Available online at www.sciencedirect.com

ScienceDirect

Computational Statistics & Data Analysis 51 (2007) 3235 – 3243

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Robust balanced measurement designs when errors are serially correlated

Chen-Tuo Liao[a], Tsai-Yu Lin[b,*]

[a]*Division of Biometry, Institute of Agronomy, National Taiwan University, Taipei 10764, Taiwan*
[b]*Department of Applied Mathematics, Feng Chia University, Taichung 40724, Taiwan*

Received 4 November 2005; received in revised form 16 August 2006; accepted 21 November 2006
Available online 14 December 2006

Abstract

This paper considers the situation in which some characteristic is to be measured on each of several specimens. For instance, it may be the concentration of lead or arsenic in water or soil samples and a laboratory may routinely analyze samples from different sources. In the measurement process, there may be some serial correlation among measurement errors, but it is hard to detect or to have a reliable estimation for this existing phenomenon. Therefore, it may be desired to make statistical inference on the true values of unknown specimens without estimating this possible correlation. To help adjust the instrument readings in a process, standards are frequently interspersed among unknown specimens at appropriate intervals. A systematic method of arranging the order of the measurements of unknown specimens and standards is provided. One is able to avoid the difficulty of estimating the possible correlation and still has good estimates of the parameters of interest using the proposed measurement designs. In addition, a simulation study is carried out to evaluate the sensitivity of the measurement designs, showing that they are robust to the existence of various error processes.

Keywords: Calibration problem; Systematic error; Random error; Design criterion

1. Introduction

The calibration of measurement procedures is commonly required in many areas. Typical application situations include industrial processes where product quality needs to be routinely monitored and measurement laboratories where customer specimen are measured on a daily basis. A measurement process is subject to errors which may be generally classified as random errors only or a combination of both random errors and systematic errors. Here *random errors* are defined to have a zero expected value and *systematic errors* are defined to be due to biases in the measurement process. In a typical measurement process, *standards*, which have known true values traceable to a national standards laboratory (e.g. NIST in the United States), are frequently used to monitor the errors. In other words, one is capable of observing the errors whenever a standard is measured. Furthermore, in common practice, the random errors are usually assumed to be independent random variables, but it is often more realistic to acknowledge that the measurement process is serially correlated. However, it may be hard to detect this kind of serial correlation when it is weak or to have a good estimate for it when the number of standards measured is small. The main interest of this study is to develop a

* Corresponding author. Tel.: +88 64 24517250x5132; fax: +88 64 24510801.
E-mail address: linty@fcu.edu.tw (T.Y. Lin).

0167-9473/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.11.032

systematic approach to arranging the order of measurements so that one is able to make accurate statistical inference on the unknowns even without reliable estimates of the serial correlation.

The calibration problems for estimation procedures have been extensively studied. The reader is referred to the textbooks of Fuller (1987) and Brown (1993). But, the issues related to the order of measurements in a calibration process are rarely discussed in the literature. Earlier literature pertaining to the *measurement design*, particularly regarding the arrangement of the order of measurements, can be found in Pepper (1973) and Perng and Tong (1977). More recently, Liao et al. (2000) consider A-optimal balanced measurement designs for an additive model under the assumption that random errors arise from a first-order autoregressive process (AR(1)).

Zhou (2001) presents a design criterion for evaluation of the robustness to possible correlation among observations in general experiments. In the same paper, the criterion has been only proven successful in estimating the slope of the simple linear regression by a simulation study. Moreover, Zhou (2001) discusses the construction of robust run order for two-level factorial designs based on the criterion, but the robustness property of the obtained designs has not been thoroughly investigated. However, her work motivates us to explore applicability of the design criterion to measurement processes.

The rest of the article is organized as follows. Section 2 first introduces the design criterion presented by Zhou (2001). Then the problem of interest in this study is formulated based on this criterion. Section 3 develops an exhaustive search method for the robust balanced measurement designs. Some practical designs are also reported. Section 4 includes a simulation study for investigating robustness of the obtained designs to various autocorrelation structures. Concluding remarks are presented in Section 5.

