Metrika (2009) 69:45–54
DOI 10.1007/s00184-008-0173-8

Minimax robust designs for field experiments

Beiyan Ou · Julie Zhou

Received: 24 July 2006 / Published online: 19 February 2008

Abstract Experimental designs for field experiments are useful in planning agricultural experiments, environmental studies, etc. Optimal designs depend on the spatial correlation structures of field plots. Without knowing the correlation structures exactly in practice, we can study robust designs. Various neighborhoods of covariance matrices are introduced and discussed. Minimax robust design criteria are proposed, and useful results are derived. The generalized least squares estimator is often more efficient than the least squares estimator if the spatial correlation structure belongs to a small neighborhood of a covariance matrix. Examples are given to compare robust designs with optimal designs.

Keywords Check variety · Covariance neighborhood · Generalized least squares · Minimax design · Optimal design · Spatial correlation

1 Introduction

Suppose there are t treatments, such as crop varieties, crop densities, or types of soil fertilizers, randomly labeled as $1, 2, \ldots, t$, and we want to compare the treatment means through a field experiment with N plots. The N plots are usually rectangular and arranged in m rows and n columns. If each treatment is replicated r times, then $rt = mn = N$. The design problem is to seek an optimal assignment of the t treatments into the N plots such that we can get the most information about the treatment means

The work was partially supported by research grants from the Natural Science and Engineering Research Council of Canada.

B. Ou · J. Zhou (✉)
Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4
e-mail: jzhou@uvic.ca

after the experiment. If there is a control in the experiment, without loss of generality, we label it as treatment 1.

For plot (i, j) at the ith row and jth column, denote y_{ij} as the response variable, and a linear model without any block effects is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{1}$$

where vector $\mathbf{y} = (y_{11}, \ldots, y_{mn})'$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_t)'$ is the treatment mean vector, the design matrix $\mathbf{X}$ $(N \times t)$ has elements 0 and 1 and each row has exactly one unity, and the error vector $\boldsymbol{\epsilon} = (\epsilon_{11}, \ldots, \epsilon_{mn})'$ has mean $\mathbf{0}$ and covariance matrix $\mathbf{R}$. One important feature about field experiments is the spatial correlation among field plots which cannot be ignored (Cressie 1993). Optimal designs have been studied extensively in the literature for various spatial correlation structures and are widely applied in agricultural experiments, environmental studies, forestry investigation, etc. Examples can be found in Fisher (1966), Cochran and Cox (1960), Kiefer and Wynn (1981), Martin (1982, 1986, 1996), Becher (1988), Cressie (1993), and Elliott et al. (1999).

In practice, however, the spatial correlation structures among field plots are never known. With the uncertainty of the correlation structures, we propose to study robust designs which do not depend on the exact correlation structures and are efficient in a neighborhood of correlation structures. In particular, we study two estimators, the ordinary least squares estimator (LSE) and the generalized least squares estimator (GLSE), and define minimax designs (Sect. 3) for the LSE and the GLSE. Through this study we try to answer the following questions:

(1) Are robust designs and optimal designs different?
(2) Which estimator is more efficient?
(3) When do we apply robust designs in practice?

Robust designs have been studied for serially correlated observations, for example, Wiens and Zhou (1996, 1999) and Zhou (2001), where the serial correlation is viewed as one-dimensional. In this paper we deal with two-dimensional correlated error processes (planar processes). Martin et al. (1993), Kiefer (1961) and Kiefer and Wynn (1984) also studied robust designs with (possible) applications in field experiments, but their correlations can be reduced to one-dimensional correlations. Recently Wiens and Zhou (2008) investigated robust estimators and designs for two-dimensional correlations, where an A-optimal design criterion is considered and a robust estimator is derived. Here we will consider another estimator and explore various robust design criteria including A-optimality and D-optimality. General results as well as new designs are obtained for robust designs. Also several neighborhoods of correlation structures are discussed in detail which give us some insight into planar processes.

The paper is organized as follows. In Sect. 2 we discuss several neighborhoods of covariance matrices and their relationships. Examples of neighborhoods for various planar processes are presented. In Sect. 3, estimation methods are addressed when the covariance matrix is not known exactly, and minimax design criteria are proposed and investigated. In Sect. 4 minimax designs are constructed and compared for a number of cases. Concluding remarks are in Sect. 5. Proofs are given in the Appendix.

2 Neighborhoods of covariance matrices

Some particular spatial correlation functions, which only depend on the horizontal and vertical distances between two plots, are used to illustrate the general methodology. Denote $cov(\epsilon_{i_1,j_1}, \epsilon_{i_2,j_2}) = \sigma^2\rho(l,s)$, where $l = |i_1 - i_2|$, $s = |j_1 - j_2|$, and $\rho(0,0) = 1$.

