varnothing_n F_x^a \right)^j \right) \left(1 - \sum_{n=-v}^{v} \theta_n F_x^n \right) \right]$$

$$= \sigma_a^2 \ \psi \ (B_x) \ \theta \ (B_x) \tag{7}$$

Stationarity of the process is ensured if the infinite series ψ (B_x) θ (B_x) is convergent, since σ_a^2 is finite, given the convergence, γ $(B_x) < \infty$. This condition introduces restrictions on the model parameters. To see this, consider a one dimensional ARMA model

$$z_{x_1} - \varnothing_{-p} Z_{x1-p} - \varnothing_{-p+1} Z_{x1+1-p} - \dots - \varnothing_q Z_{x1+q} =$$

$$= a_{x_1} - \theta_{-u} a_{x_1-u} - \theta_{-u+1} a_{x_1+1-u} - \dots - \theta_v a_{x1+v}$$

$$\left(1 - \sum_{n^1=-p}^{q} \varnothing_{n_1} F_{x_1}{}^{n^1} \right) Z_{x_1} = \left(1 - \sum_{n^1=-u}^{v} \theta_{n_1} F_{x_1}{}^{n^1} \right) a_{x_1} \tag{8}$$

For stationarity the roots of the associated polynomial to $\varnothing$ (B_{x_1}) must meet certain requirements. Multiplying $\varnothing$ $(B_{x\,1})$ in (8) by B^p we obtain

$$B^p - \varnothing_{-p} B^{p-1} - \varnothing_{-p+1} B^{p-2} \dots \varnothing_{+pq-1} B - \varnothing_{p+q}$$

whose associated polynomial is

$$w^k + \varnothing_1 w^{k-1} + \dots + \varnothing_k \tag{9}$$

where $k = p + q$. The same treatment can be applied to the MA part in (8). Assume, without loss of generality , that the polynomials (9) and

$$w^h + \Theta_1 w^{h-1} + \dots + \theta_h \tag{10}$$

have no common factor, $h = u + v$. If the process is to be weakly stationary all the roots of (9) must lie inside the unit circle, while none of the roots of (10) lies outside the unit circle, Rudin (1969) and Marden (1949). The roots of the polynomial $\sum_{j=o}^{k} \varnothing_j w^{k-j} = 0$, $\varnothing_0 = 1$

are obtained by reducing it to real and complex linear factors, such that $(w-w_1)$ $(w-w_2)$ $\dots$ $(w-w_k) = 0$, so the condition set for (9) requires the satisfaction of $| w_j | < 1$, $j = 1, 2, \dots, k$. Define $y = (w + 1)/(w - 1)$ where $w = a + bi$, a and b are real and $i = \sqrt{-1}$. It is found that $y = ((a - 1)^2 + b^2)^{-1}$ $(a^2 + b^2 - 1 - 2ib)$, or y

$= R(y) - C(y)$ with $R(y) = (a^2 + b^2 - 1)\,((a-1)^2 + b^2)^{-1}$ and $C(y) = (2ib)\,((a-1)^2 + b^2))^{-1}$, where $|w| = (a^2 + b^2)^{1/2}$ and $(a-1)^2 + b^2 > 0$. It is necessary and sufficient for $|w| < 1$ that $R(y) < 0$ since $w = (y+1)/(y-1)$, and $|w_j| < 1$ is satisfied only if $R(y_j) < 0$, $j = 1, 2, \ldots k$. Substituting $w = (y+1)/(y-1)$ into $\sum_{j=o}^{k} \varnothing_j w^{k-j} = 0$ we obtain

$$(y+1)^k + \varnothing_1 (y+1)^{k-1}(y-1) + \ldots + \varnothing_k (y-1) = 0 \tag{11}$$

the roots of (11) are y_j. Expressing (11) in ascending powers of y it is obtained that

$$s_o + s_1 y + s_2 y^2 + \ldots + s_k y^k = 0 \tag{12}$$

where $s_o = 1$ and $s_r = \sum_{j=o}^{k} \varnothing_j C_{rj}$. C_{rj} is the coefficient of y^r in $(y+1)^{k-j}(y-1)^j$ in (11). For instance, if $p + q = 2$ so $k = 2$, given $p = q$, $C_{00} = C_{02} = C_{2j} = 1$, $j = 1, 2, 3$, $C_{01} = -1$, $C_{10} = C_{12} = 2$, $C_{11} = 0$. The stationarity condition in terms of s_i's of (12) can be set for by the Routh criteria, Routh (1930), Bellman and Cookse (1963). Set matrix

$$S = \begin{bmatrix} s_1 & s_3 & s_5 & \ldots \\ s_o & s_2 & s_4 & \ldots \\ 0 & s_1 & s_3 & \ldots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{bmatrix} \tag{13}$$

the condition $|w_j| < 1$ is satisfied if the principal minors of S, s_r' $r = 1, 2, \ldots, k-1$, are positive definite. For $k = 2$, $s_1 = (2\,(1 - \varnothing_1)/(1 - \varnothing_1 + \varnothing_2)) > 0$ and $s_2 = ((1 + \varnothing_1 + \varnothing_2)/(1 - \varnothing_1 + \varnothing_2)) > 0$, it follows that $-\varnothing_1 > -1$; $-(\varnothing_1 + \varnothing_2) > -1$; $-(\varnothing_1 + \varnothing_1) < 1$, or satisfaction of all three in one condition $|\varnothing_1| + |\varnothing_2| < 1$. That is, the power series in (7) converges if the terms $(\varnothing_1 B + \varnothing_{-1} F)$ of the series for the AR part, $p = q = 1$, satisfy $|\varnothing_1 B + \varnothing_{-1} F| < 1$; convergence on $\overset{1}{\underset{i=-1}{X}}$ s_i where $s_{-i} = (B_{x_1}: |B_{x_1}| \leq 1)$ and $s_i = (F_{x_1}: |F_{x_1}| \geq 1)$. The MA part of the model needs restrictions on θ's only for the invertibility of the model to an infinite AR process. A similar argument as above proceeds for the polynomial in (10). If $u = v = 1$, then it is found that the restriction is $|\theta_1| + |\theta_2| < 1$.

The autocovariance generating function of the ARMA process represented by (1) is given in (7). The autocovariance function can also be obtained by multiplying equation (1) by $Z_{x-\iota}$, and taking expectations

$$\gamma_\iota = \sum_{n=-p}^{q} \varnothing_n \gamma_{\iota+n} \cdots - \sum_{n=-u}^{v} \theta_n \gamma_{z,a} \ldots (\iota+n) + \gamma_{z,a}(\iota) \qquad (14)$$

where $\gamma_{z,a}(\iota+n) = E(Z_{x+n}\, a_x)$ which exists if $\iota + n = x$ since Z_x and a_x are not correlated if they do not have the same location; $\gamma_{z,a}(\iota) = 0$ whenever $\iota > 0$. When $\iota = 0$, (14) reduces to

$$\gamma_o = \sum_{n=-p}^{q} \varnothing_n \gamma_n - \sum_{n=-u}^{v} \gamma_{z,a}(n) + \sigma_a^2 \qquad (15)$$

the variance of the process. From these, $\rho_\iota = \gamma_\iota / \gamma_o$ A two dimensional, $m = 2$, ARMA process with second order MA and AR parts, ARMA (2, 2: 0, 2; 0, 2) is represented by the model

$$Z_{x_1, x_2} = \varnothing_1 (Z_{x_1, x_2} + Z_{x_1, x_{2+1}})$$
$$+ a_{x_1, x_2} - \theta_{o,-1}\, a_{x_1, x_{2-1}} - \theta_{o,1}\, a_{x_1, x_2} + 1$$

The autocovariance function of the model is

$$\gamma_{\iota_1,\iota_2} = \varnothing_1 (\gamma_{\iota_1,\iota_{2-1}} + \gamma_{\iota_1,\iota_{2+1}}) - \theta_{o,-1}\, \gamma_{z,a}(\iota_1, \iota_2)$$
$$- \theta_{o,1}\, \gamma_{z,a}(\iota_1, \iota_2 + 1) + \gamma_{z,a}(\iota_1, \iota_2)$$

Then the variance is

$$\gamma_{o,o} = \varnothing_1 (\gamma_{o,-1} + \gamma_{o,1}) - \theta_{o,-1}\, \gamma_{za}(0, -1) - \theta_{o\,2}\ \gamma_{z,a}(0,1) + \sigma_a^2.$$

The autocovariances needed to estimate the model parameters are $\gamma_{o,1}$, $\gamma_{1,o}$, and $\gamma_{1,1}$ whose functions can be obtained from the function for γ_{l_1,l_2}. Dividing these by $\gamma_{o,o}$ and substituting sample estimates of autocorrelations in the resulting set of equations for $\rho_{1,o}$, $\rho_{o,1}$, and $\rho_{1,1}$, the Yule–Walker estimates of the parameters are obtained. The cross covariances are $\gamma_{z,a}(0, 0) = \sigma^2$, $\gamma_{z,a}(0, 1) = (\varnothing_1 - \theta_{o,1})\, \sigma^2_a$, $\gamma_{z,a}(0, -1) = (\Phi_1 - \Theta_{o,-1})\, \sigma^2_a$, others nonexisting because of the cutoff property of autocovariances for the MA part.

