Ann. Inst. Statist. Math.
Vol. 53, No. 2, 370–379 (2001)

STATISTICAL INFERENCE IN A LINEAR MODEL FOR SPATIALLY LOCATED SENSORS AND RANDOM INPUT *

STANISŁAW GNOT[1]**, EWARYST RAFAJŁOWICZ[2] AND AGNIESZKA URBAŃSKA-MOTYKA[1]

[1] *Pedagogical University, Institute of Mathematics, pl. Słowiański 9, 65-069 Zielona Góra, Poland*
[2] *Wrocław University of Technology, Institute of Engineering Cybernetics, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland*

(Received June 8, 1998; revised October 18, 1999)

Abstract. In the paper we consider a random linear model for observations provided by spatially located sensors measuring signals coming from one source. For this model a set of sufficient and complete statistics are found, and it is shown that the maximum likelihood estimators of unknown parameters (characteristics of the source) are functions of those statistics. The problem of nonnegative estimators of variance components of the model is shortly discussed. Comparisons of the mean squared errors of several estimators are given. Numerical example concerning hunting for defects in solar cells is considered in details.

Key words and phrases: Inverse problem, random linear model, sufficient statistics, variance components, ML estimators.

1. Introduction

Quick inspection of papers and monographs from a wide range of scientific and engineering disciplines reveals that the so called inverse problems for spatially located sources are recently the subject of intensive investigations (see e.g. Engl and Groetsch (1987) for the collection of papers on inverse problems, where also other references can be found). Roughly speaking, by the inverse problems of this class we mean attempts to estimate various characteristics of a source, e.g., its location or intensity, from indirect observations provided by sensors. The notion of a source is consider here in a rather wide sense, including sources of electromagnetic or acoustic signals or noises as well as sources of heat, mass or energy. Similarly, also the notion of a sensor is treated here in an abstract manner, i.e., it can be any device, which indirectly provides information concerning the sources of interest. The third ingredient, which is necessary to formulate linear models for inverse problem with one source of signals is a vector k with components k_j, which indicates the influence of the source on the j-th sensor, $j = 1, \dots, J$ (more detailed description is given in the next section).

Let us suppose that at discrete instants of time noisy observations from a source are available. It is also known that J sensors with given locations change its intensity at random in the intervals between observations. Assume that also the influence vector $k = (k_1, \dots, k_J)'$ of source on sensers is given. Some remarks concerning knowledge of the vector k are in order. In some cases k can be measured directly in earlier

*This work was supported by Komitet Badań Naukowych, grant No 8T11A 00313.

**Now at Technical University, Institute of Mathematics, Podgórna 50, 65-246 Zielona Góra, Poland.

carefully controlled experiments, since elements k_j are directly interpretable. The following example illustrates the other extreme, when k can be obtained from theoretical considerations. Consider a linear distributed parameter system described by a partial differential equation of the elliptic type with given boundary conditions. Such an equation is treated here as a model of diffusion considered in the steady state. Let us assume that the Green function (Stakgold (1968)) of this equation is known or its numerical approximation is available. Suppose we have a source of pollution located at a known point x and J monitoring stations placed at x_j, $j = 1, 2, \ldots, J$. Then as the influence of a pollution source acting pointwise at x as measured in j-th monitoring station we can take $k_j = k(x_j, x)$. In particular, the Green function frequently met in applications has the form: $k(x_j, x) = (\sqrt{2\pi}a)^{-1}\exp[-\|x_j - x\|^2/2a^2]$, where $a > 0$ denotes the diffussion coefficient. This coefficient is usually unknown and have to be estimation from observations or taken from standard tables. For further details concerning this system see Rafajłowicz (1995), where the above model has been considered from the point of view of D-optimality.

The aim of the statistician is to estimate the mean intensity and the variance of source as well as the variance of technical errors. More precise problem statement together with references to the theory of variance components estimation is given in the next section. Here, we concentrate on practical motivations behind estimating mean intensity and measures of variability of the process. Possible applications include water and air pollution monitoring systems, in which estimation of mean intensity of individual source can be used for identifying the most dangerous places of pollutants emission. Furthermore, knowledge of variances is important in the cases, when instantaneous increase of levels of toxic emission can be dangerous. Analogous situations arise, if instead of chemical pollutants we are faced with a source of noises or vibrations in industrial regions.