2. Design criterion and the problem of interest

Zhou (2001) proposes an experimental design criterion for evaluation of the robustness to possible correlation among the observations. Suppose the design model of interest is assumed to be

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\beta}$ is the unknown parameters vector whose OLSE (ordinary least squares estimator) is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$. The random errors $\boldsymbol{\epsilon}$ are serially correlated with covariance matrix $Var(\boldsymbol{y}) = \sigma^2\boldsymbol{P}$ for some correlation matrix $\boldsymbol{P}$. Without knowing $\boldsymbol{P}$, one may intend to use the OLSE for $\boldsymbol{\beta}$. Let d be the design used. The proposed criterion is based on the *change of variance function*, abbreviated as CVF, given by

$$\mathrm{CVF}_{\boldsymbol{a}}(d, \boldsymbol{P}) = \frac{\boldsymbol{a}'[V(\boldsymbol{P}) - V(\boldsymbol{I})]\boldsymbol{a}}{\boldsymbol{a}'V(\boldsymbol{I})\boldsymbol{a}}, \tag{2.1}$$

where $\boldsymbol{a}'\boldsymbol{\beta}$ is the parameter of interest; $V(\boldsymbol{P})$ is the covariance matrix of the $\hat{\boldsymbol{\beta}}$ under the assumption that $Var(\boldsymbol{y}) = \sigma^2\boldsymbol{P}$. That is,

$$Var(\hat{\boldsymbol{\beta}}) = V(\boldsymbol{P}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}. \tag{2.2}$$

Moreover, $V(\boldsymbol{I}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$ is the covariance matrix of $\hat{\boldsymbol{\beta}}$ assuming homogeneous variances. Then the robust run order design with respect to $\boldsymbol{P}$, denoted by d^*, is defined as the design minimizing $|\mathrm{CVF}_{\boldsymbol{a}}(d, \boldsymbol{P})|$, the absolute value of the CVF, among all possible competing designs. When multiple parameters are of interest, say $\boldsymbol{a}'_1\boldsymbol{\beta}, \boldsymbol{a}'_2\boldsymbol{\beta}, \ldots, \boldsymbol{a}'_k\boldsymbol{\beta}$, the criterion is modified as

$$d^* = \min_d \sum_{i=1}^{k} |\mathrm{CVF}_{\boldsymbol{a}_i}(d, \boldsymbol{P})|. \tag{2.3}$$

Zhou (2001) also suggests that it may be reasonable to assume MA(1), a first-order moving average, error process in construction of the robust designs for practical use.

In this study, we are interested in determining robust measurement designs when the response variable obeys the following model

$$z_i = \mu + \delta_{i,0}\tau_0 + \sum_{j=1}^{m} \delta_{i,j}\tau_j + \varepsilon_i, \tag{2.4}$$

where $i = 1, 2, \ldots, N$ and N denotes the total number of measurements; μ is the systematic error; τ_0 is the known true value of the standard; $\tau_1, \tau_2, \ldots, \tau_m$ are the true values of the m unknowns; ε_i are random errors; the indicator $\delta_{i,0}$ is 1 if the ith measurement is the standard, and 0 otherwise; $\delta_{i,j}$ is 1 if the ith measurement is unknown j, and 0 otherwise. Since τ_0 is known, let

$$y_i = \begin{cases} z_i - \tau_0 & \text{if observation } i \text{ is of the standard,} \\ z_i & \text{otherwise.} \end{cases}$$

Hence, model (2.4) can be rewritten as $y_i = \mu + \sum_{j=1}^{m} \delta_{i,j}\tau_j + \varepsilon_i$ or in matrix presentation

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{2.5}$$

where $\boldsymbol{X}$ is of size $N \times (1 + m)$ defined by

$$x_{i,k} = \begin{cases} 1 & \text{if } k = 1, \\ \delta_{i,k-1} & \text{if } k > 1, \end{cases}$$

and $\boldsymbol{\beta} = [\mu, \boldsymbol{\tau}']'$ where $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_m]'$. The main objective of this study is to search for a robust design d^* satisfying (2.3) for $\boldsymbol{a}_i'\boldsymbol{\beta} = \tau_i$, $i = 1, 2, \ldots, m$.

3. Robust balanced measurement designs

To construct a robust measurement design, we also consider that the random errors ε_i of (2.5) arise from an MA(1) process. Namely, the random errors vector $\boldsymbol{\epsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]'$ has a distribution with a zero mean vector and a covariance matrix $\sigma^2\boldsymbol{P}$, where $\boldsymbol{P} = \{p_{i,j}\}$, for $i, j = 1, 2, \ldots, N$, is a positive definite matrix with elements $p_{i,i} = 1$, $p_{i,i-1} = p_{i-1,i} = \rho = \phi/(1 + \phi^2)$ and $p_{i,j} = 0$ otherwise. Here $-1 < \phi < 1$.

Before computing the corresponding CVF under our setting, some parameters related to the measurement design are introduced below.

