A NOTE ON THE PROPERTIES OF SPATIAL YULE-WALKER ESTIMATORS

by

Sabyasachi Basu
Southern Methodist University

Gregory C. Reinsel
University of Wisconsin-Madison

Technical Report No. SMU/DS/TR-250
Department of Statistical Science

January 1991

A Note on the Properties of Spatial Yule-Walker Estimators

Sabyasachi Basu[1] and Gregory C. Reinsel[2]

[1]Department of Statistical Science, Southern Methodist Univerity, Dallas, TX 75275

[2]Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706

Summary

For two-dimensional spatial autoregressive (AR) models, asymptotic properties of the spatial Yule-Walker (YW) estimators (Tjostheim, 1978) are studied. These estimators although consistent, are shown to be asymptotically biased. Estimators from the first-order spatial bilateral AR model are looked at in more detail and the spatial YW estimators for this model are compared with the exact maximum likelihood estimators. Small sample properties of both estimators are also discussed briefly and some simulation results are presented.

Key Words: Spatial AR Models; Yule-Walker Estimator; Maximum Likelihood Estimator; Asymptotic Results.

1. Introduction

The analysis of spatial processes has received much attention in recent years and has been studied in disciplines such as geography, geology, biology and agriculture. Many of the developments have been summarized in the books by Bartlett (1975), Ripley (1981), Cliff and Ord (1981), and Upton and Fingleton (1985). We consider spatial processes on a regular rectangular grid of size $m \times n$ defined in two dimensions with sites labeled (i,j), with an associated random variable Y_{ij} defined at each site. Examples of such phenomena include data collected from satellites and from agricultural field trials.

Tjostheim (1978, 1983) examined in detail a special case of the bilateral AR models, as defined by Whittle (1954), where the value Y_{ij} defined at each site (i,j) is a (finite) autoregression on the values at the sites which are in the lower quadrant of (i,j). In two dimensions the spatial autoregressive model of order (p_1,p_2) becomes

$$Y_{ij} = \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} \, Y_{i-k,j-l} \; + \; \epsilon_{ij} \tag{1}$$

with the convention that $\alpha_{00} = 0$. The ϵ_{ij} are a collection of independent random variables with $E(\epsilon_{ij}) = 0$ and $Var(\epsilon_{ij}) = \sigma^2$. The observations Y_{ij}, $i=1,\ldots,m$, $j=1,\ldots,n$ are available for the estimation of α_{kl}. Tjostheim (1978) considered spatial Yule-Walker (YW) estimators and proved that the asymptotic distribution of the spatial YW estimators from model (1) is Gaussian. There is an error in the proof, and hence the asymptotic distribution of the spatial YW estimator, although Gaussian, contains an asymptotic bias term. We first study this asymptotic bias of the spatial YW estimator in more detail, and discuss the consequences for small to moderate grid sizes. The results are also compared with the recently proposed exact maximum likelihood estimator of Basu and Reinsel (1990) for the special case of the first-order spatial bilateral AR model.

2. Asymptotic Bias of Spatial Yule-Walker Estimators

For $s \geq 0$ and $t \geq 0$, define

$$R(s,t) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} Y_{ij} \, Y_{i+s,j+t} \qquad \text{and} \qquad R(s,-t) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=t+1}^{n} Y_{ij} \, Y_{i+s,j-t} \; ,$$

which are the sample covariances at lags (s,t) and $(s,-t)$, respectively. Note that $R(s,t) = R(-s,-t)$ and $R(-s,t) = R(s,-t)$. Guyon (1982) compares $R(s,t)$ with the unbiased estimator of $\gamma(s,t) = E(Y_{ij} Y_{i+s,j+t})$, $R_1(s,t) = mn \; R(s,t)/(m\text{-}s)(n\text{-}t)$, and claimed that the unbiased estimator is preferred. As one of the examples, Guyon (1982) considered the YW estimator for a spatial model with one parameter and showed that the estimator was asymptotically biased if $R(s, \; t)$ is used, but the bias vanishes when $R_1(s,t)$ is used instead.

