KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0214

TREND-RESISTANT AND COST-EFFICIENT CROSS-OVER DESIGNS FOR MIXED MODELS

by
L. TACK
M. VANDEBROEK

D/2002/2376/14

Trend-resistant and cost-efficient cross-over designs for mixed models

Lieven Tack
Martina Vandebroek
Katholieke Universiteit Leuven, Belgium

A mixed model approach is used to construct optimal cross-over designs. In a cross-over experiment the same subject is tested at different points in time. Consider as an example an experiment to investigate the influence of physical attributes of the work environment such as luminance, ambient temperature and relative humidity on human performance of acceptance inspection in quality assurance. In a mixed model context, the subject effects are assumed to be independent and normally distributed. Besides the induction of correlated observations within the same inspector, the mixed model approach also enables one to specify the covariance structure of the inspection data. Here, several covariance structures are considered either depending on the time variable or not. Unfortunately, a serious drawback of the inspection experiment is that the results may be influenced by an unknown time trend because of inspector fatigue due to monotony of the inspection task. In other circumstances, time trend effects can be caused by learning effects of the test subjects in behavioural and life sciences, heating or aging of material in prototype experiments, etc. An algorithm is presented to construct cross-over designs that are optimally balanced for time trend effects. The costs for using the subjects and for altering the factor levels between consecutive observations can also be taken into account. A number of examples illustrate utility of the outlined design methodology.

Keywords: optimum design of experiments; linear mixed effects model; cross-over design; run order; time trend; cost

1 Introduction

In an experiment the same unit is often tested or investigated at different points in time. The resulting experiment is called a cross-over design and the study units are commonly referred to as subjects. Consider as an example an experiment in which interest is in the effects of different advertising campaigns on sales in a number of test markets. Here, the subjects are the test markets. Fifteen markets are randomly selected and in each market four different advertising campaigns are tested spread over one year. Other examples of cross-over studies include prototype experiments in industry, clinical trials in medical sciences, animal feeding experiments and tasting trials. An important advantage of cross-over designs is that they economize on subjects. This is of particular importance when only a few subjects can be utilized for the experiment.

In this paper a mixed model approach is used to design cross-over experiments. A careful description of the use of mixed models in practice is given in the excellent books of Verbeke and Molenberghs (1997) and Brown and Prescott (1999). The simplest way a linear mixed model can be used to analyze cross-over data is to use a random effects model with subject effects fitted as random. Instead of taking fixed values, the subject effects are then realisations from a normal probability distribution. They thus give rise to another source of random variation in addition to the residual variation. This induces observations on the same subject to have a constant correlation while observations on different subjects are uncorrelated. However, the correlation between observations on the same subject is often not constant. For instance, correlation may decrease as the measurements become more widely separated in time. Such specific patterns of covariance are usually defined as a function of the time point or the number of the observation. Consequently, as opposed to the basic assumption that the error terms are independent and identically distributed, in cross-over studies the independence assumption is relaxed and more complicated data structures can be taken into account. The correlated nature of the repeated observations gives more appropriate effect estimates and standard errors because all sources of variability between subjects are excluded from the experimental error. This means that only variation within subjects enters the experimental error.

2 Literature review

For simple regression models without blocking, a number of authors have tackled the problem of designing experiments under the assumption of correlated error terms instead of relying on uncorrelated errors. Bischoff (1993) investigates in what situations an exact $\mathcal{D}$-optimal design with uncorrelated errors and constant variance is also $\mathcal{D}$-optimal when the error terms are correlated. Martin, Jones and Eccleston (1998) study the efficiency of two-level factorial designs under an AR(1) or an MA(1) error process. This approach is extended to multi-level factorials in Martin, Eccleston and Jones (1998). Kiefer and Wynn (1981) study optimal balanced block designs and Latin square designs with autocorrelated errors. Eccleston and Chan (1998) and Donev (1998a) search for optimal row-column

designs with correlated observations. The former authors present design algorithms based on simulated annealing and tabu search, whereas Donev (1998a) proposes an exchange algorithm.

Another part of the relevant literature concerns the design of cross-over trials for treatment comparisons. Hedayat and Afsarinejad (1975, 1978) derive conditions for their optimality and Cheng and Wu (1980) investigate the universal optimality of balanced cross-over trials. For an arbitrary covariance matrix, necessary and sufficient conditions for the universal optimality of cross-over designs for treatment comparisons are given by Kushner (1997a). The case of two treatments is elaborated in Matthews (1987) and Kushner (1997b) and more on the topic can be found in Matthews (1994a, 1994b). Algorithmic approaches are given in Donev and Jones (1995), Jones and Donev (1996) and Donev (1997). The design of cross-over trials for treatment comparisons and correlated error terms receives attention from Kunert (1991) and Donev (1998b).

However, the literature on the construction of cross-over experiments for response surface models is extremely concise. The only important contribution comes from Berger and Tan (1998) who apply optimal design theory to longitudinal studies in which interest is in describing the behaviour of an important response variable through time. They compute $\mathcal{D}$-optimal designs for linear regression models with random subject effects and first-order autoregressive correlations. A numerical method is used to compute the optimal time points at which to measure the response.

An important disadvantage of cross-over studies is the possibility of a systematic effect or trend that distorts the outcome of the experiment. For instance, in the advertising experiment seasonal differences in purchase behaviour will certainly influence the results. In prototype experiments, trend effects may be induced due to aging of the material over time, warm-up in laboratories, instrument drift, etc. In cross-over studies for the behavioural and life sciences, learning effects or fatigue of the test subjects may cause an important order effect. Afsarinejad (2001) accounts for the presence of a time trend and derives some rules to construct trend-free repeated measures designs for treatment comparisons. However, he does not consider the possibility of correlated error terms. In Giovagnoli and Romano (2001) an example is given of a cross-over study in software engineering—an area that has recently established a new perspective for the application of experimental design. An experiment was carried out at the Telecom Research Centre in Turin in which only one programmer was involved. Interest was in a comparison among different ways of writing the computer code for multimedia services on the web. However, a specific problem is that as the experiment progresses, the programmer's performance improves because of a learning process, especially due to greater familiarity with the tools or the language. One may expect that experience gained with a certain type of application will lead to a higher productivity for subsequent projects. Based on complete enumeration, Giovagnoli and Romano (2001) compute a $\mathcal{D}$-optimal design for this non standard design problem. Again, a serious criticism is that the authors did not take into account the possibility for correlated errors within the same programmer. Another criticism concerns

the fact that they did not balance the run order for the unwanted trend effects. A final example of the appearance of trend effects in cross-over studies is an experiment on acceptance inspection in quality assurance. The aim is to investigate the influence of a number of physical attributes of the work environment such as luminance, temperature and relative humidity on human accuracy of the inspection task. Unfortunately, the results of the inspection experiment may be influenced by an unknown time trend because of inspector fatigue due to monotony and repetition associated with the inspection task. This example will be further elaborated in section 6.1.

The literature review shows that addressing the design of trend-resistant cross-over studies for regression models offers a challenging problem still to be solved. This paper will present a linear mixed model approach and will propose a generic construction algorithm to tackle this important design problem.

3 The statistical model

The linear mixed model introduced by Laird and Ware (1982) is an extension of the fixed effects model by the incorporation of random effects. In the context of cross-over studies, the jth response value y_{ij} for subject i may be specified as

$$y_{ij} = \mathbf{f}'(\mathbf{x}_{ij})\boldsymbol{\beta} + \mathbf{g}'(t_{ij})\boldsymbol{\theta} + \gamma_i + \varepsilon_{ij}, \tag{1}$$

where $\mathbf{x}_{ij}$ is the design point for the jth measurement within subject i and $t_{ij} \in [-1, 1]$ is the time point at which the observation is taken. All factors are assumed to have coded levels between -1 and 1. The $(p \times 1)$ vector $\mathbf{f}(\mathbf{x}_{ij})$ represents the polynomial expansion of $\mathbf{x}_{ij}$ for the response model. In this paper the $(q \times 1)$ vector $\mathbf{g}(t_{ij})$ is introduced to represent the polynomial expansion for the time trend. The p fixed effects are represented by $\boldsymbol{\beta}$ and the $(q \times 1)$ vector of parameters of the polynomial time trend is denoted as $\boldsymbol{\theta}$. Finally, γ_i is the random effect for subject i and ε_{ij} denotes the error term. The random subject effects are assumed to have mean zero and constant variance σ^2_γ.

For the k_i observed responses pertaining to subject i model (1) can be rewritten as

$$\mathbf{y}_i = \mathbf{F}_i\boldsymbol{\beta} + \mathbf{G}_i\boldsymbol{\theta} + \mathbf{1}_{k_i}\gamma_i + \boldsymbol{\varepsilon}_i, \tag{2}$$

where $\mathbf{y}_i$ is the k_i dimensional vector of response values y_{ij}, with $j \in \{1, \ldots, k_i\}$, $\mathbf{F}_i$ and $\mathbf{G}_i$ are the $(k_i \times p)$ and the $(k_i \times q)$ extended design matrices for $\mathbf{f}(\mathbf{x})$ and the polynomial time trend $\mathbf{g}(t)$ respectively. The k_i dimensional vector $\mathbf{1}_{k_i}$ has elements one and the vector $\boldsymbol{\varepsilon}_i$ contains the k_i error components ε_{ij} of the ith subject. These error terms are assumed to be normally distributed with mean zero and variance-covariance matrix $\mathbf{V}_{\boldsymbol{\varepsilon}_i}$.

A large variety of covariance patterns are available for use in cross-over studies. In the general unstructured pattern, the variances of the responses differ for each time point t_{ij} and the covariances differ between each pair of time points t_{ij} and t_{ik}, with $j, k \in$

$\{1, \ldots, k_i\}$. Here, emphasis is restricted to relatively simple covariance structures in which the intrasubject variance-covariance matrix $\mathbf{V}_{\boldsymbol{\varepsilon}_i}$ only depends on its dimension k_i.