All the above mentioned examples cover the case when, usually pointwise, source of emission is (or may be) present, while its (possible) activity is measured at several locations. In such cases one is usually faced with a low level background emission spread over a certain region. In our model below the background emission is taken into account as additional errors. At least the following two questions are of interest. How to separate variability of the source emission from variability of the background emission and how to detect whether variability of the source is significant during the period of observations?

The paper is organized as follows. In Section 2 a statistical model is formulated. In Section 3 formulas for sufficient and complete statistics, unbiased estimators for the expectation of the intensivity of source and the variance components are given. It is shown that obtained estimators are the uniformly minimum variance unbiased estimators as functions of sufficient and complete statistics. The problem of maximum likelihood estimation of these parameters is considered in Section 4. An explicit form of these estimators for variance components are presented as functions of sufficient and complete statistics. Finally results of comparisons of the mean squared errors of several estimators of variances are presented. In Section 5 some numerical example concerning hunting for defects in solar cells is given to illustrate applicability of the theoretical results.

2. The linear model

To introduce the model let us denote by v, the intensity of the source, and let y_j denotes the measurement provided by the j-th sensor, $j = 1, 2, \ldots, J$. Moreover let k_j

be the influence of the source with the unit intensity on the j-th sensor. It is assumed that the vector $k = (k_1, \ldots, k_J)'$, further called the unit response vector, is given. Under these notations, the model considered in this paper has the form:

$$y = vk + \varepsilon, \tag{2.1}$$

where $y = (y_1, \ldots, y_J)'$, while $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_J)'$ denotes the vector of uncorrelated random errors having the expectations zero and the common variance σ_e^2. The intensity v is assumed to be unobserved random variable with an unknown expectation $E(v) = \mu$ and an unknown nonnegative variance σ_1^2, while σ_e^2 is assumed to be positive. Additionally we assume that v and ε_j's are uncorrelated, i.e., $E((v - \mu)\varepsilon_j) = 0$ for each j.

Under these assumptions the expectation $E(y)$ and the covariance matrix Σ of y have the form:

$$E(y) = \mu k, \tag{2.2}$$

$$\Sigma = \sigma_1^2 kk' + \sigma_e^2 I_J. \tag{2.3}$$

The main problem considered in this paper is to estimate the vector $\sigma = (\sigma_1^2, \sigma_e^2)'$ of variance components from the independent sequence of observed vectors $y^{(1)}, y^{(2)}, \ldots, y^{(n)}$, each of them having the normal distribution with the expectation and the covariance matrix given by (2.2) and (2.3), respectively. Here $y^{(i)} = (y_{i1}, y_{i2}, \ldots, y_{iJ})'$, where y_{ij} is the i-th measurement provided by the j-th sensor, $i = 1, 2, \ldots, n$.

Let $\boldsymbol{Y} = \mathrm{vec}(y^{(1)}, \ldots, y^{(n)})$. It can be easily established that $\boldsymbol{Y}$ has a normal distribution with

$$E(\boldsymbol{Y}) = 1_n \otimes E(y) = \mu \boldsymbol{X}, \tag{2.4}$$

$$\mathrm{Cov}(\boldsymbol{Y}) = \boldsymbol{\Sigma} = I_n \otimes \Sigma = \sigma_1^2 \boldsymbol{V} + \sigma_e^2 \boldsymbol{I}_N, \tag{2.5}$$

where $\boldsymbol{X} = 1_n \otimes k$, $\boldsymbol{V} = I_n \otimes kk'$, $N = n \cdot J$, while for arbitrary matrices A, B the symbol $\otimes$ denotes the Kronecker product of A and B. We denote this model by $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2 \boldsymbol{V} + \sigma_e^2 \boldsymbol{I}_N)$.

3. Unbiased estimation of parameters

3.1 *Some algebraic results*

In this section we present some algebraic results, which are useful for further considerations concerning the problem of estimation of unknown parameters in the model $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2 \boldsymbol{V} + \sigma_e^2 \boldsymbol{I}_N)$. The following two propositions give some characterization of the model, which are essential for further considerations.