For $0 \leq s \leq p_1$, $0 \leq t \leq p_2$ and $(s,t) \neq (0,0)$, using (1) we can write

$$R(s,t) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} Y_{ij} \, \epsilon_{i+s,j+t} \; + \; \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} \left(\sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} \, Y_{ij} \, Y_{i+s-k,j+t-l} \right)$$

$$= A(s,t) + \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} \, B(s,t,k,l) \tag{2}$$

where $A(s,t) = (mn)^{-1} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} Y_{ij} \, \epsilon_{i+s,j+t}$ and

$$B(s,t,k,l) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} Y_{ij} \, Y_{i+s-k,j+t-l} \; = \; R(s-k,t-l) \; + \; R^*(s,t,k,l), \tag{3}$$

with $R(s-k,t-l) = \{\sum_{i=i^*}^{m^*} \sum_{j=j^*}^{n^*} Y_{ij}\, Y_{i+s-k,\, j+t-l}\}/mn$, $i^* = \max(1,k-s+1)$, $m^* = \min(m,m-s+k)$, $j^* = \max(1,l-t+1)$, and $n^* = \min(n,n-t+l)$. So, using (3), (2) can be written as

$$R(s,t) = \frac{1}{mn}\sum_{i=1}^{m-s}\sum_{j=1}^{n-t} Y_{ij}\,\epsilon_{i+s,j+t} + \sum_{k=0}^{p_1}\sum_{l=0}^{p_2} \alpha_{kl}\, R(s-k,t-l) + \sum_{k=0}^{p_1}\sum_{l=0}^{p_2} R^*(s,t,k,l).$$

Define $\alpha = (\alpha_{01},\ldots, \alpha_{0,p_2}, \alpha_{01},\ldots, \alpha_{1,p_2},\ldots, \alpha_{p_1,0},\ldots, \alpha_{p_1,p_2})'$,

$$r = (R(0,1),\ldots, R(0, p_2), R(1,0),\ldots, R(1, p_2),\ldots, R(p_1,0),\ldots, R(p_1, p_2))',$$

$$A = (A(0,1),\ldots, A(0, p_2), A(1,0),\ldots, A(1, p_2),\ldots, A(p_1,0),\ldots, A(p_1, p_2))',$$

$$\mathrm{R} = \begin{bmatrix} R(0,0) & R(0,1) & . & . & . & R(-p_1,1-p_2) \\ R(0,1) & R(0,0) & . & . & . & R(-p_1,2-p_2) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ R(p_1,p_2-1) & R(p_1,p_2-2) & . & . & . & R(0,0) \end{bmatrix}$$

$$\mathrm{R}^* = \begin{bmatrix} R^*(0,1,0,1) & R^*(0,1,0,2) & . & . & . & R^*(0,1,p_1,p_2) \\ R^*(0,2,0,1) & R^*(0,2,0,2) & . & . & . & R^*(0,2,p_1,p_2) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ R^*(p_1,p_2,0,1) & R^*(p_1,p_2,0,2) & . & . & . & R^*(p_1,p_2,p_1,p_2) \end{bmatrix}.$$

Note that α, r, and A are $(p_1+1)(p_2+1)-1$ dimensional vectors, while R and R^* are $(p_1+1)(p_2+1)-1$ dimensional square matrices. These imply that we can write the equations (2) in matrix form as

$$r = A + \mathrm{R}\,\alpha + \mathrm{R}^*\alpha\,. \tag{4}$$

The spatial YW estimator of α is given by $\hat{\alpha}_{YW} = \mathrm{R}^{-1}\, r$ (Tjostheim, 1978), and thus from (4) we obtain

$$\hat{\alpha}_{YW} = \mathrm{R}^{-1}\, r = \mathrm{R}^{-1}\, A + \alpha + \mathrm{R}^{-1}\,\mathrm{R}^*\alpha\,.$$

This implies $\hat{\alpha}_{YW} - \alpha = \mathrm{R}^{-1} A + \mathrm{R}^{-1} \mathrm{R}^{*}\alpha$, so that

$$(mn)^{1/2} \left(\hat{\alpha}_{YW} - \alpha \right) = \mathrm{R}^{-1} (mn)^{1/2} A + \mathrm{R}^{-1} (mn)^{1/2} \mathrm{R}^{*}\alpha \,. \tag{5}$$

Writing $\gamma(0,0) = \gamma$, note that as $m, n \rightarrow \infty$, $\mathrm{R} \xrightarrow{P} \gamma\, \Gamma(\alpha)$, where

$$\gamma \Gamma(\alpha) = \begin{bmatrix} \gamma & \gamma(0,1) & \cdot & \cdot & \cdot & \gamma(-p_1, 1-p_2) \\ \gamma(0,1) & \gamma & \cdot & \cdot & \cdot & \gamma(-p_1, 2-p_2) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma(p_1, p_2-1) & \gamma(p_1, p_2-2) & \cdot & \cdot & \cdot & \gamma \end{bmatrix} .$$

Also, it can be established that $(mn)^{1/2}\mathrm{R}^{*} \xrightarrow{P} \gamma\Gamma^{*} \neq 0$, where the matrix $\gamma\Gamma^{*}$ has a particular form, under the assumption that $m/n \rightarrow c^2$, $0 < c < \infty$. This implies that

$$\mathrm{R}^{-1} (mn)^{1/2} \mathrm{R}^{*}\alpha \xrightarrow{P} \Gamma(\alpha)^{-1} \Gamma^{*} \alpha = b \,. \tag{6}$$