Most of the covariance patterns are easy to justify when the observations are equally spaced. For example, in the compound symmetry model all covariances are equal. Take as a simple example an experiment with four measurements on the ith subject (i.e. $k_i = 4$). The corresponding variance-covariance matrix then looks like

$$\mathbf{V}_{\boldsymbol{\varepsilon}_i} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}, \tag{3}$$

where σ_ε^2 denotes the experimental error variance and ρ is the coefficient of correlation between different error terms.

An obvious way to model the possibility that periods close together in time might have a higher correlation than periods far apart in time comes from a time series viewpoint. For the autoregressive model the variances are equal and the covariances decrease exponentially depending on their separation $|j - k|$. More specifically, in the first-order autoregressive model—henceforth shortly referred to as AR(1)—the error term ε_{ij} is assumed to depend upon its predecessor $\varepsilon_{i,j-1}$ as

$$\varepsilon_{ij} = \rho \varepsilon_{i,j-1} + \nu_{ij}, \tag{4}$$

where ρ is the autocorrelation parameter and the disturbance term ν_{ij} is independent and identically distributed with mean zero and constant variance σ_ν^2. It is commonly assumed that ε_{i1} has mean zero with the same variance as all other error terms. Besides, to derive the variance-covariance matrix the stationarity condition $|\rho| < 1$ is imposed. This condition ensures that the impact of what happened in the past gradually dies out. For instance, for four observations the variance-covariance matrix for the first-order autoregressive model is

$$\mathbf{V}_{\boldsymbol{\varepsilon}_i} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}. \tag{5}$$

There are also covariance patterns in which the covariances are based on the exact value of time rather than on the number of observation. Such covariance patterns are especially useful in situations where the time points are irregularly spaced. Many ways are available to define covariances from the time interval but most of them are based on an exponential decay of the covariance with the time interval between pairs of observations on the same subject. This paper restricts attention to three frequently used patterns. Firstly, for the

power model the variance-covariance matrix for four observations equals

$$\mathbf{V}_{\boldsymbol{\varepsilon}_i} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \rho^{|t_{i1}-t_{i2}|} & \rho^{|t_{i1}-t_{i3}|} & \rho^{|t_{i1}-t_{i4}|} \\ \rho^{|t_{i2}-t_{i1}|} & 1 & \rho^{|t_{i2}-t_{i3}|} & \rho^{|t_{i2}-t_{i4}|} \\ \rho^{|t_{i3}-t_{i1}|} & \rho^{|t_{i3}-t_{i2}|} & 1 & \rho^{|t_{i3}-t_{i4}|} \\ \rho^{|t_{i4}-t_{i1}|} & \rho^{|t_{i4}-t_{i2}|} & \rho^{|t_{i4}-t_{i3}|} & 1 \end{bmatrix}. \tag{6}$$

Secondly, the variance-covariance matrix for four observations under the exponential model is

$$\mathbf{V}_{\boldsymbol{\varepsilon}_i} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \exp\left(-\frac{|t_{i1}-t_{i2}|}{\rho}\right) & \exp\left(-\frac{|t_{i1}-t_{i3}|}{\rho}\right) & \exp\left(-\frac{|t_{i1}-t_{i4}|}{\rho}\right) \\ \exp\left(-\frac{|t_{i2}-t_{i1}|}{\rho}\right) & 1 & \exp\left(-\frac{|t_{i2}-t_{i3}|}{\rho}\right) & \exp\left(-\frac{|t_{i2}-t_{i4}|}{\rho}\right) \\ \exp\left(-\frac{|t_{i3}-t_{i1}|}{\rho}\right) & \exp\left(-\frac{|t_{i3}-t_{i2}|}{\rho}\right) & 1 & \exp\left(-\frac{|t_{i3}-t_{i4}|}{\rho}\right) \\ \exp\left(-\frac{|t_{i4}-t_{i1}|}{\rho}\right) & \exp\left(-\frac{|t_{i4}-t_{i2}|}{\rho}\right) & \exp\left(-\frac{|t_{i4}-t_{i3}|}{\rho}\right) & 1 \end{bmatrix} \tag{7}$$

and finally, the variance-covariance matrix for the Gaussian model is

$$\mathbf{V}_{\boldsymbol{\varepsilon}_i} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \exp\left(-\frac{|t_{i1}-t_{i2}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i1}-t_{i3}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i1}-t_{i4}|^2}{\rho^2}\right) \\ \exp\left(-\frac{|t_{i2}-t_{i1}|^2}{\rho^2}\right) & 1 & \exp\left(-\frac{|t_{i2}-t_{i3}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i2}-t_{i4}|^2}{\rho^2}\right) \\ \exp\left(-\frac{|t_{i3}-t_{i1}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i3}-t_{i2}|^2}{\rho^2}\right) & 1 & \exp\left(-\frac{|t_{i3}-t_{i4}|^2}{\rho^2}\right) \\ \exp\left(-\frac{|t_{i4}-t_{i1}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i4}-t_{i2}|^2}{\rho^2}\right) & \exp\left(-\frac{|t_{i4}-t_{i3}|^2}{\rho^2}\right) & 1 \end{bmatrix}. \tag{8}$$

If b subjects have to be tested, combining all subject-specific models (2) gives the following linear mixed model for $n = \sum_{i=1}^{b} k_i$ observations:

$$\mathbf{y} = \mathbf{F}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\theta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{9}$$

where $\mathbf{y}$, $\boldsymbol{\varepsilon}$, $\mathbf{F}$ and $\mathbf{G}$ are obtained from stacking the respective vectors $\mathbf{y}_i$ and $\boldsymbol{\varepsilon}_i$ and the extended design matrices $\mathbf{F}_i$ and $\mathbf{G}_i$ underneath each other. The matrix $\mathbf{Z}$ equals $\text{diag}[\mathbf{1}_{k_1}, \ldots, \mathbf{1}_{k_b}]$ and the vector $\boldsymbol{\gamma}$ equals $[\gamma_1, \ldots, \gamma_b]'$ with mean zero and between-subject variance-covariance matrix $\mathbf{V}_{\boldsymbol{\gamma}} = \sigma_\gamma^2 \mathbf{I}_b$. Besides, the random subject effects are assumed to be uncorrelated with the error terms. The variance-covariance matrix $\mathbf{V}_{\boldsymbol{\varepsilon}}$ of the error terms of all subjects now has a block-diagonal structure with the submatrices $\mathbf{V}_{\boldsymbol{\varepsilon}_i}$ on the diagonal. The variance-covariance matrix of the observations $\mathbf{y}$ equals

$$\text{cov}(\mathbf{y}) = \mathbf{Z}\mathbf{V}_{\boldsymbol{\gamma}}\mathbf{Z}' + \mathbf{V}_{\boldsymbol{\varepsilon}} = \sigma_\gamma^2 \, \text{diag}[\mathbf{1}_{k_1}\mathbf{1}_{k_1}', \ldots, \mathbf{1}_{k_b}\mathbf{1}_{k_b}'] + \mathbf{V}_{\boldsymbol{\varepsilon}}. \tag{10}$$

It follows that this matrix is block-diagonal with the size of the blocks equal to the number of observations on each subject. The matrix blocks of zeros off the diagonal represent zero covariances for observations belonging to different subjects.

4 Optimal run orders

Based on the linear mixed effects model (9), in this section a methodology is outlined to construct trend-resistant and cost-efficient run orders for cross-over experiments. Firstly, an optimality criterion is presented with which run orders can be computed that have an optimal balance for time trend effects. Next, some definitions are introduced to allow for cost considerations during the design phase of a cross-over study. Taking into account cost considerations is recommended since trend-resistant run orders can be difficult to perform due to the large number of factor level changes they often involve. Finally, an optimality criterion is given to incorporate both trend-resistance and cost-efficiency.

4.1 Trend-resistant run orders

This section explains how the $\mathcal{D}$-optimality criterion can be used to compute run orders that are optimally balanced for time trend effects. The $\mathcal{D}$-optimality criterion minimizes the generalized variance of the parameter estimates or, alternatively stated, it minimizes the volume of the confidence ellipsoid around the parameter estimates. This criterion is also preferred for two other reasons. Firstly, $\mathcal{D}$-optimal designs usually perform well with regard to other alphabetic optimality criteria, whereas this is often not true for other optimality criteria. An extra reason is the invariability of the $\mathcal{D}$-optimality criterion to a linear transformation of the design matrix or the factor levels. This implies that, as opposed to for instance the $\mathcal{A}$- and the $\mathcal{E}$-optimality criterion, the ordering of designs with respect to the $\mathcal{D}$-optimality criterion does not depend on the coding of the variables.