3.1. Design parameters

A measurement design will be described in terms of "batches". A *batch* is defined to be (i) the set of all measurements of unknowns between two successive measurements of the standard; or (ii) the set of all measurements (if any) of unknowns before the first measurement of a standard; or (iii) the set of all measurements (if any) of unknowns after the last measurement of a standard. Empty batches are permissible. As a consequence, $b - 1$ observations of a standard give rise to b batches. Moreover, let $T = \sum_{i=1}^{m} t_i$, where t_i denotes the number of unknown i and $b - 1$ denotes the number of the standard. Thus, the total number of measurements $N = T + (b - 1)$. The following definitions are related to the arrangement of measurements.

- b_0 — the number of empty batches.
- e_i — the number of times unknown i is the first observation plus the number of times unknown i is the last observation. $e_i = \delta_{1,i} + \delta_{N,i}$ and $\sum_{i=1}^{m} e_i = 0$, 1 or 2.
- q_i — the number of times two consecutive observations are of the unknown i. $q_i = \sum_{k=1}^{N-1} \delta_{k,i}\delta_{k+1,i}$.
- $r_{i,j}$ — the number of times unknowns i and j occur as neighboring pairs in the design.

The following example is given to illustrate the above definitions.

Example 3.1. Consider the following sequence of the observations.

$$U_1 U_3 S\, U_1 U_3 U_3 U_3 U_2 U_1 U_2 U_2 S\, S,$$

where U_i denotes unknown i and S denotes the standard. In this case, there are $b=4$ batches. Of these four batches $b_0=2$ are empty. And $t_1=3$, $t_2=3$, $t_3=4$, so $N=T+(b-1)=10+3=13$. Also

$e_1=1$ (the first observation is of U_1).
$e_2=0$ (neither the first nor the last observation is of U_2).
$e_3=0$ (neither the first nor the last observation is of U_3).
$q_1=0$ (no two consecutive observations are of U_1).
$q_2=1$ (due to observations (10 and 11)).
$q_3=2$ (due to the pairs of observations (5 and 6) and (6 and 7)).
$r_{1,2}=2$ (due to the pairs of observations (8 and 9) and (9 and 10)).
$r_{1,3}=2$ (due to the pairs of observations (1 and 2) and (4 and 5)).
$r_{2,3}=1$ (due to observations (7 and 8)).

3.2. Calculating the CVF

From (2.2), based on the MA(1) error process, the covariance matrix of $\hat{\boldsymbol{\tau}}$, denoted by $V_{\boldsymbol{P}}(\hat{\boldsymbol{\tau}})=\sigma^2\{\theta_{i,j}\}$, for $i, j = 1, 2, \dots, m$, is given by

$$\theta_{i,j}=\begin{cases}(h_{i,i}+g)\rho+\dfrac{1}{N-T}+\frac{1}{t_i} & \text{if } i=j,\\[2ex] (h_{i,j}+g)\rho-\dfrac{1}{N-T} & \text{if } i\neq j,\end{cases}$$

where

$$h_{i,i}=\frac{2}{(N-T)t_i}\left(\sum_{k\neq i}^{m} r_{i,k}+2q_i+e_i\right)+\frac{2q_i}{t_i^2},$$

$$g=\frac{2}{(N-T)^2}\left(\sum_{k=1}^{m}\sum_{l\neq k}^{m} r_{k,l}+\sum_{k=1}^{m} q_k+\sum_{k=1}^{m} e_k-N-1\right),$$

$$h_{i,j}=\frac{1}{(N-T)t_i}\left(\sum_{k\neq i}^{m} r_{i,k}+2q_i+e_i\right)+\frac{1}{(N-T)t_j}\left(\sum_{k\neq j}^{m} r_{j,k}+2q_j+e_j\right)+\frac{r_{i,j}}{t_i t_j}.$$

Similarly, let $V_{\boldsymbol{I}}(\hat{\boldsymbol{\tau}})=\sigma^2\{\lambda_{i,j}\}$, for $i, j=1, 2, \dots, m$, denoting the covariance matrix of $\hat{\boldsymbol{\tau}}$ under the assumption of homogeneous variance. Then we have

$$\lambda_{i,j}=\begin{cases}\dfrac{1}{N-T}+\dfrac{1}{t_i} & \text{if } i=j,\\[2ex] \dfrac{-1}{N-T} & \text{if } i\neq j.\end{cases}$$

Hence, the CVF of (2.1) associated with τ_i can be equivalently expressed as

$$\mathrm{CVF}_{\tau_i}(d,\boldsymbol{P})=\frac{\theta_{i,i}-\lambda_{i,i}}{\lambda_{i,i}}. \tag{3.1}$$