Also, from Lemmas (8.1)–(8.4) of Tjostheim (1978), we obtain that

$$\mathrm{R}^{-1} (mn)^{1/2} A \xrightarrow{D} N\!\left(0, \frac{\sigma^2}{\gamma} \Gamma(\alpha)^{-1}\right). \tag{7}$$

So from (5)–(7) and Slutsky's theorem we obtain,

$$(mn)^{1/2} \left(\hat{\alpha}_{YW} - \alpha \right) \xrightarrow{D} N\!\left(b, \frac{\sigma^2}{\gamma} \Gamma(\alpha)^{-1}\right). \tag{8}$$

It should be noted that Tjostheim (1978) incorrectly arrived at the conclusion that $b = 0$ in the above distributional result (8) concerning the spatial YW estimator $\hat{\alpha}_{YW}$, while this result indicates that the estimator has an asymptotic bias of the order $(mn)^{-1/2}$, with $E(\hat{\alpha}_{YW}) \approx \alpha + (mn)^{-1/2} b$.

3. Yule-Walker Estimator for the First-Order Spatial Bilateral AR Model

A special case which is of interest is the first-order spatial bilateral AR model. This model has been studied by Basu and Reinsel (1990), and Martin (1979, 1990) studied the special case of

$\alpha_{11} = -\alpha_{10}\alpha_{01}$. In this case we have $p_1 = 1$ and $p_2 = 1$, and the model is

$$Y_{ij} = \alpha_{01}\, Y_{i,j-1} + \alpha_{10}\, Y_{i-1,j} + \alpha_{11}\, Y_{i-1,j-1} + \epsilon_{ij} \tag{9}$$

for $i = 1,\ldots,m$ and $j = 1,\ldots,n$. Recall that $i^* = \max(1,k-s+1)$, $m^* = \min(m,m-s+k)$, $j^* = \max(1,l-t+1)$, and $n^* = \min(n,n-t+l)$ and (s,t) and (k,l) take on only the values (0,1), (1,0) and (1,1) for the spatial YW estimation in this model. We will obtain explicit expressions for all the relevant terms $B(s,t,k,l)$, $R^*(s,t,k,l)$, and $I^*(s,t,k,l)$ for the special case of first-order spatial bilateral AR model. From (3), the expressions for $B(s,t,k,l)$ are as follows:

$$mn\ B(0,1,0,1) = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} Y_{ij}^2 - \sum_{i=1}^{m} Y_{in}^2 \ ,$$

$$mn\ B(0,1,1,0) = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i-1,j+1} = \sum_{i=2}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i-1,j+1} + \sum_{j=1}^{n-1} Y_{1j}Y_{0,j+1},$$

$$mn\ B(0,1,1,1) = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i-1,j} = \sum_{i=2}^{m}\sum_{j=1}^{n} Y_{ij}Y_{i-1,j} - \Big(\sum_{i=1}^{m} Y_{in}Y_{i-1,n} - \sum_{j=1}^{n} Y_{1j}Y_{0j}\Big)$$

$$mn\ B(1,0,0,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i+1,j-1} = \sum_{i=1}^{m-1}\sum_{j=2}^{n} Y_{ij}Y_{i+1,j-1} + \sum_{i=1}^{m-1} Y_{i1}Y_{i+1,0}\ ,$$

$$mn\ B(1,0,1,0) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} Y_{ij}^2 - \sum_{j=1}^{n} Y_{mj}^2 \ ,$$

$$mn\ B(1,0,1,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i,j-1} = \sum_{i=1}^{m}\sum_{j=2}^{n} Y_{ij}Y_{i,j-1} + \Big(\sum_{i=1}^{m} Y_{i1}Y_{i0} - \sum_{j=1}^{n} Y_{mj}Y_{m,j-1}\Big)$$

$$mn\ B(1,1,0,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}Y_{i+1,j} = \sum_{i=1}^{m-1}\sum_{j=1}^{n} Y_{ij}Y_{i+1,j} - \sum_{i=1}^{m-1} Y_{in}Y_{i+1,n}\ ,$$

$$mn\ B(1,1,1,0) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}Y_{i,j+1} = \sum_{i=1}^{m}\sum_{j=1}^{n-1} Y_{ij}Y_{i,j+1} - \sum_{j=1}^{n-1} Y_{mj}Y_{m,j+1},$$

$$mn\ B(1,1,1,1) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} Y_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} Y_{ij}^2 - \Big(\sum_{i=1}^{m} Y_{in}^2 + \sum_{j=1}^{n} Y_{mj}^2 - Y_{mn}^2\Big),$$