Based on Searle (1971) it can readily be shown that, in the absence of time trend effects $\boldsymbol{\theta}$, the information matrix on the parameter estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ of model (9) is given by

$$\mathbf{M} = \begin{bmatrix} \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} + \mathbf{V}_{\boldsymbol{\gamma}}^{-1} \end{bmatrix}. \tag{11}$$

The $\mathcal{D}$-optimal run order $\delta_{\mathcal{D}}$ is found by maximizing the information on the parameter estimates of interest, namely $\hat{\boldsymbol{\beta}}$. Mathematically, the corresponding criterion value equals

$$\mathcal{D} = \frac{\begin{vmatrix} \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} + \mathbf{V}_{\boldsymbol{\gamma}}^{-1} \end{vmatrix}}{\left|\mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} + \mathbf{V}_{\boldsymbol{\gamma}}^{-1}\right|}. \tag{12}$$

However, when the outcome of the experiment is influenced by time trend effects $\boldsymbol{\theta}$, the information matrix on the parameter estimates $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\gamma}}$ in model (9) equals

$$\mathbf{M} = \begin{bmatrix} \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} + \mathbf{V}_{\boldsymbol{\gamma}}^{-1} \end{bmatrix}. \tag{13}$$

If interest is restricted to the important parameter estimates $\hat{\boldsymbol{\beta}}$, whereas the q parameters modeling the time dependence are treated as nuisance parameters, the $\mathcal{D}_t$-optimal run

order $\delta_{\mathcal{D}_t}$ that minimizes the generalized variance of the parameter estimates $\hat{\boldsymbol{\beta}}$ is found by maximizing the criterion value

$$\mathcal{D}_t = \frac{\begin{vmatrix} \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{F}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{F} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z}+\mathbf{V}_{\boldsymbol{\gamma}}^{-1} \end{vmatrix}}{\begin{vmatrix} \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{G}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} & \mathbf{Z}'\mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{Z}+\mathbf{V}_{\boldsymbol{\gamma}}^{-1} \end{vmatrix}}. \tag{14}$$

Optimality criterion (14) is in fact nothing else but an extension of the $\mathcal{D}_s$-optimality criterion of Atkinson and Donev (1992) to the linear mixed effects model. Based on Atkinson and Donev (1996), the $\mathcal{D}_t$-optimal run order $\delta_{\mathcal{D}_t}$ and the $\mathcal{D}$-optimal run order $\delta_{\mathcal{D}}$ are compared to each other by means of the so-called trend factor

$$\mathrm{TF}(\delta_{\mathcal{D}_t}) = \left\{ \frac{\mathcal{D}_t(\delta_{\mathcal{D}_t})}{\mathcal{D}(\delta_{\mathcal{D}})} \right\}^{1/p}. \tag{15}$$

This trend factor is a positive number between zero and one and reaches its maximum value 1 if all time trend components are orthogonal to the factorial effects. Alternatively stated, this condition boils down to the requirement

$$\begin{bmatrix} \mathbf{F}' \\ \mathbf{Z}' \end{bmatrix} \mathbf{V}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{G} = \mathbf{0}. \tag{16}$$

The $\mathcal{D}_t$-optimal run order $\delta_{\mathcal{D}_t}$ is then said to be completely trend-free for the postulated time trend $\mathbf{g}(t)$. Note that the power $1/p$ in (15) ensures the unit of variance, independence of the dimension of the model and proportionality to the design size n. For instance, this means that a $\mathcal{D}_t$-optimal run order with trend factor 0.5 has to be replicated twice in order to obtain the same amount of information about the important effects $\boldsymbol{\beta}$ as with the $\mathcal{D}$-optimal run order in the absence of the time trend.

4.2 Cost considerations

In Tack and Vandebroek (2001, 2002) it is shown that trend-robust run orders for the linear fixed effects model can be costly to perform since they involve a large number of factor level changes from one observation to the next one. Therefore, it is worthwhile to take account of cost considerations during the design phase. The importance of allowing for costs also follows from the fact that industrial experimentation counts for a large part of total innovation time and cost. This cost also includes the opportunity cost of time or time-to-market, a variable which has gained importance in today's global economy since it has a major impact on total profit.

Here, the general cost approach of Tack and Vandebroek (2001) is adopted and extended to the optimum design problem for cross-over studies. Two cost concepts will be introduced. Firstly, the measurement cost $c_m(\mathbf{x}_i)$ at design point $\mathbf{x}_i$ is defined as the cost that is

associated with the factor levels constituting design point $\mathbf{x}_i$. Examples include the cost for using a particular subject, the cost of material, the cost for spending time during the measurement, etc. For instance, in a clinical trial for controling blood pressure, applying a new hypertensive drug A may be more expensive than a standard drug B. The measurement costs $c_m(\mathbf{x}_i)$ are based on prior knowledge or can be estimated from previous experimentation. If a rough guess is not satisfactory, the following method can be used to get a more reliable cost estimate. The aim is to have an expert to produce three cost estimates: an optimistic estimate c_{opt}, a realistic estimate c_{real} and a pessimistic estimate c_{pes}. The expected cost then equals $(c_{\text{opt}} + 4c_{\text{real}} + c_{\text{pes}})/6$. The total measurement cost C_m of an experiment simply equals the weighted sum of the measurement costs at the different design points. Mathematically,

$$C_m = \sum_{i=1}^{d} n_i c_m(\mathbf{x}_i), \tag{17}$$

where d denotes the number of distinct design points and n_i is the number of replicates at design point $\mathbf{x}_i$.

Secondly, the transition cost $c_t(\mathbf{x}_i, \mathbf{x}_j)$ between two consecutive measurements at $\mathbf{x}_i$ and $\mathbf{x}_j$ is also allowed for. In cross-over trials the transition cost is often proportional to the time between two consecutive treatments. For instance, the larger the temperature change in the inspection experiment, the longer the waiting time between two consecutive measurements, resulting in a higher transition cost. The transition costs $c_t(\mathbf{x}_i, \mathbf{x}_j)$ are again based on prior knowledge or can be estimated from previous experimentation. The total transition cost C_t of an experiment equals the weighted sum

$$C_t = \sum_{i=1,j=1}^{d} n_{(i,j)} c_t(\mathbf{x}_i, \mathbf{x}_j), \tag{18}$$

where $n_{(i,j)}$ is the number of transitions from design point $\mathbf{x}_i$ to design point $\mathbf{x}_j$ in a particular run order. Note that contrary to the total measurement cost C_m, the total transition cost C_t depends on the specific run sequence. The total cost C of an experiment now equals the sum of the total measurement cost and the total transition cost, or equivalently,

$$C = C_m + C_t. \tag{19}$$

4.3 Trend-resistant and cost-efficient run orders

When both safeguard against destructive temporal trend effects and cost-efficiency are calculated for, the $(\mathcal{D}_t, C)$-optimal run order $\delta_{(\mathcal{D}_t,C)}$ maximizes the amount of information

per unit cost. Mathematically, the $(\mathcal{D}_t, C)$-optimality criterion equals

$$(\mathcal{D}_t, C) = \frac{\mathcal{D}_t^{1/p}}{C} = \frac{\begin{vmatrix} \mathbf{F}'\mathbf{V}_\varepsilon^{-1}\mathbf{F} & \mathbf{F}'\mathbf{V}_\varepsilon^{-1}\mathbf{G} & \mathbf{F}'\mathbf{V}_\varepsilon^{-1}\mathbf{Z} \\ \mathbf{G}'\mathbf{V}_\varepsilon^{-1}\mathbf{F} & \mathbf{G}'\mathbf{V}_\varepsilon^{-1}\mathbf{G} & \mathbf{G}'\mathbf{V}_\varepsilon^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_\varepsilon^{-1}\mathbf{F} & \mathbf{Z}'\mathbf{V}_\varepsilon^{-1}\mathbf{G} & \mathbf{Z}'\mathbf{V}_\varepsilon^{-1}\mathbf{Z} + \mathbf{V}_\gamma^{-1} \end{vmatrix}^{1/p}}{\begin{vmatrix} \mathbf{G}'\mathbf{V}_\varepsilon^{-1}\mathbf{G} & \mathbf{G}'\mathbf{V}_\varepsilon^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{V}_\varepsilon^{-1}\mathbf{G} & \mathbf{Z}'\mathbf{V}_\varepsilon^{-1}\mathbf{Z} + \mathbf{V}_\gamma^{-1} \end{vmatrix}^{1/p} C}. \tag{20}$$

In a similar way as (15), the trend factor of the $(\mathcal{D}_t, C)$-optimal run order $\delta_{(\mathcal{D}_t,C)}$ is defined as

$$\mathrm{TF}(\delta_{(\mathcal{D}_t,C)}) = \left\{ \frac{\mathcal{D}_t(\delta_{(\mathcal{D}_t,C)})}{\mathcal{D}(\delta_{\mathcal{D}})} \right\}^{1/p}. \tag{21}$$

4.4 Estimation of the variance components

Not surprizingly the optimality of a cross-over design depends on the unknown covariance structure. Henceforth, let σ_ε^2, σ_γ^2 and ρ denote the guesses or point estimates of the true but unknown parameters σ_ε^{2*}, σ_γ^{2*} and ρ^* respectively. These guesses or estimates have to be used in the optimality criteria (12), (14) and (20) and may sometimes be available from a pilot study or a similar investigation.

It should however be pointed out that the estimates of the variance components based on prior experimentation are generally biased downwards because the estimation methods assume that the fixed effects are known, rather than being estimated from the data. Residual or restricted maximum likelihood overcomes this problem by automatically adjusting for the loss in degrees of freedom corresponding to the estimation of the fixed effects. More on the analysis of cross-over designs can be found in the excellent textbook of Vonesh and Chinchilli (1997).

Using estimates of the variance components as approximations to the true variance components generally leads to an estimated variance-covariance matrix of the important effects that is biased downwards. This is due to the fact that the variability introduced by not working with the true variance components is not taken into account in the approximation for the true variance-covariance matrix (see Kackar and Harville, 1984).

The next section describes our algorithm to compute trend-resistant and cost-efficient run orders for cross-over studies under the linear mixed effects model.

5 The design construction algorithm

The literature offers a number of methods to compute optimal designs for correlated observations. For instance, Eccleston and Chan (1998) present algorithms based on simulated annealing and tabu search to construct optimal row-column designs. Donev (1998a)

does the same but uses an exchange algorithm. Design algorithms for the construction of cross-over designs for treatment comparisons can be found in Donev and Jones (1995), Jones and Donev (1996) and Donev (1997, 1998b). Goos and Vandebroek (2001a, 2001b) develop exchange algorithms to compute $\mathcal{D}$-optimal regression designs in the presence of random block effects and for split-plot experiments respectively. Tack and Vandebroek (2002) provide an exchange algorithm to construct trend-resistant and cost-efficient regression designs with either fixed or random block effects.

Here, the latter exchange algorithm will be extended to construct optimal cross-over designs for the linear mixed model. An exchange algorithm computes an optimal design by searching over a predefined list of candidate points. This list of candidate or support points consists of a coarse grid which covers the entire experimental region. It is usually computed as in Atkinson and Donev (1992). They assume the design region to be a hypercube and the grid points are chosen with two, three, four or more equi-spaced levels per factor depending on the fact whether the response model respectively contains linear, quadratic, cubic or higher order terms. For instance, for a full quadratic model in two factors the grid points are $(0,0)$, $(\pm 1,0)$, $(0,\pm 1)$ and $(\pm 1,\pm 1)$. Alternatively, the experimenter can specify the set of candidate points himself. This often occurs when the design region is hyperspherical, restricted due to technical constraints on the factor levels or when a finer grid is necessary. After specifying or constructing the candidate set of design points, the design problem is the combinatorial one of selecting with replacement n design points out of the list of d candidate points in order to optimize a user-specified optimality criterion. The number of observations n is usually dictated by technical or economical considerations.