It may be inefficient and costly to search for robust measurement designs through a complete enumeration, even for moderate values of design parameters. Therefore, we primarily consider the situation where the parameters $\tau_1, \tau_2, \dots, \tau_m$ are *equally important*, and hence we need all $\mathrm{CVF}_{\tau_i}(d,\boldsymbol{P})$ of (3.1) to be equal. The following lemma describes conditions that must be satisfied by the design parameters e_i, q_i, t_i and $r_{i,j}$ in order for a measurement design to have equal $\mathrm{CVF}_{\tau_i}(d,\boldsymbol{P})$ for every unknown. A measurement design is said to be *balanced* hereinafter provided it has the same $\mathrm{CVF}_{\tau_i}(d,\boldsymbol{P})$ for $i=1, 2, \dots, m$.

Lemma 3.1. *A measurement design is* balanced *for all values of* ρ, *where* $\rho = \phi/(1+\phi^2)$ *and* $-1 < \phi < 1$, *if and only if, for all i and j,* $t_i = t$, $q_i = q$, $e_i = e$ *and* $r_{i,j} = r$. *Also the* CVF *in this case is given by*

$$\mathrm{CVF}(d, \boldsymbol{P}) = \frac{\rho\left\{[m(m-1)(2q-r) - 2(N+1)]t^2 - 2N[(m-1)(2q-r) - e]t + 2N^2 q\right\}}{t(N-mt)[N-(m-1)t]}. \tag{3.2}$$

The proof involves straightforward algebra and is omitted.

Consequently, the problem of searching for d^* of (2.3) is reduced to finding the combinations of the design parameters which minimize the absolute value of (3.2) for given *N* and *m*. Importantly, the search for the robust design is free of the autocorrelation parameter ρ. We present the steps of exhaustive search for the robust balanced measurement designs in the following.

3.3. Searching for the desired designs

Proposition 3.1 gives some constraints on the design parameters, which are direct consequences of the definitions.

Proposition 3.1. *For fixed values of N and m, the following describes some interrelationships among the design parameters.*

(i) $r = 0$ *for* $m = 1$.
(ii) $\binom{m}{2}r + mq = mt - (b - b_0)$.
(iii) $0 \leqslant e \leqslant \max\{0, 3-m\}$.
(iv) $0 \leqslant q \leqslant t-1$.
(v) $\max\{2-me, b-mt\} + mt - b \leqslant mq + \binom{m}{2}r \leqslant mt - 1$.

It is now possible to enumerate all possible combinations of the design parameters satisfying the above constraints in a practical range of N and m. These combinations are called eligible designs hereinafter. The following steps are implemented to search for the robust balanced measurement designs.

Step 1: *Input N and m.*
Step 2: *Calculate* $b = (N+1) - mt$ *for* $1 \leqslant t \leqslant \lfloor (N-1)/m \rfloor$, *where* $\lfloor a \rfloor$ *is the largest integer less than or equal to a.*
Step 3: *For each combination of b and t generated from Step* 2, *search all possible combinations of e, q and r satisfying* (i), (iii), (iv) *and* (v) *of Proposition* 3.1.
Step 4: *For each combination of* (b, t, e, q, r) *generated from Step* 3, *calculate* $b_0 = \binom{m}{2}r + mq - mt + b$. *This is according to* (ii) *of Proposition* 3.1.
Step 5: *Compute the absolute value of* (3.2) *with* $\rho = 1$ *for each eligible design and select the robust design* d^* *which has the minimum one.*

Some robust balanced measurement designs are reported in Table 1. It is worth noting that the resulting designs may not be unique for a set of given *N* and *m*. In practice, one may determine the final design according to the measurement cost of unknowns and standard. More interestingly, it is not uncommon to obtain *ideal designs*, i.e. those which have zero $\mathrm{CVF}(d, \boldsymbol{P})$ in (3.2).

3.4. Arranging the order of measurements

For fixed values of *N* and *m* and a given robust balanced measurement design with parameters (t, b, e, q, r, b_0), we would directly use the algorithm presented in Liao et al. (2000) to arrange the order of measurements. Their algorithm is mainly based on a special class of Latin squares which is considered by Kiefer and Wynn (1981) and Cheng (1983) in the study of *equineighbored designs*. Structurewise, a difference between these two classes of designs is that the measurement designs usually turn out to be non-binary and unequal block-size designs if considering the "batches", defined in Section 3.1, as the classical blocks.

