Review of random and group-screening designs

Kleijnen, J.P.C.

Published in:
Communications in statistics: Part A: Theory and methods

Publication date:
1987

COMMUN. STATIST.-THEORY METH., 16(10), 2885-2900 (1987)

REVIEW OF RANDOM AND GROUP-SCREENING DESIGNS

Jack P.C. Kleijnen

Department of Information Systems and Auditing
Catholic University Brabant
5000 LE Tilburg
The Netherlands

Key Words and Phrases: resolution III, resolution IV; multi-stage; regression analysis; bias; conditional expectation, aggregation; simulation

ABSTRACT

Experiments with real systems, but especially with simulated systems, may involve hundreds of factors. However, only a few factors are really important. Detecting these important factors requires special designs such as random and group-screening designs. Random designs are simple, but they yield biased estimators. Group-screening is based on aggregation. The assumptions of group-screening are discussed in detail, and seem not very restrictive. References to applications of both design types are given.

1. INTRODUCTION

We shall discuss statistical designs for experiments with either real or simulated systems involving a great many factors. Through these designs we hope to detect the few factors which are really important: screening. Such designs are useful in the preliminary, exploratory phase of an investigation. Note that a different application area is experiments to isolate defective items in a population; see other papers in this Special Issue.

At the beginning of an investigation, extremely many factors may be important; for example, k = 1000. We assume that actually only a few factors are really important; for instance, k' = 10. Note that we do not know the exact value of k' nor do we know which factors are truly important. (if k' is known that "search linear models" apply; see Srivastava, 1976.) We assume that only a few factors are important, because we look for a parsimonious explanation, i.e., we do not wish to report that "everything depends on everything else".

In simulation (as opposed to real-life experimentation) we are indeed able to perform experiments with a great many factors, since we are in full control of all factors: all we have to do is let the computer program read in its inputs. In experiments with non-simulated systems it is virtually impossible to control hundreds of factors, and little attention is paid to screening designs. In simulation, applications of such designs do exist; see Section 3.3. Note that simulation includes both "systems simulation" and "Monte Carlo simulation", i.e., simulation of dynamic systems (such as queuing networks) and simulation of statistical "systems" (such as t tests and regression analyses). In academic studies the dynamic systems are often so small that screening is not necessary; in practical studies the systems are often large but the statistical know-how is missing. In Monte Carlo simulation "the shoemaker's children often go barefooted".

If we have (say) k = 1000 potentially important factors, then standard designs (such as 2^{k-p} and one-factor-at-a-time designs) would require a number of factor combinations n at least 1001. It is impractical to run so many combinations. So in the present paper we add the condition $n < k$.

There are several solutions for the problem of too many factors. A degenerated solution assumes that the output is insensitive to most inputs and concentrates on a few factors selected intuitively. Such an approach means that the limitations of the resulting conclusions are unknown. The linear search model (assuming a known upper limit for k') allows $n < k$ but is not discussed here; see Ghosh (1987) for recent references. Bettonvil (1987) also gives a procedure with $n < k$. Two other approaches are random designs and group-screening, reviewed in the next sections, using only elementary mathematics and statistics and supplying an up-to-data list of references.

Before proceeding we define some basic symbols. The design matrix $\underset{\sim}{D}$ is an $n \times k$ matrix with elements d_{ij} which specify the value of factor k in run i with integers $j = 1,\ldots,k$ $(k > 1)$ and $i = 1,\ldots,n$ $(n > 1)$. The matrix of independent variables $\underset{\sim}{X}$ is an $n \times q$ matrix. In a first-order regression model (that is, a model with k main effects β_j besides the intercept β_0) $\underset{\sim}{X}$ equals $\underset{\sim}{D}$ augmented with a column vector $\underset{\sim}{e}$ of n one's (corresponding to β_0), or $\underset{\sim}{X} = (\underset{\sim}{e},\underset{\sim}{D})$, so that $q = 1+k$. If the regression model is augmented with two-factor interactions $\beta_{jj'}$ $(j,j' = 1,\ldots,k$ and $j < j')$, then the matrix $(\underset{\sim}{e},\underset{\sim}{D})$ is augmented with the columns formed by the cross-products $x_{ij}x_{ij'}$ so that $q = 1 + k + k(k-1)/2$.