PROPOSITION 3.1. *In the model $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2 \boldsymbol{V} + \sigma_e^2 \boldsymbol{I}_N)$ the following two conditions hold*

(a) $\boldsymbol{PV} = \boldsymbol{VP}$,

(b) $\boldsymbol{V}^2 = k'k\boldsymbol{V}$,

where

$$\boldsymbol{P} = \frac{1}{nk'k} 1_n 1_n' \otimes kk'$$

is the orthogonal projector on $R(\boldsymbol{X})$.

PROPOSITION 3.2. *For an arbitrary symmetric and idempotent matrix $\boldsymbol{P}$ and for symmetric matrix $\boldsymbol{V}$ such that $\boldsymbol{V}^2 = \alpha\boldsymbol{V}$ for some scalar α, the conditions*

$$\boldsymbol{PV} = \boldsymbol{VP} \tag{3.1}$$

and

$$(\boldsymbol{MVM})^2 = \alpha\boldsymbol{MVM}, \tag{3.2}$$

are equivalent. Here $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{P}$.

PROOF. It is clear that under $\boldsymbol{V}^2 = \alpha\boldsymbol{V}$ the condition (3.1) implies (3.2). Now let $\boldsymbol{V}^2 = \alpha\boldsymbol{V}$ and $(\boldsymbol{MVM})^2 = \alpha\boldsymbol{MVM}$ for some α. Then we have

$$\begin{aligned}\boldsymbol{MV}^2\boldsymbol{M} &= \boldsymbol{MV}(\boldsymbol{P}+\boldsymbol{M})\boldsymbol{VM}\\ &= \boldsymbol{MVPVM} + \boldsymbol{MVMVM} = \alpha\boldsymbol{MVM}.\end{aligned}$$

On the other hand $\boldsymbol{MVMVM} = (\boldsymbol{MVM})^2 = \alpha\boldsymbol{MVM}$. It follows that $\boldsymbol{MVPVM} = \boldsymbol{0}$. In consequence $\boldsymbol{MVP} = \boldsymbol{PVM} = 0$, and since $\boldsymbol{V} = \boldsymbol{PVP} + \boldsymbol{MVP} + \boldsymbol{PVM} + \boldsymbol{MVM}$ we have $\boldsymbol{V} = \boldsymbol{PVP} + \boldsymbol{MVM}$ and finally $\boldsymbol{VP} = \boldsymbol{PV}$. □

Remark 3.1. In the model $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2\boldsymbol{V} + \sigma_e^2\boldsymbol{I}_N)$, following Propositions 3.1 and 3.2 we have $\boldsymbol{PV} = \boldsymbol{VP}$, and for the linear space $\mathcal{V} = sp\{\boldsymbol{MVM}, \boldsymbol{M}\}$ has the following property $\boldsymbol{A} \in \mathcal{V}$ implies that also $\boldsymbol{A}^2 \in \mathcal{V}$. It has been proved by Seely (1977) and also by Zmyślony ((1981), Theorem 1) that in such a case there exist minimal sufficient and complete statistics for the family of the normal distributions of $\boldsymbol{Y}$.

3.2 *Minimal sufficient and complete statistics*

Let y_{ij} be the j-th component of $y^{(i)}$. From Seely ((1977), Theorem 2.7 and Example 2.4) we find that one of the possible representation of the minimal sufficient and complete statistics in the model $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2\boldsymbol{V} + \sigma_e^2\boldsymbol{I}_N)$ are

$$\begin{aligned}
\boldsymbol{T} &= (1_n \otimes k)'\boldsymbol{Y} = \sum_{i=1}^{n} k'y^{(i)} = \sum_{i=1}^{n}\sum_{j=1}^{J} k_j y_{ij},\\
\boldsymbol{T}_1 &= \frac{1}{(n-1)k'k}\boldsymbol{Y}'\boldsymbol{M}(I_n \otimes kk')\boldsymbol{MY}\\
&= \frac{1}{(n-1)k'k}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{J} k_j y_{ij}\right)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{J} k_j y_{ij}\right)^2\right],\\
\boldsymbol{T}_2 &= \frac{1}{N-n}[\boldsymbol{Y}'\boldsymbol{MY} - (n-1)\boldsymbol{T}_1]\\
&= \frac{1}{N-n}\left[\sum_{i=1}^{n}\sum_{j=1}^{J} y_{ij}^2 - \frac{1}{nk'k}\left(\sum_{i=1}^{n}\sum_{j=1}^{J} k_j y_{ij}\right)^2 - (n-1)\boldsymbol{T}_1\right]\\
&= \frac{1}{N-n}\left[\sum_{i=1}^{n}\sum_{j=1}^{J} y_{ij}^2 - \frac{1}{k'k}\sum_{i=1}^{n}\left(\sum_{j=1}^{J} k_j y_{ij}\right)^2\right].
\end{aligned}$$