5.1 Description of the algorithm

The input to our algorithm consists of the number of observations n, the number of subjects b, the number of measurements k_i per subject (with $i \in \{1, \ldots, b\}$), the number of factors m, the order and the number of parameters p of the response function, the polynomial expansion for the response function $\mathbf{f}(\mathbf{x})$, the order and the number of parameters q of the time trend, the polynomial expansion for the time trend $\mathbf{g}(t)$, cost information and the list of time points at which the important response can be measured. A covariance pattern and estimates σ_ε^2, σ_γ^2 and ρ of the true parameters σ_ε^{2*}, σ_γ^{2*} and ρ^* are also needed. The experimenter also has the possibility to impose a constraint on the number of replications for each design point $\mathbf{x}_i$. Such constraints are usually due to the scarcity of resources. Finally, the optimality criterion must be specified too. Choices are among the $\mathcal{D}$-, the $\mathcal{D}_t$- and the $(\mathcal{D}_t, C)$-optimality criterion.

The exchange algorithm proceeds in three phases. In the first phase the experimenter is at liberty to include n_s $(0 \leq n_s < n)$ design points with corresponding time points. These points represent data already available and the design problem then boils down to the augmentation or the repair of a previously failed experiment. A starting run order is now completed by the assignment of $n_0 - n_s$ $(n_s \leq n_0 \leq n)$ randomly chosen design

points to arbitrarily chosen time points within the subjects. To avoid singular information matrices, the number n_0 must satisfy $n_0 \geq p + q + b$.

In the second phase the starting run order is augmented to n trials. This is done by sequentially adding $n - n_0$ design points at time points still available. These additions are made at the largest improvement of the criterion value and within the constraints k_i on the number of observations per subject.

The third phase is called the optimization phase. This final phase comprises an iterative improvement of the run order by evaluating three different exchanges or interchanges of design points and/or time points. Firstly, the effect on the criterion value is investigated of replacing any design point $\mathbf{x}_i$ within a subject by a new design point $\mathbf{x}_j$ from the list of candidate points. Secondly, the effect on the optimality criterion is examined of moving a design point from any time point t_{ij} within a subject i to a free time point t_{ik} within the same subject. Thirdly, the interchange of design point $\mathbf{x}_{ij}$ at time point t_{ij} within subject i and design point $\mathbf{x}_{kl}$ at time point t_{kl} within another subject k is evaluated. The design change leading to the largest improvement of the criterion value will be executed and the process is repeated till no further improvement in the criterion value can be obtained. Note that during the optimization phase none of the n_s observations specified in the first phase can be removed from the run order or assigned to another subject or time point.

The probability of finding the global optimum, instead of getting stuck at a local optimum, can be increased by repeating the algorithmic search several times from different starting designs or 'tries'. This number of tries ν is a user-specified constant. The appendix outlines the design algorithm in detail.

5.2 Computational aspects

In the exchange algorithms of Fedorov (1972), Wynn (1972), Mitchell (1974), Cook and Nachtsheim (1980) and Atkinson and Donev (1989, 1992) for the linear model with uncorrelated errors the effect of a design change can readily be calculated since the information matrix can be written as a sum of outer products. This is also possible in the approaches of Tack and Vandebroek (2001, 2002) and Goos and Vandebroek (2001a, 2001b). More specifically, the criterion value after the addition or the deletion of a design point is easily updated by simply adding or subtracting outer products of the polynomial expansions of the design points.

However, in case of correlated error terms the information matrix can no longer be written as a sum of outer products. Updating the determinant and the inverse of the information matrix then becomes computationally more prohibitive. In Chasalow (1992) and Donev (1998b) the problem is solved by using the Cholesky decomposition (see theorem 14.5.11 of Harville, 1997) of the variance-covariance matrix of the observations. Cholesky decomposition enables one to write the variance-covariance matrix as the product of a unique lower triangular matrix and its transpose. The information matrix can then be written

as a sum of outer products, which makes the use of handy update formulae still possible. Since the variance-covariance matrix of the observations remains unchanged during the optimization phase, the decomposition has to be carried out only once.

Unfortunately, in this paper Cholesky decomposition is no longer useful. This is because the decomposition has to be carried out each time the elements of the variance-covariance matrix of the observations are affected by a design change. For instance, if the covariance structure is based on the spatial patterns (6), (7) or (8) and if an already present design point is assigned to a new time point, the variance-covariance matrix of the observations has to be updated and a new Cholesky decomposition is needed. We are therefore no longer convinced of the computational advantages of using Cholesky decomposition. Instead, the criterion values (12), (14) or (20) will be calculated from scratch after each design change. This facilitates the computer code because there is no need to implement difficult update formulae.

6 Examples and practical applications

This section presents two examples in order to carefully illustrate the developed design methodology. The inspection experiment is discussed in the first example, in which the aim is to demonstrate the benefit of taking into account cost considerations during the design phase of a cross-over study. The experiment will also clarify how to use the exchange algorithm of section 5 in a practical setting. The aim of the second example is to investigate how robust the optimal run orders are for misspecification of the variance components, the correlation ρ^* and the covariance pattern.

6.1 The inspection experiment

This example concerns an experiment in acceptance inspection in quality assurance. Inspection has always been a vital part of a manufacturing process. The purpose of inspection, i.e. the control of outcoming products, is to accept or reject an item based on specified quality characteristics. An inspector examines all outcoming items one at a time and measures important quality characteristics to determine if they conform to specified purchasing standards. Unfortunately, inspection is not always perfect because it may give false results due to monotony and repetition associated with the inspection task. It can create boredom and fatigue in inspectors and cause rapid degradation of the inspector's performance as the number of inspected items increases. Besides these mental issues, the heterogeneity of the group of inspectors may also influence the quality of the inspection task. Think of the variation in terms of age, experience, physical characteristics and cognitive abilities. Physical issues in the work environment form another critical element in human performance in inspection tasks. Important physical issues include the surroundings such as lighting and temperature. The appropriate strategy then is to design the environment and the workload to fit the human inspectors and not the other way around.

Therefore, an experiment will be set up in an attempt to decide on the most legible attributes of an inspection environment. The production process to be inspected consists of an oil groove cutting process and the subsequent soldering of lead wires. The outcoming items have to be checked for both the grooves and the quality of the soldering step. In the cutting process, tool changes and damaged bits can lead to items missing grooves. Also items with grooves cut in the wrong places must be discovered and removed. Finally, soldering defects such as undone lead wires make an item unacceptable. The aim of the experiment is to improve the probability of succesful human performance of the inspection task. Therefore, the experiment involves a number of inspectors and the influence of two physical factors on human performance is investigated: the distance between the inspector and the outcoming items (x_1) and the ambient temperature (x_2). The corresponding factor levels are shown in table 1. Each combination of levels of the two controllable factors constitutes an experimental condition. A measurement then comprises the control by a particular inspector of a fixed number of outcoming items for a certain combination of distance x_1 and temperature x_2. The number of nonconforming items detected by the inspector is compared to the true error rate of the checked items and a performance ratio is calculated. All measurements on the same inspector are performed one after the other. This induces the results being influenced by an unwanted time trend due to inspector fatigue over the course of the experiment.

Table 1: Description of the inspection experiment

factor	levels	coded levels
distance x_1 (m)	0,1; 0,5; 0,9	-1, 0, 1
temperature x_2 (°C)	16, 20, 24	-1, 0, 1

The input to the exchange algorithm consists of the number of tries $\nu = 30000$, the number of factors $m = 2$, the number of observations $n = 18$, the number of inspectors $b = 3$, the number of measurements per inspector $k_1 = k_2 = k_3 = 6$, the design points $(0,0)$, $(0,\pm 1)$, $(\pm 1,0)$ and $(\pm 1,\pm 1)$ and the full second-order polynomial expansion $\mathbf{f}(\mathbf{x}) = [1, x_1, x_2, x_1x_2, x_1^2, x_2^2]'$. The transition cost for altering the distance x_1 or the temperature x_2 from ± 1 to 0 or vice versa is assumed to be twice as high as when the particular factor level remains the same. The transition between -1 and $+1$ or vice versa is three times as expensive. Besides, the error terms are assumed to follow the AR(1) model with $\rho \in \{-0.9; -0.5; 0.0; 0.5; 0.9\}$. The guesses of the variance components are based on previous experimentation and equal $\sigma_\varepsilon^2 = 0.9$ and $\sigma_\gamma^2 = 0.1$. $\mathcal{D}_t$- and $(\mathcal{D}_t, C)$-optimal run orders will now be computed for different postulated time trends and 6 equally spaced time points between -1 and 1 within each block. These time trends are of first, second, third or fourth order and will be denoted as $\mathbf{g}_1(t)$, $\mathbf{g}_2(t)$, $\mathbf{g}_3(t)$ and $\mathbf{g}_4(t)$ respectively. Finally, no restriction is imposed on the number of replicates per design point.

As an illustration, the computed $\mathcal{D}_t$-optimal run orders with the postulated time trend

$\mathbf{g}_1(t)$ are displayed in table 2. A striking result is that the number of factor level changes increases with the autocorrelation coefficient ρ. A similar conclusion could be drawn for the $\mathcal{D}_t$-optimal run orders computed with other postulated time trends and/or other values for σ^2_ε, σ^2_γ or ρ.

