# ON THE COSTWISE OPTIMALITY OF HIERARCHICAL MULTIRESPONSE RANDOMIZED BLOCK DESIGNS UNDER THE TRACE CRITERION*

J. N. SRIVASTAVA AND L. L. MCDONALD

(Received July 16, 1969)

## 1. Summary

Consider the class of general incomplete multiresponse (GIM) designs in which the set of units is divided into blocks of equal size, such that in any block the same subset of responses is measured on each unit. It is shown that with respect to the trace criterion and a reasonable cost restriction, the subclass of hierarchical multiresponse (HM) designs is complete in the sense that given any GIM design, there exists an HM design such that the cost involved under the two designs is the same, but the trace of the covariance matrix of the estimates of the parameters under the HM design is less than or equal to the similar quantity under the GIM design. Our results also establish the important fact that there is a large class of situations where the standard multiresponse model (under which all responses are measured on each unit) should not be used. Furthermore, the nonlinear programming problem associated with obtaining the optimum HM design is stated and solved.

## 2. Introduction and preliminaries

Responsewise incomplete multiresponse experiments are needed when it is either physically impossible, uneconomic, or otherwise inadvisable to study all responses or characteristics on each experimental unit. For a review of the literature on incomplete multiresponse designs, the interested reader is referred to the illustrative (though necessarily inexhaustive) list of references at the end.

Consider a multiresponse experiment with $p$ responses (say $V_1, \cdots, V_p$), $v$ treatments (say $\tau_1, \cdots, \tau_v$) and a set $S^*$ of experimental units. Suppose further that $S^*$ is divided into a set $S$ of blocks such that each block is of size $v$ and such that each block contains a set of units homo-

* This research was wholly supported by U.S.A.F. (Aerospace Research Laboratories) Contract No. F33615-67-C-1436.

genous with respect to each response. For any subset of $k$ responses (say $V_{i_1}, \cdots, V_{i_k}$), $1 \leqq k \leqq p$, let $S(i_1, \cdots, i_k)$ denote the subset (possibly empty) of $S$, such that for any block $B \in S(i_1, \cdots, i_k)$, and any unit $U \in B$, the set of responses measured on $U$ is exactly $(V_{i_1}, \cdots, V_{i_k})$. Such a design may be called a general incomplete multiresponse randomized block (GIM(RB)) design. Let $U_r = \{S(i_1, \cdots, i_k) \mid r \in (i_1, \cdots, i_k)\}$ denote the totality of all blocks such that response $V_r$ is measured on each unit of each such block. If there exists a permutation $(r_1, \cdots, r_p)$ of $(1, 2, \cdots, p)$ such that $U_{r_1} \supseteq U_{r_2} \supseteq \cdots \supseteq U_{r_p}$, then the above GIM(RB) design will be called a hierarchical multiresponse randomized block (HM(RB)) design.

Let $\tau_{jr}$ ($j=1, \cdots, v$; $r=1, \cdots, p$) denote the "true effect" of the treatment $\tau_j$ for the response $V_r$. Suppose $U_r$ has $n_r$ blocks in it. Consider $\bar{y}_{jr}$, the mean of the $n_r$ "observed yields" of the $j$th treatment from these $n_r$ blocks. Assume all the units of $S^*$ to be independent. Also let $\sigma_{rr}$ denote the variance of an observation on the $r$th response $V_r$ on any unit (same for all units), and $\sigma_{rs}$ the covariance between the observations on $V_r$ and $V_s$. Consider the quantity

$$Q = (v-1) \sum_{r=1}^{p} \sigma_{rr}/n_r \,. \tag{2.1}$$

To interpret $Q$, first recall that for any response $V_r$, only linear contrasts of the form $\left(\sum_{j=1}^{v} u_{jr}\tau_{jr}\right)$ with $\left(\sum_{j=1}^{v} u_{jr}=0\right)$ are estimable. Consider the set of $(v-1)p$ linear functions $\left(\sum_{j=1}^{v} u_{jri}\tau_{jr}\right)$, $(r=1, 2, \cdots, p;\ i=1, \cdots, (v-1))$, with

$$\sum_{j=1}^{v} u_{jri}=0\,, \qquad \sum_{j=1}^{v} u_{jri}u_{jri'}=0 \quad (i \neq i'), \tag{2.2}$$

for all permissible $r$, $i$ and $i'$.

