Acta Math. Univ. Comenianae
Vol. LXIII, 2(1994), pp. 247–254

MEAN SQUARED ERRORS OF PREDICTION BY KRIGING IN LINEAR MODELS WITH $AR(1)$ ERRORS

F. ŠTULAJTER

1. Introduction

Kriging, in the scientific literature, is used as a name for the theory of prediction in random processes (random fields) with an unknown mean value and, possibly, with an unknown covariance function. M. Stein in a series of articles (1988), (1990a), (1990b) and (1990c) studies the case when the unknown covariance function of the observed process is misspecified, but not estimated from the data. Limit theory of prediction of time series with estimated parameters has been studied by many authors including Bhansali (1981), Fuller and Hasza (1981), Kunimoto and Yamamoto (1985) and Toyooka (1982). Some of these authors have assumed that the mean value of the observed process is zero.

Harville (1985), Harville and Jeske (1992) and Zimmerman and Cressie (1992) studied properties and approximations of the mean squared error of prediction with unbiasedly estimated parameters in the case when a covariance function depends linearly on unknown parameters.

The main aim of this paper is to derive an approximate expression for the mean square error of a predictor with estimated parameters which is based on a finite observation of a stochastic process following a linear regression model with $AR(1)$ errors. In this case the dependence of covariance function on unknown parameters is nonlinear.

2. Kriging Predictors in a Linear Regression Model

Let $\mathbf{X} = (X(1), \dots, X(n))'$ be a finite observation of length n of a stochastic process $X = \{X(t);\ t \in T\}$ with the mean function $m(t) = \sum_{i=1}^{k} \beta_i f_i(t);\ t \in T$ where $f_1, \dots, f_k$ are known functions and $\beta = (\beta_1, \dots, \beta_k)'$ are unknown regression parameters and with some covariance function $R(s,t);\ s,t \in T$. Then we can write

$$\mathbf{X} = F\beta + \varepsilon$$

Received December 13, 1993.
1980 *Mathematics Subject Classification* (1991 *Revision*). Primary 62M20, 62M10.

with $E[\varepsilon] = 0$, $E[\varepsilon\varepsilon'] = \Sigma$, where $\Sigma_{ij} = R(i,j)$; $i,j = 1,2,\dots,n$. Let us assume that Σ is a positive definite $n \times n$ matrix.

Let U be a predicted random variable (for example $U = X(n+1)$) with $E_\beta[U] = f'\beta$, where f is a given vector, with a finite variance $D[U]$ and with a known vector r of covariances between $\mathbf{X}$ and U: $r = (\mathrm{Cov}\,(X(1);U),\dots,\mathrm{Cov}\,(X(n);U))'$.

Then the kriging predictor U^* of U based on $\mathbf{X}$ is given by

$$U^* = f'\beta^* + r'\Sigma^{-1}(\mathbf{X} - F\beta^*) \tag{1}$$

where $\beta^* = (F'\Sigma^{-1}F)^{-1}F'\Sigma^{-1}\mathbf{X}$ is the best linear unbiased estimator (BLUE) for β with the covariance matrix $\Sigma_{\beta^*} = (F'\Sigma^{-1}F)^{-1}$. The mean square error of the predictor U^* is given by

$$E[U^* - U]^2 = D[U] - r'\Sigma^{-1}r + \|f - F'\Sigma^{-1}r\|^2_{\Sigma_{\beta^*}} \tag{2}$$

where $\|g\|^2_A = g'Ag$ denotes a norm defined by a positive definite matrix A.

The kriging predictor U^* is in fact the best linear unbiased predictor (BLUP) of U based on $\mathbf{X}$ (see Harville (1990)).

The practical use of (1) is limited, since we usually do not know the vector r and the matrix Σ.