For the first order nearest neighbor correlation structure NN1(ρ), we have $\rho(1,0) = \rho(0,1) = \rho$, $\rho(l,s) = 0$ for $l + s > 1$, where $\rho \in [0, 0.25]$. For a special two-dimensional moving average correlation structure MA1(γ) (Martin 1986), $\epsilon_{i,j} = \gamma u_{i-1,j} + \gamma u_{i+1,j} + \gamma u_{i,j-1} + \gamma u_{i,j+1} + u_{i,j}$, where $\gamma \in [0, 0.25]$ and $u_{i,j}$ are i.i.d. with mean 0 and variance σ_u^2. The correlation function $\rho(1,0) = \rho(0,1) = \rho$, $\rho(1,1) = \gamma\rho$, $\rho(2,0) = \rho(0,2) = \frac{1}{2}\gamma\rho$, $\rho(l,s) = 0$ for $l + s > 2$, with $\rho = 2\gamma/(1 + 4\gamma^2)$.

Another commonly used correlation structure is the doubly geometric DG(λ) (Martin 1979), and the correlation function is $\rho(l,s) = \lambda^{l+s}$, where $0 < \lambda < 1$. A similar structure is the discrete exponential correlation DE(λ), and the correlation function (Becher 1988) is $\rho(l,s) = \lambda^{\sqrt{l^2+s^2}}$.

The GLSE is usually used to estimate $\boldsymbol{\theta}$, where $\hat{\boldsymbol{\theta}}_G = \left(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}$, and $cov(\hat{\boldsymbol{\theta}}_G) = \left(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right)^{-1}$. On the other hand if $\mathbf{R}$ is unknown, the LSE can be used, $\hat{\boldsymbol{\theta}}_{LS} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$, and $cov(\hat{\boldsymbol{\theta}}_{LS}) = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{R}\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}$. In practice, researchers often have some idea about the spatial correlation function of the error process for a particular field, but not be able to specify $\mathbf{R}$ exactly. Therefore we need flexibility in specifying $\mathbf{R}$ and use the available information about $\mathbf{R}$ in the design and estimation of field experiments. This motivates the definition of two neighborhoods of covariance matrices below.

For a real symmetric $N \times N$ matrix $\mathbf{A} = (a_{ij})$, the matrix norms $||\mathbf{A}||_p$, $p = 1, 2$, are defined by

$$||\mathbf{A}||_1 = \max_{1\le j\le N}\sum_{i=1}^{N}|a_{ij}|, \quad ||\mathbf{A}||_2 = \max\{|v| : \ v \text{ is an eigenvalue of } \mathbf{A}\}.$$

For any two symmetric matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, the Loewner ordering (Pukelsheim 1993, p. 12) $\mathbf{A}_1 \ge \mathbf{A}_2$ means that $\mathbf{A}_1 - \mathbf{A}_2$ is non-negative definite. Two neighborhoods of a given covariance matrix $\mathbf{R}_0 > \mathbf{0}$ are defined as follows, for $\alpha \ge 0$,

(1) $\mathcal{R}_{1,p,\alpha} = \{\mathbf{R} : \ ||\mathbf{R} - \mathbf{R}_0||_p \le \alpha, \mathbf{R}' = \mathbf{R} \ge \mathbf{0}\}, \quad p = 1, 2,$
(2) $\mathcal{R}_{2,\mathbf{K},\alpha} = \{\mathbf{R} : \ \mathbf{0} \le \mathbf{R} \le \mathbf{R}_0 + \alpha\mathbf{K}, \mathbf{K} \ge \mathbf{0}, \mathbf{R}' = \mathbf{R}\}.$

It is obvious that neighborhood $\mathcal{R}_{1,p,\alpha}$ includes $\mathbf{R}$ with small deviation from $\mathbf{R}_0$. Parameter α controls the size of the neighborhood, and $\mathbf{R}_0$ can be viewed as the center. If $\alpha = 0$, the neighborhood reduces to the single matrix $\mathbf{R}_0$. Neighborhood $\mathcal{R}_{2,\mathbf{K},\alpha}$ is introduced in Wiens and Zhou (2008). In particular, we consider $\mathbf{K} = \mathbf{I}$ or $\mathbf{K} = \mathbf{R}_0$ to study robust designs here. The following result is from Wiens and Zhou (2008, Remark 3 of Sect. 2).

Lemma 1 *For any* $\mathbf{R} \in \mathcal{R}_{1,p,\alpha}$, $p = 1, 2$, $\mathbf{0} \le \mathbf{R} \le \mathbf{R}_0 + \alpha\mathbf{I}_N$. *Also we have* $\mathbf{R}_0 + \alpha\mathbf{I}_N \in \mathcal{R}_{1,p,\alpha}$.

From Lemma 1, we have $\mathcal{R}_{1,p,\alpha} \subset \mathcal{R}_{2,\mathbf{I},\alpha}$, which means that neighborhood $\mathcal{R}_{2,\mathbf{I},\alpha}$ is bigger than $\mathcal{R}_{1,p,\alpha}$. All the neighborhoods contain the covariance matrices from error processes which are close to the one specifying $\mathbf{R}_0$. For example, consider a 6×6 plots and $\mathbf{R}_0$ is the covariance matrix from DG(0.5) with $\sigma^2 = 1$. Numerical computation verifies that $\mathcal{R}_{2,\mathbf{I},0.25}$ contains all the covariance matrices from DG(λ) with $\sigma^2 = 1$ and $\lambda \in [0.36, 0.51]$, DE(λ) with $\sigma^2 = 1$ and $\lambda \in [0.41, 0.44]$, and other error processes as well.