It is well known that the estimate $\left(\sum_{j=1}^{v} u_{jri}\bar{y}_{jr}\right)$ of $\left(\sum_{j=1}^{v} u_{jri}\tau_{jr}\right)$ is free from block effects. Let $V$ denote the $(v-1)p \times (v-1)p$ variance-covariance matrix of the set of all estimates $\left(\sum_{j=1}^{v} u_{jri}\bar{y}_{jr}\right)$, $(r=1, \cdots, p;$ $i=1, \cdots, (v-1))$. Then it can be checked that

$$Q = \operatorname{tr} V \,. \tag{2.3}$$

This shows that $Q$ depends only on the GIM(RB) design used (indeed, only on the integers $n_1, \cdots, n_p$ arising from the design) and not on any particular values of $u_{jri}$ so long as (2.2) is satisfied, i.e. so long as any maximal set of orthogonal contrasts is used. The value of $Q$ corresponding to any GIM(RB) design $D$ could be denoted by $Q(D)$. Then, (2.3) gives a method of comparing two designs $D$ and $D^*$. Thus, we may say

$D^*$ is better than $D$ w.r.t. the 'trace criterion' if $Q(D^*)<Q(D)$.

Now, given a design $D$, we can clearly obtain $D^*$ such that $Q(D^*)<Q(D)$, by just increasing the $n_r$. However, in practice this cannot be done since only a limited number of units may be available. In other words, the cost consideration comes in. Thus, let $\psi_0$ denote the initial cost of making available one block of $v$ experimental units, and $\psi_r(r=1, \cdots, p)$ the cost of measuring the response $V_r$ on all the $v$ units of any block. Then, clearly, for any GIM(RB) design $D$, the associated cost $\psi(D)$ is given by

$$\psi(D)=\psi_0 n_0+\psi_1 n_1+\cdots+\psi_p n_p\,, \tag{2.4}$$

where $n_0$ is the number of distinct blocks in $D$ (if $D$ is hierarchical then $n_0=\max(n_1, \cdots, n_p)$).

For any $\psi'>0$, let $\{\psi'\}$ denote the class of all GIM(RB) designs $D$ such that $\psi(D)\leqq\psi'$. Thus, if a total sum of money $\psi'$ is given, then we must choose a design from $\{\psi'\}$. At this stage, a good basis of choice may be provided by the above trace criterion. This motivates

DEFINITION 2.1. A design $D^*\in\{\psi'\}$ is defined to be at least as good as $D\in\{\psi'\}$ relative to the trace criterion if

$$\psi(D^*)\leqq\psi(D) \quad\text{and}\quad Q(D^*)\leqq Q(D)\,; \tag{2.5}$$

if one of the inequalities is strict, $D^*$ is said to be better than $D$.

In the following sections, we consider the problem of selection of the best design in $\{\psi'\}$. Also, $D(n_0; n_1, \cdots, n_p)$ will indicate a GIM(RB) design with $n_i(i=0, \cdots, p)$ being as above.

## 3. Optimality of HM(RB) designs

Let $D, D^*\in\{\psi'\}$ where

$$D=D(n_{10}; n_{11}, \cdots, n_{1p}) \tag{3.1a}$$

$$D^*=D(n_{20}; n_{21}, \cdots, n_{2p})\,, \tag{3.1b}$$

where $n_{ir}$ are positive integers, and $n_{i0}\geqq\max(n_{i1}, \cdots, n_{ip})$. Then, from (2.1) and (2.4), it follows that $D^*$ is at least as good as $D$ if $(n_{11}, \cdots, n_{1p})=(n_{21}, \cdots, n_{2p})$ and $n_{10}\geqq n_{20}$.