The properties of the estimator

$$\hat{U} = f'\hat{\beta} + r'\Sigma^{-1}(\mathbf{X} - F\hat{\beta}), \tag{3}$$

where $\hat{\beta} = (F'F)^{-1}F'\mathbf{X}$ is the least squares estimator (LSE) of β were studied by Štulajter (1991). It was shown that

$$E[\hat{U} - U]^2 = D[U] - r'\Sigma^{-1}r + \|f - F'\Sigma^{-1}r\|^2_{\Sigma_{\hat{\beta}}} \tag{4}$$

where $\Sigma_{\hat{\beta}} = (F'F)^{-1}F'\Sigma F(F'F)^{-1}$. Since U^* is the BLUP for U, it is clear that

$$\|f - F'\Sigma^{-1}r\|_{\Sigma_{\beta^*}} \le \|f - F'\Sigma^{-1}r\|_{\Sigma_{\hat{\beta}}}.$$

Let us assume now that the errors $\varepsilon(t)$; $t = 1,2,\dots$ form an $AR(1)$ process with parameters σ^2 and ρ, $|\rho| < 1$; that means $\varepsilon(t+1) = \rho\varepsilon(t) + e(t)$ for $t = 1,2,\dots$, where $E[e(t)] = 0$, $E[e(s)e(t)] = \sigma^2\delta_{st}$. Then the observed process X is covariance stationary with the covariance function $R(t) = \sigma^2\frac{\rho^t}{1-\rho^2}$, $t = 0,1,\dots$.

Let $U = X(n+1)$, then

$$\Sigma^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & -\rho & 0 & \dots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & 0 & -\rho & 1 \end{pmatrix},$$

$r'\Sigma^{-1} = (0,0,\dots,\rho)'$ and the estimator $\hat{X}(n+1)$ given by (3) can be rewritten as

$$\hat{X}(n+1) = f'\hat{\beta} + \rho(X(n) - (F\hat{\beta})_n) \tag{5}$$

where $(F\hat{\beta})_n$ denotes the n-th coordinate of the vector $F\hat{\beta}$. Next we get

$$E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2 + \|f - F'\Sigma^{-1}r\|^2_{\Sigma_{\hat{\beta}}}. \tag{6}$$

Example 1. Let X be a stationary process with an unknown constant mean value β. Then $F = (1,\dots,1)'$, $f = 1$, $(F'F)^{-1} = \frac{1}{n}$, $F'\Sigma F = n(R(0) + 2\sum_{t=1}^{n}(1-\frac{t}{n})R(t))$ and we get from (5) and (6) that

$$X(n+1) = \hat{\beta} + \rho(X(n) - \hat{\beta}) \qquad \text{and}$$
$$E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2 + (1-\rho)^2\left(\frac{R(0)}{n} + \frac{2}{n}\sum_{t=1}^{n}\left(1 - \frac{t}{n}\right)R(t)\right),$$

where $\hat{\beta} = \frac{1}{n}\sum_{t=1}^{n} X(t)$ is the LSE of the unknown (constant) mean value β. It is easy to prove that

$$\lim_{n\to\infty} E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2 \quad \text{for every } \rho \in (-1,1).$$

Example 2. Let X be a covariance stationary $AR(1)$ process with a linear trend $E_\beta[X(t)] = \beta_1 + \beta_2 t$; $t = 1,2,\dots$. Then

$$F = F_n = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \end{pmatrix}',$$
$$f = f_n = \begin{pmatrix} 1 \\ n+1 \end{pmatrix} \qquad \text{and}$$
$$g_n = f_n - F_n'\Sigma_n^{-1}r_n \begin{pmatrix} 1 \\ n+1 \end{pmatrix} - \begin{pmatrix} \rho \\ n\rho \end{pmatrix}$$

depend on n. Again, $\hat{X}(n+1) = f_n'\hat{\beta} + \rho(X(n) - (F\hat{\beta})_n)$, $\hat{\beta} = (F'F)^{-1}F'\mathbf{X}$ and $E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2 + \|f_n - F_n'\Sigma_n r_n\|^2_{\Sigma_{\hat{\beta}_n}}$. Our aim is to show that, again, $\lim_{n\to\infty} E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2$. This result follows from the next theorem.