Table 1. ANOVA table for the minimal sufficient and complete statistics.

Source of variation	Sum of square	d.f.	Mean square	Expected mean of square
between replications	$(n-1)\boldsymbol{T}_1$	$n-1$	$\boldsymbol{T}_1$	$k'k\sigma_1^2+\sigma_e^2$
within replications	$(N-n)\boldsymbol{T}_2$	$N-n$	$\boldsymbol{T}_2$	σ_e^2
total	$\boldsymbol{Y'MY}$	$N-1$		

Moreover $\boldsymbol{T}_1$,$\boldsymbol{T}_2$ are independently distributed and

$$\frac{(n-1)\boldsymbol{T}_1}{k'k\sigma_1^2+\sigma_e^2}\sim\chi^2_{n-1},\qquad \frac{(N-n)\boldsymbol{T}_2}{\sigma_e^2}\sim\chi^2_{N-n}.$$

The above result is presented in Table 1.

3.3 *Uniformly minimum variance unbiased estimation*

In the model under consideration the least squared (ANOVA) estimator $\hat{\mu}=\frac{1}{nk'k}\boldsymbol{T}$ of μ is the uniformly minimum variance unbiased estimator of μ as a function of sufficient and complete statistic $\boldsymbol{T}$. For the same reason

$$\hat{\sigma}^2_{eUMV}=\boldsymbol{T}_2,\qquad \hat{\sigma}^2_{1UMV}=\frac{1}{k'k}(\boldsymbol{T}_1-\boldsymbol{T}_2)$$

and $f_1\hat{\sigma}^2_{1UMV}+f_2\hat{\sigma}^2_{eUMV}$ are the uniformly minimum variance unbiased estimators for singular variances σ_e^2, σ_1^2 and for an arbitrary linear combination $f_1\sigma_1^2+f_2\sigma_e^2$, respectively. There are some unconveniences connected with the using of $\hat{\sigma}^2_{1UMV}$, since the estimator can take negative values with positive probability. Several positive definite biased estimators of σ_1^2 have been considered in literature. Gnot and Kleffe (1983) have proposed the biased estimator of the form

$$\hat{\sigma}^2_{1NN}=\frac{n-1}{(n+1)k'k}\boldsymbol{T}_1,$$

and proved that $\hat{\sigma}^2_{1NN}$ is admissible for σ_1^2 in the class of nonnegative estimators with respect to the mean squared error loss function. Note that this estimator is nonnegative definite by construction and

$$E(\hat{\sigma}^2_{1NN})=\frac{n-1}{n+1}\left(\sigma_1^2+\frac{1}{k'k}\sigma_e^2\right)$$

is close to σ_1^2 if n and $k'k$ are sufficiently large. Mathew *et al.* (1992) specified a sufficient condition under which $\hat{\sigma}^2_{1NN}$ has a uniformly smaller mean squared error than unbiased estimator $\hat{\sigma}^2_{1UMV}$ in general models with two variance components. Simulating study of the mean squared error of $\hat{\sigma}^2_{1NN}$ can be found in Gnot *et al.* (1994).

Remark 3.2. We can also use the statistics $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ to construct a test for testing H: $\sigma_1^2=0$ vs K: $\sigma_1^2>0$. Test at a significant level α reject H if $F=\boldsymbol{T}_1/\boldsymbol{T}_2$ excceed the critical value F_α of a central F distribution with $n-1$ and $N-n$ degrees of freedom. It follows from Mathew (1989) that this is the uniformly most powerful invariant test for testing H vs K.