2. RANDOM DESIGNS

Random designs have been known for many years; an elaborate discussion appeared in Technometrics in 1959; for example Satterthwaite (1959). In random designs each factor has only 2 levels, denoted by +1 and -1. These factors are qualitative (that is, +1 and -1 are mnemonics for "On" and "Off" respectively: nominal scale, no inter- or extrapolation; see Kleijnen, 1987). We specify the design matrix $\underset{\sim}{D}$ by sampling these 2 values, with equal probabilities:

$$P(d_{ij} = +1) = 0.5 \qquad (i = 1,\ldots,n)\ (j = 1,\ldots,k)\ (n < k)$$

$$P(d_{ij} = -1) = 0.5 \qquad (1)$$

This solution is extremely simple, and has indeed been applied, not only to real-life experiments (see Technometrics, 1959) but also to systems simulation; see Cyert and March (1963) , Maas et al. (1962), Rosenkranz and Bürgisser (1976).

Eq. (1) can be refined by imposing the condition that each column of the design matrix has an equal number of +1 and -1 values:

$$\sum_{i=1}^{n} d_{ij} = 0 \qquad (j = 1,\ldots,k) \qquad (2)$$

We may call a design that meets eq. (2) partially balanced (a fully balanced design has orthogonal columns; see the comment on eq. 4). We can realize this condition by sampling from eq. (1) without replacement; obviously n should be an even number. (Eq. 2 is also met by classical designs such as 2^{k-p} designs, but not by one-factor-at-a-time designs.)

We may happen to sample identical columns, for example, d_{i1} may equal d_{i2} for all i values. Intuitively, if we vary the factors 1 and 2 simultaneously, we cannot estimate their individual effects. Algebraically, if we assume a first-order regression model, we find that the matrix of independent variables $\underset{\sim}{X}$ is collinear, because 2 columns are identical; hence the Least

Squares estimator $\hat{\underset{\sim}{\beta}}$ does not exist (moreover, $\underset{\sim}{X}$ is collinear because $n < k+1$; see the comment below eq. 4). Of course we can compute the following k <u>intuitive</u> estimators, where y_i denotes the response of run i:

$$\tilde{\beta}_j = \frac{\sum_{i=1}^{n} x_{ij}\, y_i}{n} \qquad (j = 1,\ldots,k) \quad (n < k) \tag{3}$$

It is simple to prove that these estimators have the following conditional expected values (we shall discuss conditional expectations in the next paragraph):

$$E(\tilde{\beta}_j | \underset{\sim}{X}) = \beta_j + \frac{1}{n} \sum_{j' \neq j} \beta_{j'} \sum_{i=1}^{n} x_{ij}\, x_{ij'} \qquad (j,j' = 1,\ldots,k) \tag{4}$$

An $n \times k$ matrix with $n < k$ has at most n orthogonal columns. Consequently, in eq. (4) the expression $\sum_i x_{ij}\, x_{ij'}$ can not be zero for all pairs (j,j'), that is, the intuitive estimators are <u>biased</u>. For example, if (as in the beginning of this paragraph) the columns 1 and 2 of $\underset{\sim}{D}$ are identical and orthogonal to all other $q-2 = k-1$ columns of $\underset{\sim}{X}$, then

$$E(\tilde{\beta}_1 | \underset{\sim}{X}) = E(\tilde{\beta}_2 | \underset{\sim}{X}) = \beta_1 + \beta_2 \tag{5}$$

The intuitive estimators $\tilde{\beta}_j$ of eq. (3) are unbiased, if we do not <u>condition</u> on $\underset{\sim}{X}$: $E(\tilde{\beta}_j) = \beta_j$. In other words, if we were to repeat the experiment (i.e. if we sampled $\underset{\sim}{D}$ and hence $\underset{\sim}{X}$ more than once, with given n and k values and $n < k$), then $\tilde{\beta}_j$ would be unbiased. Actually the basic assumption is that we cannot afford such extensive experimentation; in other words, we sample $\underset{\sim}{X}$ only once, with $n < k$, and for <u>any</u> realization of $\underset{\sim}{X}$ the estimators $\tilde{\beta}_j$ are biased!

The older literature gives several significance tests per factor. However, since the intuitive estimators $\tilde{\beta}_j$ are biased, we do not discuss these tests, but refer to Kleijnen (1974/1975, pp. 378-382).