where for each of the expressions above, following notation of (3), the first term on the right-hand side corresponds to $mn\ R(s-k,t-l)$, while the remaining terms equal $mn\ R^*(s,t,k,l)$. It is shown in Basu and Reinsel (1990) that for model (9), $\gamma(1,0) = \lambda\gamma$, $\gamma(0,1) = \mu\gamma$, $\gamma(1,-1) = \lambda\mu\gamma$, where λ and μ are solved using the relations

$$\alpha_{10}\,\lambda + \alpha_{01}\,\mu^{-1} + \alpha_{11}\,\lambda\,\mu^{-1} = 1 = \alpha_{10}\,\lambda^{-1} + \alpha_{01}\,\mu + \alpha_{11}\,\lambda^{-1}\,\mu\ , \tag{10}$$

with $|\lambda| < 1$ and $|\mu| < 1$. Supposing that $m/n \to c^2$, $0 < c < \infty$, as $m,n \to \infty$, and noting that $(mn)^{-1/2} = (m/n)^{-1/2} n^{-1} = (n/m)^{-1/2} m^{-1}$, we obtain from the above relations, $(mn)^{1/2} \mathrm{R}^* \xrightarrow{P} \gamma I^*$ and $\mathrm{R} \xrightarrow{P} \gamma I(\alpha)$ where

$$\gamma I(\alpha) = \begin{bmatrix} \gamma & \gamma(1,-1) & \gamma(1,0) \\ \gamma(1,-1) & \gamma & \gamma(0,1) \\ \gamma(1,0) & \gamma(0,1) & \gamma \end{bmatrix} = \gamma \begin{bmatrix} 1 & \lambda\,\mu & \lambda \\ \lambda\mu & 1 & \mu \\ \lambda & \mu & 1 \end{bmatrix} \tag{11a}$$

$$\gamma I^* = \gamma \begin{bmatrix} -c & \frac{1}{c}\,\lambda\,\mu & (\frac{1}{c} - c)\,\lambda \\ c\,\lambda\,\mu & -\frac{1}{c} & (c - \frac{1}{c})\,\mu \\ -c\,\lambda & -\frac{1}{c}\,\mu & (c + \frac{1}{c}) \end{bmatrix} . \tag{11b}$$

So from (6) and (11a-b), using the relation $I(\alpha)^{-1} I^* \alpha = b$, we obtain

$$b = \begin{bmatrix} -c & \dfrac{2\,\lambda\,\mu}{c(1-\lambda^2)} & \dfrac{2\,\lambda}{c(1-\lambda^2)} \\ \dfrac{2\,c\,\lambda\,\mu}{1-\mu^2} & -\dfrac{1}{c} & \dfrac{2\,c\,\mu}{1-\mu^2} \\ -\dfrac{2\,c\,\lambda\,\mu^2}{1-\mu^2} & -\dfrac{2\,\lambda^2\,\mu}{c(1-\lambda^2)} & -\left(\dfrac{1+\lambda^2}{c(1-\lambda^2)} + \dfrac{c(1+\mu^2)}{1-\mu^2}\right) \end{bmatrix} \begin{bmatrix} \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix} . \tag{12}$$

Now using the relations (10) and simplifying equation (12), we obtain the elements of the asymptotic bias vector $b = (b_{01}, b_{10}, b_{11})$ in (6) explicitly as

$$b_{01} = -c\,\alpha_{01} + \frac{2\lambda}{c(1-\lambda^2)}\left(\mu\,\alpha_{10} + \alpha_{11}\right) = -\left(c + \frac{2}{c(1-\lambda^2)}\right)\alpha_{01} + \frac{2\,\mu}{c(1-\lambda^2)} ,$$

$$b_{10} = -\frac{1}{c}\,\alpha_{10} + \frac{2c\mu}{(1-\mu^2)}\left(\lambda\,\alpha_{01} + \alpha_{11}\right) = -\left(\frac{1}{c} + \frac{2c}{(1-\mu^2)}\right)\alpha_{10} + \frac{2\,c\lambda}{1-\mu^2} ,$$

$$b_{11} = -\left(\frac{c(1+\mu^2)}{1-\mu^2} + \frac{1+\lambda^2}{c(1-\lambda^2)}\right)\alpha_{11} - \frac{2c\lambda\mu^2}{1-\mu^2}\,\alpha_{01} - \frac{2\,\lambda^2\mu}{c(1-\lambda^2)}\,\alpha_{10} .$$

Hence we find that the asymptotic bias of the spatial YW estimator for this model may be particularly large when the roots of the characteristic equations (10), λ or μ, have absolute value close to one.