Liao et al. (2000) provide detailed construction steps separately for three cases: (A) $r = 0$, (B) $r > 0$ with $(m-1)r$ being even and (C) $r > 0$ with $(m-1)r$ being odd. Even the algorithm is considered for $m \geqslant 3$, it can easily be modified

Table 1
Some robust balanced measurement designs

Design	m	N	NED	(t, b, e, q, r, b_0)
$d_{11,1}$–$d_{11,2}$	1	10	74	(3,8,0,1,0,6) (7,4,2,4,0,1)
$d_{12,1}$–$d_{12,2}$		20	299	(1,20,1,0,0,19) (19,2,1,18,0,1)
$d_{13,1}$–$d_{13,2}$		50	1874	(7,44,0,1,0,38) (43,8,2,36,0,1)
$d_{14,1}$–$d_{14,2}$		100	7499	(33,68,0,11,0,46) (67,34,2,44,0,11)
$d_{21,1}$[a]	2	10	50	(4,3,0,2,3,2)
$d_{22,1}$[a]–$d_{22,7}$[a]		20	375	(4,13,0,1,1,8) (4,13,1,0,4,9) (5,11,0,0,7,8) (6,9,0,0,6,9) (8,5,0,6,1,2) (8,5,0,0,14,3) (8,5,1,2,8,1)
$d_{23,1}$[a]–$d_{23,10}$[a]		50	5500	(10,31,0,3,0,17) (10,31,1,2,3,18) (16,19,0,0,24,11) (20,11,0,12,8,3) (20,11,0,6,21,4) (20,11,0,0,34,5) (20,11,1,14,2,1) (20,11,1,8,15,2) (20,11,1,2,28,3) (24,3,0,12,23,2)
$d_{24,1}$[a]–$d_{24,26}$[a]		100	42875	(10,81,1,1,1,64) (20,61,0,5,4,35) (20,61,0,1,21,44) (20,61,1,4,7,36) (20,61,1,0,24,45) (24,53,0,6,11,28) (25,51,0,8,7,24) (25,51,0,5,17,28) (25,51,0,2,27,32) (25,51,1,7,9,24) (25,51,1,4,19,28) (25,51,1,1,29,32) (30,41,0,3,35,22) (30,41,1,9,17,16) (32,37,0,16,6,11) (40,21,0,26,11,4) (40,21,0,20,24,5) (40,21,0,14,37,6) (40,21,0,8,50,7) (40,21,0,2,63,8) (40,21,1,28,5,2) (40,21,1,22,18,3) (40,21,1,16,31,4) (40,21,1,10,44,5) (40,21,1,4,57,6)(48,5,0,36,21,2)
$d_{31,1}$[a]–$d_{31,2}$[a]	3	20	22	(4,9,0,0,3,6) (5,6,0,1,3,3)
$d_{32,1}$[a]–$d_{32,2}$[a]		50	372	(5,36,0,0,3,30) (10,21,0,4,1,6)
$d_{33,1}$[a]–$d_{33,7}$[a]		100	3128	(20,41,0,6,5,14) (25,26,0,16,1,2) (25,26,0,11,7,5) (25,26,0,6,13,8) (25,26,0,1,19,11) (30,11,0,18,9,2) (32,5,0,16,15,2)
$d_{41,1}$	5	20	2	(1,16,0,0,0,11)
$d_{42,1}$		50	40	(7,16,0,1,2,6)
$d_{43,1}$		100	330	(13,36,0,1,4,16)
$d_{51,1}$	10	50	4	(1,41,0,0,0,31)
$d_{52,1}$		100	29	(1,91,0,0,0,81)

In the table, the abbreviation NED denotes the number of eligible designs for the settings.
[a]Indicates the ideal designs whose CVF values are equal to zero.

for the simpler situations $m = 1$ and 2. Note that when $m \geqslant 3$, $e = 0$, but e can be non-zero for $m = 1$ and 2. In the following, we just present the measurement sequences for some designs of Table 1 obtained from the algorithm. The interested reader may consult Liao et al. (2000) for details regarding the process.

Example 3.2. For design $d_{11,2}$, $m = 1$, $N = 10$, $t = 7$, $b = 4$, $e = 2$, $q = 4$, $r = 0$ and $b_0 = 1$. One measurement order obtained from Case (A) is given by

$$UUUSSUUSUU.$$

Example 3.3. For design $d_{21,1}$, $m = 2$, $N = 10$, $t = 4$, $b = 3$, $e = 0$, $q = 2$, $r = 3$ and $b_0 = 2$. From Case (C), one required sequence is given by

$$S \quad U_2 \quad U_2 \quad U_2 \quad U_1 \quad U_2 \quad U_1 \quad U_1 \quad U_1 \quad S.$$