Theorem 1. *Let $\mathbf{X}$ and U fulfil the conditions given in the beginning of this paragraph and let $g_n = f_n' - F_n'\Sigma_n^{-1}r_n$. If $g_n'(F_n'F_n)^{-1}g_n = O(1/n)$ and if $\lim_{n\to\infty}\frac{\|\Sigma_n\|}{n} = 0$, where $\|\cdot\|$ denotes the Euclidean matrix norm, then*

$$\lim_{n\to\infty} E[\hat{U}_n - U]^2 = D[U] - \lim_{n\to\infty} r_n'\Sigma_n^{-1}r_n. \tag{7}$$

Proof. Since $r_n'\Sigma_n^{-1}r_n$; $n = 1,2,\dots$ is non decreasing and bounded by $D[U]$, it is enough to prove that $\lim_{n\to\infty}\|g_n\|_{\Sigma_{\hat{\beta}_n}} = 0$ if the conditions of the theorem are fulfilled. Using the Schwarz inequality we get $\|g_n\|^2_{\Sigma_{\hat{\beta}_n}} \le \|\Sigma_n\| g_n'(F_n'F_n)^{-1}g_n$, from which the theorem follows. □

Example 2 (continuation). For the linear trend matrix F_n given in the Example 2 we get (see Štulajter (1991)) $(F_n'F_n)^{-1} = \frac{1}{n}G_n$, where $G_n = \begin{pmatrix} \frac{2(2n+1)}{n-1} & -\frac{6}{n-1} \\ -\frac{6}{n-1} & \frac{12}{n^2-1} \end{pmatrix}$. Thus

$$g_n' G_n g_n = (1-\rho)^2 \left[\frac{2(2n+1)}{n-1} - 12\frac{1+n(1-\rho)}{(n-1)(1-\rho)} + 12\frac{(1+n(1-\rho))^2}{(n^2-1)(1-\rho)^2} \right],$$

$$\lim_{n\to\infty} g_n' G_n g_n = 4(1-\rho)^2 \qquad \text{and}$$

$$\lim_{n\to\infty} \frac{\|\Sigma_n\|}{n} = \lim_{n\to\infty} \left(\frac{R^2(0)}{n} + \frac{2}{n}\sum_{t=1}^{n} \left(1-\frac{t}{n}\right) R^2(t) \right)^{1/2} = 0$$

from which we have that $\lim_{n\to\infty} E[\hat{X}(n+1) - X(n+1)]^2 = \sigma^2$.

Remark. The predictor $\hat{U}$ given by (3) for which the condition (7) holds can be called adaptive, since the right hand side of (7) is equal to the limit of the mean square error of the best linear predictor of U based on the random process X with mean value equal to zero.

3. Kriging Predictors with Estimated Parameters in a Regression Model with $AR(1)$ Errors

As we can see from (5) the predictor $\hat{X}(n+1)$ depends only on the last observation $X(n)$ of $\mathbf{X}$ and can be written in the form

$$\hat{X}(n+1) = f'\hat{\beta} + \frac{R(1)}{R(0)}(X(n) - (F\hat{\beta})_n), \tag{8}$$

since for the $AR(1)$ process $\frac{R(1)}{R(0)} = \rho$. Our aim is now to substitute suitable estimates $\hat{R}(0)$ and $\hat{R}(1)$ for the unknown $R(0)$ and $R(1)$ respectively and to consider the predictor

$$\hat{\hat{X}}(n+1) = f'\hat{\beta} + \frac{\hat{R}(1)}{\hat{R}(0)}(X(n) - (F\hat{\beta})_n). \tag{9}$$

The problem of estimating an unknown covariance function of stationary errors in a linear regression model was considered in Štulajter (1991), where it was shown that the estimators

$$\hat{R}(t) = \frac{1}{n-t} - \sum_{s=1}^{n-t}(X(s+t) - (F\hat{\beta})_{s+t})(X(s) - (F\hat{\beta})_s) \tag{10}$$

are consistent estimators of $R(t)$ for every fixed t if $\lim_{t\to\infty} R(t) = 0$ and X is a Gaussian process. The estimates $\hat{R}(\cdot)$ can be written in the form $\hat{R}(t) = \varepsilon' C(t)\varepsilon$, where $C(t)$; $t = 0, 1, \ldots, n-1$ are symmetric $n \times n$ matrices (see Štulajter (1989)).