The model expressed as an MA process is

$$z_{x_1,x_2} = (1 - \theta_{o,-1}\, B_{x_2}{}^{-1} - \theta_{o,1}\, B_{x_2}) . \left(\sum_{j=o}^{\infty} \sum_{k=o}^{j} \binom{j}{k} \varnothing^{j}{}_{1}\, B_{x_2}{}^{j-2k}\, a_{x_1,x_2} \right)$$

which has finite variance if $| \varnothing_1 | < 1/2$, the condition for stationarity. The same model as an AR process is

$$a_{x_1,x_2} = (1 - \varnothing_1 (B_{x_2} + B_{x_2}{}^{-1})) \left(\sum_{j=o}^{\infty} \sum_{k=o}^{j} \binom{j}{k} \theta^{j-k}{}_{o,-1}\, \theta^{k}{}_{o,1}\, z_{x_1,x_2-j-2k} \right)$$

For the invertibility the condition on the parameters is $| \theta_{o,1} | + | \theta_{o,-1} | < 1$.

Identification and Estimation

The first step in the fit of an ARMA family model to an observed m–dimensional process is to determine the values of m, p, q, u, and v. Although the process is seemingly m–dimensional, the dimension of an appropriate model can be $m' < m$. Nonzero values of $p = (p_1, p_2, \ldots, p_m)$, $q = (q_1, q_2, \ldots, q_m)$, $u = (u_1, u_2, \ldots, u_m)$, and $v = (v_1, v_2, \ldots, v_m)$ can be determined by the cut–off properties of autocorrelation function, a.c.f., and partial autocorrelation function, p.a.c.f. To express sample autocorrelation let N_i denote the number of equidistant points on x_i, $i = 1, 2, \ldots, m$, so there are $\prod_{i=1}^{m} N_i$ observations. Then the sample autocorrelation is $r_l = c_l / c_o$ where $c_l = (1/\Pi N_i) \sum_{x=1}^{N-l} (Z_x . Z_{x+l})$ and $C_o = (1/\Pi\ N_i) \sum_{x=1}^{N-2} Z_x{}^{2'}$ where $N = (N_1, N_2, \ldots, N_m)$. r_l, c_l, and c_o are unbiased estimators of ρ_l, γ_l, and γ_o respectively, Jenkins and Watts (1968).

Bartlett (1946) suggests an approach to the determination of variance and covariance estimates of autocorrelations if it is shown that the sample observations on a random variable or their sum are from a normal process. A dependent sequence $\{Z_x\}$ should meet certain regularity conditions and assumptions to have central limit property, Serfling (1968, 1980). If $\{Z_x\}$ is considered as ordered like $Z_{1,\,1,\,\ldots,\,1}, \ldots, Z_{N_1, N_2, \ldots, N_m}$ and denoted by Z_n, $n = ((1, 1, \ldots, 1), (2, 1, \ldots, 1), \ldots, (N_1, N_2, \ldots, N_m))$, assumptions can be stated as

Constant Mean : $E(Z_n) \equiv 0$

Unifoım Convergence of
Normed Sum Expectation : $E\left(b^{-1/2} \sum_{a+1}^{a+b} Z_n\right)^2 = A^2 > 0$ (16)

Uniform Boundedness : $E \mid Z_n \mid^{2+\delta} \leq M$

where $b \in |0, {}_iIIN_i|$, $\delta > 0$ and $M < \infty$.

Under these assumptions the sequence $\{Z_x\}$ has central limit property, that is,

$$\lim_{b \to \infty} \{bA)^{-1/2} \sum_{1}^{b} Z_n < z\} = (2\pi)^{-1/2} \int_{\infty}^{z} \exp(-t^2/2)\, dt.$$

The first assumption is satisfied. Second assumption is satisfied if $\sum_{l=0}^{\infty} c_l$ converges and $E \mid Z_n \mid^{2+\delta} \leq M < \infty$ for some l, c_l is the autocovariance estimator. We can see that

$$E\left(b^{-1/2} \sum_{a+1}^{a+b} Z_n\right)^2 = E(H_a) = b^{-1} \sum_{a+1}^{a+b} E(Z_n^2)$$

$$+ 2b^{-1} \sum_{i=1}^{b-1} \sum_{j=i+1}^{b} E(Z_{n+i} + Z_{n+j})$$

$$= c_o + 2b^{-1} \sum_{i=1}^{b-1} \sum_{j=i+1}^{b} c_{j-i}$$

$$= c_o + 2b^{-1} \sum_{i=1}^{b-1} c_i$$

where $c_i = \sum_{j=1}^{i-1} cj$. $\{c_b\}$ converges to $\sum_{j=1}^{\infty} c_j$, so $b^{-1} \sum_{i=1}^{b} c_i$ converges to $\sum_{j=1}^{\infty} c_j$, therefore the second assumption in (16) holds.

To show that the third assumption in (16) holds is equivalent to saying that

$$E \mid E\,(H_a \mid \zeta \mid \leq T_1\,(b) \tag{17}$$

$$E \mid E\,(H_{a^2}, \mid \zeta) - E\,(H_{a^2}) \mid \leq T_2\,(b) \tag{18}$$

are satisfied, where ζ is the sigma algebra generated by Z_X.A condition on the moments of sums $\sum_{a+1}^{a+b} Z_n$ is $E \mid H_a \mid^{2+\varepsilon} = 0\,(b^{-\alpha})$, uniformly, as $b \to \infty$ for some $\varepsilon > 0$ and $\alpha > 0$. Under assumption (16) this condition holds when $\alpha = 0$ for $\varepsilon = 0$, and $\alpha = 1 + (1/2)\,\varepsilon$ for $0 < \varepsilon < \delta$, $\delta > 0$. By this, T_1 (b) and T_2 (b) equal to $0\,(b^{-\alpha})$, so mean deviation expressed by (17) and (18) are bounded uniformly to the observed spatial series. Since Z_n is asymptotically normal, AN (μ, σ_n^2), a real valued function of Z's, differentiable with $g'\,(z) \neq 0$, also has the property that $g\,(Z_n)$ is AN $(g(z = \mu), [g'\,(Z = \mu)\,]^2 \sigma_n^2)$. Now suppose that $Z_n = (Z_{1,\,1}, \ldots_1 \ldots, Z_{N_1,\,N_2} \ldots, {}_{Nm})$, if there are $N = \Pi_i N_i$ observations $Z_n = (Z_1, Z_2, \ldots, Z_N)$. If Z_n is AN $(\mu, b_n^2 \Sigma)$ where μ is mean vector and Σ is covariance matrix, $g\,(Z_n) = (g_1(Z_n), \ldots, g_s\,(Z_n))$ a real valued function of Z_n which has nonzero differentials at $z = \mu$, define $D = [\partial g_i / \partial Z_j \mid_{z=\mu}]$, an s x N matrix, then

$$g\,(Z_n) \text{ is AN } \Big(g(\mu),\; N^{-1} \sum_{j=1}^{N} \sigma_{ii}\;(\partial g \,/\, \partial Z_i \mid_{z=\mu}) \cdot (\partial g \,/\, \partial z_j \mid_{z=\mu})\Big).$$

These results can be applied to the sample autocorrelation. Let (Z_x, Z_{x+l}) be asymptotically normally distributed pairs. Then $r_l = \hat{\rho}_l = g\,(V)$ where $V = (\bar{Z}_x, \bar{Z}_{x+l}, N^{-1} \Sigma\, Z_x^2, N^{-1} \Sigma\, Z_{x+l}^2, N^{-1} \Sigma\, Z_x Z_{x+l})$. So, we can write $g\,(y_1, y_2, y_3, y_4, y_5) = (y_5 - y_1\, y_2) \cdot [(y_3 - y_1^2)^{1/2}\, (y_4 - y_2^2)^{1/2}]^{-1}$. The V vector is AN $(E\,(V), N^{-1}\,\Sigma)$ where Σ is 5 x 5 covariance matrix of $(Z_x, Z_{x+l}, Z_x^2, Z_{x+l}^2, Z_x Z_{x+l})$. From asymptotic distribution of g (Z_N), shown above, it follows that r_l is AN $(\rho_l, N^{-1}\, d\, \Sigma\, d')$ where $d = (\partial g / \partial y_1 \mid_{y=E(V)}, \ldots, \partial g / \partial y_5 \mid_{y=E(V)})$. By this finding, significance of r_l values can be determined for the purpose of determination of dimensions and orders of dependency at each dimension for the MA part. For a two dimensional process utilizing asymptotic distribution of r_l we can write