The neighborhoods allow variations in the correlation parameters as well as in the variance σ^2. Suppose that $\mathbf{R}_0$ is the covariance matrix from NN1(ρ_0) with variance σ_0^2. For $m,\ n > 2$, neighborhoods $\mathcal{R}_{1,1,\alpha}$ and $\mathcal{R}_{2,\mathbf{I},\alpha}$ include all the covariance matrices from NN1(ρ) processes with ρ and σ^2 satisfying $\rho \in [0, 0.25]$, $\sigma^2 > 0$ and $|\sigma^2 - \sigma_0^2| + 4|\sigma^2\rho - \sigma_0^2\rho_0| \le \alpha$. Similarly, if $\mathbf{R}_0$ is the covariance matrix from MA1(γ_0) with variance σ_0^2 and $m,\ n > 4$, neighborhoods $\mathcal{R}_{1,1,\alpha}$ and $\mathcal{R}_{2,\mathbf{I},\alpha}$ include all the covariance matrices from MA1(γ) processes with $\sigma^2 > 0$ and $\gamma \in [0, 0.25]$ satisfying $|\sigma^2 - \sigma_0^2| + 8 \mid \frac{\sigma^2\gamma}{1+4\gamma^2} - \frac{\sigma_0^2\gamma_0}{1+4\gamma_0^2} \mid + 12 \mid \frac{\sigma^2\gamma^2}{1+4\gamma^2} - \frac{\sigma_0^2\gamma_0^2}{1+4\gamma_0^2} \mid \le \alpha$.

3 Minimax designs

Instead of specifying $\mathbf{R}$ exactly in model (1), we can specify $\mathbf{R}$ in a neighborhood of $\mathbf{R}_0$, such as $\mathcal{R}_{1,p,\alpha}$ or $\mathcal{R}_{2,\mathbf{K},\alpha}$, and study robust designs. These designs are robust against misspecification of the covariance matrix. Without knowing $\mathbf{R}$ exactly, we cannot compute $\hat{\boldsymbol{\theta}}_G$. Since $\mathbf{R}$ belongs to a neighborhood centered at $\mathbf{R}_0$, it is reasonable to modify the GLSE as $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0} = \left(\mathbf{X}'{\mathbf{R}_0}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'{\mathbf{R}_0}^{-1}\mathbf{y}$, and $cov(\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}) = \left(\mathbf{X}'{\mathbf{R}_0}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'{\mathbf{R}_0}^{-1}\mathbf{R}\mathbf{R}_0^{-1}\mathbf{X}\left(\mathbf{X}'{\mathbf{R}_0}^{-1}\mathbf{X}\right)^{-1}$. Estimator $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is also proposed in Martin (1986) to study optimal designs.

The matrix means ϕ_q ($q \in [-\infty, 1]$) (Pukelsheim 1993, Chap. 6) are often used as scalar functions mapping non-negative definite matrices into the real line and have the following properties:

(P1) They are *isotonic*, i.e., $\phi_q(\mathbf{B}_1) \ge \phi_q(\mathbf{B}_2)$ if $\mathbf{B}_1 \ge \mathbf{B}_2$.
(P2) They are *positive homogeneous*, i.e., $\phi_q(\delta\mathbf{B}) = \delta\phi_q(\mathbf{B})$ for all $\delta > 0$, $\mathbf{B} \ge \mathbf{0}$.

Some commonly used matrix means defined for an $s \times s$ matrix $\mathbf{B}$ are

$$\phi_q(\mathbf{B}) = \begin{cases} \frac{1}{s}\text{trace}(\mathbf{B}), & \text{for } q = 1; \\ (\det(\mathbf{B}))^{1/s}, & \text{for } q = 0. \end{cases}$$

A minimax approach is used to define robust designs here. Let $\hat{\boldsymbol{\theta}}$ be either $\hat{\boldsymbol{\theta}}_{LS}$ or $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$.

A minimax design for $\mathbf{C}\hat{\boldsymbol{\theta}}$ over a neighborhood $\mathcal{R}$ is a solution to the following minimax problem,

$$\min_{\mathbf{X}} \max_{\mathbf{R}\in\mathcal{R}} \phi_q(cov(\mathbf{C}\hat{\boldsymbol{\theta}})),$$

where matrix $\mathbf{C}$ ($s \times t$) is constant. For comparative experiments $\mathbf{C}$ is a contrast matrix.