We started with an example in which 2 columns were identical; see eq. (5). Then it is simple to prove that their estimated correlation coefficient $\hat{\rho}_{12}$ equals +1. Obviously, if $\hat{\rho}_{12} = -1$ then the design is equally bad: eq. (5) becomes $E(\tilde{\beta}_1|\underset{\sim}{X}) = -E(\tilde{\beta}_2|\underset{\sim}{X}) = \beta_1-\beta_2$. In general, we may add the following condition to sampling without replacement from P(d) in eq. (1): reject a sampled column j' if the correlation with a preceding column j is too high in absolute value, i.e., sample from eq. (1) until $|\hat{\rho}_{jj'}| \ll 1$.

Note: Since a design $\underset{\sim}{D}$ with $n < k$ cannot yield an $\underset{\sim}{X}$ with all k+1 columns orthogonal, we might minimize the maximum non-orthogonality or correlation between any 2 columns of $\underset{\sim}{X}$. This criterion leads to the non-random supersaturated designs of Booth and Cox (1962) (also see Morris and Mitchell, 1983, p. 348). We do not further discuss their results, because their designs yield biased estimators and we do not know any applications.

Summary: Random designs are extremely simple, since all we have to do is sample (without replacement) the elements of the design matrix $\underset{\sim}{D}$. Unfortunately, if the number of runs n does not exceed the number of factors k, then the estimated factor effects are biased.

3. GROUP-SCREENING DESIGNS

Because random designs certainly yield biased results, we now discuss group screening, proposed by Watson, Patel, and Li in the early sixties; see Kleijnen (1974/1975) for exact references. In group screening we aggregate individual factors into groups of factors; if a group is not significant, then we conclude that all its members are unimportant. In Section 3.1 we shall consider an example; in Section 3.2 we shall discuss the assumptions of this example; in Section 3.3 we shall present applications.

3.1. Example of Group Screening

TABLE 1
Group-screening Example

Run	Group factor		Individual factor						Expected Responose
i	w_1	w_2	d_1	...	d_{50}	d_{51}	...	d_{100}	$E(y)$
1	-	-	-	...	-	-	...	-	$\beta_0 - \beta_1 - \beta_2$
2	-	+	-	...	-	+	...	+	$\beta_0 - \beta_1 - \beta_2$
3	+	-	+	...	+	-	...	-	$\beta_0 + \beta_1 + \beta_2$
4	+	+	+	...	+	+	...	+	$\beta_0 + \beta_1 + \beta_2$

In the example of Table 1 we have $k = 100$ individual factors, and we form 2 groups of 50 individual factors each. The number of groups is denoted by G; here $G = 2$. To investigate the 2 group factors w_1 and w_2 we use a 2^2 design; a group factor has value -1, if all its (individual) members are -1; and the group factor has value +1, if all its members are +1. The last column of Table 1 follows from the assumption that in this example all individual factors have zero effects, except for the factors 1 and 2 (in practice we do not know that only β_1 and β_2 are nonzero, besides β_0).

The least squares estimator of the first-order effects $\underset{\sim}{\gamma}$ of the group factors is

$$\hat{\underset{\sim}{\gamma}} = (\underset{\sim}{V}'\underset{\sim}{V})^{-1}\underset{\sim}{V}'\underset{\sim}{y} \tag{6}$$

where $\underset{\sim}{V} = (\underset{\sim}{e},\underset{\sim}{W})$ with $\underset{\sim}{W} = (w_{ig})$ and $g = 1,\ldots,G$; In Table 1 we have an orthogonal $\underset{\sim}{V}$ so that eq. (6) reduces to

$$E(\hat{\gamma}_1) = E(\frac{-y_1 - y_2 + y_3 + y_4}{4}) = \beta_1 + \beta_2 \tag{7a}$$

and

$$E(\hat{\gamma}_2) = 0. \tag{7b}$$

In eq. (7a) the individual effects do not cancel out, or $E(\hat{\gamma}_1) \neq 0$, if the following 2 assumptions hold:

(1) We know the signs of the individual main effects, i.e., if the individual factor does have an effect then we know the direction of the effect. (For example, if the number of servers is important in a queuing system, then that number has a negative effect on the waiting times.) Hence we can define the 2 levels of the individual factors such that their main effects are either positive or zero: $\beta_j > 0$ for $j = 1,\ldots,k$.