$$\begin{aligned} \mathrm{Var}\,(r_{l_1,\,l_2}) = (N_1 N_2)^{-1} \sum_{V_1=-K_1}^{K_1} \sum_{V_2=-K_2}^{K} [\rho_{V_1,O}\, \rho_{O,V_2} \\ + \rho_{l_1+v_1,\, l_2+v_2}\, \rho_{l_1+v_1,\, l_2-v_2}\, \rho_{l_1-v_1,\, l_2+v_2} \\ \rho_{l_1-v_1,\, l_2-v_2} + \rho_{l_1,\, l_2}^2\, (\rho_{v_1,\, v_2} + (1/2)\, \rho_{v_1,\, O}^2 \end{aligned} \tag{19}$$

$$+ (1/2)\ \rho^2_{o,v_2}) - 2\rho_{\iota_1, \iota_2} (\rho_{v_1,o}\ \rho_{v_1+\iota_1, v_2+\iota_2}$$
$$+ \rho_{-v_1, -v_2}\ \rho_{o, v_2+\iota_2})]$$

where K is a large number such that $K > N_k$, $i = 1, 2, \ldots, m$, and $(N_1\ N_2)$ should be replaced by $\Pi_i (N_i - \iota_i)$ if N_i's are small, or if ι_i's are large. For $\iota_i > q'$, q' a sufficiently large number, $r_l = 0$, then (19) reduces to

$$\mathrm{Var}\ (r_{\iota_1, \iota_2}) = (N_1\ N_2)^{-1} \left(1 + 2 \sum_{v_1=-q'_1}^{q'_1} \sum_{v_2=-q'_2}^{q'_2} \rho_{v_1, o}\rho_{o, v_2}\right).$$

Cut–off property of a.c.f. can be stated that if a spatial process is purely MA, then all r_l vanishes for those with $l < u$ and $v, u = (u_1, u_2, \ldots, u_m)$, $v = (v_1, v_2, \ldots, v_m)$.

Utilizing the partial autocorrelation function, p.a.c.f., the dimensions and order of dependencies at each dimension for AR part can be determined. This is by the cut–off property of p.a.c.f. As Cramer (1946), Lawrence (1976, 1979), and Hannan (1970) discussed , the partial autocorrelation of, say, $z_{x_1, x_2, \ldots, xm}$ and $Z_{x_1, x_2+l_2, x_3, \ldots, xm}$ detects the correlation between the two which is not due to the linear dependence of both on the intervening values $Z_{x_1, x_{2+1}, x_3, \ldots, xm}, \ldots, Z_{x_1, x_2+l_2-1, \ldots xm}$. If dependence of $Z_{x_1, x_2+l_2, \ldots, xm}$ on intervening values is defined as the best linear estimate in the expected mean square sense

$$E(Z_{x+l} - Z'_{x+l})^2 = E\ (Z_{x+l} - \alpha_1\ Z_{x+l-1} - \ldots - \alpha_l\ Z_{x+1})$$

$$Z'_{x+l} = \alpha_1\ Z_{x+l-1} + \ldots + \alpha_l\ Z_{x+1} \tag{20}$$

where $\iota = (0, \iota_2, 0, \ldots, 0)$, m–element lag vector, and α_i's are mean squares regression coefficients, autocorrelation equations are written from (20). They are

$$\rho_i = \alpha_1\rho_i + \alpha_2\rho_{i-1} + \ldots + \alpha_{l-1}\ \rho_{l-i-1} \tag{21}$$

$1 < i \leq l - 1$. So α_i's are functions of l and i. If ρ_i equations, as in (21) are expressed in matrix form

$$\tilde{\rho}_{l-1} = \tilde{P}_{l-1}\ \tilde{\alpha}_{l-1} \tag{22}$$

with $\tilde{\rho}_{l-1} = (\rho_1, \rho_2, \ldots, \rho_{l-1})$, $\tilde{\alpha}_{l-1} = (\alpha_1, \alpha_2, \ldots, \alpha_{l-1})$ and

$$\widetilde{P}_{l-1} = \begin{bmatrix} \rho_0 & \rho_1 & \rho_2 \cdots & \rho_{l-1} \\ \rho_1 & \rho_0 & \rho_2 \cdots & \rho_{l-3} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \rho_{l-2} & \rho_{l-3} & \rho_{l-4} & \rho_0 \end{bmatrix}$$

$l - i = (0, l_2 - i, 0, \ldots, 0)$, $\rho_j = \rho_{0, j, 0, \ldots, 0}$ $\alpha_j = \alpha_{0, j, 0, \ldots, 0}$

Equation (22) yields solution to α_{l-1}

$$\tilde{\alpha}_{l-1} = \widetilde{P}^{-1}_{l-1}\ \tilde{\rho}_{l-1} \tag{23}$$

It can be verified that the partial variance is Var $(Z_{x+l} - Z'_{x+1})$ Var $(Z_x - Z_x') = 1 - \alpha_1\rho_1 - \ldots - \alpha_{l-1}\ \rho_{l-1}$ and partial covariance is Cov $(Z_{x+l} - Z'_{x+l})\ ((Z_x - Z'_x) = \rho_l - \alpha_1\ \rho_{l-1} - \ldots - \alpha_{l-1}\ \rho_1$ By definition, then, the partial autocorrelation function is

$$w_\iota = (\rho_1 - \alpha_1\rho_{1-1} - \ldots - \alpha_{\iota-1}\rho_1)\ (1 - \alpha_1\rho_1 - \ldots - \alpha_{\iota-1}\rho_{\iota-1})^{-1} \tag{24}$$

where α_i, as suggested by (23) ,is

$\alpha_i =$ Det ($\widetilde{P}_{\iota-1}$ with ith column replaced by $\tilde{\rho}_{\iota-1}$). (Det $(\widetilde{P}_{\iota-1}))^{-1}$ and $w_\iota = w_{0, l_2, 0, \ldots, 0}$ The same procedure should follow for all other $w_{\iota_1, \iota_2, \ldots, \iota_m}$.

There is an important matter about the estimation of p.a.c.f. Determination of intervening variables must be made in such a way that the representation of an m–dimensional process should be in linear equations, as in (20) which reflects the dependency of variables of concern on the intervening variables, Gebizlioğlu (1981). For instance, if $m = 2$, drawing a line between two points on the plane, and taking the variables at those points falling on the line as the only intervening variables would be erroneous.

If $\theta_n = 0$ in (1) the resulting stochastic difference equation is an AR process model with order $p_i + q_i$ on each x_i. The cut–off property of p.a.c.f. is that the stationary spatial series (Z_x) is from an AR process of order $p_i + q_i$ in each x_i if its p.a.c.f.'s are zero beyond $p_i + q_i$, $i = 1, 2, \ldots, m$. To see this consider the Hilbert Space of real random variables Z_x with zero mean and finite second order moments, with expected product as inner product, Cramer and Leadbetter (1967), and let $H_{k \cdot x}$ be the subspace spanned by

$(Z_{k+1}, \ldots, Z_{x-1})$ for $X > k + i$, $k = (k_1, k_2, \ldots, k_m)$, $X = (x_1, x_2, \ldots, x_m)$. Let Z^*_k and Z^*_x be the respective projections of Z_k and Z_x on $H_{k, x}$. Consider the AR model

$$Z_X = \sum_{n=-p}^{a} \varnothing_n \, Z_{X+n} = a_X$$

This equation represents a good fit if it denotes the unique decomposition of Z_X into the sum of its projection on an orthogonal distance, a_X to $H_{X-(p+q)-1,X}$. By assumption $E(a_X) = 0$ and $E(a^2_X) > 0$. To establish that $a_X \perp a_k$ for all $k \neq X$ define H^*_X as the subspace spanned by all a_k for $k < X$. Then $H^*_X = UH_{X-k,X}$. $a_X \perp H_{X-(p+q)-1,X}$ by construction. According to $H_{X-(p+q)-1}$ we can write

$$Z_{X-(p+q)-1} = Z^*_{X-(p+q)-1} + W_{X-(p+q)-1}'$$

then $a_X \perp Z_{X-(p+q)-1}$ and $a_X \perp W_{X-(p+q)-1}'$ because p.a.c.f. $w_{X-(p+q)-1} = 0$, partial autocorrelation coefficients of order higher than the actual models' order on each dimension X_i. Therefore $a_X \perp Z_{X-(p+q)-1}'$ implying that

$a_X \perp H_{X-(p+q)-2,X}$

Utilizing the results of Ouenoille (1947, 1949, 1958), Jenkins (1954, 1956), and Daniels (1956), and central limit property, it can be shown that w_l are asymptotically normally distributed random variables with variance $\mathrm{Var}(w_l) = (\prod_{i=1}^{m} (N_k - l_1))^{-1}$. Two different models of AR type fit to the same series can be tested for the goodness of fit by approximate χ^2 distribution. Let $a = (s_1, s_2, \ldots)$ stand for the dimensions of a model, asymptotically

$$Ts = \sum_{j=1}^{k} (N - s)\,(w_{s(j)} - w_{s(\cdot)}) \sim \chi^2_{k-1} \tag{25}$$

where $j = p+q$, $k>N$, N is the total number of observations,

$s = \prod_i s_i$, $s_i \neq 0$, and $w_{s(\cdot)} = \sum_{j=1}^{k} w_{s(j)}/(N - s)$.