When $q = 0$ and 1, the designs are called D-optimal and A-optimal minimax designs respectively. Denote loss functions

$$l_{\mathrm{OLS},q,\mathcal{R}}(\mathbf{X}) = \max_{\mathbf{R}\in\mathcal{R}} \phi_q(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{LS})), \quad l_{\mathrm{GLS},q,\mathcal{R}}(\mathbf{X}) = \max_{\mathbf{R}\in\mathcal{R}} \phi_q(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0})).$$

For a design $\mathbf{X}$, define $\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ to be a more efficient estimator over $\mathcal{R}$ than $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ if $l_{\mathrm{GLS},q,\mathcal{R}}(\mathbf{X}) \le l_{\mathrm{OLS},q,\mathcal{R}}(\mathbf{X})$. For simplicity, we use $\mathcal{R}_1, \mathcal{R}(\mathbf{I})$ and $\mathcal{R}(\mathbf{R}_0)$ to denote neighborhoods $\mathcal{R}_{1,p,\alpha}$, $\mathcal{R}_{2,\mathbf{I},\alpha}$ and $\mathcal{R}_{2,\mathbf{R}_0,\alpha}$ respectively. The following results for the loss functions are useful to study and construct minimax designs.

Theorem 1 *The loss functions over $\mathcal{R}_1$, $\mathcal{R}(\mathbf{I})$ and $\mathcal{R}(\mathbf{R}_0)$ are*

$$l_{\mathrm{OLS},q,\mathcal{R}_1}(\mathbf{X}) = l_{\mathrm{OLS},q,\mathcal{R}(\mathbf{I})}(\mathbf{X}) = \phi_q(\mathbf{D}_1 + \alpha\mathbf{D}_2), \tag{2}$$

$$l_{\mathrm{OLS},q,\mathcal{R}(\mathbf{R}_0)}(\mathbf{X}) = \phi_q(\mathbf{D}_1 + \alpha\mathbf{D}_1) = (1+\alpha)\phi_q(\mathbf{D}_1), \tag{3}$$

$$l_{\mathrm{GLS},q,\mathcal{R}_1}(\mathbf{X}) = l_{\mathrm{GLS},q,\mathcal{R}(\mathbf{I})}(\mathbf{X}) = \phi_q(\mathbf{D}_3 + \alpha\mathbf{D}_4), \tag{4}$$

$$l_{\mathrm{GLS},q,\mathcal{R}(\mathbf{R}_0)}(\mathbf{X}) = \phi_q(\mathbf{D}_3 + \alpha\mathbf{D}_3) = (1+\alpha)\phi_q(\mathbf{D}_3), \tag{5}$$

where

$$\mathbf{D}_1 = \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}', \quad \mathbf{D}_2 = \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}',$$
$$\mathbf{D}_3 = \mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}', \quad \mathbf{D}_4 = \mathbf{C}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-2}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{C}'.$$

Remarks (1) In all cases, the maximum loss over a neighborhood is reached at the largest member (in the sense of Loewner ordering) of the neighborhood. Theorem 1 implies that the minimax designs over $\mathcal{R}_1$ and over $\mathcal{R}(\mathbf{I})$ are the same.

(2) The minimax designs over $\mathcal{R}(\mathbf{R}_0)$ are the same as the optimal designs if $\mathbf{R}_0$ is assumed to be the exact covariance matrix for the error process, since then $cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}) = \mathbf{D}_1$ and $cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}) = cov(\mathbf{C}\hat{\boldsymbol{\theta}}_G) = \mathbf{D}_3$. Detailed discussion for finding optimal designs for various $\mathbf{R}_0$ for $\hat{\boldsymbol{\theta}}_{LS}$ and $\hat{\boldsymbol{\theta}}_G$ can be found in Martin (1986). Furthermore, the minimax designs over $\mathcal{R}(\mathbf{I})$ with $\alpha = 0$ are optimal designs. Thus, when we construct the minimax designs over $\mathcal{R}(\mathbf{I})$ in Sect. 4 for various values of α, they represent the minimax designs over $\mathcal{R}_1$ and over $\mathcal{R}(\mathbf{I})$ if $\alpha > 0$ or the optimal designs and minimax designs over $\mathcal{R}(\mathbf{R}_0)$ if $\alpha = 0$.

(3) From the Gauss-Markov theorem (Pukelsheim 1993, p. 20), $\mathbf{D}_1 \ge \mathbf{D}_3$ and $\mathbf{D}_4 \ge \mathbf{D}_2$. Thus $l_{\mathrm{GLS},q,\mathcal{R}(\mathbf{R}_0)}(\mathbf{X}) \le l_{\mathrm{OLS},q,\mathcal{R}(\mathbf{R}_0)}(\mathbf{X})$ for all $\mathbf{X}$, and estimator $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is often more efficient over $\mathcal{R}(\mathbf{R}_0)$ than $\hat{\boldsymbol{\theta}}_{LS}$. However it is not clear if $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is more efficient over $\mathcal{R}(\mathbf{I})$ than $\hat{\boldsymbol{\theta}}_{LS}$ in general. Some specific results are given below.

Theorem 2 *For a design $\mathbf{X}$ on a field and a covariance matrix $\mathbf{R}_0$, there exists an $\alpha_0 > 0$ such that $\phi_q(\mathbf{D}_3 + \alpha\mathbf{D}_4) < \phi_q(\mathbf{D}_1 + \alpha\mathbf{D}_2)$, for all $0 \le \alpha \le \alpha_0$. Also there exists an $\alpha_1 > 0$ such that $\phi_q(\mathbf{D}_1 + \alpha\mathbf{D}_2) < \phi_q(\mathbf{D}_3 + \alpha\mathbf{D}_4)$, for all $\alpha \ge \alpha_1$.*

Roughly speaking $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is more efficient over $\mathcal{R}(\mathbf{I})$ than $\hat{\boldsymbol{\theta}}_{LS}$ if the neighborhood size is small. Otherwise $\hat{\boldsymbol{\theta}}_{LS}$ can be more efficient than $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$.