DEFINITION 3.1. Let $[\psi']$ be the subclass of HM(RB) designs contained in the class $\{\psi'\}$. Then $[\psi']$ is said to be complete w.r.t. $\{\psi'\}$ if, for any $D\in\{\psi'\}$, there exists a $D^*\in[\psi']$ such that $D^*$ is at least as good as $D$.

THEOREM 3.1. *The subclass* $[\psi']$ *is complete* w.r.t. $\{\psi'\}$.

PROOF. Let $D \in \{\psi'\}$. Consider the $p$-tuple, $(n_{11}, \cdots, n_{1p})$, and order the elements such that (say) $n_{1i_1} \geqq n_{1i_2} \geqq \cdots \geqq n_{1i_p}$. We can assume, without loss of generality, that $i_1=1, \cdots, i_p=p$. Construct the design $D^* \in [\psi']$ in the following manner. First select any subset $U_1^*$, containing $n_{11}$ blocks of the $n_{10}$ blocks in $D$, and measure response $V_1$ on these blocks. Note that $n_{11} \leqq n_{10}$. Select any subset $U_2^*$, containing $n_{12}$ blocks of $U_1^*$, and measure response $V_2$ on these blocks; and so on. Finally, select any subset $U_p^*$, containing $n_{1p}$ blocks of $U_{p-1}^*$, and measure response $V_p$ on these blocks. Thus $D^*=D(n_{11}; n_{11}, n_{12}, \cdots, n_{1p}) \in [\psi']$, and by the above discussion, $D^*$ is at least as good as $D$. This completes the proof.

Theorem 3.1 clearly points out that in designing multiresponse RB experiments (where it is possible to measure any response on any experimental unit, but resources are limited), one need consider only HM(RB) designs.

## 4. Determination of the optimum HM(RB) design

To obtain the optimal design we assume that the variances $(\sigma_{11}, \cdots, \sigma_{pp})$ are known (in practice 'a priori' estimates of $\sigma_{11}, \cdots, \sigma_{pp}$ will have to be used), and that the 'costs' $(\psi_0; \psi_1, \cdots, \psi_p)$ are given. In order to find the optimum design $D^*$, the problem is to minimize the nonlinear function $Q(D)$ subject to the linear restraints.

$$(4.1) \qquad \psi'=\sum_{r=0}^{p} \psi_r n_r , \qquad n_0 \geqq n_r \geqq 0, \ (r=1, \cdots, p) .$$

We are thus faced with a nonlinear programming problem, with the r.h.s. of (2.1) as the objective function, and (4.1) as the constraints. Further, the $n_r$ must be positive integers. We however ignore this here and assume that the optimal $n_r$ will be rounded off to integral values.

Let $D^*$ be the optimal design. (The uniqueness of the optimal design is a consequence of the convexity of the objective function in this problem.) Let the values of $(n_0, n_1, \cdots, n_p)$ for $D^*$ be $(m_0, m_1, \cdots, m_p)$. Since $D^*$ is a HM(RB) design, we have $m_0=\max(m_1, \cdots, m_p)$.

THEOREM 4.1. *If $\sigma_{ii}/\sigma_{jj} \geqq \psi_i/\psi_j$, then in $D^*$, we must have*

$$m_i \geqq m_j , \qquad (i \neq j), \ (i, j=1, \cdots, p) .$$

PROOF. Assume that $m_j > m_i$, and consider first the case where $m_0 > m_j$. Select $m^*$ such that

$$(4.2) \qquad (\sigma_{ii}/m_i)+(\sigma_{jj}/m_j)=(\sigma_{ii}+\sigma_{jj})/m^* .$$