For low order w_l, $\mathrm{Var}(w_l) \simeq N^{-1}$. Assume there are $N_1 N_2 = 100$ observations on Z_{X_1, X_2}, and it is found that $w_{2,0} = .6638$. $(N^{-1})^{1/2} = 0.1$, so $w_{2,0}$ is beyond three standard errors of w_l. If other w_l's are not significant, then the suggested model belongs to a second order bilateral AR model in two dimensions,

$$Z_{X_1,X_2} = \phi_{2,0} Z_{X_1+2,X_2} + \phi_{-2,0}\, Z_{-X_1-2,X_2} + a_{X_1,X_2} \quad ,$$

given that autocorrelations die out rapidly and are not significant. If both autocorrelations and partial autocorrelations are significant at some spatial lags, then an ARMA model is suggested. Orders of AR and MA parts are found as explained above. Another way of looking at this problem is by the spectral density of the process prepresented by(1),

$$f(\lambda) = 2^{-1} (h(e^{i\lambda}))^{-1} g(e^{i\lambda}) Gg(e^{i\lambda}) (h^*(e^{i\lambda}))^{-1} \tag{26}$$

where $i = \sqrt{-1}$, h^* is the conjugate, G is the covariance matrix of a_x found by (5), and

$$h(e^{i\lambda}) = I - \sum_{n=-p}^{q} \phi_n (e^{i\lambda})^n \tag{27}$$

$$g(e^{i\lambda}) = I + \sum_{n=-u}^{v} \theta_n (e^{i\lambda})^n.$$

To make an adequate fit some conditions on ϕ_n, θ_n and G must be imposed as Rosanov (1967) and Hannan (1969) suggested. These conditions are (i) that, G is nonsingular, (ii) that, $\Gamma(e^{i\lambda}) = (h(e^{i\lambda}))^{-1} g(e^{i\lambda})$ is analytic and nonsingular in the unit circle, (iii) that the polynominal matrices $g(e^{i\lambda})$ and $h(e^{i\lambda})$ are nonzero and have no common left factor other than matrices with constant determinants.

If an ARMA model is identified by determination of p, q, u, and v, the next step is the estimation of parameters ϕ_n and θ_n. Assume that the process under consideration is

$$Z_{x_1,x_2} = \phi_1 Z_{x_1-1,x_2} + \phi_2 Z_{x_1+1,x_2} + \theta a_{x_1-1,x_2+1} + a_{x_1,x_2} \tag{28}$$

and there are $N = N_1 N_2$ observations on the plane with layout

$$\begin{bmatrix} Z_{1,1} & Z_{1,2} & \cdots & Z_{1,N_1} \\ Z_{2,1} & Z_{2,2} & & Z_{2,N_1} \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ Z_{N_2,1} & Z_{N_2,2} & & Z_{N_2,N_1} \end{bmatrix} = \begin{bmatrix} Z_1 & Z_2 & \cdots & Z_{N_1} \\ Z_{N_1+1} & Z_{N_1+2} & \cdots & Z_{2N_2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ Z_{(N_2-1)N_1+1} & & \cdots & Z_{N_1N_2} \end{bmatrix} = \underset{Z(N)}{} \tag{29}$$

Now, let $Z_n = (Z_1, Z_2, \ldots, Z_{N_1}.., \ldots, Z_{N_1N_2})$, then the N x N matrix of $\varnothing$ parameters will be

$$\begin{bmatrix} 0 & \phi_2 & 0 \dots & 0 & 0 & 0 \\ \phi_1 & 0 & \phi_2 \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \phi_1 & 0 & \phi_2 \\ 0 & 0 & 0 & 0 & \phi_1 & 0 \end{bmatrix} = \phi\ (N) \tag{30}$$

Similarly, letting $a_n = (a_1, a_2, \dots, a_{N_1}, a_{N_1} + 1, \dots a_{N_1 N_2})$, the matrix of parameters will be

$$\left[\begin{array}{ccc} & 0 & \\ \hline \theta_1^2 & 0 & 0 \\ 0 & \theta_2^2 & 0 \\ \cdot & \cdot & \\ \cdot & \cdot & \\ \cdot & \cdot & \\ 0 & 0 & \dots\ \theta^2_k \end{array}\right] = \left[\begin{array}{c} \theta^1(N) \\ \hline \\ \theta^2(N) \\ \\ \end{array}\right] = \theta\ (N) \tag{31}$$

where θ^1 (N) is N_1 x N null matrix, and θ^2 (N) is $(N-N_1)$ x N matrix whose submatrices $\theta^2{}_j$ are N_1 x N_1 matrices, $j = 1, \dots, k$, $k = N_2-1$, of the form

$$\begin{bmatrix} 0 & 0 & 0 \dots & 0 & 0 \\ \theta & 0 & 0 \dots & 0 & 0 \\ 0 & \theta & 0 \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \dots & \theta & 0 \end{bmatrix} \tag{32}$$

The process expressed in (28) can be re-expressed as

$$Z_n = \phi\ (N)\ Z_n + \theta\ (N) a_n + a_n \tag{33}$$

Let $\beta = (\theta, \phi)$ be the vector of parameters and G is unknown.

We will demand that the parameters β lie within a parameter space $\tilde{\Omega}$ determined by the conditions that the roots of characteristic equations ϕ (B_x) and θ (B_x) should satisfy convergence criteria mentioned in section (2.1). This assures that $\{a_x\}$ series can be recovered as shown in (5). For a true model $Z_x - \hat{Z}_x = a_x$, $\hat{Z}_x$ is the predicted value of random variable Z'_x and a_x will have the ergodicity conditions shown by Hannan (1970). Assuming joint normality of a_n, the likelihood of parametes β and G is

$$L(\beta, G) = (2\pi)^{-N/2} \mid G \mid^{N/2} \exp \left[- \frac{1}{2} \left(\sum_{n=1}^{N} a_n \; G^{-1} a_n \right) \right] \tag{34}$$

where $\mid G \mid$ is the determinant of G, the covariance matrix of a_n. The logarithm of the likelihood is

$$LL(\beta, G) = -(N/2) \log (2\pi) - (1/2) \, NF(\beta, G) \tag{35}$$

where F (β, G) is the objective function to be minimized with respect to elements of β and G. Although this objective function is motivated by the normality assumtion, it may be used when this assumption is not valid. LS estimation is such a case in which expected sum of squares of residuals are minimized. Another method of estimation, minimum variance prediction error, MVPE, yields equivalent estimates by minimizing the variance of $\varepsilon_x = Z_x - \hat{Z}_x$, $\varepsilon_x = a_x$ if the model is the correct one.

The derivatives of F with respect to β_k, $k = 1, 2, \ldots, e$, e is the number of parametes, are

$$2N^{-1} \sum_{n=1}^{N} a_n^k G^{-1} a_n \tag{36}$$

and the derivatives with respect to G^{-1} are

$$- G_{l,s} + N^{-1} \sum_{n=1}^{N} \qquad a_l a_s \tag{37}$$

where $G_{l,s} = E\,(a_l a_s)$.
The zero of (37) for fixed β is obtained as

$$G_{l,s} = N^{-1} \sum_{n=1}^{N} a_l a_s \tag{38}$$

This is the conditional estimate of G, and is denoted by $\tilde{G}$ which is the sample variance matrix of a_x and

$$\tilde{G} = \tilde{G}(\beta) = N^{-1} \sum_{n=1}^{N} a_n a_n'. \tag{39}$$

Using this the second term in F (β, G) becomes $N^{-1} \sum_{n=1}^{N}$

$a_n' G^{-1} a_n = \text{trace}\,(G^{-1}\tilde{G})$, so

$$F\,(\beta\; G) = \log\,|\,G\,| + \text{trace}\,(G^{-1}\tilde{G}) \tag{40}$$

and inf F $(\beta, G) = \log\,|G| + y$, then β is obtained by minimizing F

(β, G) with respect to β only. That is $|\,\tilde{G}\,| = |\,N^{-1} \sum_{n=1}^{N} a_n a_n'\,|$

can be used as the objective function. This result depicts the equivalence of MVPE and ML, LS estimations.