4 Examples of minimax designs

Examples of A-optimal and D-optimal minimax designs for several cases are constructed and compared. Some results are analytical, and some are obtained through a simulated annealing algorithm. Simulated annealing algorithms are efficient to find optimal (robust) designs, see Haines (1987), Elliott et al. (1999), Fang and Wiens (2000) and Zhou (2001). The algorithm in Elliott et al. (1999) is applied in our computation, and new designs are generated by exchanging the treatments at two randomly selected plots. Note that minimax designs are not unique.

An A-optimal minimax design for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ over $\mathcal{R}(\mathbf{I})$ with a contrast matrix $\mathbf{C} = \mathbf{I}_t - t^{-1}\mathbf{1}_t \cdot \mathbf{1}_t'$ minimizes, for a field with $N = rt$ and $\mathbf{X}'\mathbf{X} = r\mathbf{I}_t$,

$$l_{\mathrm{OLS},1,\mathcal{R}(\mathbf{I})}(\mathbf{X}) = \frac{1}{r^2 t}\mathrm{trace}(\mathbf{X}'\mathbf{R}_0\mathbf{X}) - \frac{1}{r^2 t^2}\mathbf{1}_N'\mathbf{R}_0\mathbf{1}_N + \frac{\alpha(t-1)}{rt},$$

where $\mathbf{1}_N$ $(N \times 1)$ is a vector of ones. Therefore A-optimal minimax designs for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ do not depend on α and only minimize the trace$(\mathbf{X}'\mathbf{R}_0\mathbf{X})$.

Theorem 3 *For any covariance matrix* $\mathbf{R}_0$ *with variance* σ^2 *and nonnegative correlation function, trace*$(\mathbf{X}'\mathbf{R}_0\mathbf{X}) \geq rt\sigma^2$.

Remarks (1) If $\mathbf{R}_0$ is the covariance matrix from NN1(ρ_0) with σ^2, $\rho_0 > 0$ and $t \geq 2$, $\min_{\mathbf{X}} \mathrm{trace}(\mathbf{X}'\mathbf{R}_0\mathbf{X}) = rt\sigma^2$, and the A-optimal minimax design for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ assigns different treatments for any two nearest neighbors. As a special case with $m = n = t = r$, any Latin square design is an A-optimal minimax design.

(2) Suppose $\mathbf{R}_0$ is the covariance matrix from MA1(γ_0) with $\gamma_0 > 0$, then $\min_{\mathbf{X}}$ trace $(\mathbf{X}'\mathbf{R}_0\mathbf{X}) = rt\sigma^2$, if there exists a design assigning different treatments for any two moving average neighbors. A Latin square design is called a Knight's move Latin square design if all the plots with the same treatment can be traversed by a series of Knight's moves as on a chessboard. As a special case with $m = n = t = r$, any Knight's move Latin square design is an A-optimal minimax design for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$.

(3) Similar results are obtained for A-optimal designs for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ in Kiefer and Wynn (1981) and Martin (1986).

An A-optimal minimax design for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ over $\mathcal{R}(\mathbf{I})$ minimizes $l_{\mathrm{GLS},1,\mathcal{R}(\mathbf{I})}(\mathbf{X}) = \frac{1}{t}$ trace $\left((\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1} + \alpha(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}_0^{-2}\mathbf{X}(\mathbf{X}'\mathbf{R}_0^{-1}\mathbf{X})^{-1}\right)$, and usually depends on α.

Example 4.1 Consider a field with 5×5 plots testing $t = 5$ treatments, and $\mathbf{R}_0$ is the covariance matrix from MA1(0.2) with $\sigma^2 = 1$. A Knight's move Latin square design in Fig. 1a is an A-optimal minimax design for $\hat{\boldsymbol{\theta}}_{LS}$ for all $\alpha \geq 0$, and is denoted by **d**=[1 2 3 4 5; 3 4 5 1 2; 5 1 2 3 4; 2 3 4 5 1; 4 5 1 2 3]. A-optimal minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ are in Fig. 1b for $\alpha = 0$ and 0.1, and in Fig. 1c for $\alpha = 0.3$ and 0.5. The design for $\alpha = 0.7$ is $\mathbf{d}_1$= [5 1 5 1 4; 1 3 1 4 5; 2 5 2 1 3; 5 2 3 2 4; 3 4 2 4 3].