Thus if we form a new design $D'$ with the same $m$'s as in $D^*$ but with

$m_i$ and $m_j$ replaced by $m^*$, we have that $Q(D^*)=Q(D')$. Also

$$\phi(D^*)-\phi(D')=\phi_i m_i+\phi_j m_m-(\phi_i+\phi_j)m^* . \tag{4.3}$$

This difference is positive if and only if

$$(\phi_i+\phi_j)(\sigma_{ii}+\sigma_{jj})(\sigma_{ii}/m_i+\sigma_{jj}/m_j)^{-1}<\phi_i m_i+\phi_j m_j . \tag{4.4}$$

Using the assumption that $(m_i-m_j)$ is negative we find that inequality (4.8) is equivalent to

$$m_i/m_j<\phi_j\sigma_{ii}/\phi_i\sigma_{jj} . \tag{4.5}$$

Since $m_i/m_j<1$ by assumption, and $\sigma_{ii}/\sigma_{jj}\geqq\phi_i/\phi_j$ implies $\phi_j\sigma_{ii}/\phi_i\sigma_{jj}\geqq 1$, we find that $\phi(D^*)>\phi(D')$. Since $Q(D^*)=Q(D')$, this implies that $D'$ is better than $D^*$, which is a contradiction since $D^*$ was supposed to be optimum. Hence in this first case, we have $m_i\geqq m_j$.

Next, assume that $m_0=m_j>m_i$. Again select $m^*$ as in equation (4.2). Forming $D'$ as before, we have $Q(D^*)=Q(D')$. However, the difference in the costs is now given by

$$\phi(D^*)-\phi(D')=\phi_0 m_0+\phi_i m_i+\phi_j m_j-\phi_0 m_0^*-(\phi_i+\phi_j)m^* , \tag{4.6}$$

where $m_0^*=\max(m^*, m_1, \cdots, m_{i-1}, m_{i+1}, \cdots, m_{j-1}, m_{j+1}, \cdots, m_p)$. But, $m_0=\max(m_1, m_2, \cdots, m_p)$ and $m^*\leqq\max(m_i, m_j)$. Therefore $m_0^*\leqq m_0$ and the difference in (4.6) is positive if

$$(\phi_i m_i+\phi_j m_j)>(\phi_i+\phi_j)m^* . \tag{4.7}$$

Inequality (4.7) is the same as inequality (4.4), and the proof is completed by using the same argument as above.

We now have a method of finding out the order between the $m$'s of the optimum HM(RB) design $D^*$. Indeed, $m_i\leqq m_j$ if $\sigma_{ii}/\sigma_{jj}\leqq\phi_i/\phi_j$ and $m_i\geqq m_j$ otherwise. Thus, we can establish that in $D^*$ we have (say): $m_0=m_{i_1}\geqq m_{i_2}\geqq\cdots\geqq m_{i_p}$, where $(i_1, \cdots, i_p)$ is some permutation of $(1, 2, \cdots, p)$. In the following we will assume, without loss of generality, that

$$\sigma_{ii}/\sigma_{jj}\geqq\phi_i/\phi_j , \qquad (i<j),\ (i, j=1, \cdots, p) , \tag{4.8}$$

and hence $m_0=m_1\geqq m_2\geqq\cdots\geqq m_p$.

THEOREM 4.2. *Let $k$ be the smallest integer in the set $2, \cdots, p$ such that*

$$\sigma_{k-1}^*/\sigma_{kk}>\phi_{k-1}^*/\phi_k \quad\text{and}\quad \sigma_{k-2}^*/\sigma_{k-1,k-1}^*\leqq\phi_{k-2}^*/\phi_{k-1} , \tag{4.9}$$

*where*

$$\sigma_i^*=\sigma_{11}+\cdots+\sigma_{ii} , \quad \phi_i^*=\phi_0+\phi_1+\cdots+\phi_i , \qquad (i=1, \cdots, p) . \tag{4.10}$$

*Then* $m_0 = m_1 = \cdots = m_{k-1} > m_k \geqq \cdots \geqq m_p$.