The conditional estimates of β are also found by solving

$$N^{-1} \sum_{n=1}^{N} a_n^k G^{-1} a_n = 0 \tag{41}$$

Linearizing a_n about some chosen point β_o we can write

$$a_n\,(\beta) = a_n\,(\beta_o) + \sum_{k=1}^{e} a_n^k\, \delta\beta + 0\,\|\,\delta\beta\,\| \tag{42}$$

where $\delta\beta = \beta - \beta_o$. Ignoring the term $\|\delta\beta\|$ the linear equation for $\delta\beta$ is $A\delta\beta = -f$, where A is a k x k matrix with elements $A_{ks} = N^{-1}$

$\sum_{n=1}^{N} a_n^k G^{-1} a^s{}_n$, and f is a half of the first derivative vector of F with

respect to β; $f_k = N^{-1} \sum_{n=1}^{N} a_n^{k\prime}\, G^{-1} a_n$. If the ARMA model in (1)

contains no MA terms, the corrected parameters $\beta_1 = \beta_o + \delta\beta$ would give the solution to (41). Otherwise, replace β_o by β_1 and repeat linearization until sequence $\beta_o, \beta_1, \beta_2 \ldots$ converges to conditional estimate $\beta = \beta\,(G)$.

To estimate β and G simultaneously set $G_t = G\,(\beta_t)$ and $\beta_{t+1} = \beta\,(G)$, $t = 0, 1, \ldots$ is the number of steps, then perform conditional minimization at each step until prameters β_t, G_t converge to their overall minimum values $\hat{\beta}$, $\hat{\sigma}_a^2$. Newton–Raphson or Margquardt (1963) estimation methods should be used in computations. Yule–Walker estimates of β can be used as starting values of β in the iteration.

It can be shown that estimates $\hat{\beta}$ obtained by ML methods are consistent and have asympotic joint normal distributions with mean β and convariance matrix $N^{-1}C^{-1}$ where C is the matrix with elements $C_{ks} = (1/2)\ E\ (\partial^2 F/\partial\beta_k\beta s)$ evaluated at β. The estimates $\hat{\beta}$ are asymptotically uncorrelated with $\hat{\sigma}^2_a$ which consistently estimate σ^2_a. Accuracy of the estimates as well as consistency can be shown, for instance, by setting Cramer–Rao bounds, Wilks (1962); Zacks (1970). Fisher information matrix, which will be shown in the next section, provides bounds on the estimates.

AR AND MA MODELS

The general AR model which represents an m–dimensional autoregressive process in space is

$$Z_x = \sum_{n=-p}^{q} \Phi_n Z_{x+n} + a_x \tag{43}$$

whose characteristic equation is $\Phi\ (B_x) = 1 - \sum_{n=-p}^{q} \Phi_n F_x^{n}$, so (43)

is re–expressed as

$$\varnothing(B_x)\ Z_x = a_x. \tag{44}$$

Restrictions on the values of Φ are needed for stationarity, that is, convergence of power series $\psi\ (B_x)$ in

$$Z_x = \Phi^{-1}\ (B_x)\ a_x = \left(\sum_{j=o}^{\infty} \sum_{n=-p}^{q} \Phi_n F_x^n\right)^j a_x = \psi\ (B_x)\ a_x \tag{45}$$

requires that $|\ \Sigma\ \Phi_n F_x^n\ | < 1$, $\Phi_o = 0$. Given that $\psi\ (B_x)$ converges, the variance of the process

$$\Gamma(B_x) = \sigma_a^2\ \psi\ (B_x) \tag{46}$$

is finite. $E\ (Z_x Z_{x+l}) = \gamma_\iota = \sigma^2_z \rho_\iota$, where γ_ι can be obtained from the autocovariance generating function

$$\Gamma(B_x) = \sum_{C=-\infty}^{\infty} \gamma_c B_x^c \tag{47}$$

where γ_c is the coefficient of both B_x and B_x^{-c}. Therefore theautocovariance al lag ι is symmetric to the origin, so is $\rho\iota = \gamma\iota/\gamma_o$

Consider a simple bilateral , one dimensional AR process.

$$Z_{x_1} = \Phi_1 Z_{x_{1-1}} + \Phi_2 Z_{x_{1+1}} + a_x \tag{48}$$

Multiply (48) with $Z_{x_1+l_1}$ and take expectations, the first two auto-correlation functions, as a result, are

$$\rho_{-1} = \phi_1 + \phi_2\rho_2$$
$$\rho_1 = \phi_1\rho_2 + \phi_2$$

Yule–Walker equations. From these we obtain that $\phi_1 = (1-\rho_2^2)^{-1}(\rho_{-1}-\rho_2\rho_1)$, $\phi_2 = (1-\rho_2^2)^{-1}(\rho_1-\rho_2\rho_{-1})$. Since $\rho_{-1} = \rho_1$ by weak stationarity $\phi_1 = \phi_2 = \phi$. Yule–Walker estimates are asymptotically LS estimates which are

$$\hat{\hat{\phi}}_1 = \left[\sum_{n=1}^{N_1} Z_n Z_{n-1} \sum_{n=1}^{N_1} Z_{n+1}^2 - \left(\sum_{n=1}^{N_1} Z_{n+1}Z_{n-1}\right)\left(\sum_{n=1}^{N_1} Z_n Z_{n+1}\right)\right].$$

$$\left[\sum_{n=1}^{N} Z_{n+1}^2 \sum_{n=1}^{N} Z_{n-1}^2 - \left(\sum_{n=1}^{N} Z_{n-1}Z_{n+1}\right)^2\right]^{-1}$$

$$\hat{\phi}_1 = \left[\sum_{n=1}^{N_1} Z_n Z_{n+1} \sum_{n=1}^{N_1} Z_{n-1}^2 - \left(\sum_{n=1}^{N_1} Z_{n+1}Z_{n-1}\right)\left(\sum_{n=1}^{N_1} Z_n Z_{n-1}\right)\right].$$

$$\left[\sum_{n=1}^{N_1} Z^2_{n+1} \sum_{n=1}^{N_1} Z^2_{n-1} - \left(\sum_{n=1}^{N_1} Z_{n-1}Z_{n+1}\right)\right]^{-1}.$$

Asymptotically $\hat{\hat{\phi}}_1 = \hat{\phi}_2$. Therefore AR model in (48) should actually be

$$Z_{x_1} = \phi\,(x_{1-1} + Z_{x_{1+1}}) + a_x\,. \tag{49}$$

This property extends to all AR models. MA representation of (49) is

$$Z_{x_1} = \sum_{s=0}^{\infty} \sum_{k=0}^{s} \binom{s}{k} \phi^s a_{x_1+2k-s} \tag{50}$$

The expected mean square for a finite j is

$$E\left(Z_{x_1} - \sum_{s=0}^{j-1} \sum_{k=0}^{s} \binom{s}{k} \phi^s a_{x_1+2k-s}\right)^2 = E\left(\sum_{s=0}^{j} \binom{j}{s} \phi^j Z_{x_1-j+2s}\right)^2$$

where $\sum_{s=0}^{j} \binom{j}{s} \phi^j = (2\phi)^j$ and as $j \to \infty$ convergence requires $|\phi| < 1/2$. Further restrictions on $\varnothing$ are introduced by the condition that the autocorrelation matrix of the process

$$\begin{bmatrix} \rho_0 & \rho_2 & \rho_1 \\ \rho_2 & \rho_0 & \rho_1 \\ \rho_{-1} & \rho_1 & \rho_0 \end{bmatrix} \tag{51}$$

is positive definite: $\rho_0 = 1$, $\rho_{-1} = \rho_1$, so $|\rho_2| < 1$, $(1 + 2\rho_2\rho^2_1) > (2\rho^2_1 + \rho^2{}_2)$. Apply these to Yule–Walker equations and determine further restrictions on $\varnothing$. The autocovariance and autocorrelation function for (49) are $\gamma_\iota = \varnothing\ (\gamma_{\iota-1} + \gamma_{\iota+1})$ $\gamma_0 = \sigma^2_a\ (1\text{–}2\ \ \varnothing\rho_1)^{-1}$, and $\rho_\iota = \varnothing\ (\rho_{\iota-1} + \rho_{\iota+1})$. From the homogenous second order difference equation $\rho_\iota - \varnothing\ (\rho_{\iota-1} + \rho_{\iota+1}) = 0$ the values of $\varnothing$ can be determined. The general solution to the difference equation is $\rho_\iota = c_1\ (k_1)^\iota + c_2\ (k_2)^\iota$ with

$$k_1 = 2^{-1}\ ((1/\varnothing) + ((1/\varnothing^2) - 4)^{1/2})$$
$$k_2 = 2^{-1}\ ((1/\varnothing) - ((1/\varnothing^2) - 4)\ ^{1/2})$$

and $c_1 = (k_2\text{–}r_1)/(k_2\text{–}k_1)$, $c_2 = (r_1\text{–}k_1)/k_2\text{–}k_1)$, with $r_1 = \hat{\rho}^2$. From c_1 and c_2, $\varnothing$ is found. The ψ weights in (45) can be found. Express (45) as

$$Z_{x_1} = \sum_{j=-\infty}^{\infty} \psi_j a_x\ \psi_j\ a_{x_1+}{}^{j'}\ \psi_0 = 1 + \varnothing\ (\psi_1 + \psi_{-1}),\ \psi_j =$$

$$\varnothing\ (\psi_{j-1} + \psi_{j+1}),\ \psi_{-j} = \varnothing\ (\psi_{-j-2} + \psi_{-j+1}),\ \psi_s = \psi_{-s} \text{ then}$$

$$\psi_0 = \sum_{j=0}^{\infty} \binom{2j}{j}\ \varnothing\ 2^j$$

$$\psi_j = \sum_{i=t}^{\infty} \binom{2i\text{-}1}{i\text{-}t/2}\ \varnothing^{2i},\ \text{i is even integer,}$$

$$\psi_j = \sum_{i=t}^{\infty} \binom{2i\text{-}1}{i-(t/2+.5)}\ \varnothing\ ^{2i-1},\ \text{i is odd integer.}$$