These estimators are "nonparametric", while the covariance function R of our model is "parametric", it depends nonlinearly on the parameter $\theta = (\sigma^2, \rho)'$.

To estimate this parameter let us consider the nonlinear regression model

$$\hat{R}(t) = R_\theta(t) + (\hat{R}(t) - R_\theta(t)); \quad t = 0, 1 \tag{11}$$

with the parametric function $R_\theta(t) = \sigma^2 \frac{\rho^t}{1-\rho^2}$; $t = 0, 1$. Now we prove the following lemma.

Lemma 1. *The estimator* $\hat{\theta} = (\hat{\sigma}^2, \hat{\rho})' = \left(\frac{\hat{R}(0)^2 - \hat{R}(1)^2}{\hat{R}(0)}, \frac{\hat{R}(1)}{\hat{R}(0)}\right)'$ *is the least squares estimator of* $\theta = (\sigma^2, \rho)'$ *in the nonlinear regression model* (11).

Proof. We are looking for $\arg\min_{\theta \in \Theta} [(R_\theta(0) - \hat{R}(0))^2 + (R_\theta(1) - \hat{R}(1))^2] = \arg\min_{\theta \in \Theta} k(\theta)$. It is easy to show that $\hat{\theta}$ satisfies the normal equations

$$\frac{\partial}{\partial \sigma^2} k(\theta)|_{\hat{\theta}} = 0$$
$$\frac{\partial}{\partial \rho} k(\theta)|_{\hat{\theta}} = 0$$

and that $k(\cdot)$ has its minimum at $\hat{\theta}$. □

Using this result we can write the predictor $\hat{\hat{X}}(n+1)$ in the form

$$\hat{\hat{X}}(n+1) = f'\hat{\beta} + \hat{\rho}(X(n) - (F\hat{\beta})_n),$$

where $\hat{\rho} = \frac{\hat{R}(1)}{\hat{R}(0)}$ is the least squares estimator of ρ.

Now we'll investigate properties of a predictor which approximate the predictor $\hat{\hat{X}}(n+1)$. We shall proceed as follows: the least squares estimator $\hat{\rho}$ will be approximated by some random variable $\tilde{\rho}$ and instead of the estimator $\hat{\hat{X}}(n+1)$ we'll consider its approximation $\tilde{X}(n+1)$ given by

$$\tilde{X}(n+1) = f'\hat{\beta} - \tilde{\rho}(X(n) - (F\hat{\beta})_n). \tag{12}$$

The approximation $\tilde{\rho}$ of $\hat{\rho}$ can be obtained in the following manner. The nonlinear regression model (11) can be written in the form

$$\hat{R} = R_\theta + (\varepsilon' C \varepsilon - R_\theta),$$

where $\hat{R} = (\hat{R}(0), \hat{R}(1))' = \varepsilon' C \varepsilon = (\varepsilon' C(0)\varepsilon, \varepsilon' C(1)\varepsilon)'$, $R_\theta = (R_\theta(0), R_\theta(1))'$ and $C(0)$ and $C(1)$ are symmetric $n \times n$ matrices.

It was shown in Štulajter (1992) that the LSE $\hat{\theta}$ can be well approximated by $\tilde{\theta} = \theta + \overline{\theta}$, where

$$\begin{aligned}\overline{\theta} = A(\varepsilon' C \varepsilon - R(\theta)) + (J'J)^{-1} \big[&(\varepsilon' C \varepsilon - R_\theta)' N (\varepsilon' C \varepsilon - R_\theta) \\ &- \frac{1}{2} J' (\varepsilon' C \varepsilon - R_\theta)' D (\varepsilon' C \varepsilon - R_\theta)\big],\end{aligned} \tag{13}$$

where $J = \frac{\partial R_\theta}{\partial \theta}$ is a 2×2 matrix, $A = (J'J)^{-1}J'$, and N and D are arrays which are given in Štulajter (1991).