Table 2: $\mathcal{D}_t$-optimal run orders for $\mathbf{g}_1(t)$ and varying values of ρ

		$\rho = -0.9$		$\rho = -0.5$		$\rho = 0.0$		$\rho = 0.5$		$\rho = 0.9$	
subject	t	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
1	-1.0	-1	0	0	-1	1	1	1	-1	-1	-1
1	-0.6	-1	0	0	-1	1	-1	-1	-1	-1	1
1	-0.2	1	-1	1	1	0	0	0	1	1	0
1	0.2	1	-1	1	1	0	0	-1	-1	-1	1
1	0.6	0	1	-1	0	-1	-1	1	0	0	-1
1	1.0	0	1	-1	0	-1	1	-1	1	1	1
2	-1.0	0	-1	-1	1	1	1	-1	1	1	-1
2	-0.6	0	-1	-1	1	0	-1	0	-1	0	1
2	-0.2	-1	1	0	0	-1	1	1	0	1	-1
2	0.2	-1	1	0	0	-1	-1	-1	1	-1	-1
2	0.6	1	0	1	-1	1	-1	0	-1	0	0
2	1.0	1	0	1	-1	1	0	1	1	-1	-1
3	-1.0	1	1	1	0	-1	0	1	1	1	1
3	-0.6	1	1	1	0	-1	1	-1	0	0	-1
3	-0.2	0	0	-1	-1	1	-1	1	1	-1	0
3	0.2	0	0	-1	-1	-1	-1	1	-1	1	1
3	0.6	-1	-1	0	1	0	1	0	0	-1	0
3	1.0	-1	-1	0	1	1	1	1	-1	1	-1

All computed run orders are compared to each other in table 3. The comparison is made in terms of the amount of information obtained on the important parameters $\boldsymbol{\beta}$, i.e. the $\mathcal{D}_t$-criterion value. The first panel refers to the $\mathcal{D}_t$-optimal run orders, whereas the second panel relates to the $(\mathcal{D}_t, C)$-optimal run orders. For instance, the $\mathcal{D}_t$-criterion value of the $(\mathcal{D}_t, C)$-optimal run order with $\rho = 0.9$ and a first-order time dependence $\mathbf{g}_1(t)$ equals 48.9. As a matter of fact, the $(\mathcal{D}_t, C)$-optimal run orders are less informative than their $\mathcal{D}_t$-optimal counterparts. It also turns out that, for a given optimality criterion and given order of the time trend, more information is gathered if the observations are strongly autocorrelated, either negatively or positively.

In terms of the trend factors (15) and (21), the computed $\mathcal{D}_t$- and $(\mathcal{D}_t, C)$-optimal run orders are compared to each other in table 4. A quick glance at the table reveals that for a given autocorrelation coefficient ρ the trend factor decreases with the order of the postulated time dependence. This is not unexpected since the more complicated the time

Table 3: $\mathcal{D}_t$-criterion values of the optimal run orders

	$\mathcal{D}_t$-optimality				$(\mathcal{D}_t, C)$-optimality			
ρ	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$
-0.9	64.8	53.8	52.1	48.1	58.0	50.0	48.7	45.3
-0.5	15.3	12.7	12.5	11.1	14.1	11.2	11.2	9.9
0.0	8.5	7.7	7.7	7.0	7.6	6.3	6.2	5.5
0.5	13.7	12.5	12.4	11.9	12.5	11.6	11.5	10.7
0.9	52.3	50.3	50.1	49.2	48.9	47.7	47.3	45.6

trend, the more difficult it is to attain an optimal protection against unwanted time trend effects. It is also worthy of mention that in most cases the $\mathcal{D}_t$-optimal run orders become better balanced for trend effects as the coefficient ρ grows larger. Alternatively stated, the more the observations belonging to the same subject are positively autocorrelated, the less distortion of the results by the unwanted trend effects.

Table 4: Trend factors of the optimal run orders

	$\mathcal{D}_t$-optimality				$(\mathcal{D}_t, C)$-optimality			
ρ	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$
-0.9	0.970	0.804	0.779	0.719	0.868	0.747	0.728	0.678
-0.5	0.981	0.816	0.801	0.708	0.903	0.717	0.715	0.637
0.0	0.999	0.903	0.899	0.823	0.890	0.743	0.731	0.646
0.5	0.999	0.912	0.905	0.864	0.908	0.842	0.834	0.778
0.9	0.999	0.963	0.958	0.940	0.936	0.913	0.905	0.873

It can also be seen from table 4 that the temporal trend causes much more harm to the estimates in case of the $(\mathcal{D}_t, C)$-optimality criterion than in case of the $\mathcal{D}_t$-optimality criterion. For instance, for a linear time dependence $\mathbf{g}_1(t)$ and $\rho = 0.9$, the trend factor of the $(\mathcal{D}_t, C)$-optimal run order only equals 0.936, whereas the trend factor of the $\mathcal{D}_t$-optimal run order amounts to 0.999. However, the $(\mathcal{D}_t, C)$-optimal run orders demonstrate superiority with respect to the amount of information obtained per unit cost. This is clearly illustrated in table 5, which compares the computed run orders in terms of their $(\mathcal{D}_t, C)$-criterion value. It also follows that autocorrelation is very beneficial to the $(\mathcal{D}_t, C)$-criterion value of the optimal run orders. The highest $(\mathcal{D}_t, C)$-criterion values, i.e. the lowest cost of information, are obtained for strongly negative values of the autocorrelation parameter ρ. Finally, the cost of information increases with the order of the postulated time dependence.

Table 5: $(\mathcal{D}_t, C)$-criterion values of the optimal run orders

	$\mathcal{D}_t$-optimality				$(\mathcal{D}_t, C)$-optimality			
ρ	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$	$\mathbf{g}_1(t)$	$\mathbf{g}_2(t)$	$\mathbf{g}_3(t)$	$\mathbf{g}_4(t)$
-0.9	1.410	1.120	1.085	1.002	1.416	1.190	1.160	1.079
-0.5	0.333	0.265	0.272	0.230	0.371	0.320	0.319	0.284
0.0	0.142	0.119	0.128	0.121	0.200	0.176	0.173	0.158
0.5	0.196	0.176	0.180	0.175	0.223	0.203	0.201	0.191
0.9	0.746	0.719	0.726	0.723	0.829	0.796	0.789	0.774

6.2 Robustness against misspecification

This section studies the robustness of the $\mathcal{D}_t$-optimal designs against misspecification of the variance components σ_ε^{2*} and σ_γ^{2*}, the parameter ρ^* and the covariance pattern of the error terms. More specifically, consider an experiment with two factors x_1 and x_2, the polynomial expansion $\mathbf{f}(\mathbf{x}) = [1, x_1, x_2, x_1x_2, x_1^2, x_2^2]'$, a first-order polynomial time trend $\mathbf{g}_1(t)$, design points taken from the 3^2 factorial, no restriction on the number of replicates per design point and three subjects with six measurements each. Within each subject there are 6 equally spaced time points between -1 and 1. $\mathcal{D}_t$-optimal run orders will be computed for various combinations of the error structure, the variance components σ_ε^2 and σ_γ^2 and the parameter ρ. In order to preserve valid comparisons and without loss of generality, the total variance $\sigma_\varepsilon^2 + \sigma_\gamma^2$ is set equal to 1. The following patterns for the variance-covariance matrix $\mathbf{V}_{\boldsymbol{\varepsilon}_i}$ are used:

- E1: independent errors with constant variance
- E2: compound symmetry error structure (3)
- E3: autoregressive error terms (5)
- E4: the power model (6)
- E5: the exponential model (7)
- E6: the Gaussian model (8)

Figure 1 displays the criterion values of all $\mathcal{D}_t$-optimal run orders computed for several combinations of the error structure E_i and the subject variance σ_γ^2. Note that the error variance σ_ε^2 is then automatically set equal to $1 - \sigma_\gamma^2$. The results are restricted to cases in which $\rho = 0.5$. More specifically, figure 1 relates to 54 run orders resulting from the combination of 9 subject variances $\sigma_\gamma^2 \in \{0.1; 0.2; \ldots; 0.8; 0.9\}$ and 6 error structures E_i. Each line then refers to the $\mathcal{D}_t$-optimal run orders computed for a particular error structure (see the legend of the figure) and differing subject variances σ_γ^2 on the horizontal axis. For instance, the criterion value (on a per parameter basis) of the $\mathcal{D}_t$-optimal run

order under the assumption of independent error terms E1 (solid line) and $\sigma_\gamma^2 = 0.5$ (see the horizontal axis) is equal to 12.03. A striking result is that for each line (i.e. for any covariance pattern E_i) the optimal run orders become more informative as the subject variance σ_γ^2 grows larger. In addition, for a given subject variance σ_γ^2 on the horizontal axis, the largest criterion value corresponds with the covariance pattern defined by the power model (E4). The opposite is true for error terms assumed to be uncorrelated, namely error structure E1.

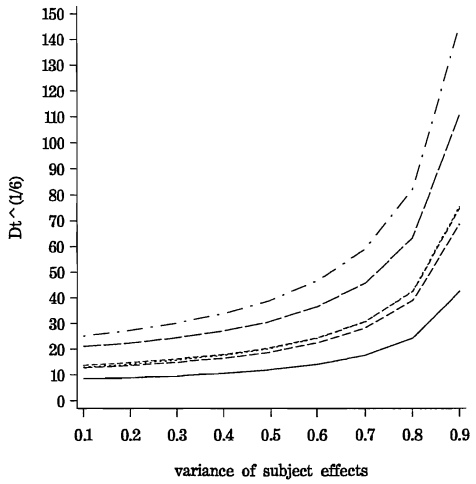

Figure 1: $\mathcal{D}_t$-criterion values of the optimal run orders for $\rho = \rho^* = 0.5$ and different error structures and subject variances σ_γ^2

——— (E1) independent error terms
····· (E2) compound symmetry
- - - - - - - (E3) autoregressive error terms
·—·—·— (E4) power model
– – – – – (E5) exponential model
— — — (E6) Gaussian model

In a similar way, figure 2 depicts the criterion values of all $\mathcal{D}_t$-optimal run orders computed for several combinations of the error structure E_i and parameter ρ, but with fixed $\sigma_\varepsilon^2 = \sigma_\gamma^2 = 0.5$. In particular, it concerns 45 $\mathcal{D}_t$-optimal run orders resulting from the combination of 9 values of ρ (i.e. $\rho \in \{0.1; 0.2; \ldots; 0.8; 0.9\}$) and 5 error structures E_i. No results are displayed for the $\mathcal{D}_t$-optimal run orders with independent error terms E1 since for this error structure the parameter ρ is not defined. Each line refers to the $\mathcal{D}_t$-optimal run orders computed for a particular error structure (see the legend of the figure) and

differing parameters ρ on the horizontal axis. For instance, the criterion value (on a per parameter basis) of the $\mathcal{D}_t$-optimal run order with error terms according to the Gaussian model E6 (see the legend of the figure) and $\rho = 0.9$ (see the horizontal axis) is equal to 447. Not surprizingly, it holds for each line (i.e. for any covariance pattern E_i) that the larger the value of ρ, the more informative the associated optimal run order. The increase in the criterion value for large values of ρ is especially pronounced for an error structure that follows the Gaussian model E6.