By the simple spatial AR process in (48) all properties, including Yule–Walker and LS estimation, are shown for spatial AR models. This finding extends to all other AR models in all dimensions. To discuss ML *estimation of* AR *parameters,* $\varnothing$, *consider a two dimensional* multilateral process (AR (2, 2: 1, 1;1,1)

$$Z_{x_1,x_2} = \varnothing_1(Z_{x_1-1,x_2-1} + Z_{x_1+1,x_2+1})$$
$$+ \varnothing_2\ (Z_{x_1-1,x_2+1} + Z_{x_2+1,x_2-1}) + a_{x_1,x_2} \tag{52}$$

with variance $\sigma_z^2 = \gamma_0 = \sigma_a^2 (1-2\rho_{1,1}\phi_1-2\rho_{-1,1}\phi_2)$, and Yule–Walker equations

$$\rho_{l_1,l_2} = \phi_1 (\rho_{l_1-1},\ _{l_2-1} + \rho_{l_1+1},\ _{l_2+1}) + \phi_2 (\rho_{l_1-1},\ _{l_2+1} + \rho_{l_1+1,l_2-1}) \tag{53}$$

$\iota_1 = \iota_2 \neq 0$. Yule–Walker estimates of ϕ_2, ϕ_2 are

$$\begin{aligned} \phi_1 &= A^{-1}(\rho_{1,1} (\rho_{2,2} + 1) - \rho_{1,-1} (\rho_{0,2} + \rho_{1,c})) \\ \phi_2 &= A^{-1} (\rho_{1,-1} (\rho_{-2,2} + 1) - (\rho_{0,2} + \rho_{2,0})^2) \\ A &= (1 + \rho_{2,2}) (1 + \rho_{-2,2}) - (\rho_{0,2} + \rho_{2,0}\ ), \end{aligned} \tag{54}$$

ρ_ι should be replaced by sample estimates $r_\iota = \hat{\rho}_\iota$.

Assume there are $N = N_1N_2$ observations on the plane, and a_n are i.i.d. normal random variables with zero mean, and σ_a^2 variance, with joint density

$$f\ (a) = (2\pi\sigma^2{}_a)^{N-/2} \exp \left(- \sum_{n=1}^{N} a^2{}_n /2\sigma_a{}^2\right) \tag{55}$$

The likelihood function, then, is

$$L\ (Z) = (2\pi\sigma^2{}_a)^{-N/2} \exp\ (-M /z\sigma_a{}^2) \tag{56}$$

where $M = \sum_{x_1=1}^{N1} \sum_{x_2=1}^{N2} (Z_{x_1},\ _{x_2} - \hat{Z}_{x_1},\ _{x_2})$, $\hat{Z}_{x_1},\ _{x_2}$ is estimated Z

by the model in (52). The log–likelihod

$$LL = -\ (N/2) \log\ (2\pi\sigma_a^2) - (M/2\sigma_a^2) \tag{57}$$

yields score equations

$$\partial LL/\partial\phi_1 = -\sigma_a{}^{-2} \sum_{x_1=1}^{N1} \sum_{x_1=1}^{N2} (Z_{x_1}\ _{-1,x_2-1} + Z_{x_1 1+},\ _{x_2+1})\ M$$

$$\partial LL/\partial\phi_2 = -\sigma_a{}^{-2} \sum_{x_1=1}^{N1} \sum_{x_1=1}^{N2} (Z_{x_1-1},\ _{x_2+1} + Z_{x_1+1},\ _{x_2-1})\ M.$$

Equalizing these to zero yields

$$r_{1,1} = \phi_1 (r_{0,0} + r_{\cdot,3}) + \phi_2 (r_{0,2} + r_{2,0})$$

$$r_{-1,2} = \phi_{\cdot} (r_{0,2} + r_{2,2}) + \phi\ (r_{0,2} + r_{-2,2})$$

which in turn yields estimates $\hat{\phi}_1$ and $\hat{\phi}_2$. Note that ML estimates and Yule–Walker estimates are approximately equal, so are LS estimates.

The variance and covariance of $\hat{\phi}_1$ and $\hat{\phi}_2$ are found from the Fisher information matrix $- [E (S (Z_{x_1,x_2} \mid \Omega))]$ where S is a 2 x 2 matrix with elements $S_{11} = \partial LL/\partial^2\phi_1$, $S_{22} = \partial LL/\partial^2\phi_2$, $S_{12} = S_{21} = \partial LL/\partial\phi_1\partial\phi_2$.

Let $S^{-1} = - [E (S (Z_{x_1,x_2} \mid \Omega))]^{-1}$, then

$$S^{-1}_{11} = \mathrm{Var}\ (\hat{\phi}_1) = \sigma^2_a\ (1 + r_{-2,2})\ T^{-1}$$

$$S^{-1}_{22} = \mathrm{Var}\ (\hat{\phi}_2) = \sigma^2_a\ (1 + r_{2,2})\ T^{-1}$$

$$S^{-1}_{12} = S^{-1}_{21} = \mathrm{Cov}\ (\hat{\phi}_1\ \hat{\phi}^2) = - \sigma^2_a\ (r_{o,2} + r_{2,o})\ T^{-1}$$

$$T = N_1N_2\sigma^2_z\ [(1 + r_{2,2}))\ (1 + r_{-2,2}) - (r_{o,2} + r_{2,o})^2].$$

$\hat{\sigma}^2_a$ is obtained from the relation $\sigma^2_a/\sigma^2_z = (1-2r_{1,1}\ \phi_1 - 2r_{-1,1}\phi_2)^{-1}$. The general MA model for an m-dimensional spatial MA process is

$$Z_x = \sum_{n=-p}^{q} \theta_n a_{x+n} + a_x'\ \theta_o = 0. \tag{58}$$

The simplest one dimensional, bilateral MA model is MA (2: 1,1)

with variance $\gamma_o = \sigma^2_z = \sigma^2_a\ (1 + \theta^2_{-1} + \theta^2_1)$, and autocorrelation functions to obtain θ

$$\rho_1 = \rho_{-1} = - (\theta_{-1} + \theta_1)\ (1 + \theta^2_{-1} + \theta^2_1)^{-1}$$

$$\rho_2 = \rho_{-2} = - (\theta_{-1}\theta_1)\ (1 + \theta^2_{-1} + \theta^2_1)^{-1},$$

$\rho_{l_1} = \rho_{-l_1} = 0$ if $|\ l_1| > 2$ by the cut-off property of autocorrelation for MA models. Using the autocorrelation function (s) the parameter θ can be estimated by MVPE method, by replacing ρ_l with their sample estimates, r_l.

The characteristic equation of (59) is $\theta\ (B_x) = 1-\theta_{-1}\ B-\theta_1\ B^{-1}$, so the model can be expressed as an infinite AR process, if the invertibility condition $|\ \theta_{-1}\ | + |\ \theta_1\ | < 1$ is satisfied,

$$a_x = \theta^{-1}\ (B_x) = \pi\ (B_x)\ Z_x$$

$$= \sum_{j=o}^{\infty} (\theta_{-1}B_{x_1} + \theta_1 F_{x_1})\ ^j Z_{x_1} \tag{60}$$

$$= \pi_o Z_x + \pi_{-1} Z_{x-1} + \pi_1 Z_{x_1} + \dots$$

To see the ML estimation method for MA models consider, for simplicity, $\theta_{-1} = \theta_1 = \theta$ in (59) so $\sigma^2_z = \sigma^2_a\,(1 + 2\theta^2)$, $\rho_1 = -2\theta(1 + 2\theta^2)^{-1}$, $\rho_2 = \theta^2\,(1 + 2\theta^2)$, thus $\theta = 2\rho_2/\rho_1$, and invertibility condition is $|\,\theta\,| < 1/2$. Assume there are N observations on the line; then the model $Z_{x_1} = -\ \theta\ (a_{x_1-1} + a_{x_1+1}) + a_{x_1}$ in matrix form is

$$Z_{x_1} = (I - \theta Y)a_{x_1} = Va_{x_1} \tag{61}$$

where $Z_{x_1} = (Z_1, Z_2, \ldots, Z_n)$, $a_{x_1} = (a_1, a_2, \ldots, a_N)$, I is N x N identitiy matrix and Y is N x N matrix of form

$$\begin{bmatrix} 1 & -1 & 0 \ldots & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \ldots\ldots & -1 & 1 & -1 \\ 0 & 0 & 0 \ldots\ldots & 0 & -1 & 1 \end{bmatrix}$$