Remark. Since $\varepsilon' C \varepsilon$ converges, as $n \to \infty$, in probability to R_θ if ε is a Gaussian $AR(1)$ process, the estimator $\tilde{\theta}$ converges in probability to θ if ε is a Gaussian $AR(1)$ process.

Thus we can approximate $\hat{\rho}$ by $\tilde{\rho} = \rho + \overline{\rho}$, where $\overline{\rho}$ contains only linear combinations of quadratic forms in ε and linear combinations of products of two such quadratic forms.

Since $X(n) - (F\hat{\beta})_n = (M\varepsilon)_n$, where $M = I - F(F'F)^{-1}F'$ we can write

$$\tilde{X}(n+1) = f'\hat{\beta} + (\rho + \overline{\rho})(M\varepsilon)_n \tag{14}$$

and we see that

$$E_\beta[\tilde{X}(n+1)] = f'\beta \quad \text{for all } \beta$$

if the errors ε are such that all the first, the third and the fifth moments are equal to zero, which is fulfilled if ε are e.g. normally distributed.
Since

$$\tilde{X}(n+1) = f'\beta + f'P_1\varepsilon + (\rho + \overline{\rho})(M\varepsilon)_n,$$

where $P_1 = (F'F)^{-1}F'$ and since $X(n+1) = f'\beta + \varepsilon_{n+1}$, we can write

$$\begin{aligned} E[X(n+1) - \tilde{X}(n+1)]^2 &= E[\varepsilon_{n+1} - f'P_1\varepsilon - (\rho + \overline{\rho})(M\varepsilon)_n]^2 \\ &= E[X(n+1) - \hat{X}(n+1)]^2 - 2E[\overline{\rho}(M\varepsilon)_n(\varepsilon_{n+1} \\ &\qquad - f'P_1\varepsilon - \rho(M\varepsilon)_n)]E[\overline{\rho}(M\varepsilon)_n]^2. \end{aligned}$$

We can see from (13) that $\tilde{\theta}$ can be written in the form

$$\begin{aligned} \tilde{\theta} &= \theta + ABS(\overline{\theta}) + QUAD(\overline{\theta}) + QUAR(\overline{\theta}), \quad \text{where} \\ ABS(\overline{\theta}) &= -AR_\theta + (J'J)^{-1}(R'_\theta N R_\theta - \frac{1}{2}J'R'_\theta D R_\theta) \\ QUAD(\overline{\theta}) &= A\varepsilon' C\varepsilon - (J'J)^{-1}(2R'_\theta N\varepsilon' C\varepsilon - J'R'_\theta D\varepsilon' C\varepsilon) \\ QUAR(\overline{\theta}) &= (J'J)^{-1}(\varepsilon' C\varepsilon N \varepsilon' C\varepsilon - \frac{1}{2}J'\varepsilon' C\varepsilon D\varepsilon' C\varepsilon). \end{aligned}$$

We'll use only the terms $ABS(\overline{\rho})$ and $QUAD(\overline{\rho})$ in the sequal, and we'll neglect the terms of higher power then four by computing the mean square error. Then we can write:

$$\begin{aligned} &E[\overline{\rho}(M\varepsilon)_n(\varepsilon_{n+1} - f'P_1\varepsilon - \rho(M\varepsilon)_n)] \\ &\quad \doteq ABS(\overline{\rho})E[(M\varepsilon)_n(\varepsilon_{n+1} - f'P_1\varepsilon - \rho(M\varepsilon)_n)] \\ &\qquad + E[(QUAD(\overline{\rho}))(M\varepsilon)_n(\varepsilon_{n+1} - f'P_1\varepsilon - \rho(M\varepsilon)_n)] \end{aligned} \tag{15}$$