Example 3.4. For design $d_{31,2}$, $m = 3$, $N = 20$, $t = 5$, $b = 6$, $e = 0$, $q = 1$, $r = 3$ and $b_0 = 3$. A measurement order satisfying these parameters obtained from Case (B) is given by

$$S\,U_3U_1U_3U_1U_1SS\,U_1U_2U_1U_2U_2S\,U_2U_3U_2U_3U_3\,S.$$

Example 3.5. For design $d_{42,1}$, $m=5$, $N=50$, $t=7$, $b=16$, $e=0$, $q=1$, $r=2$ and $b_0=6$. A desired sequence of measurements for this design is given by

$$\begin{array}{cccccccccccccccccccc} S & U_5 & U_1 & U_4 & U_1 & S & U_1 & U_5 & U_5 & S & S & U_1 & U_2 & U_5 & U_2 & S & U_2 & U_1 & U_1 & S \\ S & U_4 & U_5 & U_3 & U_5 & S & U_5 & U_4 & U_4 & S & S & U_2 & U_3 & U_1 & U_3 & S & U_3 & U_2 & U_2 & S \\ S & U_3 & U_4 & U_2 & U_4 & S & U_4 & U_3 & U_3 & S. \end{array}$$

Example 3.6. For design $d_{51,1}$, $m=10$, $N=50$, $t=1$, $b=41$, $e=0$, $q=0$, $r=0$ and $b_0=31$. A robust run order of the measurements satisfying these design parameters is given by

$$\begin{array}{cccccccccccccccccccc} S & S & S & S & U_1 & S & S & S & U_2 & S & S & S & S & U_3 & S & S & S & U_4 & S & S \\ S & S & U_5 & S & S & S & U_6 & S & S & S & S & U_7 & S & S & S & U_8 & S & S & S & S \\ U_9 & S & S & S & S & U_{10} & S & S & S & S. \end{array}$$

4. Simulation study

To evaluate performance of the obtained robust balanced measurement designs, the following simulation study is conducted by constructing a $100(1-\alpha)\%$ confidence interval for τ_i. Recall that the motivation of the robustness study is to avoid the difficulty of obtaining a reliable estimate for possible correlation among the measurements, and still able to make statistical inference on the unknown parameters under the usual assumption of the homogeneity of variance. Therefore, we calculate a $100(1-\alpha)\%$ using (4.1) for each simulated data set:

$$\hat{\tau}_i \pm t_{\alpha/2,N-m-1}\hat{\sigma}\sqrt{\frac{1}{N-mt}+\frac{1}{t}}, \tag{4.1}$$

where $\hat{\tau}_i$ is the OLSE of τ_i, $t_{\alpha/2,N-m-1}$ is the $100(1-\alpha/2)$th percentile of the Student's t distribution with $N-m-1$ degrees of freedom and $\hat{\sigma}$ is the square root of MSE (mean square error).

The robust designs obtained in the present study are derived by assuming an MA(1) error process. Hence, we first evaluate the robustness to this error process. The simulated coverage probability for 90% confidence interval of (4.1) is computed for τ_1 using 10,000 simulated data sets for each simulation setting. The results of some designs listed in Table 1 along with a "bad" design are displayed in Table 2. The bad design whose design parameters $(t, b, e, q, r, b_0) = (3, 12, 2, 9, 9, 0)$ has the largest absolute value of CVF among all eligible designs with $m=3$ and $N=20$.

For most of the settings, the robust balanced measurements are successful in maintaining the simulated coverage probabilities close to the nominal level 0.90. By contrast, the bad design $d_{31,b}$, compared with $d_{31,2}$, cannot have satisfactory performance.

Table 2
The simulated coverage probability for an MA(1) error process with $\varepsilon_i = \omega_i + \phi\omega_{i-1}$

Design	ρ_1						
	0.4	0.3	0.2	0	−0.2	−0.3	−0.4
$d_{11,2}$	0.8747	0.8928	0.8978	0.9057	0.9086	0.9055	0.9127
$d_{13,2}$	0.8981	0.8989	0.9022	0.8983	0.9042	0.9004	0.9089
$d_{21,1}$	0.8873	0.8847	0.8922	0.9040	0.9085	0.9128	0.9172
$d_{24,26}$	0.9022	0.8985	0.9015	0.8954	0.8972	0.9040	0.9032
$d_{31,2}$	0.8891	0.8955	0.8965	0.8966	0.9001	0.9053	0.9026
$d_{33,7}$	0.9013	0.9002	0.8952	0.8973	0.9016	0.9031	0.9029
$d_{41,1}$	0.9071	0.9081	0.9030	0.9017	0.8939	0.9030	0.8965
$d_{43,1}$	0.8995	0.9023	0.8952	0.8989	0.9031	0.9005	0.8976
$d_{52,1}$	0.9001	0.9015	0.9019	0.8996	0.9067	0.9036	0.8970
$d_{31,b}$ [a]	0.8013	0.8339	0.8522	0.8908	0.9355	0.9560	0.9715

The first autocorrelation $\rho_1 = \phi/(1+\phi^2)$.
[a] $d_{31,b}$ denotes the "bad" design in the class of designs with $m=3$ and $N=20$.