(2) The k individual factors have main effects only. Consequently, the effects of the individual factors within a group cannot compensate each other.

To test the significance of the estimated group effects $\hat{\gamma}_g$ we need variance estimators. We obtain estimates $\hat{\sigma}_i^2 = \hat{var}(y_i)$ through the use of replications (called subruns in system simulation; see Kleijnen, 1974/ 1975, 1987). These $\hat{\sigma}_i^2$ yield $\hat{var}(\hat{\gamma}_g)$. An unimportant group factor - such as w_2 in the example (see eq. 7b) - has only α probability of being significant, say $\alpha = 5\%$. We eliminate all individual factors within a non-significant group, before we proceed to the second phase of the experiment (not shown in Table 1). In that phase we investigate the individual factors within the significant groups only. In the example, the estimator $\hat{\gamma}_1$ has a higher chance of being significant as $\beta_1 + \beta_2$ increases; see eq. (7a) and the definition of the power function of a test. We emphasize that after the 4 runs of Table 1 we do not know that the individual factors 1 and 2 cause the significance of group factor 1; all we know is that one or more factors within group 1 are important.

If after the first phase of the experiment we declare many individual factors to be unimportant, which levels then should these factors have in the next phase of the experiment? If the factors have exactly zero effects, then the choice of their values has no effect. Actually their effects may be non-zero, yet so small that the effects are masked by the noise plus important effects of the other factors (β error of test). We recommend keeping the "unimportant" factors at fixed levels during future experimentation, provided our main goal is to estimate the effects of the important factors: if important factors happen to be non-significant, then their effects are confounded with β_0 and not with the effects of the other individual factors. If our goal, however, is to predict the response variable y, then we randomly vary the levels of the non-significant factors. Also see Kleijnen (1974/1975, pp. 395, 417-419) and Olivi (1984).

3.2. Assumptions of Group Screening

One assumption of group screening was that we know the signs of the main effects. How restrictive is this assumption?

(1) In some applications we have so much qualitative insight into the system that we do know these signs; see the queuing example below eq. (7).

(2) For the example of Table 1, eq. (7a) proves that we miss the important individual factors 1 and 2, only if they belong to the same group, if their effects have opposite signs, and if these

effects are of the same magnitude so that the experimental noise masks the effects. The smaller the group size, the smaller the chances of such "pathological" behavior. (Mauro and Smith, 1982, p. 83, assume that "factors are grouped randomly"; in practice, however, we have prior knowledge and we use that knowledge to group factors; also see Shubik, 1975, p. 108.)

(3) If we believe a priori that a specific factor is important, then we can test that factor individually.

A second assumption was that the individual factors have first-order effects only. Now we consider a regression model with two-factor interactions. For example, factor 1 interacts with the important factor 2 and with the "unimportant" factor 100 or $\beta_{1,2} \neq 0$, $\beta_{1,100} \neq 0$ and (as before) $\beta_1 > 0$, $\beta_2 > 0$, $\beta_0 \neq 0$. Using the analogue of eq. (2), namely $\sum_i w_{ig} = 0$, and also $w_{ig}^2 = 1$ (so that $\sum_i w_{ig}^3 = 0$), $w_{i1} = x_{i1} = \ldots = x_{i50}$ and $w_{i2} = x_{i51} = \ldots = x_{i100}$ we obtain the same result as eq. (7a), which assumed main effects only:

$$E(\hat{\gamma}_1) = E\left(\frac{\sum_{i=1}^{n} w_{i1}\, y_i}{n}\right) = \frac{1}{n} \sum_i w_{i1}\, (\beta_0 + \beta_1\, w_{i1} +$$

$$+ \beta_2\, w_{i1} + \ldots + \beta_{1,2}\, w_{i1}\, w_{i1} + \ldots + \beta_{1,100}\, w_{i1}\, w_{i2})$$

$$= \frac{1}{n} (\beta_0 \sum_i w_{i1} + \beta_1 \sum_i w_{i1}^2 + \beta_2 \sum w_{i1}^2 + \ldots +$$

$$+ \beta_{1,2} \sum_i w_{i1}^3 + \ldots + \beta_{1,100} \sum_i w_{i1}^2\, w_{i2}) = \beta_1 + \beta_2 \qquad (8)$$