where $ABS(\overline{\rho})$ depends only on θ and $QUAD(\overline{\rho}) = a_0(\theta)\varepsilon' C(0)\varepsilon + a_1(\theta)\varepsilon' C(1)\varepsilon$. It is easy to show that

$$E[(M\varepsilon)_n \varepsilon_{n+1}] = m_n' r, \quad \text{where } r = (R_\theta(n), \dots, R_\theta(1))'$$

and m_n' is the n-th row of the matrix M,

$$E[(M\varepsilon)_n f' P_1 \varepsilon] = m_n' \Sigma P_1' f \qquad \text{and} \qquad E[\rho (M\varepsilon)_n^2] = \rho m_n' \Sigma m_n.$$

For computing the second expectation in (15) we need to compute

$$E[\varepsilon' C(T)\varepsilon (M\varepsilon)_n (\varepsilon_{n+1} - f' P_1 \varepsilon - \rho (M\varepsilon)_n)],$$

where $C(t)$ is a symmetric $n \times n$ matrix. This can be done as follows. We can write: $\varepsilon' C(t)\varepsilon = \varepsilon'(n+1) C(t)_{n+1} \varepsilon(n+1)$, where $C(t)_{n+1}$ is the $(n+1)\times(n+1)$ matrix, $C(t)_{n+1} = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}$, $\varepsilon(n+1) = (\varepsilon_1, \dots, \varepsilon_{n+1})'$, $(M\varepsilon)_n \varepsilon(n+1) = m_n' \varepsilon \varepsilon_{n+1} = \varepsilon'(n+1) B_{n+1} \varepsilon(n+1)$, where m_n' denotes the n-th row of the matrix M and B_{n+1} is the $(n+1)\times(n+1)$ matrix, $B_{n+1} = \frac{1}{2}\begin{pmatrix} 0 & m_n \\ m_n' & 0 \end{pmatrix}$. Thus $\varepsilon' C \varepsilon (M\varepsilon)_n \varepsilon_{n+1} = \varepsilon'(n+1) C_{n+1} \varepsilon(n+1) \varepsilon'(n+1) B_{n+1} \varepsilon(n+1)$, where C_{n+1} and B_{n+1} are symmetric $(n+1)\times(n+1)$ matrices.

By analogy every other product of two linear (in ε) forms can be written as a quadratic form $\varepsilon' B \varepsilon$ with some symmetric matrix B and we can use the expression

$$E[\varepsilon' C \varepsilon \varepsilon' B \varepsilon] = 2\,\mathrm{tr}\,(C\Sigma B\Sigma) + \mathrm{tr}\,(C\Sigma)\,\mathrm{tr}\,(B\Sigma)$$

which holds (see Štulajter (1989)) for every random vector ε which is $N(0,\Sigma)$ distributed.

It remains to express $E[\overline{\rho}(M\varepsilon)_n]^2$ as

$$E[\overline{\rho}(M\varepsilon)_n]^2 \doteq ABS(\overline{\rho}) E[(M\varepsilon)_n^2] + 2ABS(\overline{\rho}) E[QUAD(\overline{\rho})(M\varepsilon)_n^2]$$

and to compute the expectations by the same manner as before.

Thus we are able to write an approximate expression for the mean square error $E[\tilde{X}(n+1) - X(n+1)]^2$ for the case when the $AR(1)$ process is Gaussian. A closed form of this expression is rather complicated and we'll not write it.

Since $\tilde{\theta}$ is a good approximation for $\hat{\theta}$ (see Štulajter (1992)) $E[\hat{\hat{X}}(n+1) - X(n+1)]^2$ can be well approximated by the same expression as $E[\tilde{X}(n+1) - X(n+1)]^2$.