1	2	3	4	5
3	4	5	1	2
5	1	2	3	4
2	3	4	5	1
4	5	1	2	3

(a)

1	2	3	4	3
2	1	2	3	4
5	2	3	5	3
1	5	1	4	5
5	1	4	2	4

(b)

4	1	3	1	3
1	5	1	3	5
2	1	2	4	3
3	2	4	5	4
2	5	2	4	5

(c)

Fig. 1 **a** Minimax design for $\hat{\boldsymbol{\theta}}_{LS}$, **b** minimax design for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ with $\alpha = 0$ and 0.1, **c** minimax design for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ with $\alpha = 0.3$ and 0.5

The minimum loss functions at the minimax designs for $\hat{\boldsymbol{\theta}}_{LS}$ ($\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$) are 0.20 (0.1454), 0.22 (0.1748), 0.26 (0.2316), 0.30 (0.2878), and 0.34 (0.3417) for $\alpha = 0$, 0.1, 0.3, 0.5, and 0.7 respectively. It indicates (see also Example 4.3) that $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is more efficient than $\hat{\boldsymbol{\theta}}_{LS}$ for small α, and $\hat{\boldsymbol{\theta}}_{LS}$ is more efficient than $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ for large α. The crossover point is approximately at $\alpha = 0.7$. A-optimal minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ and $\hat{\boldsymbol{\theta}}_{LS}$ are different for MA1 processes. The design in Fig. 1b is an optimal design in Martin and Eccleston (2001).

Example 4.2 Consider the same setting as in Example 4.1, but $\mathbf{R}_0$ is the covariance matrix from DE(0.5) with $\sigma^2 = 1$. It turns out that the Knight's move Latin square design in Fig. 1a is an A-optimal minimax design for $\hat{\boldsymbol{\theta}}_{LS}$ for $\alpha = 0$, 0.1 and 0.6 and for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ for $\alpha = 0.1$ and 0.6. An A-optimal minimax design for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ for $\alpha = 0$ is given by $\mathbf{d}_2$=[1 4 3 1 5; 3 2 5 4 2; 5 4 3 1 3; 2 1 2 5 4; 4 5 3 1 2], which is not a Latin square design. The minimum loss functions at the minimax designs for $\hat{\boldsymbol{\theta}}_{LS}$ ($\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$) are 0.3164 (0.2818), 0.3364 (0.3050) and 0.4364 (0.4210) for $\alpha = 0$, 0.1 and 0.6 respectively.

D-optimal minimax designs for $\hat{\boldsymbol{\theta}}_{LS}$ and $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ over $\mathcal{R}(\mathbf{I})$ minimize $l_{\mathrm{OLS},0,\mathcal{R}(\mathbf{I})}(\mathbf{X})$ and $l_{\mathrm{GLS},0,\mathcal{R}(\mathbf{I})}(\mathbf{X})$ respectively, and they usually depend on α.

Example 4.3 Consider the same setting as in Example 4.1. A Latin square design $\mathbf{d}_3$=[1 2 3 4 5; 2 3 4 5 1; 3 4 5 1 2; 4 5 1 2 3; 5 1 2 3 4] is a D-optimal minimax design for $\hat{\boldsymbol{\theta}}_{LS}$ for $\alpha = 0$, 0.1, 0.3 and 0.5, which is different from the A-optimal minimax design for $\hat{\boldsymbol{\theta}}_{LS}$ in Fig. 1a. A D-optimal minimax design for $\hat{\boldsymbol{\theta}}_{LS}$ for $\alpha = 0.7$ is in Fig. 1a. A D-optimal minimax design for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ for $\alpha = 0$, 0.1 and 0.3 is in Fig. 1c, and designs $\mathbf{d}_4$=[3 1 5 1 5; 1 3 1 5 4; 2 1 2 4 5; 3 2 4 3 4; 2 5 2 4 3] and $\mathbf{d}_5$=[5 1 4 1 3; 4 5 1 3 4; 3 1 2 1 3; 5 2 4 2 5; 3 5 2 4 2] are optimal for $\alpha = 0.5$ and 0.7 respectively. The D-optimal minimax designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ are not Latin square designs and are different from that for $\hat{\boldsymbol{\theta}}_{LS}$. The minimum loss functions for $\hat{\boldsymbol{\theta}}_{LS}$ ($\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$) are 0.1506 (0.0937), 0.1792 (0.1318), 0.2303 (0.1994), 0.2775 (0.2623) and 0.3198 (0.3213) for α =0, 0.1, 0.3, 0.5 and 0.7 respectively.

If we are interested in comparing the control mean, θ_1, with the others, the numbers of replicates received by the t treatments can be different. In particular, the number

Table 1 Minimum loss functions for A-optimal minimax designs in Example 4.4

Design	$l_{\mathrm{OLS},1,\mathcal{R}(\mathbf{I})}$		$l_{\mathrm{GLS},1,\mathcal{R}(\mathbf{I})}$	
$(r_1, r_2, \ldots, r_8)$	NN1(0.15)	DG(0.15)	NN1(0.15)	DG(0.15)
$(3, 3, \ldots, 3)$	0.7429	0.7368	0.6974	0.7163
$(10, 2, \ldots, 2)$	0.6493	0.6542	0.6212	0.6362
$(17, 1, \ldots, 1)$	1.2236	1.2207	1.1667	1.1725

of replicates for the control, r_1, may be different from the others, $r_2, \ldots, r_t$. If the other $t-1$ treatments are randomly assigned to labels $2, \ldots, t$, it is reasonable to have $r_2 = \cdots = r_t$. The following example shows some minimax designs using a contrast matrix $\mathbf{C} = (\mathbf{1}_{t-1} : -\mathbf{I}_{t-1})$.