Table 3
The simulated coverage probability for an AR(1) error process with $\varepsilon_i = \omega_i + \psi\varepsilon_{i-1}$

Design	ρ_1						
	0.9	0.5	0.3	0	−0.3	−0.5	−0.9
$d_{11,2}$	0.9293	0.9074	0.8931	0.8995	0.9190	0.9382	0.9820
$d_{13,2}$	0.9823	0.9178	0.9023	0.9023	0.9057	0.9114	0.9773
$d_{21,1}$	0.8351	0.8725	0.8895	0.8970	0.9121	0.9202	0.9555
$d_{24,26}$	0.8464	0.9041	0.9006	0.9006	0.9082	0.9093	0.9209
$d_{31,2}$	0.9275	0.9049	0.8975	0.8966	0.9111	0.9233	0.9258
$d_{33,7}$	0.7799	0.8907	0.8987	0.8942	0.9040	0.9008	0.8630
$d_{41,1}$	0.9381	0.9055	0.9034	0.9019	0.9014	0.9015	0.9423
$d_{43,1}$	0.9134	0.9188	0.9043	0.8985	0.9037	0.9112	0.9774
$d_{52,1}$	0.9024	0.8967	0.8984	0.8988	0.9021	0.9003	0.9177
$d_{31,b}$[a]	0.7092	0.7865	0.8306	0.9029	0.9577	0.9763	0.9958

The first autocorrelation $\rho_1 = \psi$.
[a] $d_{31,b}$ denotes the "bad" design in the class of designs with $m = 3$ and $N = 20$.

Table 4
The simulated coverage probability for an MA(2) error process with $\varepsilon_i = \omega_i + \phi_1\omega_{i-1} + \phi_2\omega_{i-2}$

Design	(ϕ_1, ϕ_2)					
	(0.6,0.3) [0.54,0.21]	(−0.6,−0.3) [−0.29−0.21]	(−0.6,0.3) [−0.54,0.21]	(0.5,0.1) [0.44,0.08]	(−0.5,−0.1) [−0.36,−0.08]	(−0.5,0.1) [−0.44,0.08]
$d_{11,2}$	0.9182	0.8858	0.9410	0.8931	0.9031	0.9214
$d_{13,2}$	0.9040	0.8913	0.9149	0.9019	0.8982	0.9140
$d_{21,1}$	0.8655	0.9189	0.9109	0.8790	0.9197	0.9154
$d_{24,26}$	0.8891	0.9035	0.8952	0.8906	0.9026	0.8969
$d_{31,2}$	0.9046	0.8951	0.9155	0.8971	0.9070	0.9096
$d_{33,7}$	0.8930	0.9059	0.8922	0.8953	0.9014	0.9036
$d_{41,1}$	0.9135	0.8938	0.9025	0.9090	0.8961	0.8979
$d_{43,1}$	0.9149	0.8944	0.9139	0.9028	0.9017	0.9073
$d_{52,1}$	0.9032	0.9016	0.8973	0.9011	0.8970	0.9004
$d_{31,b}$[a]	0.7511	0.9621	0.9866	0.7907	0.9692	0.9770

The first and second autocorrelations $\rho_1 = (\phi_1 + \phi_1\phi_2)/(1 + \phi_1^2 + \phi_2^2)$ and $\rho_2 = \phi_2/(1 + \phi_1^2 + \phi_2^2)$ are displayed in the square brackets.
[a] $d_{31,b}$ denotes the "bad" design in the class of designs with $m = 3$ and $N = 20$.