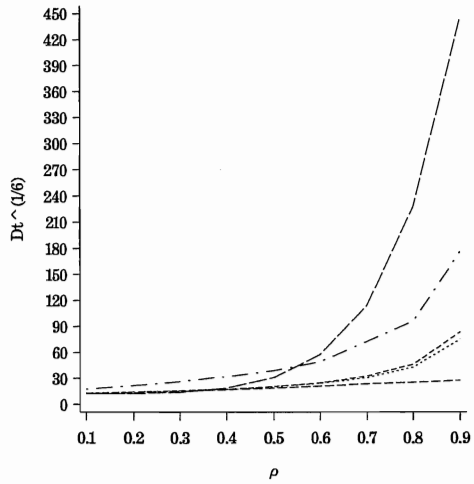

Figure 2: $\mathcal{D}_t$-criterion values of the optimal run orders for $\sigma_\varepsilon^2 = \sigma_\gamma^2 = 0.5$ and different error structures and values of ρ

..... (E2) compound symmetry
- - - - - - - (E3) autoregressive error terms
·—·—·- (E4) power model
– – – – – (E5) exponential model
— — — (E6) Gaussian model

The trend factors (15) of all computed run orders are given in table 6 and table 7. Table 6 refers to the run orders computed with $\rho = 0.5$ (i.e. the run orders used in figure 1), whereas in table 7 $\sigma_\gamma^2 = \sigma_\varepsilon^2 = 0.5$ is assumed (see the run orders used in figure 2). For instance, the $\mathcal{D}_t$-optimal run order under the assumptions $\sigma_\gamma^2 = 0.1$, $\sigma_\varepsilon^2 = 0.9$, $\rho = 0.5$ and uncorrelated errors (E1) has trend factor 0.9995. Both tables obviously show the complete trend-resistance of the $\mathcal{D}_t$-optimal run orders under the compound symmetry error structure (E2). Generally speaking, all run orders demonstrate a good safeguard against unwanted trend effects. It also follows that, at least for this example, low subject

variances σ_γ^2 in table 6 or high values for ρ in table 7 are beneficial for the trend factors of the optimal run orders with autocorrelated error terms (E3) or errors that follow the power model (E4). The opposite is true if the covariance pattern follows the Gaussian model (E6).

Table 6: Trend factors of the $\mathcal{D}_t$-optimal run orders for several error structures ($\rho = 0.5$)

σ_γ^2	σ_ε^2	E1	E2	E3	E4	E5	E6
0.1	0.9	0.9995	1.0000	0.9986	0.9996	0.9976	0.9771
0.2	0.8	0.9995	1.0000	0.9983	0.9994	0.9977	0.9769
0.3	0.7	0.9995	1.0000	0.9979	0.9992	0.9977	0.9771
0.4	0.6	0.9995	1.0000	0.9975	0.9990	0.9978	0.9774
0.5	0.5	1.0000	1.0000	0.9973	0.9988	0.9976	0.9778
0.6	0.4	1.0000	1.0000	0.9966	0.9987	0.9977	0.9780
0.7	0.3	1.0000	1.0000	0.9965	0.9985	0.9976	0.9782
0.8	0.2	1.0000	1.0000	0.9965	0.9983	0.9976	0.9784
0.9	0.1	1.0000	1.0000	0.9963	0.9981	0.9976	0.9786

Table 7: Trend factors of the $\mathcal{D}_t$-optimal run orders for several error structures ($\sigma_\gamma^2 = \sigma_\varepsilon^2 = 0.5$)

ρ	E2	E3	E4	E5	E6
0.1	1.0000	0.9970	0.9980	0.9993	1.0000
0.2	1.0000	0.9975	0.9977	0.9960	0.9992
0.3	1.0000	0.9979	0.9986	0.9972	0.9957
0.4	1.0000	0.9980	0.9985	0.9982	0.9920
0.5	1.0000	0.9973	0.9988	0.9976	0.9778
0.6	1.0000	0.9960	0.9990	0.9977	0.9721
0.7	1.0000	0.9986	0.9991	0.9973	0.9659
0.8	1.0000	0.9989	0.9993	0.9973	0.9593
0.9	1.0000	0.9992	0.9995	0.9987	0.9531

In the remainder of this section, we will have a close look at the robustness of the optimal run orders for misspecification of σ_γ^{2*}, σ_ε^{2*} or ρ^*. We have computed how the optimal run orders discussed in the previous paragraphs perform for several true variance components σ_ε^{2*} (σ_γ^{2*} is then set to $1 - \sigma_\varepsilon^{2*}$). Without loss of generality, the results will be restricted to fixed $\rho = \rho^* = 0.5$ and error terms following the AR(1) process. They are shown in figure 3. The vertical axis displays the decrease in the trend factor (in percent) caused by possible misspecification of the true error variance σ_ε^{2*} as σ_ε^2 ($\sigma_\gamma^2 = 1 - \sigma_\varepsilon^2$). The true error variances σ_ε^{2*} are marked out on the horizontal axis. More specifically, each line in

figure 3 refers to one particular run order, namely the $\mathcal{D}_t$-optimal run order computed with $\rho = 0.5$ and the assumed error variance σ_ε^2 as shown in the legend of the figure. For instance, the solid line refers to the $\mathcal{D}_t$-optimal run order computed with $\rho = \rho^* = 0.5$ and $\sigma_\varepsilon^2 = 0.1$. The line shows how this run order performs if the true error variance σ_ε^{2*} (on the horizontal axis) differs from the assumed value $\sigma_\varepsilon^2 = 0.1$. For instance, if the true error variance equals $\sigma_\varepsilon^{2*} = 0.6$ (on the horizontal axis), whereas the assumed one equals $\sigma_\varepsilon^2 = 0.1$, then this run order is about 0.13% less trend-resistant than if the error variance would have been correctly specified as $\sigma_\varepsilon^2 = \sigma_\varepsilon^{2*} = 0.9$. It turns out that the more the true error variance σ_ε^{2*} on the horizontal axis deviates from the assumed variance $\sigma_\varepsilon^2 = 0.1$, the lower the performance of that run order. For instance, if the true error variance is $\sigma_\varepsilon^{2*} = 0.9$ rather than 0.1, then the loss in the trend factor amounts to about 0.38%, which is much more than 0.13%. This observation also holds for the optimal run orders computed with another assumed error variance σ_ε^2 (see the other lines in the figure).

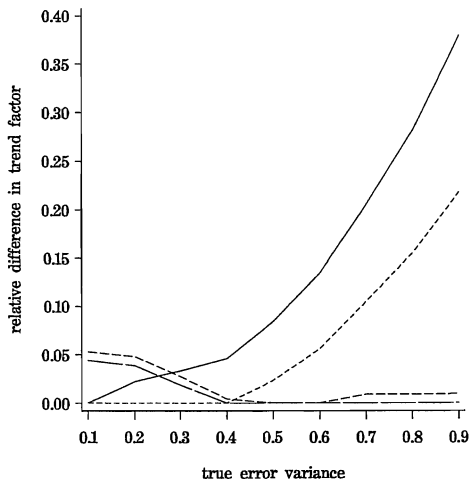

Figure 3: Relative difference in trend factor for several assumed error variances, error structure E3 and $\rho = \rho^* = 0.5$

——— $\sigma_\varepsilon^2 = 0.1$
- - - - - - - $\sigma_\varepsilon^2 \in \{0.2; 0.3; 0.4\}$
– – – – – $\sigma_\varepsilon^2 \in \{0.5; 0.6\}$
—— —— $\sigma_\varepsilon^2 \in \{0.7; 0.8; 0.9\}$

In fact, the losses in the trend factor due to misspecification of the variance components are very small as compared to the decreases caused by misspecification of the autocorrelation coefficient ρ^*. The results for misspecification of ρ^* are given in figure 4, in which each

line refers to the $\mathcal{D}_t$-optimal run order computed with $\sigma_\varepsilon^2 = \sigma_\varepsilon^{2*} = \sigma_\gamma^2 = \sigma_\gamma^{2*} = 0.5$ and a particular coefficient ρ shown in the legend of the figure. The true autocorrelation coefficients ρ^* are marked out on the horizontal axis. A quick glance at the figure learns that the losses in the trend factor due to misspecification of ρ^* increase with the difference between the assumed correlation coefficient ρ and the true one ρ^*. For instance, suppose that the assumption $\rho = 0.6$ is made. The legend refers to the corresponding line in the figure. If the true autocorrelation coefficient equals $\rho^* = 0.4$ (on the horizontal axis), then the loss in the trend factor due to a wrong specification of ρ equals 0.64%. However, if the true autocorrelation coefficient equals $\rho^* = 0.2$ (on the horizontal axis), then the loss in the trend factor is much larger, namely 1.76%.