Example 4.4 Consider a field with 3×8 plots comparing $t = 8$ treatments. A-optimal minimax designs are obtained for $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$ and $\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$. Designs with various distributions of replicates in Table 1 are compared, and minimax designs with replicates $(r_1, r_2, \ldots, r_8) = (10, 2, \ldots, 2)$ have the smallest loss functions for various error processes. Some results are presented in Table 1 with $\mathbf{R}_0$ from NN1(0.15) and DG(0.15) processes with $\sigma^2 = 1$ and $\alpha = 0.2$.

A-optimal minimax designs are presented in Fig. 2: (a) for NN1(0.15) and $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$, (b) for NN1(0.15) and $\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$, (c) for DG(0.15) and $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$, and (d) for DG(0.15) and $\mathbf{C}\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$. The layout for the control (treatment 1) in the field plots are all different in Fig. 2. For NN1(0.15) and $\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}$, as pointed out by a referee, the A-optimal minimax design minimizes $\{(t-1)(r_1+r_2)(1+\alpha) + 2[n_{11}(t-1)r_2/r_1 + \sum_{i=2}^{t} n_{ii}r_1/r_2 - \sum_{i=2}^{t} n_{1i}]\rho\}/\{r_1 r_2(t-1)\}$, where n_{ij} denotes the number of adjacencies from treatment j to treatment i. For $(r_1, r_2, \ldots, r_8) = (10, 2, \ldots, 2)$, the minimax design has $n_{11} = \sum_{i=2}^{t} n_{ii} = 0$ and $\sum_{i=2}^{t} n_{1i} = 33$ as in Fig. 2a.

5 Conclusion

By introducing the neighborhoods of covariance matrices, we can study robust designs for field experiments. Several neighborhoods are discussed, and minimax robust design criteria are proposed using the matrix means as scalar functions. This paper provides a general methodology to construct minimax designs which are robust against misspecification of two-dimensional spatial error processes, and gives several explicit robust designs.

In practice, if the covariance matrix $\mathbf{R}$ of the error process belongs to a small neighborhood of $\mathbf{R}_0$, i.e., we have fairly accurate information about the correlation of the error process, the GLSE $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ is more efficient than the LSE $\hat{\boldsymbol{\theta}}_{LS}$.

Some optimal designs assuming $\mathbf{R}_0$ to be the exact covariance matrix of the error process are also minimax designs. In particular, A-optimal designs for $\hat{\boldsymbol{\theta}}_{LS}$ are A-optimal minimax designs for $\hat{\boldsymbol{\theta}}_{LS}$ for any neighborhood size. The minimax de-

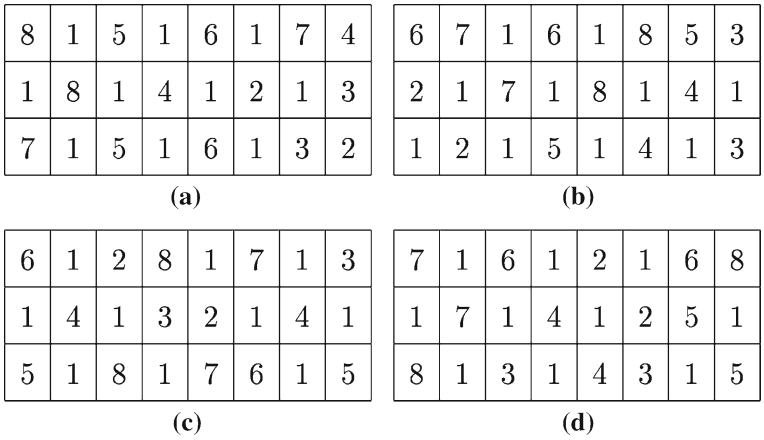

Fig. 2 A-optimal designs for Example 4.4

signs for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$ may be different from the optimal designs for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$, see Examples 4.1 and 4.2. Minimax designs can be constructed using a simulated annealing algorithm. Analytical results about minimax designs need further investigation, especially for $\hat{\boldsymbol{\theta}}_{\mathbf{R}_0}$. For practical applications, we recommend minimax designs. If we have very accurate information about $\mathbf{R}$, optimal designs can still be applied. Numerical computation shows that in many cases the minimax designs with small α are the same or very close to the optimal designs.

Acknowledgements The authors would like to thank two referees and the Editor for their helpful comments and suggestions which greatly improved the presentation of the paper. They also thank Professor Mike Wilmut for reading a previous version of the article and helpful suggestions.