Table 5
The simulated coverage probability for an AR(2) error process with $\varepsilon_i = \omega_i + \psi_1\varepsilon_{i-1} + \psi_2\varepsilon_{i-2}$

Design	(ϕ_1, ϕ_2)					
	(0.6,0.3) [0.86,0.81]	(−0.6,−0.3) [−0.46,−0.02]	(−0.6,0.3) [−0.86,0.81]	(0.5,0.1) [0.56,0.38]	(−0.5,−0.1) [−0.45,0.13]	(−0.5,0.1) [−0.56,0.38]
$d_{11,2}$	0.9352	0.9099	0.9773	0.9156	0.9243	0.9545
$d_{13,2}$	0.9786	0.9083	0.9799	0.9339	0.9063	0.9218
$d_{21,1}$	0.8527	0.9305	0.9461	0.8674	0.9237	0.9256
$d_{24,26}$	0.8419	0.9089	0.9162	0.8935	0.9091	0.9116
$d_{31,2}$	0.9203	0.9077	0.9093	0.9100	0.9135	0.9200
$d_{33,7}$	0.7574	0.9181	0.8598	0.8875	0.9088	0.8938
$d_{41,1}$	0.9215	0.8945	0.9388	0.9135	0.8926	0.9017
$d_{43,1}$	0.9068	0.9061	0.9801	0.9275	0.9029	0.9199
$d_{52,1}$	0.9033	0.9029	0.9232	0.9014	0.9035	0.8996
$d_{31,b}$[a]	0.7473	0.9875	0.9836	0.7842	0.9786	0.9750

The first and second autocorrelations $\rho_1 = \psi_1/(1 - \psi_2)$ and $\rho_2 = \psi_1\rho_1 + \psi_2$ are displayed in the square brackets.
[a] $d_{31,b}$ denotes the "bad" design in the class of designs with $m = 3$ and $N = 20$.

To check sensitivity of the obtained robust balanced measurement designs to other error processes than MA(1), the same simulation study is repeated for AR(1), AR(2) and MA(2) error processes. The results are displayed in Tables 3–5.

From Tables 3–5, almost all the obtained robust balanced measurement designs, regardless of the number of unknowns and the total number of measurements, are robust against the various error processes provided that the absolute value of the first or the second autocorrelation is not quite large. For example, from Table 3, some of designs cannot have satisfactory performance when the first autocorrelation is fixed at 0.9 or -0.9, but all the designs perform quite well for the remaining first autocorrelation values. Similar results can be found in Table 5. Notice that the comparison results between $d_{31,2}$ and $d_{31,b}$ in the tables indicate the importance of design choice for a measurement process.

5. Concluding remarks

In this article, the approach we have taken is to ignore the correlation structure of errors and use the ordinary least squares estimators for the unknowns. To compensate for this, we choose an appropriate design under which the variance of the estimates under the working model are close to that under the correct model. Also, the final judgment of the goodness of the chosen designs is according to the actual coverage of the confidence intervals for the unknowns. On the other hand, it might be of interest to consider a criterion for the robust measurement designs that directly based on such coverage probabilities along with the expected lengths of confidence intervals.

We have proposed a systematic method of constructing the robust balanced measurement designs based on the assumption of the additive model with an MA(1) error structure. The resulting designs are shown to be only minimally impacted by the error structure and are robust to various assumed structures, indicating that they can be recommended for practical use. Nonetheless, in practice, it is not uncommon to encounter more complicated calibration models, such as nonlinear and multiplicative models. For example see Lin and Liao (2005). We will investigate a possible extension to these calibration models in a future study.

Acknowledgments

The authors thank an associate editor and a referee for their useful comments that result in a much improved article. Thanks are also due to Professor Hari K. Iyer of Colorado State University, USA for his helpful discussions on the issues of measurement design. This work is partially supported by National Science Council of ROC via contract NSC-92-2118-M-002-007.

References

Brown, P.J., 1993. Measurement Regression and Calibration. Oxford University Press, New York.
Cheng, C.S., 1983. Construction of optimal balanced incomplete block designs for correlated observations. Ann. Statist. 11, 204–209.
Fuller, W.A., 1987. Measurement Error Models. Wiley, New York.
Kiefer, J., Wynn, H.P., 1981. Optimal balanced block and latin square designs for correlated observations. Ann. Statist. 9, 737–757.
Liao, C.T., Taylor, C.H., Iyer, H.K., 2000. Optimal balanced measurement designs when errors are correlated. J. Statist. Plann. Inference 84, 295–321.
Lin, T.Y., Liao, C.T., 2005. Optimal allocation of measurements in a linear calibration process. Metrika 61, 157–168.
Pepper, M.P.G., 1973. A calibration of instruments with non-random errors. Technometrics 15, 587–599.
Perng, S.K., Tong, Y.L., 1977. Optimal allocation of observations in inverse linear regression. Ann. Statist. 5, 191–196.
Zhou, J., 2001. A robust criterion for experimental designs for serially correlated observations. Technometrics 43, 462–467.