PROOF. The second inequality of (4.9) implies that

$$\frac{\sigma_{11} + \cdots + \sigma_{k-3,k-3} + \sigma_{k-2,k-2}}{\sigma_{k-1,k-1}} \cdot \frac{\sigma_{k-1,k-1}}{\sigma_{k-2,k-2}} \leqq \frac{\phi_0 + \cdots + \phi_{k-3} + \phi_{k-2}}{\phi_{k-1}} \cdot \frac{\phi_{k-1}}{\phi_{k-2}} \tag{4.11}$$

since by the assumption (4.8), $\sigma_{k-1,k-1}/\sigma_{k-2,k-2} \leqq \phi_{k-1}/\phi_{k-2}$. Thus

$$\sigma^*_{k-3}/\sigma_{k-2,k-2} \leqq \phi^*_{k-3}/\phi_{k-2} . \tag{4.12}$$

Continuing in a similar manner we find that

$$\sigma^*_1/\sigma_{22} \leqq \phi^*_1/\phi_2, \ \cdots, \ \sigma^*_{k-3}/\sigma_{k-2,k-2} \leqq \phi^*_{k-3}/\phi_{k-2} . \tag{4.13}$$

Under the assumption that $m_1 > m_2$, we have

$$\phi(D^*) = \phi^*_1 m_1 + \sum_{i=2}^{p} \phi_i m_i , \quad \text{and} \quad Q(D^*) = (v-1) \sum_{i=1}^{p} \sigma_{ii} m_i^{-1} . \tag{4.14}$$

Thus, by Cauchy's inequality,

$$\phi(D^*) \cdot Q(D^*) = (v-1)\left(\phi^*_1 m_1 + \sum_{i=2}^{p} \phi_i m_i\right)\left(\sum_{i=1}^{p} \sigma_{ii} m_i^{-1}\right) \geqq (v-1)\left(\sqrt{\phi^*_1 \sigma_{11}} + \sum_{i=2}^{p} \sqrt{\phi_i \sigma_{ii}}\right)^2 , \tag{4.15}$$

the lower bound being attained when $\sqrt{\phi^*_1 m_1} = \alpha \sqrt{\sigma_{11}/m_1}$, $\sqrt{\phi_i m_i} = \alpha \sqrt{\sigma_{ii}/m_i}$, $(i = 2, \cdots, p)$, where $\alpha$ is an arbitrary constant. We must select $\alpha$ such that the "cost equation" in (4.14) is satisfied, i.e.

$$\alpha = \gamma^{-1} \phi(D^*) , \tag{4.16}$$

where

$$\gamma = \sqrt{\phi^*_1 \sigma_{11}} + \sum_{i=2}^{p} \sqrt{\phi_i \sigma_{ii}} . \tag{4.17}$$

Thus,

$$m_1 = \sqrt{\sigma_{11}}\, \phi(D^*)/\sqrt{\phi^*_1}\, \gamma , \qquad m_i = \sqrt{\sigma_{ii}}\, \phi(D^*)/\sqrt{\phi_i}\, \gamma , \quad (i = 1, \cdots, p) . \tag{4.18}$$

From (4.18), $m_1 > m_2$ is equivalent to $\sigma^*_1/\sigma_{22} > \phi^*_1/\phi_2$, which is a contradiction of (4.13). Thus $m_1 \leqq m_2$, and we conclude that $m_1 = m_2$ in $D^*$.

To complete the proof, repeat the above $(k-2)$ times where at the $i$th step, $(i = 1, 2, \cdots, k-2)$, we assume $m_1 = m_2 = \cdots = m_i > m_{i+1} \geqq m_{i+2} \geqq \cdots \geqq m_p$, and arrive at a contradiction as before. We conclude that $m_1 = m_2 = \cdots = m_k \geqq \cdots \geqq m_p$. Thus we can write

$$(4.19)\qquad Q(D^*)=(v-1)\left(\sigma^*_{k-1}m_1^{-1}+\sum_{i=k}^{p}\sigma_{ii}m_i^{-1}\right),$$