4. ML estimators of variance components

We start with a looking for the maximum likelihood (ML) estimators of μ and of the variance components σ_1^2 and σ_e^2 in our model $\mathcal{N}(\boldsymbol{Y}, \mu\boldsymbol{X}, \boldsymbol{\Sigma} = \sigma_1^2\boldsymbol{V} + \sigma_e^2\boldsymbol{I}_N)$. Following Rao and Kleffe ((1988), p. 232, Theorem 9.2.1) or Searle *et al.* ((1992) p. 237) the ML estimators of σ_1^2 and σ_e^2 are solutions of the following nonlinear equation system

$$\begin{pmatrix} \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{V}) & \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}) \\ \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}) & \mathrm{tr}(\boldsymbol{\Sigma}^{-2}) \end{pmatrix} \begin{pmatrix} \sigma_1^2 \\ \sigma_e^2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}'\boldsymbol{G}\boldsymbol{V}\boldsymbol{G}\boldsymbol{Y} \\ \boldsymbol{Y}'\boldsymbol{G}^2\boldsymbol{Y} \end{pmatrix}, \tag{4.1}$$

where

$$\boldsymbol{G} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1} = (\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M})^{+}.$$

Remark 4.1. We assume that $\boldsymbol{MVM} \neq 0$, i.e. that $\boldsymbol{G} \neq \boldsymbol{M}$. The assumption is necessary to avoid trivialities and contradictions.

Generally there is no analytical expressions for the ML estimators, and some iterative procedures to obtain a solution of the above nonlinear equation system are proposed in the literature (cf. Searle *et al.* (1992), Section 8). However, since $\boldsymbol{PV} = \boldsymbol{VP}$ and $\boldsymbol{V}^2 = \boldsymbol{k}'\boldsymbol{k}\boldsymbol{V}$ (see Proposition 3.1), following Szatrowski (1980) and Szatrowski and Miller (1980) in our model, the ML estimator of μ is $\hat{\mu} = \frac{1}{nk'k}\boldsymbol{T}$, and coincides with the ANOVA estimators, while the ML estimators of σ_1^2 and σ_e^2 can be presented in an explicit form. It is because under the conditions $\boldsymbol{V}^2 = k'kV$ we have

$$\boldsymbol{\Sigma}^{-1} = s_1\boldsymbol{V} + s_2\boldsymbol{I}_N, \tag{4.2}$$

where

$$s_1 = -\frac{\sigma_1^2}{\sigma_e^2(k'k\sigma_1^2 + \sigma_e^2)}, \qquad s_2 = \frac{1}{\sigma_e^2}.$$

Moreover using Proposition 3.2 we get $\boldsymbol{G} = (\boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M})^{+} = s_1\boldsymbol{MVM} + s_2\boldsymbol{M}$. Thus the equation system (4.1) reduces to the following linear one

$$\begin{pmatrix} \mathrm{tr}(\boldsymbol{V}^2) & \mathrm{tr}(\boldsymbol{V}) \\ \mathrm{tr}(\boldsymbol{V}) & N \end{pmatrix} \begin{pmatrix} \sigma_1^2 \\ \sigma_e^2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}'\boldsymbol{MVMY} \\ \boldsymbol{Y}'\boldsymbol{MY} \end{pmatrix}. \tag{4.3}$$

From the above we find that

$$\hat{\sigma}_{1ML}^2 = \frac{N\boldsymbol{Y}'\boldsymbol{MVMY} - k'kn\boldsymbol{Y}'\boldsymbol{MY}}{(k'k)^2 n(N-n)}, \tag{4.4}$$

$$\hat{\sigma}_{eML}^2 = \frac{k'k\boldsymbol{Y}'\boldsymbol{MY} - \boldsymbol{Y}'\boldsymbol{MVMY}}{k'k(N-n)}. \tag{4.5}$$

We easily find that $\hat{\sigma}_{1ML}^2$ and $\hat{\sigma}_{eML}^2$ can be expressed as functions of the minimal and sufficient statistics as follows

$$\hat{\sigma}_{1ML}^2 = \frac{1}{k'k}\left(\frac{n-1}{n}\boldsymbol{T}_1 - \boldsymbol{T}_2\right), \qquad \hat{\sigma}_{eML}^2 = \boldsymbol{T}_2,$$

where T_1 and T_2 are the minimal and sufficient statistics defined in Section 3.2. It follows that $\hat{\sigma}^2_{eML}$ coincides with the uniformly minimum variance unbiased estimator, while $\hat{\sigma}^2_{1ML}$ is biased estimator of σ_1^2, since

$$E(\hat{\sigma}^2_{1ML}) = \frac{n-1}{n}\sigma_1^2 - \frac{1}{nk'k}\sigma_e^2.$$