As a special case of model (9), consider the multiplicative AR(1) model where $\alpha_{11} = -\alpha_{10}\alpha_{01}$. In this special case, (12) can be simplified and the asymptotic biases of the spatial YW estimator reduce to

$$b_{01} = -c\,\alpha_{01}\,, \quad b_{10} = -\tfrac{1}{c}\,\alpha_{10}\,, \quad b_{11} = -\left(c + \tfrac{1}{c}\right)\alpha_{01}\alpha_{10}\,.$$

This simple result indicates that the biases can be quite substantial when the grid sizes m and n are small or moderate.

4. Comparison of Spatial YW Estimator with the ML Estimator and Concluding Remarks

In this section, we first give some numerical illustrations of the values of the theoretical asymptotic biases b from (12) for the spatial YW estimator for different values of $c = \lim(m/n)^{1/2}$. The values of c are chosen so that the biases from the simulation results, given later, can directly be compared to the corresponding asymptotic values. For the choice of $(m,\, n)$, c ranges from 0.5 to 1.0. The three sets of α-values chosen are (a) $\alpha_{01} = 0.2$, $\alpha_{10} = 0.3$, $\alpha_{11} = 0.2$, (b) $\alpha_{01} = 0.3$, $\alpha_{10} = 0.5$, $\alpha_{11} = 0.1$, and (c) $\alpha_{01} = 0.7$, $\alpha_{10} = 0.8$, $\alpha_{11} = -0.6$. The values of the α's are chosen to represent a range of values for λ and μ.

Table 1 gives the values of the asymptotic biases of the spatial YW estimator. It is seen that magnitude of the bias is dependent on the value of c. For example, for small values of λ and μ, that is, for $\alpha_{01} = 0.2$, $\alpha_{10} = 0.3$, and $\alpha_{11} = 0.2$, when c is changed from 0.5 to 1.0, the bias b_{01} changes from 0.436 to 0.068 and the bias b_{10} changes from -0.503 to -0.107 for the spatial YW estimator. For moderate values of λ and μ, that is, for $\alpha_{01} = 0.3$, $\alpha_{10} = 0.5$, and $\alpha_{11} = 0.1$, the magnitude of the biases are larger and changes (with changing c) are also larger. For large values of λ and μ, the biases are very large for the spatial YW estimator estimator. Also, it is seen that the biases are larger when c has a smaller value, and they decrease in magnitude, in general, as $c \to 1$.

Next we look at some simulation results for small to moderate grid sizes and compare the exact ML estimators (Basu and Reinsel, 1990) and the spatial YW estimators. The simulations are performed for the grid sizes (8×8), (20×20), (8×10), (16×20), (6×10), (15×25), (6×15), (12×30), (5×20), and (10×40), and for the three different sets of α-values. 500 replications are obtained for each of the three sets of α-values as well as for each of the grids chosen. The averages of the estimates of the parameter values from the 500 replications are presented in Tables 2 and 3 for the spatial YW and the exact ML estimators, respectively. The standard errors of the average estimates over the 500 replications are of the order of 0.002 for both the estimators. It is seen that the spatial YW estimates are more biases than the exact ML estimates and the biases increase as the values of λ and μ increase.

The average values of the exact ML estimates were always very close to the true values of the parameters for all grid sizes and for all values of the parameters.

The biases of the estimators from each replication are calculated as $(mn)^{1/2}(\hat{\alpha} - \alpha)$ and the average biases over the 500 replications are given in Table 4 for the spatial YW estimator. The standard errors of the average biases over the 500 replications are of the order of 0.04 for the estimators. These values can then be compared with the corresponding asymptotic values in Table 1. The agreement between the two sets of values is reasonably good although not excellent, but it does tend to improve for the larger grid sizes. As the values of λ and μ increase, the fact that the biases increase also tends to hold. The lack of agreement may be due in part to the contributions to the bias of the spatial YW estimator from terms of higher order.

To examine further the lack of agreement between the asymptotic biases and those obtained from the simulation results for the spatial YW estimators, we use the identity from equation (3.6) of Lewis and Reinsel (1988),

$$\mathrm{R}^{-1} = \Gamma^{-1} - \Gamma^{-1}(\mathrm{R} - \Gamma)\Gamma^{-1} + \Gamma^{-1}(\Gamma - \mathrm{R})(\mathrm{R}^{-1} - \Gamma^{-1})$$

and approximate R^{-1} by the first two terms. Now using (5), since $E(A) = 0$ we approximate the biases of the spatial YW estimator from the small to moderate grid sizes as