$$\phi(D^*)=\phi^*_{k-1}m_1+\sum_{i=k}^{p}\phi_i m_i .$$

Using Cauchy's inequality as before, the minimum of $Q(D^*)$, for $\phi(D^*)=\phi$ (say), is attained when

$$(4.20)\qquad m_0=m_1=\cdots=m_{k-1}=\sqrt{\sigma^*_{k-1}}\,\phi/\sqrt{\phi^*_{k-1}}\,\gamma^*,$$
$$m_i=\sqrt{\sigma_{ii}}\,\phi/\sqrt{\phi_i}\,\gamma^*,\qquad (i=k,\cdots,p),$$

where

$$(4.21)\qquad \gamma^*=\sqrt{\phi^*_{k-1}\sigma^*_{k-1}}+\sum_{i=k}^{p}\sqrt{\phi_i\sigma_{ii}} .$$

By (4.9), $m_i>m_k$, $(i=0,\cdots,k-1)$, and by the assumption $\sigma_{ii}/\sigma_{i+1,i+1}\geqq\phi_i/\phi_{i+1}$, we see that $m_i\geqq m_{i+1}$, $(i=k,\cdots,p-1)$. This completes the proof.

We have also established

COROLLARY 4.1. *If the $p$ responses are numbered such that the inequalities in* (4.8) *are satisfied and $k$ is defined as in Theorem* 4.2, *then the values of $m_0, m_1, \cdots, m_p$, for the optimum* HM(RB) *design, are given by the equations in* (4.20).

The 'standard multiresponse model', i.e. when all responses are measured on each experimental unit, is applicable only when $m_1=m_2=\cdots=m_p$. Thus, the above establishes the important fact that in a large class of situations the optimum design (with respect to the trace criterion, and under the cost restriction) is strictly hierarchical, and it is not advisable to force the design to satisfy the standard multiresponse model. With respect to other criteria too (like the determinant criterion), the authors have partial results which also favor the HM model. But the development of those is on different lines and should be considered elsewhere.

The above studies point out to the need for more extensive research on various aspects of the HM model whose theory is still in its infancy.

As a final remark, we may say that the use of the above results depends on the variances $\sigma_{ii}$ $(i=1,\cdots,p)$ of the $p$ responses, which are usually unknown. Thus it is assumed that in practice, 'a priori' estimates of $\sigma_{ii}$, $\sqrt{\sigma_{ii}}$, $\sigma_{ii}/\sigma_{jj}$ and $\sigma^*_i/\sigma_{i+1,i+1}$, $(i,j=1,\cdots,p)$ are available (either from a pilot experiment or otherwise). Such estimation problems will be considered in a later communication.

Thanks are due to the late Professor S. N. Roy, who long ago suggested to one of the authors (Srivastava) an inquiry into the advisability

of using the incomplete multiresponse designs in comparison to the standard ones.

COLORADO STATE UNIVERSITY

## REFERENCES

[1] J. Kiefer, "Optimum experimental designs," *J. Roy. Statist. Soc., Ser. B*, 21 (1959), 273-319.

[2] I. P. Monahan, "Incomplete-variable designs in multivariate experiments," Unpublished thesis, Virginia Polytechnic Institute, Blacksburg, Virginia, 1961.

[3] S. N. Roy, R. Gnanadesikan and J. N. Srivastava, *Analysis and Design of Certain Multiresponse Experiments*, (Under print, Pergamon Press), 1969.

[4] S. N. Roy and J. N. Srivastava, "Hierarchical and $p$-block multiresponse designs and their analysis," *Sankhyā*, Mahalanobis Volume, 1964, 419-428.

[5] J. N. Srivastava, "On the extension of Gauss-Markov Theorem to complex multivariate linear models," *Ann. Inst. Statist. Math.*, 19 (1967), 417-437.

[6] I. M. Trawinski and R. E. Bargmann, "Maximum likelihood estimation with incomplete multivariate data," *Ann. Math. Statist.*, 35 (1964), 647-657.