Remark 4.2. Let us note that $\hat{\sigma}^2_{eML}$ is positive by construction, while $\hat{\sigma}^2_{1ML}$ can take negative values with positive probability. It has been shown by Gnot *et al.* (1997) that if $\hat{\sigma}^2_{1ML} < 0$, then the ML estimators of σ_1^2 and σ_2^2 under nonnegativity restrictions have the form $\hat{\sigma}^2_{1NML} = 0$, while

$$\hat{\sigma}^2_{eNML} = \frac{\boldsymbol{Y'MY}}{N} = \frac{1}{N}\left[\sum_{i=1}^{n}\sum_{j=1}^{J} y_{ij}^2 - \frac{1}{nk'k}\left(\sum_{i=1}^{n}\sum_{j=1}^{J} k_j y_{ij}\right)^2\right]$$

coincides with the unrestricted ML estimator of σ_e^2 in the model $E(y) = \mu k$, $\Sigma = \sigma_e^2 I_J$, i.e. in the new model, after dropping σ_1^2. For the balanced one-way random model this problem has been considered by Searle *et al.* ((1992), Section 3.7).

4.1 *Comparison of the mean squared errors*

In this section we compare the mean squared error of the maximum likelihood estimator

$$\hat{\sigma}^2_{1ML} = \frac{1}{k'k}\left(\frac{n-1}{n}\boldsymbol{T}_1 - \boldsymbol{T}_2\right)$$

with the variance of the uniformly minimum variance unbiased estimator

$$\hat{\sigma}^2_{1UMV} = \frac{1}{k'k}\left(\boldsymbol{T}_1 - \boldsymbol{T}_2\right)$$

and with the mean squared error of the nonnegative estimator

$$\hat{\sigma}^2_{1NN} = \frac{n-1}{(n+1)k'k}\boldsymbol{T}_1.$$

Using the results from Section 3 after straightforward calculations we find that the mean squared errors (MSE's) as sums of the appropriate square of biases and variances are given by the following formulas

$$\frac{1}{\sigma_e^4}MSE(\hat{\sigma}^2_{1UMV}) = \frac{2}{n-1}\rho^2 + \frac{4}{(n-1)k'k}\rho + \frac{2}{(n-1)(k'k)^2} + \frac{2}{(N-n)(k'k)^2},$$

$$\frac{1}{\sigma_e^4}MSE(\hat{\sigma}^2_{1NN}) = \frac{2}{n+1}\rho^2 + \frac{n-1}{(n+1)(k'k)^2},$$

$$\frac{1}{\sigma_e^4}MSE(\hat{\sigma}^2_{1ML}) = \frac{2n-1}{n^2}\rho^2 + \frac{2(2n-1)}{n^2k'k}\rho + \frac{2n-1}{n^2(k'k)^2} + \frac{2}{(N-n)(k'k)^2},$$

where $\rho = \sigma_1^2/\sigma_e^2$. Since for each $n > 1$

$$\frac{2}{n+1} < \frac{2n-1}{n^2} < \frac{2}{n-1},$$

we find that $\hat{\sigma}^2_{1ML}$ is uniformly better than $\hat{\sigma}^2_{1UMV}$, with respect to the mean squared error loss function. Let us however note that both risks are very close each to other if n is sufficiently large.

5. Numerical example-hunting for defects in solar cells

5.1 *Description of the experiment*

To illustrate potential applicability of the results we consider searching for defects in semiconductor devices. Our real-life data come from measurements of a solar cell, but the approach seems to be applicable to other types of semiconductor devices.

A solar cell performs well if there are no defects in its internal structure i.e., if donors and acceptors are uniformly spread over areas that they should occupy. A possible defect can be detected by passing various types of electrical signals through a device at hand and registrating the temperature distributions arising at its surface. It is expected that defective parts of the semiconductor responded to various input signals by the local temperature variability, which is larger than those generated by nondefective parts of the semiconductor. Defetive parts can also have locally higher temperature and for this reason they are called "hot spots". This can be explained by the fact that nonuniform distribution of donors and/or acceptors lead to local changes of the resistance, leading to changes in the heat generation.