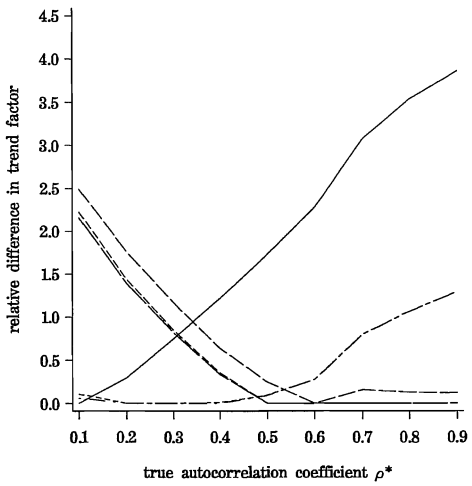

Figure 4: Relative difference in trend factor for several values of ρ^*, error structure E3 and $\sigma_\varepsilon^2 = \sigma_\varepsilon^{2*} = \sigma_\gamma^2 = \sigma_\gamma^{2*} = 0.5$

——— $\rho = 0.1$
- - - - - - - $\rho \in \{0.2; 0.3\}$
- — · — · - $\rho = 0.4$
– – – – – $\rho = 0.5$
— — — $\rho = 0.6$
—— — $\rho \in \{0.7; 0.8; 0.9\}$

With the exception of error terms that follow the Gaussian model, comparable losses were observed for the other error structures. In case of the Gaussian model, the decreases in the trend factor due to misspecification of the autocorrelation coefficient ρ^* are more pronounced. Figure 5 demonstrates that in this case the loss can amount to 40%.

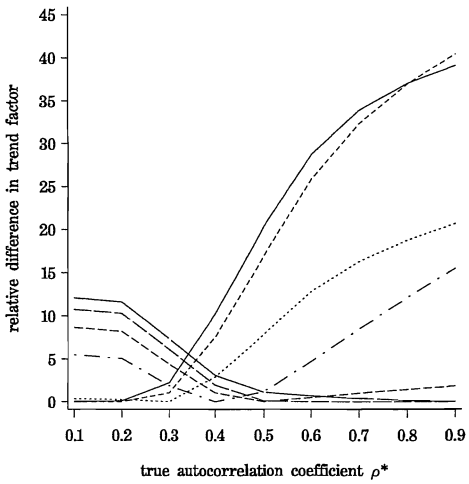

Figure 5: Relative difference in trend factor for several values of ρ^*, error structure E6 and $\sigma_\varepsilon^2 = \sigma_\varepsilon^{2*} = \sigma_\gamma^2 = \sigma_\gamma^{2*} = 0.5$

——— $\rho^* = 0.1$
- - - - - - - $\rho^* = 0.2$
· · · · · · · · · · $\rho^* = 0.3$
- — · — · - $\rho^* = 0.4$
– – – – – $\rho^* = 0.5$
— — — $\rho^* \in \{0.6; 0.7; 0.8\}$
——— – $\rho^* = 0.9$

As a final illustration, it is worthwhile to investigate the robustness of the computed run orders against misspecification of the covariance pattern. In order to clarify the importance of such an investigation, suppose that for a specific design problem at hand a $\mathcal{D}_t$-optimal run order is computed under the assumption of autocorrelated error terms. If, however, the subsequent data analysis and statistical inference would reveal that the error terms are uncorrelated rather than autocorrelated, then one may wonder to what extent the computed run order is still useful as compared to an optimal run order under the correct assumption about the error terms. Here, the results are confined to the $\mathcal{D}_t$-optimal run orders with $\sigma_\varepsilon^2 = \sigma_\gamma^2 = \rho = 0.5$, but the conclusions can be extended to other parameter settings. These parameters are assumed to be correctly specified, i.e. $\sigma_\varepsilon^{2*} = \sigma_\gamma^{2*} = \rho^* = 0.5$. Table 8 displays the results. Each column refers to the $\mathcal{D}_t$-optimal run order computed for a particular assumption about the error structure and gives the losses in the trend factor due to possible misspecification of the covariance

pattern. For instance, if the error terms are assumed to be uncorrelated (E1), whereas they in fact decay according to the AR(1) model (E3), the computed run order is 7.30% less trend-resistant than the optimal run order that one would obtain if the error structure would have been specified correctly. It follows that the assumption of uncorrelated error terms (E1) or compound symmetry (E2) is least robust against misspecification, since these assumptions involve large decreases in the trend factor if neither E1 nor E2 is the true covariance pattern (see the first two columns). Much better safeguard against misspecification of the error structure is obtained if the error terms are assumed to follow the AR(1) model (E3), the power model (E4) or the exponential model (E5). In these cases, the losses are much lower (see the columns with headings E3, E4 and E5). For instance, misspecifying the true error structure E1 as E3 leads to a loss of about 3.39%, whereas in the opposite case a misspecification of the true error structure E3 as E1 gives a loss of 7.30%. In addition, it can also be seen from the tabulated results that no harm is caused if the error terms are postulated to be uncorrelated (E1), whereas in fact they have a compound symmetric pattern (E2), or vice versa. Finally, the loss in the trend factor is also negligible if the AR(1), the power or the exponential model is assumed and if the true error structure is one of these three covariance patterns.

Table 8: Relative difference in trend factor for several error structures and $\sigma_\gamma^2 = \sigma_\gamma^{2*} = \sigma_\varepsilon^2 = \sigma_\varepsilon^{2*} = \rho = \rho^* = 0.5$

		assumed covariance pattern					
true covariance pattern		E1	E2	E3	E4	E5	E6
independent error terms	E1	0.00	0.00	3.39	3.86	1.88	8.67
compound symmetry	E2	0.00	0.00	3.74	4.23	2.09	9.41
autoregressive error terms	E3	7.30	14.94	0.00	0.03	0.00	2.35
power model	E4	11.09	19.78	0.00	0.00	0.36	1.86
exponential model	E5	6.56	13.79	0.19	0.23	0.00	2.69
Gaussian model	E6	29.21	39.09	4.17	5.10	5.31	0.00

Summing up, the results of the robustness study have revealed that little harm is induced by misspecifying the variance components, whereas the opposite is true if the correlation parameter is misspecified. In the latter case, assuming a large value for the correlation parameter is recommended to guard against possible misspecification. In addition, good protection against misspecification of the error structure is obtained if the error terms are modeled according to the AR(1), the power or the exponential model. On the contrary, destructive loss can result from neglecting the possibility for correlated error terms.

7 Conclusion

An exchange algorithm was presented to compute optimal cross-over designs for linear mixed effects models. Under this model, correlation among the repeated measures is induced either indirectly through the subject specific random effects or directly through specification of the intrasubject variance-covariance matrix. This paper has focused on covariance structures that readily fit into the cross-over scheme. The covariance patterns are usually chosen to depend on a variable such as the time point or the number of the observation. The algorithm enables the experimenter to construct run orders of cross-over studies that are optimally balanced for polynomial time trend effects. A practical application has clearly shown that taking into account cost considerations leads to substantial decreases in the cost per unit information obtained. In addition, the results have revealed that misspecification of the parameter ρ can lead to considerable losses in the trend-resistance. Finally, ignoring the possibility for correlated error terms can be very destructive for the optimality of the computed run order. Therefore, an optimal protection against misspecification of the covariance pattern is obtained if the error terms are assumed to follow the AR(1), the power or the exponential model.

Appendix. The design algorithm

In the outline of the exchange algorithm the value of the user-specified optimality criterion is written as $\mathcal{Q}$. The $\mathcal{D}$-, the $\mathcal{D}_t$- and the $(\mathcal{D}_t, C)$-optimality criterion are different possibilities. The set P stands for the list of d distinct candidate points. The set B contains the b test subjects and the number of observations per subject are denoted as $k_1, \ldots, k_b$. The list of time points is given by $T = \{T_1, \ldots, T_b\}$ where T_r denotes the list of time points within subject r. After each design change the list of time points T has to be updated. Let the initial list of time points be written as T^*. The current number of measurements per subject is given by the series $k_{(1)}, \ldots, k_{(b)}$ and n_i represents the current number of replicates at design point $\mathbf{x}_i$. The set $R = \{R_1, \ldots, R_b\}$ represents the run order, where $R_r = \{(\mathbf{x}_i, t_j)\}$ is the run order for subject r. The optimal run order and the corresponding criterion value will be written as $R_{\text{opt}} = \{R_{1_{\text{opt}}}, \ldots, R_{b_{\text{opt}}}\}$ and $\mathcal{Q}_{\text{opt}}$ respectively. In order to preserve clearness, the possibilities to include $n_s \geq 0$ design points in the starting run order and to impose a restriction on the number of replicates for each design point are omitted from the outline. After reading the input, the algorithm proceeds as follows:

1. Set $\mathcal{Q}_{\text{opt}} = 0$.
2. Repeat ν times:
   (a) Set $\forall i \in P : n_i = 0$, $\forall r \in B : k_{(r)} = 0$, $R = \{\}$, $T = T^*$ and $\mathcal{Q} = 0$.
   (b) Randomly choose n_0 with $p + q + b \leq n_0 \leq n$.
   (c) Repeat n_0 times:

i. Randomly choose $r \in B$.

ii. If $T_r \neq \{\}$ and $k_{(r)} < k_r$ then $k_{(r)} = k_{(r)} + 1$, else go to i.

iii. Randomly choose $k \in T_r$.

iv. Randomly choose $i \in P$.

v. $n_i = n_i + 1$.

vi. $R_r = R_r \cup \{(\mathbf{x}_i, t_k)\}$.

vii. $T_r = T_r \setminus \{t_k\}$.

viii. Update $\mathcal{Q}$.

(d) Repeat $n - n_0$ times:

i. Determine $i \in P$, $r \in B$ with $k_{(r)} < k_r$ and $k \in T_r$ with largest effect on $\mathcal{Q}$.

ii. $n_i = n_i + 1$ and $k_{(r)} = k_{(r)} + 1$.

iii. $R_r = R_r \cup \{(\mathbf{x}_i, t_k)\}$.

iv. $T_r = T_r \setminus \{k\}$.

v. Update $\mathcal{Q}$.