$$E[(mn)^{1/2}(\hat{\alpha}_{\mathrm{YW}} - \alpha)] \approx [\, \Gamma^{-1} - \Gamma^{-1}(E(\mathrm{R}) - \Gamma)\Gamma^{-1}]\,(mn)^{1/2}E[\mathrm{R}^*]\,\alpha = b^* \tag{13}$$

and hope that b^* might provide a somewhat more accurate approximation to the biases than b for smaller values of m and n. The values of b^* are presented in Table 5 for the grid sizes and the three different sets of α-values considered above. The values of b^* approach the asymptotic values b as m and n get large. For cases (a) and (c), the differences between the asymptotic biases in equation (12) and the small grid approximation in (13) are rather small and both approximations give values close to the simulated values. For case (b), the asymptotic biases from (12) are not in good agreement with the simulated biases, but the biases obtained using (13) differ substantially from the values obtained through (12) and those from (13) are in much better agreement with the simulated biases.

In conclusion, we have obtained the asymptotic distribution of the spatial YW estimator, and shown the spatial YW estimator to be asymptotically biased. The first-order spatial bilateral AR model is studied in more detail, and for the first-order model, the spatial YW estimators are also compared with the exact maximum likelihood estimators. Simulation results show that the bias of the spatial YW estimators for small to moderate grid sizes is substantial.

References

Bartlett, M.S. (1975). *The Statistical Analysis of Spatial Pattern.* Chapman and Hall, London.

Basu, S. and Reinsel, G. C. (1990). Analysis of first-order spatial bilateral ARMA models. *Technical Report No.* 862. Department of Statistics, University of Wisconsin—Madison.

Cliff, A.D. and Ord, J.K. (1981). *Spatial Processes: Models and Applications.* Pion, London.

Guyon, X. (1982). Parameter estimation for a stationary process on a *d*-dimensional lattice. *Biometrika*, 69, 95-105.

Lewis, R.A. and Reinsel, G.C. (1988). Prediction error of multivariate time series with misspecified models. *Journal of Time Series Analysis*, 9, 43-57.

Martin, R.J. (1979). A subclass of lattice processes applied to a problem of planar sampling. *Biometrika*, 66, 209-217.

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

Tjostheim, D. (1978). Statistical spatial series modelling. *Advances in Applied Probability*, 10, 130-154.

Tjostheim, D. (1983). Statistical spatial series modelling II : Some further results on unilateral processes. *Advances in Applied Probability*, 15, 562-584.

Upton. G.J.G. and Fingleton, B. (1985). *Spatial Data Analysis by Example, Volume I.* John Wiley, New York.

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

Table 1. Asymptotic biases of spatial Yule-Walker estimators obtained from (12) for different sets of α-values and for different values of $c = \lim(m/n)^{1/2}$.

(a) $\alpha_{01} = 0.2$, $\alpha_{10} = 0.3$, $\alpha_{11} = 0.2$, $\lambda = 0.387$, $\mu = 0.314$

c	b_{01}	b_{10}	b_{11}
1.000	0.068	−0.107	−0.564
0.894	0.121	−0.163	−0.573
0.775	0.191	−0.238	−0.594
0.632	0.297	−0.352	−0.645
0.500	0.436	−0.503	−0.738

(b) $\alpha_{01} = 0.3$, $\alpha_{10} = 0.5$, $\alpha_{11} = 0.1$, $\lambda = 0.664$, $\mu = 0.549$

c	b_{01}	b_{10}	b_{11}
1.000	0.590	−0.030	−1.049
0.894	0.726	−0.139	−1.092
0.775	0.916	−0.282	−1.169
0.632	1.217	−0.494	−1.319
0.500	1.629	−0.765	−1.561

(c) $\alpha_{01} = 0.7$, $\alpha_{10} = 0.8$, $\alpha_{11} = -0.6$, $\lambda = 0.756$, $\mu = 0.624$

c	b_{01}	b_{10}	b_{11}
1.000	−1.057	−0.945	1.560
0.894	−1.025	−1.024	1.590
0.775	−1.003	−1.145	1.657
0.632	−1.007	−1.357	1.811
0.500	−1.063	−1.672	2.084

Table 2. Average estimated value of parameters from spatial YW estimation from 500 replications for each set of α-values.