The experiment was performed as follows:

1. Measurements of the surface temperature of the cell were made at nodes of a rectangular grid with coordinates denoted further as $(j_1\Delta_1, j_2\Delta_2)$ $j_1 = 1, 2, \ldots, J_1$, $j_2 = 1, 2, \ldots, J_2$ or numbered in the lexicographical order as $j = 1, 2, \ldots, J$, $J = J_1 \cdot J_2$.

2. Possible locations of the heat source are restricted to the nodes of the above mentioned grid.

3. At most one hot spot (heat source) is present in the cell and, if it is present, its position, denoted by $(j_1^\star\Delta_1, j_2^\star\Delta_2)$, does not change between all the series of measurements.

5.2 *Assumptions about the model and numerical results*

The following two assumptions were made in order to simplify the problem:

1. The time between subsequent changes of input signals (currents) was sufficiently large that the steady state temperature distribution over the cell surface was attained. The assumption is necessary to be sure the $y^{(1)}, y^{(2)}, \ldots, y^{(n)}$ are independent (see Section 2).

2. The response k_{j_1,j_2} at grid point $(j_1\Delta_1, j_2\Delta_2)$ for the unit source at $(j_1^\star\Delta_1, j_2^\star\Delta_2)$ has the following form:

$$k_{j_1,j_2} = \exp[-c((j_1 - j_1^\star)\Delta_1^2 + (j_2 - j_2^\star)\Delta_2^2)], \tag{5.1}$$

where coefficient c is a constant, which is treated as given. In the example below we used $c = 10^{-4}$ (in fact, value of c was taken from observations). Rearranging (5.1) for $j_1 = 1, 2, \ldots, J_1$, $j_2 = 1, 2, \ldots, J_2$ in the lexicographical order we form vector k (for details see Stakgold (1968)).

Under these assumptions each series $y^{(i)}$, $i = 1, 2, \ldots, n$ can be interpreted as the measurements of the temperature surface, which depends only on the two spatial coordinates, and we can assume, that $y^{(1)}, y^{(2)}, \ldots, y^{(n)}$ are normally distributed independent random vectors. The expectation and the covariance matrix of $\boldsymbol{Y} = \text{vec}(y^{(1)}, \ldots, y^{(n)})$ are assumed to have the forms given by (2.4) and (2.5).

In the numerical example $N = 5616$ observations of a cell were made and digitally recorded. Observations were arranged into $n = 117$ series as follows: each series were measured on 6×8 grid, i.e., $J_1 = 6$, $J_2 = 8$, $J = 48$, with the step sizes $\Delta_1 = 12$,

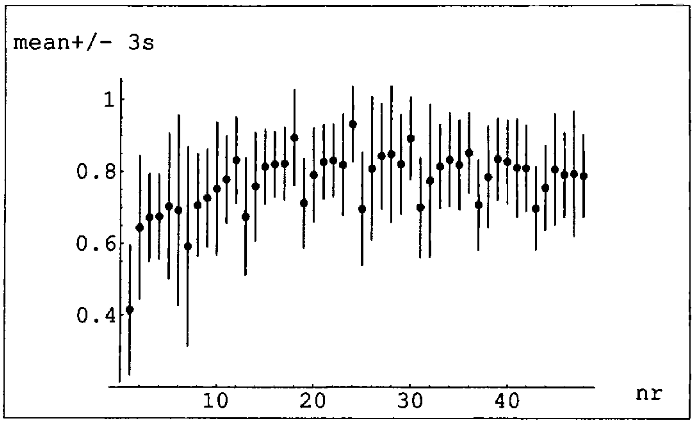

Fig. 1. Empirical means ±3 empirical dispersions at linearly ordered measurement points.

Table 2. Estimates of σ_1 and σ_e for different possible positions of the heat source. ⋆ replaces negative values of the corresponding $\hat{\sigma}_1^2$ (values of the order -10^{-5} were found).