(e) Consider the exchanges and interchanges.

i. Set $\triangle = 1$.

ii. Find the best exchange of design points:
$\forall\, r \in B, \forall\, \mathbf{x}_i \in R_r, \forall\, j \in D,\ \mathbf{x}_i \neq \mathbf{x}_j$:
compute the effect $\triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_j,t_k)}$ on $\mathcal{Q}$ of deleting $(\mathbf{x}_i, t_k)$ and adding $(\mathbf{x}_j, t_k)$ in subject r.
If $\triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_j,t_k)} > \triangle$ then $\triangle = \triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_j,t_k)}$, $S = 1$ and store i, j, k, r.

iii. Find the best exchange of time points:
$\forall\, r \in B, \forall\, t_k \in R_r, \forall\, l \in T_r, t_k \neq t_l$:
compute the effect $\triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_i,t_l)}$ on $\mathcal{Q}$ of deleting $(\mathbf{x}_i, t_k)$ and adding $(\mathbf{x}_i, t_l)$ in subject r.
If $\triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_i,t_l)} > \triangle$ then $\triangle = \triangle^E_{r,(\mathbf{x}_i,t_k),(\mathbf{x}_i,t_l)}$, $S = 2$ and store i, k, l, r.

iv. Find the best interchange between subjects:
$\forall\, r, s \in B$ with $s > r, \forall\, (\mathbf{x}_i, t_k) \in R_r, \forall\, (\mathbf{x}_j, t_l) \in R_s,\ \mathbf{x}_i \neq \mathbf{x}_j$:
compute the effect $\triangle^I_{r,(\mathbf{x}_i,t_k),s,(\mathbf{x}_j,t_l)}$ on $\mathcal{Q}$ of interchanging $(\mathbf{x}_i, t_k)$ in subject r and $(\mathbf{x}_j, t_l)$ in subject s.
If $\triangle^I_{r,(\mathbf{x}_i,t_k),s,(\mathbf{x}_j,t_l)} > \triangle$ then $\triangle = \triangle^I_{r,(\mathbf{x}_i,t_k),s,(\mathbf{x}_j,t_l)}$, $S = 3$ and store i, j, k, l, r, s.

(f) If $\triangle > 1$ then

i. If $S = 1$ then $R_r = R_r \setminus \{(\mathbf{x}_i, t_k)\} \cup \{(\mathbf{x}_j, t_k)\}$, $n_i = n_i - 1$ and $n_j = n_j + 1$.

ii. If $S = 2$ then $R_r = R_r \setminus \{(\mathbf{x}_i, t_k)\} \cup \{(\mathbf{x}_i, t_l)\}$ and $T_r = T_r \setminus \{l\} \cup \{k\}$.

iii. If $S = 3$ then $R_r = R_r \setminus \{(\mathbf{x}_i, t_k)\} \cup \{(\mathbf{x}_j, t_k)\}$ and $R_s = R_s \setminus \{(\mathbf{x}_j, t_l)\} \cup \{(\mathbf{x}_i, t_l)\}$.

iv. Update $\mathcal{Q}$.

v. Go to step (e).

(g) If $\mathcal{Q} \geq \mathcal{Q}_{\text{opt}}$, then $\mathcal{Q}_{\text{opt}} = \mathcal{Q}$, $R_{\text{opt}} = R$.

3. Write $\mathcal{Q}_{\text{opt}}$ and R_{opt}.

The algorithm is implemented in Fortran 77 and makes use of the Netlib library of Bell Labs. More specifically, it uses the random number generator 'rand' and the respective routines 'svert', 'sfact' and 'sdet' to compute the inverse, the factorization and the determinant of a symmetric matrix.

References

Afsarinejad, K. (2001). Trend-free repeated measurement designs, *In* A. Atkinson, P. Hackl and W. Müller (eds), *mODa 6 - Advances in Model-Oriented Design and Analysis*, Heidelberg: Physica-Verlag, pp. 1–12.

Atkinson, A. and Donev, A. (1989). The construction of exact $\mathcal{D}$-optimum experimental designs with application to blocking response surface designs, *Biometrika* **76**: 515–526.

Atkinson, A. and Donev, A. (1992). *Optimum Experimental Designs*, Oxford: Clarendon Press.

Atkinson, A. and Donev, A. (1996). Experimental designs optimally balanced for trend, *Technometrics* **38**: 333–341.

Berger, M. and Tan, F. (1998). Optimal designs for repeated measures experiments, *Kwantitatieve Methoden* **59**: 45–67.

Bischoff, W. (1993). On $\mathcal{D}$-optimal designs for linear models under correlated observations with an application to a linear model with multiple response, *Journal of Statistical Planning and Inference* **37**: 69–80.

Brown, H. and Prescott, R. (1999). *Applied Mixed Models in Medicine*, Chichester: John Wiley & Sons.

Chasalow, S. (1992). *Exact Response Surface Designs with Random Block Effects*, Ph.D. dissertation, University of California, Berkeley.

Cheng, C.-S. and Wu, C.-F. (1980). Balanced repeated measurements designs, *The Annals of Statistics* **8**: 1272–1283.

Cook, R. and Nachtsheim, C. (1980). A comparison of algorithms for constructing exact $\mathcal{D}$-optimal designs, *Technometrics* **22**: 315–324.

Donev, A. (1997). An algorithm for the construction of crossover trials, *Journal of the Royal Statistical Society, Series C,* **46**: 288–295.

Donev, A. (1998a). Construction of non-standard row-column designs, *In* R. Payne and P. Green (eds), *COMPSTAT 1998, Proceedings in Computational Statistics*, Harpenden: IAC-Rothamsted, pp. 275–280.

Donev, A. (1998b). Crossover designs with correlated observations, *Journal of Biopharmaceutical Statistics* **8**: 249–262.

Donev, A. and Jones, B. (1995). Construction of $\mathcal{A}$-optimum cross-over designs, *In* C. Kitsos and W. Müller (eds), *MODA4 - Advances in Model-Oriented Data Analysis*, Heidelberg: Physica-Verlag, pp. 165–171.

Eccleston, J. and Chan, B. (1998). Design algorithms for correlated data, *In* R. Payne and P. Green (eds), *COMPSTAT 1998, Proceedings in Computational Statistics*, Harpenden: IAC-Rothamsted, pp. 41–52.

Fedorov, V. (1972). *Theory of Optimal Experiments*, New York: Academic Press.

Giovagnoli, A. and Romano, D. (2001). Experiments in software engineering: some implications on design, *In Proceedings of the ENBIS conference 2001 (CD-rom).*

Goos, P. and Vandebroek, M. (2001a). $\mathcal{D}$-optimal response surface designs in the presence of random block effects, *Computational Statistics and Data Analysis* **37**: 433–453.

Goos, P. and Vandebroek, M. (2001b). Optimal split-plot designs, *Journal of Quality Technology* **33**: 435–450.

Harville, D. (1997). *Matrix Algebra from a Statistician's Perspective*, New York: Springer-Verlag.

Hedayat, A. and Afsarinejad, K. (1975). Repeated measures designs I, *In* J. Srivastava (ed.), *A Survey of Statistical Design and Linear Models*, Amsterdam: North-Holland, pp. 229–242.

Hedayat, A. and Afsarinejad, K. (1978). Repeated measures designs II, *The Annals of Statistics* **6**: 619–628.

Jones, B. and Donev, A. (1996). Modelling and design of cross-over trials, *Statistics in medicine* **15**: 1435–1446.

Kackar, R. and Harville, D. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models, *Journal of the American Statistical Association* **79**: 853–862.

Kiefer, J. and Wynn, H. (1981). Optimum balanced block and Latin square designs for correlated observations, *The Annals of Statistics* **9**: 737–757.

Kunert, J. (1991). Cross-over designs for two treatments and correlated errors, *Biometrika* **78**: 315–324.

Kushner, H. (1997a). Optimal repeated measurements designs: the linear optimality equations, *The Annals of Statistics* **25**: 2328–2344.

Kushner, H. (1997b). Optimality and efficiency of two-treatment repeated measurements designs, *Biometrika* **84**: 455–468.

Laird, N. and Ware, J. (1982). Random effects models for longitudinal data, *Biometrics* **38**: 963–974.

Martin, R., Eccleston, J. and Jones, G. (1998). Some results on multi-level factorial designs with dependent observations, *Journal of Statistical Planning and Inference* **73**: 91–111.

Martin, R., Jones, G. and Eccleston, J. (1998). Some results on two-level factorial designs with dependent observations, *Journal of Statistical Planning and Inference* **66**: 363–384.

Matthews, J. (1987). Optimal crossover designs for the comparison of two treatments in the presence of carryover effects and auto-correlated errors, *Biometrika* **74**: 311–320. Correction (1988), **75**: 396.

Matthews, J. (1994a). Multi-period crossover trials, *Statistical Methods in Medical Research* **3**: 383–405.

Matthews, J. (1994b). Modelling and optimality in the design of crossover studies for medical applications. *Journal of Statistical Planning and Inference* **42**: 89–108.

Mitchell, T. (1974). Computer construction of $\mathcal{D}$-optimal first-order designs, *Technometrics* **16**: 211–220.

Searle, S. R. (1971). *Linear Models*, John Wiley & Sons, Inc., New York.

Tack, L. and Vandebroek, M. (2001). $(\mathcal{D}_t, C)$-optimal run orders, *Journal of Statistical Planning and Inference* **98**: 293–310.

Tack, L. and Vandebroek, M. (2002). Trend-resistant and cost-efficient block designs with fixed or random block effects, *to appear in Journal of Quality Technology.*

Verbeke, G. and Molenberghs, G. (1997). *Linear Mixed Models in Practice: A SAS-Oriented Approach*, New York: Springer-Verlag.

Vonesh, E. and Chinchilli, V. (1997). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, New York: Marcel Dekker.

Wynn, H. (1972). Results in the theory and construction of $\mathcal{D}$-optimum experimental designs, *Journal of the Royal Statistical Society, Series B,* **34**: 133–147.