Remark. The approach described can be used also for covariance functions which we get after a reparametrization of $AR(1)$ model. For example if the errors have covariance function $R_\theta(t) = \sigma^2 e^{-\alpha t}$ then the predictor given by (8) can be regarded as one in which the residual correction term is based only on the last observation. In this case $R(1)/R(0) = e^{-\alpha}$, where α is the only unknown parameter. This parameter can be estimated from $\hat{R} = (\hat{R}(0), \dots, \hat{R}(m))'$ using the nonlinear regression model

$$\hat{R} = R_\theta + (\hat{R} - R_\theta)$$

where $R_\theta = (R_\theta(0), \dots, R_\theta(m))'$ and m is a number, $m < n$. The problem of choosing m is open (usually $m \leq n/2$). The approximation $\tilde{\alpha}$ for for α is given by (13) and and $\hat{R}(1)/\hat{R}(0) = e^{-\hat{\alpha}}$ can be approximated using $\tilde{\alpha}$ and the Taylor series expansion of the function e^{-t} at the point α, the true value of the parameter.

References

Bhansali R. J., *Effect of not knowing the order of autorgression on the mean squared error of prediction I*, J. Amer. Stat. Assoc. **78** (1981), 588–597.

Fuller W. A and Hasza D. P., *Properties of predictors for autoregressive time series*, J. Amer. Stat. Assoc. **76** (1981), 155–161.

Harville D. A., *Maximum likelihood approaches to variance component estimation and to related problems*, J. Amer. Stat. Assoc. **57** (1977), 320–338.

______, *Decomposition of prediction error*, J. Amer. Stat. Assoc. **80** (1985), 132–138.

______, *BLUP (best linear unbiased prediction) and beyond*, In Advances in Statistical methods for Genetic Improvement of Livestock (D. Gianola and K. Hammond, eds.), Springer-Verlag, New York, 1990, pp. 239-276.

Harville D. A. and Jeske D. R., *Mean squared error of estimation or prediction under a general linear model*, J. Amer. Stat. Assoc. **87** (1992), 724–731.

Journel A., *Kriging in the terms of predictions*, J. Inter. Assoc. Math. Geol. **9** (1977), 563–586.

Journel A. and Huijbregts C., *Mining Geostatistics*, Academic, New York, 1978.

Kunitomo N. and Yamamoto T., *Properties of predictors in misspecified autoregressive time series models*, J. Amer. Stat. Assoc. **80** (1985), 941–950.

Lewis R. and Reinsel G. C., *Prediction of multivariate time series by autoregression model fitting*, J. Multivar. Anal. **16** (1985), 393–411.

Stein M. L., *Asymptotically efficient prediction of a random field with a misspecified covariance function*, Ann. Stat. **16** (1988), 55–63.

______, *Uniform asymptotic optimality of linear predictors of a random field using an incorrect second order structure*, Ann. Stat. **18** (1990a), 850–872.

______, *Bounds on the efficiency of linear predictors using an incorrect covariance function*, Ann. Stat. **18** (1990b), 1116–1138.

______, *A comparison of generalised cross validation and modified maximum likelihood for estimating the parameters of a stochastic process*, Ann. Stat. **18** (1990c), 1139–1157.

Štulajter F., *Estimation in Stochastic Processes*, Alfa, Bratislava, 1989. (Slovak)

______, *Consistency of linear and quadratic least squares estimators in regression models with covariance stationary errors*, Appl. Math. **36** (1991), 149–155.

______, *Mean square error matrix of an approximate least squares estimator in a nonlinear regression model with correlated errors*, Acta Math. Univ. Comen **LXI 2** (1992), 251–261.

Toyooka Y., *Prediction error in a linear model with estimated parameters*, Biometrica **69** (1982), 453–459.

Zimmerman D. L. and Cressie N., *Mean squared error in spatial linear model with estimated covariance parameters*, Ann. Inst. Statist. Math. **44** (1992), 27–43.

F. Štulajter, Department of Probability and Statistics, Faculty of Mathematics and Physics, Comenius University, 842 15 Bratislava, Slovakia