Appendix: Proofs

Proof of Theorem 1: We will prove the result for $l_{\mathrm{OLS},q,\mathcal{R}_1}(\mathbf{X})$, and the other results can be proved similarly. For $\mathcal{R}_1$, all $\mathbf{R} \le \mathbf{R}_0 + \alpha\mathbf{I}_N$ and $\mathbf{R}_0 + \alpha\mathbf{I}_N \in \mathcal{R}_1$ (Lemma 1), then $cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{LS}) = \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' \le \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{R}_0 + \alpha\mathbf{I}_N)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}' = \mathbf{D}_1 + \alpha\mathbf{D}_2$. Since ϕ_q is isotonic, we have $l_{\mathrm{OLS},q,\mathcal{R}_1}(\mathbf{X}) = \max_{\mathbf{R}\in\mathcal{R}_1} \phi_q(cov(\mathbf{C}\hat{\boldsymbol{\theta}}_{LS})) = \phi_q(\mathbf{D}_1 + \alpha\mathbf{D}_2)$, which is part of (2). □

Proof of Theorem 2: Function $\phi_q(\mathbf{D}_1+\alpha\mathbf{D}_2)$ is continuous in α, since the eigenvalues of $\mathbf{D}_1 + \alpha\mathbf{D}_2$ are continuous in α (Horn and Johnson 1985, p. 540). The function $\phi_q(\mathbf{D}_3+\alpha\mathbf{D}_4)$ is also continuous in α. Consider $f(\alpha) = \phi_q(\mathbf{D}_3+\alpha\mathbf{D}_4)-\phi_q(\mathbf{D}_1+\alpha\mathbf{D}_2)$. Since $f(0) < 0$ and $f(\alpha)$ is continuous, there exists an $\alpha_0 > 0$ such that $f(\alpha) < 0$ for all $0 \le \alpha \le \alpha_0$, which is the first half of the result. The second half can be proved similarly by considering the function $\phi_q(\mathbf{D}_4 + \frac{1}{\alpha}\mathbf{D}_3) - \phi_q(\mathbf{D}_2 + \frac{1}{\alpha}\mathbf{D}_1)$. □

Proof of Theorem 3: Note that $\mathrm{trace}(\mathbf{X}'\mathbf{R}_0\mathbf{X}) = \mathrm{trace}(\mathbf{X}'(\sigma^2\mathbf{I}_t + \mathbf{R}_0 - \sigma^2\mathbf{I}_t)\mathbf{X}) = \sigma^2 rt + \mathrm{trace}(\mathbf{X}'(\mathbf{R}_0 - \sigma^2\mathbf{I}_t)\mathbf{X}) \ge \sigma^2 rt$, since all elements of $\mathbf{X}$ and $\mathbf{R}_0 - \sigma^2\mathbf{I}_t$ are nonnegative. □

References

Becher H (1988) On optimal experimental design under spatial correlation structures for square and non-square plot designs. Commun Stat Simul Comput 17:771–780
Cressie N (1993) Statistics for spatial data. Wiley, New York
Cochran WG, Cox GM (1960) Experimental designs. Wiley, New York
Elliott LJ, Eccleston JA, Martin RJ (1999) An algorithm for the design of factorial experiments when the data are correlated. Stat Comput 9:195–201
Fang Z, Wiens DP (2000) Integer-valued minimax robust designs for estimation and extrapolation in heteroscedastic, approximately linear models. J Am Stat Assoc 95:807–818
Fisher RA (1966) The design of experiments, 8th edn. Oliver and Boyd, Edinburgh
Haines LM (1987) The application of the annealing algorithm to the construction of the exact optimal designs for linear regression models. Technometrics 29:439–447
Horn RA, Johnson CA (1985) Matrix analysis. Cambridge University Press, Cambridge
Kiefer J (1961) Optimum experimental designs. V. With applications to systematic and rotatable designs. In: Proceeding of 4th Berkeley Symposium. Math. Statist. and Prob., vol I, pp 381–405. University of California Press, Berkeley, California
Kiefer J, Wynn HP (1981) Optimum balanced block and Latin square designs for correlated observations. Ann Stat 9:737–757
Kiefer J, Wynn HP (1984) Optimum and minimax exact treatment designs for one-dimensional autoregressive error processes. Ann Stat 12:414–450
Martin RJ (1979) A subclass of lattice processes applied to a problem in planar sampling. Biometrika 66:209–217
Martin RJ (1982) Some aspects of experimental design and analysis when errors are correlated. Biometrika 69:597–612
Martin RJ (1986) On the design of experiments under spatial correlation. Biometrika 73:247–277 (Correction 75:396, 1988)
Martin RJ (1996) Spatial experimental design. Design and Analysis of Experiments [Handbook of Statistics 13], pp 477–514
Martin RJ, Eccleston JA, Gleeson AC (1993) Robust linear block designs for a suspected LV model. J Stat Plan Inference 34:433–450
Martin RJ, Eccleston JA (2001) Optimal and near-optimal designs for dependent observations. Stat Appl 3:101–116
Pukelsheim F (1993) Optimal design of experiments. Wiley, New York
Wiens DP, Zhou J (1996) Minimax regression designs for approximately linear models with autocorrelated errors. J Stat Plan Inference 55:95–106
Wiens DP, Zhou J (1999) Minimax designs for approximately linear models with AR(1) errors. Can J Stat 27:781–794
Wiens DP, Zhou J (2008) Robust estimators and designs for field experiments. J Stat Plan Inference 138: 93–104
Zhou J (2001) Integer-valued, minimax robust designs for approximately linear models with correlated errors. Commun Stat Theory Methods 30:21–39