If a's are i.i.d normal random variables with zero mean, and variance σ^2_a, the likelihood of Z_{x_1} is

$$L = (2\pi\sigma^2_a)^{-N/2} \exp\,[(-Z_x\,(VV')^{-1}Z_{x_1})\,(\sigma^2_a)^{-1}]\quad |\,V\,|^{-1} \tag{62}$$

where $|\,V\,|^{-1}$ is the Jacobian of transformation from a_x to Z_x' since $\varepsilon_x = a_{x_1} = Z_{x_1} - \hat{Z}_{x_1}$ and $\hat{Z}_{x_1}$ is predicted Z_{x_1} by (61). Note that $\varepsilon_{x_1} = a_{x_1}$ if the model is the right identified one. To evaluate $|\,V\,|$, and inverse of the covariance matrix of Z_x eigenvalues of Y can be used. If the model is correct, all eigenvalues are distinct, so all eigenvectors are unique. Then $Y = UKQ$, where K is the diagonal matrix of eigenvalues, U is the matrix of eigenvectors and Q is a matrix such that $Q^{-1} = U$. By Cayley–Hamilton theorem $V = U\,(I-\theta K)\,Q = UCQ$. Then the term log $|\,V\,|$ in log of the likelihood in (62),

$$LL = -(N/2)\,(\log\,(2\pi) + \log\sigma^2_\varepsilon) - \log\,|V| - (Z'_{x_1}\,(VV')^{-1}Z_{x_1}) \tag{63}$$

is simplified as

$$\log\ |\,V\,| = \sum_{i=1}^{N}\ (\log n_i) = \sum_{i-1}^{N}\ \log\,(1 + \theta\lambda_i)$$

where n_i is the ith element of C and λ_i is the i th diagonal element of K. Then, the simplification of $Z'_{x_1}\,(VV')^{-1}Z_{x_1}$ is

$$Z'_{x_1}\,(VV')^{-1}Z_{x_1} = (QZ_{x_1})'\,(I-H)\,U'U\,(I-H)\,(QZ_{x_1}) \tag{64}$$

where $H = I - C^{-1}$ is a diagonal matrix with elements $h_i = 1 - (1 + \theta\lambda_i)^{-1}$

After these simplifications on (62) the resulting equation can be evaluated at those values of θ which meet invertibility conditions. It is found that the maximum likelihood estimator of σ_a^2 is $\hat{\sigma}_a^2 = N^{-1} Z'_{x_1} (VV')^{-1} Z_{x_1}'$ and substituting (64) into (63) we obtain

$$LL = 2 \log |V| + N \log (Z_{x_1}' (VV')^{-1} Z_{x_1}) + K$$

where K is the constant of log–likelihood. ML estimator of θ is that value of θ which minimizes

$$LL = N\log \{ |V|^{2/N} Z'_{x_1} (VV')^{-1} Z_{x_1} \} + K.$$

To find that value, Newton–Raphson iterative procedure can be utilized, Gebizlioğlu (1981). Properties and procedures of estimation for higher order and higher dimensional MA models are similar. Patterns of matrices of spatial observations and matrices like Y in (61), and $\varnothing$ (N) and θ (N) in (30) and (31) make it harder to do simplification applied to likelihood and log–likelihood functions for higher order AR, ARMA models as well as MA models. Power spectrum of AR and MA models can be used as tools to detect the wavelengths measured with respect to spatial dimensions, that is, periodicity in several dimensions.

A two dimensional AR process, say AR (1, 1: 0, 1; 0, 1)

$$Z_{x_1,x_2} = \phi_{0,0} Z_{x_1,x_2} + \phi_{1,0} Z_{x_{1+1},x_1} + \phi_{0,1} Z_{x_1,x_{2+1}} + a_{x_1,x_2}$$

with $\sigma_z^2 = \sigma_a^2 (1-\phi_{0,0}\, \rho_{0,0} - \phi_{1,0}\, \rho_{1,0} - \phi_{0,1}\, \rho_{0,1})$, has power spectrum $P = 2\sigma_a^2 \, | 1-\phi_{0,0}-\phi_{1,0} e^{-i2\Pi f} - \phi_{0,1} e^{-i2\Pi f} |^{-1/2}$, $|f| < 0.5$, where $i = \sqrt{-1}$, and $| 1-\phi_{0,0}-\phi_{1,0} e^{-i2\Pi f} - \phi_{0,1} e^{-i2\Pi f} |^{-1/2}$ is obtained by substituting $B_{x_1,x_2} = e^{i2\Pi f}$ in the characteristic equation of the model, $|f| < .5$. Similarly the power spectrum of an MA model

$$Z_{x_1,x_2} = \theta_{0,1}\, a_{x_1,x_{2+1}} + \theta_{1,0}\, a_{x_{1+1},x_2} + a_{x_1,x_2}$$

is

$$P = 2\sigma_a^2 \, | 1-\theta_{0,1}\, e^{-i2\Pi f} - \theta_{1,0} e^{-i2\Pi f} |^2.$$

For a process observed in two dimensions, the spectral estimate is

$$P_{s,t} = (4\Pi)^{-1} \sum_{l1=-L}^{L} \sum_{l2=-L}^{L} C_{l1,l2}\, g_{l1,l2} \cos\{(\Pi / F) (t_{l1} + s_{l2})\}$$

for grequencies –.5 to .5 cycles for each sampling interval . t, s = –F. – F + 1, ..., 0, 1, ... F. F is the number of frequency estimeates.

L is the maximum lag in both dimensions, and C_{l_1, l_2} is the sample covariance estimate, g_{l_1, l_2} is a smoothing function. Adapting Bartlett's window, Jenkins and Watt (1968), to two dimensions $g_{l_1, l_2} = 1-h/L$ if $0 > h > L$, zero otherwise, where $h = (_{l1}{}^2 + {}_{l2}{}^2)^{1/2}$ distance between points (x_1, x_2) and (x_{1+l1}, x_{2+l2}). The power spectrum definitions and spectral estimates given here can easily be extended to highei dimensions and order.

VALIDATION OF MODELS

Among the several procedures which do not require specification of an alternative hypothesis, one is the validation tests based on comparison of various characteiistics of models an data. As discussed above, spontaneously, comparison of autocovariance and autocorrelations of data with those of the models derived theoretically indicates if the fit is reasonable Similarly, theoretical spectra of fitted models which are compatible with spectral estimates of data reveal if the fit is correct.

A validation test with null and an alternative hypothesis is based on the residuals. Suppose that Z_X is postulated to depend on Z_{X+l}. The exact form of this dependency is hypothesized to be characterized by an m–dimensional family of functions, that is, ARMA models of appropriate class. Let IF denote such a family of functions with M parametes. The hypothesized class of models are such that $E(Z_X) = IF(X_{Xl+}, \tilde{\Omega})$ where $\tilde{\Omega}$ is the space of M parametes. To test the hypothesis, a sample is used and estimators of parameters $\tilde{\Omega} = (\varnothing_1, \varnothing_2, \ldots,. \varnothing_k, \theta_1, \ldots \theta_y)$, $y + k = M$, are obtained. Then $\hat{IF} = IF(Z_{X+l}; \hat{\tilde{\Omega}})$ is the estimated Z_X function. If $E(\hat{IF}) = IF(Z_{X+l}{}' \tilde{\Omega})$ then the model is adequate, that is, if $e_X = Z_X - E(IF)$ then the hypothesis, actually, is $E(e_X) = 0$. Under the stated assumptions and regularity conditions (see Section 2), and with the assumption that $e_X \sim AN(0, \sigma^2{}_a)$, then $R^2 = 1-(\sigma^2{}_a/Var(Z_X))$ can be used to test that the right model is $IF(Z_{X+l}, \tilde{\Omega})$, against it is $IF(Z_{X+i} \tilde{\Omega}_S)$. The underlying theory is asymptotic and rests on the likelihood ratio tests. Let $\tilde{\Omega}_S$ be a subset of $\tilde{\Omega}$. Consider $\tilde{\Omega}_S = 0$, or $\tilde{\Omega}_S = \tilde{\Omega}_S^*$, a specified set of values of s number of parameters, $s \leq M$. Denote the maximum value of likelihood function with respect to $\tilde{\Omega}$ by $L(\tilde{\Omega})$. $L(\tilde{\Omega}_S)$ is the maximum value of the likelihood with respect to $\tilde{\Omega}_S$. The ratio $L(\tilde{\Omega}_S)/L(\tilde{\Omega})$

provides a measure on how well Ω_S fits the observations against another fit with $\tilde{\Omega}$. The corresponding F– test is with the statistic (N–K–1) $C^{-2}(R^2-R_S^2)$ $(1-R^1)$ which is approximately F distributed with parameters (C, N–K–1), where K is the number of parmeters in a fitted model, while C=K–s is the difference between the number of parametres in $\tilde{\Omega}$ and $\tilde{\Omega}$s.