Case	Position	$\hat{\mu}$	$\hat{\sigma}_{eUMV} = \hat{\sigma}_{eML}$	$\hat{\sigma}_{1UMV}$	$\hat{\sigma}_{1ML}$	F
1	(40, 48)	0.836	0.076	0.0077	0.0076	1.42
2	(56, 36)	0.861	0.098	⋆	⋆	0.87
3	(56, 60)	0.88	0.069	0.0078	0.0078	1.47

$\Delta_2 = 8$, respectively, measured in a conventional units dictated by the measuring device. Before actual data processing, observations were normalized to $(0, 1]$ interval, i.e., each observation was divided by the constant maximal range of the measuring device.

A rough overview of the collected observations is provided by Fig. 1, which the grid points are arranged linearly and for each of them the empirical mean ±3 empirical dispersion are shown.

During the experiment it was not sure, whether a heat source is present in the cell. In the case of its presence, its position was not known precisely, but the following grid points (coordinates in the conventional units) (40, 48), (56, 36), (56, 60) are suspected to be "hot spots". For each of the above mentioned points the variance componenets were estimated as if only one heat source had been active. Simultaneously, the F-statistic was used to verify the hypothesis that the source variance at a given point is zero, i.e., the source is not present there.

For comparisons, the variance components were calculated not only by the maximum likelihood estimates, but also by all other methods mentioned in the paper. The results of calculations, using the Mathematica system, are summarized in Table 2.

Calculated values of F-statistic were compared with the critical value 1.227, corresponding to $\alpha = 0.05$. The hypothesis $\sigma_1^2 = 0$ is not rejected only in case 2, what suggests the presence of the hot spot at point (56, 36).

Acknowledgements

The authors are grateful to the referees for their useful comments and remarks, which substantially improved the paper. The authors express sincere thanks to Dr. R. Szeloch from the Institute of Electron Technology, Wrocław University of Technology for providing data used in his paper.

REFERENCES

Engl, H. W. and Groetsch, C. W. (ed.) (1987). *Inverse and Ill-posed Problems*, Academic Press, New York.

Gnot, S. and Kleffe, J. (1983). Quadratic estimation in mixed linear models with two variance components, *J. Statist. Plann. Infererence*, **8**, 267–279.

Gnot, S., Trenkler, G. and Stemann, D. (1994). Nonnegative admissible invariant quadratic estimation in linear models with two variance components, *Proceedings*, *International Conference LINSTAT'93*, 129-137, Kluwer Academic Publishers, Boston.

Gnot, S., Urbańska-Motyka, A. and Zmyślony, R. (1997). Nonnegative maximum likelihood estimation in a linear model of variance components, *Discuss. Math. Algebra Stochastic Methods*, **17**, 215–224.

Mathew, T. (1989). Optimum invariant tests in mixed linear models with two variance components (ed. Y. Dodge), *Statistical Data Analysis and Inference*, 381-388, North-Holand, Amsterdam.

Mathew, T., Sinha, B. K. and Sutradhar, B.C. (1992). Nonnegative estimation of variance components in unbalanced mixed models with two variance components, *J. Multivariate Anal.*, **42**, 77–101.

Rafajłowicz, E. (1995). Optimal measurement schedules and sensors' location for estimating intensities of random sources, *Proc. II Int. Symposium on Methods and Models in Automation and Robotics*, Międzyzdroje, Poland, 30 August–2 September 1995, 169–174.

Rao, C. R. and Kleffe, J. (1988). *Estimation of Variance Components and Applications*, Elsevier, North Holland, Amsteldam.

Searle, S. R., Casella, G. and McCulloch, Ch. E. (1992). *Variance Components*, Wiley, New York.

Seely, J. (1977). Minimal sufficient statistics and completeness for multivariate normal families, *Sankhyā*, *Ser. A*, **39**, 170–185.

Stakgold, I. (1968). *Boundary Value Problems of Mathematical Physics*, MacMillan, New York.

Szatrowski, T. H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances, *Ann. Statist.*, **8**, 803–810.

Szatrowski, T. H. and Miller, J. J. (1980). Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance, *Ann. Statist.*, **8**, 811–819.

Zmyślony, R. (1981). Completeness for a family of normal distributions, Mathematical Statistics, *Banach Center Publ.*, **6**, 355–357.