(a) $\alpha_{01}= 0.2,\ \alpha_{10}= 0.3,\ \alpha_{11}= 0.2,\ \lambda = 0.387,\ \mu = 0.314$

m	n	c	$\hat{\alpha}_{01}$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
8	8	1.000	0.186	0.262	0.128
8	10	0.894	0.199	0.263	0.137
6	10	0.775	0.200	0.243	0.126
6	15	0.632	0.214	0.245	0.136
5	20	0.500	0.229	0.235	0.134
20	20	1.000	0.201	0.293	0.171
16	20	0.894	0.205	0.290	0.166
15	25	0.775	0.206	0.286	0.168
12	30	0.632	0.211	0.279	0.170
10	40	0.500	0.217	0.270	0.165

(b) $\alpha_{01}= 0.3,\ \alpha_{10}= 0.5,\ \alpha_{11}= 0.1,\ \lambda = 0.664,\ \mu = 0.549$

8	8	1.000	0.317	0.454	0.014
8	10	0.894	0.332	0.459	0.013
6	10	0.775	0.344	0.426	0.006
6	15	0.632	0.366	0.420	0.006
5	20	0.500	0.390	0.406	0.001
20	20	1.000	0.321	0.489	0.057
16	20	0.894	0.325	0.483	0.048
15	25	0.775	0.322	0.480	0.049
12	30	0.632	0.346	0.464	0.048
10	40	0.500	0.364	0.459	0.039

(c) $\alpha_{01}= 0.7,\ \alpha_{10}= 0.8,\ \alpha_{11}= -0.6,\ \lambda = 0.756,\ \mu = 0.624$

8	8	1.000	0.561	0.654	−0.405
8	10	0.894	0.583	0.673	−0.434
6	10	0.775	0.571	0.625	−0.395
6	15	0.632	0.602	0.641	−0.423
5	20	0.500	0.606	0.623	−0.418
20	20	1.000	0.645	0.748	−0.522
16	20	0.894	0.643	0.738	−0.513
15	25	0.775	0.651	0.736	−0.519
12	30	0.632	0.651	0.725	−0.509
10	40	0.500	0.653	0.713	−0.502

Table 3. Average estimated value of parameters from exact ML estimation from 500 replications for each set of α-values.

(a) $\alpha_{01}= 0.2$, $\alpha_{10}= 0.3$, $\alpha_{11}= 0.2$, $\lambda = 0.387$, $\mu = 0.314$

m	n	c	$\hat{\alpha}_{01}$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
8	8	1.000	0.183	0.282	0.191
8	10	0.894	0.190	0.283	0.191
6	10	0.775	0.184	0.276	0.194
6	15	0.632	0.191	0.283	0.198
5	20	0.500	0.196	0.284	0.196
20	20	1.000	0.199	0.299	0.199
16	20	0.894	0.200	0.300	0.196
15	25	0.775	0.198	0.299	0.198
12	30	0.632	0.198	0.297	0.201
10	40	0.500	0.198	0.295	0.199

(b) $\alpha_{01}= 0.3$, $\alpha_{10}= 0.5$, $\alpha_{11}= 0.1$, $\lambda = 0.664$, $\mu = 0.549$

m	n	c	$\hat{\alpha}_{01}$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
8	8	1.000	0.277	0.474	0.102
8	10	0.894	0.286	0.486	0.097
6	10	0.775	0.274	0.481	0.100
6	15	0.632	0.293	0.480	0.091
5	20	0.500	0.293	0.487	0.094
20	20	1.000	0.299	0.493	0.100
16	20	0.894	0.294	0.496	0.097
15	25	0.775	0.297	0.496	0.098
12	30	0.632	0.296	0.491	0.104
10	40	0.500	0.300	0.497	0.100

(c) $\alpha_{01}= 0.7$, $\alpha_{10}= 0.8$, $\alpha_{11}= -0.6$, $\lambda = 0.756$, $\mu = 0.624$

m	n	c	$\hat{\alpha}_{01}$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
8	8	1.000	0.675	0.765	−0.565
8	10	0.894	0.675	0.780	−0.575
6	10	0.775	0.672	0.772	−0.570
6	15	0.632	0.680	0.784	−0.578
5	20	0.500	0.681	0.786	−0.582
20	20	1.000	0.694	0.795	−0.594
16	20	0.894	0.693	0.793	−0.591
15	25	0.775	0.696	0.794	−0.595
12	30	0.632	0.696	0.794	−0.593
10	40	0.500	0.698	0.797	−0.596

Table 4. Average biases (average values of $(mn)^{1/2}(\hat{\alpha}-\alpha)$) from 500 replications of YW estimators from simulation results for different grid sizes and for each set of α-values.