For a large number of observations, N, the statistic

$$\lambda = -2\log\ L(\tilde{\Omega}_S)/L(\tilde{\Omega})$$

is distributed as χ^2 distribution. If λ is high, then the fit with $\tilde{\Omega}$ is not good, so the hypothesis is rejected.

If e_X's are approximately normal then log-likelihoods will be dominated by $\Sigma_X\, e_X\,(\Omega)^2$, where $e_X\,(\Omega)$ is the residuals at location x. Then

$$\lambda = (\Sigma_X\ e_X\ (\tilde{\Omega}_S)^2 - \Sigma_X\ e_X\ (\tilde{\Omega})^2)\sigma^2{}_a \tag{65}$$

which directly leads to

$$\lambda = (R^2 - R_S^2)\ [\sigma_a^2\ \ \Sigma_X\ \ (Z_X - \bar{Z}_X)]^{-1} \tag{66}$$

where $R^2 = 1 - (\Sigma_X\, e_X\, (\Omega)^2\, (\Sigma_X\, (Z_X - \bar{Z}_X))^{-2}$

and $R_S^2 = 1 - (\Sigma_X\ e_X\ (\tilde{\Omega}_S)^2)(\Sigma_X\,(Z_X - \hat{Z}_X))^{-2}$.

If the estimated value of σ_a^2, which can be obtained by ML estimation, is substituted in (66) then

$$\lambda = N(R^2 - R_S^2)\ (1 - R^2)^{-1}.$$

The corresponding F-test is with the statistic (N–K–1) C^{-1} $(R^2–R^2)$ $(1–R^2)$ which is approximately F distributed with parameters (C, N–K–1), where K is the number of parameters in fitted model, while C= K–s is the difference between the number of parameters in $\tilde{\Omega}$ and $\tilde{\Omega}_S$.

CONCLUSION

Purely spatial ARMA family models are discussed with emphasis on correlational properties, weak stationarity, and estimation. Central limit properties for spatial series are established, so that existing estimation and validation techniques are valid. It is shown that, in the estimation distributional properties are not needed, and ML, LS and Yule–Walker estimates are approximately convergent.

In many experimental situations spatial data may be far from meeting stationarity assumptions. In these cases one can utilize the suggesttions, among others, of Patankar (1954), Norcliffe (1977), and Mitchell (1974) to remove the trend along spatial axes, and to stabilize the variance. Introduction of differencing filters into the models should necessarily be undertaken to analyze nonstationary spatial data.

REFERENCES

AROIAN, L.A. (1979). Multivariate autoregressive time series in M.Dimensions. Proceedings of the American Statistical Association, 558–590.

AROIAN, L.A. (1980). Time series in M-Dimensions: definitions, problems, and prospects. *Communications in Statistics*, B9, 5, 453–465.

AROIAN, L.A. (1981). Time series in M-Dimensians: maximum likelihood estimation for autoregressive models. Monograph AES 8109, Union College and University, New York.

AROIAN, L.A. and GEBIZLİOĞLU, Ö.L. (1980). Time series in M-Dimensions: Spatial Models. Proceedings of the American Statistical Association, 320–325.

BARTLETT, M.S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society*, B8, 27–41.

BARTLETT, M.S. (1 74). The statistical analysis of spatial pattern. *Advances in Applied Probability*, 6, 336–358.

BATCHELOR, L.D. and REED, H.S. (1918). Relation of the variability of yields of fruit trees to the accuracy of field trials. *Journal of Agricultural Research*, 12, 245–283.

BELLMAN, R. and COOKE, K.L. (1963). *Differential-Difference Equations.* Academic Press, New York.

BENNETT, R.J. (1979). *Spatial Time Series: Analysis, Forecasting, Control.* Pion, London.

BESAG, J.E. (1972). On the correlational structure of two-dimensional stationary processes. *Biometrika*, 59, 43–58.

BESAG, J.E. (1947). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Stdtistical Society*, B36, 192–236.

CLIFF, A.D. and ORD, J.K. (1973). *Spatial Autocorrelation.* Pion, London.

CLIFF, A.D. and ORD, J.K. (1975). Model building and analysis of spatial pattern in human geography. *Journal of the Royal Statistical Society*, B37, 297–384.

CRAMER, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton.

CRAMER, H. and LEADBETTER, M.R. (1967). *Stationary and Related Stochastic Processes.* Wiley, New York.

DANIELS, H.E. (1956).The approximate distribution of serial correlation coefficients.*Biometrikal*, 43, 169–185.

GEBIZLİOĞLU, Ö.L. (1981). Time Series in M-Dimensions: Spatial Models. Ph.D. Dissertation, Union College and University, New York.

GEBİZLİOGLU, O.L. (1982). Multidimelsional random processes and spatial ARMA family models. *METU Studies in Development* (Special Issue), 19–54.

HANNAN, E.J. (1969). The identification of vector mixed auto-regressive maving average systems. *Biometrika*, 56, 223–225.

HANNAN, E.J. (1970). *Multiple Time Series.* Wiley, New York.

HEINE, V. (1955). Models for two-dimensional stationary stochastic processes. *Biömetrika*, 42, 170–128.

JENKINS, G.M. (1954). Tests of hypothesis in linear autoregressive model I. *Biömetrika*, 41, 405–419.

JENKINS, G.M. (1956). Tests of hypothesis in linear autoregressive model II. *Biometrika*, 43, 186–199.

JENKINS, G.M. and WATTS, D.G. (1968). *Spectral Analysis and Its Applications.* Holden-Day, San Francisco.

LAWRENCE, A.J. (1976). On conditional and partial correlation. *The American Statistician*, 70, 146–149.

LAWRENCE, A.J. (1979). Partial and multiple correlation sor time series. *The American Statisticion*, 33, 127–130.

MARDEN, M. (1949). The Geometry of Zeros of a Polynamial in a Camplex Variable. *American Mathematical Society*, New York.

MARQUARDT, D.W. (1963). An algorithm for least squares estimation of non-linear parameters. *Journal of the Society of Industrial and Applied Mathematics*, 11, No. 2.

MERCER, W.B. and HALL, A.D. (1911). The experimental error of field trials. *Journal of Agricultural Science*, 4, 107–132.

MITCHELL, B. (1974). Three approaches to resolving problems arising from assumption violation during statistical analysis in geographical research. *Chaiers de Geographie de Quebec*, 18, 507–523.

NORCLIFFE, G. (1977). *Inferential Statistics for Geographers.* Hutchinson, London.

OPRIAN, C.A., TANEJA, V., VOSS, D.A., AROIAN, L.A. (1980). General considerations and interrelationships between MA and AR models; time series in M-dimensions, the ARMA model. *Communications in Statistics*, B9, 5, 515–532.

ORD, J.K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70, 120–126.

PATANKAR, V.N. (1954). The goodness of fit of frequency distributions obtained from stochactic processes. *Biometrika*, 41, 450–462.

QUENOUILLE, M.H. (1947). A large sample test for the goodness of fit of autoregressive schemes. *Journal of the Royal Statistictisal Society*, A110, 823–129.

QUENOUILLE, M.H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society*, B11, 68–84.

QUENOUILLE, M.H. (1958). The camparison of correlations in time series. *Journal of the Royal Statististical Society*, B20, 158–164.

RIPLEY, B.D. (1981). *Spatial Statistics.* Wiley, New York.

ROZANOV, Y.A. (1967). *Stationary Random Processes.* Holden-Day, San Frdncisco.

ROUTH, E.J. (1930). *Dynamics of System of Rigid Bodies.* MacMillan, London.

RUDIN, W. (1969). *Function Theory in Polydiscs.* Benjamin, New York.

SERFLING, R.J. (1968). Contributions to central limit theory for dependent variables. *The Annals of Mathematical Statistics,* 39, 1158-1125.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

TANEJA, V.S. and AROIAN, L.A. (1980). Time series in M-dimensions: autoregressive models. *Communications in Statistics,* B9, 5, 491–513.

VOSS, D.A., OPRIAN, C.A., AROIAN, L.A. (1980). Moving average models: time series in M- dimensions. *Communications in Statistics,* B9, 5, 467–489.

WHITTLE, P. (1954). On stationary processes in the plane. *Biometrika,* 41, 434–449.

WHITTLE, P. (1963). Stochastic processes in several dimensions. *Bulletin of the International Statistical Institute,* 40(1), 974–994.

WILKS, S.S. (1962). *Mathematical Statistics.* Wiley, New York.

ZACKS, S. (1970). *The Theory of Statistical Inference.* Wiley, New York.