(*a*) $\alpha_{01}= 0.2,\ \alpha_{10}= 0.3,\ \alpha_{11}= 0.2,\ \lambda = 0.387,\ \mu = 0.314$

m	n	c	$\hat{b}_{01}$	$\hat{b}_{10}$	$\hat{b}_{11}$
8	8	1.000	−0.114	−0.303	−0.574
8	10	0.894	−0.012	−0.332	−0.567
6	10	0.775	−0.001	−0.443	−0.571
6	15	0.632	0.134	−0.526	−0.607
5	20	0.500	0.285	−0.646	−0.658
20	20	1.000	0.015	−0.148	−0.571
16	20	0.894	0.096	−0.173	−0.608
15	25	0.775	0.108	−0.281	−0.613
12	30	0.632	0.217	−0.406	−0.571
10	40	0.500	0.338	−0.612	−0.707

(*b*) $\alpha_{01}= 0.3,\ \alpha_{10}= 0.5,\ \alpha_{11}= 0.1,\ \lambda = 0.664,\ \mu = 0.549$

m	n	c	$\hat{b}_{01}$	$\hat{b}_{10}$	$\hat{b}_{11}$
8	8	1.000	0.132	−0.367	−0.691
8	10	0.894	0.285	−0.368	−0.775
6	10	0.775	0.341	−0.577	−0.727
6	15	0.632	0.623	−0.758	−0.894
5	20	0.500	0.900	−0.944	−0.990
20	20	1.000	0.420	−0.229	−0.864
16	20	0.894	0.438	−0.298	−0.925
15	25	0.775	0.656	−0.398	−0.984
12	30	0.632	0.867	−0.682	−0.992
10	40	0.500	1.278	−0.826	−1.230

(*c*) $\alpha_{01}= 0.7,\ \alpha_{10}= 0.8,\ \alpha_{11}= -0.6,\ \lambda = 0.756,\ \mu = 0.624$

m	n	c	$\hat{b}_{01}$	$\hat{b}_{10}$	$\hat{b}_{11}$
8	8	1.000	−1.113	−1.172	1.564
8	10	0.894	−1.049	−1.136	1.486
6	10	0.775	−1.001	−1.357	1.591
6	15	0.632	−0.930	−1.506	1.682
5	20	0.500	−0.939	−1.769	1.823
20	20	1.000	−1.107	−1.043	1.569
16	20	0.894	−1.015	−1.118	1.555
15	25	0.775	−0.957	−1.244	1.578
12	30	0.632	−0.929	−1.433	1.723
10	40	0.500	−0.932	−1.733	1.955

Table 5. Theoretical approximation to biases from small to moderate grid sizes for the spatial Yule-Walker estimators obtained from (13) for different grid sizes and for each set of α-values.

(*a*) $\alpha_{01}=0.2$, $\alpha_{10}=0.3$, $\alpha_{11}=0.2$, $\lambda=0.387$, $\mu=0.314$

m	n	c	b^*_{01}	b^*_{10}	b^*_{11}
8	8	1.000	0.010	−0.146	−0.486
8	10	0.894	0.062	−0.195	−0.499
6	10	0.775	0.113	−0.272	−0.503
6	15	0.632	0.213	−0.378	−0.555
5	20	0.500	0.323	−0.526	−0.626
20	20	1.000	0.044	−0.123	−0.532
16	20	0.894	0.091	−0.179	−0.535
15	25	0.775	0.158	−0.252	−0.556
12	30	0.632	0.254	−0.365	−0.599
10	40	0.500	0.378	−0.515	−0.681

(*b*) $\alpha_{01}=0.3$, $\alpha_{10}=0.5$, $\alpha_{11}=0.1$, $\lambda=0.664$, $\mu=0.549$

m	n	c	b^*_{01}	b^*_{10}	b^*_{11}
8	8	1.000	0.284	−0.158	−0.673
8	10	0.894	0.402	−0.242	−0.717
6	10	0.775	0.456	−0.386	−0.665
6	15	0.632	0.680	−0.570	−0.757
5	20	0.500	0.851	−0.831	−0.770
20	20	1.000	0.458	−0.084	−0.887
16	20	0.894	0.556	−0.192	−0.895
15	25	0.775	0.719	−0.325	−0.954
12	30	0.632	0.937	−0.532	−1.027
10	40	0.500	1.228	−0.797	−1.153

(*c*) $\alpha_{01}=0.7$, $\alpha_{10}=0.8$, $\alpha_{11}=-0.6$, $\lambda=0.756$, $\mu=0.624$

m	n	c	b^*_{01}	b^*_{10}	b^*_{11}
8	8	1.000	−0.773	−0.853	1.140
8	10	0.894	−0.740	−0.954	1.194
6	10	0.775	−0.581	−1.069	1.114
6	15	0.632	−0.564	−1.309	1.283
5	20	0.500	−0.455	−1.634	1.402
20	20	1.000	−0.967	−0.918	1.423
16	20	0.894	−0.904	−0.997	1.419
15	25	0.775	−0.886	−1.124	1.478
12	30	0.632	−0.812	−1.338	1.578
10	40	0.500	−0.789	−1.658	1.776