More generally Kleijnen (1974/1975, p. 398) and (1975, p. 492) proved that two-factor interactions do not bias the main effects, if the group factors are investigated in a design of resolution IV or higher. Designs of different resolution are discussed in Kleijnen (1974/1975) and (1987); also see Bettonvil and Kleijnen (1987). Resolution IV designs imply $n \geq 2G$. Table 1 is an example of such a design: n = 2G = 4. [If we examine the group factors in a design of resolution III then a two-factor interaction $\beta_{jj'}$ creates bias for the estimator of a main effect $\beta_{j''}$ only if the factors j, j' and j" belong to 3 different groups. Consequently, if all factors interact, then main effect estimators are biased indeed. But if only some factors interact, then we should place these "related" factors within the same group.] We are mislead even if the design is of resolution IV, if all individual effects are zero except for the interactions between 2 factors belonging to the same group (in a case study such a situation did exist; see Kleijnen, et al. 1979).

The literature gives more assumptions for group-screening. These assumptions, however, are not essential, since they are used only to derive the <u>optimal</u> sizes of the groups, i.e., to minimize the number of runs taken over all stages of screening; see Kleijnen (1974/1975, p. 396), Mauro (1984), Ottieno and Patel (1984) and Samuels (1978).

Instead of two-stage group screening, we may apply <u>multi-stage</u> group screening , i.e., we continue to use some grouping of individual factors which are not eliminated at a stage, until finally no factors are aggregated. In practice we have <u>prior knowledge</u> and we use that knowledge to form groups of related factors; such a procedure implicitly determines the group sizes; we stop screening when all groups have reduced to size 1.

<u>Summary</u>: In group screening, a number of assumptions are needed only to derive optimal group sizes. There are 2 important assumptions, namely known signs and no interactions. These assumptions do not seem to be so restrictive as to preclude application of group screening to practical experiments.

3.3. Applications of Group Screening in Systems Simulation

Applications of group screening are rare yet; see Section 1. Nevertheless, especially in experiments with simulated systems as opposed to real systems, this design type may be very useful. Therefore we list some applications in systems simulation and describe one application in a little more detail.

(1) Rooda and Van der Schilde (1982) simulate maritime transport and distribution by seagoing barges. They apply two-stage group screening to 29 individual factors. These 29 factors are aggregated into 8 groups, and these 8 groups are studied in a 2^{8-4} design, so that n = 16 = 2G (resolution IV design). They analyse the experimental data, using Least Squares. The first-order regression model expressed in the 8 group factors, is validated by comparing the simulation response y to the regression predictor $\hat{y}$; see Kleijnen (1983). Significance tests for the estimated group effects $\hat{\gamma}_g$ lead to the elimination of 4 groups. In the second stage they investigate the remaining 12 individual factors in a 2^{12-8} design: n = 16 > 12 (resolution III). They again use Least Squares, and so on.

(2) Mihram (1972) and Nolan and Mastroberti (1972) simulate a strategic airlift system. Mihram (1972, pp. 399-400) states that interactions "would tend to conceal the true significance of other factors in the subset"; however, we proved that in resolution IV designs two-factor interactions do not conceal the additive main effects of the factors within a "subset" or group.

(3) Schatzoff and Tillman (1975) examine a computer system.

(4) De Hoogh (1982) studies a dike (storm-surge barrier) in The Netherlands.

All these applications concern simulations of dynamic systems, so complicated that it is unknown which factors are really important. Consequently we cannot prove that group screening works in these applications! Therefore we perform the following small Monte Carlo simulation. We create the "real system"

$$y = \beta_0 + + \sum_{1}^{88} \beta_j x_j + e \qquad (9)$$

where all factors have zero effects, except for 7 factors. Next we apply group screening to this system. Indeed the screening technique detects the 7 important effects! Mauro and Burns (1984) perform a similar Monte Carlo experiment with 100 individual factors and they vary the total number of runs between 20 and 84.

4. CONCLUSION

In random designs we treat all factors equally and try to minimize correlations between factors. In group screening, factors within a group have maximum correlation +1 and between groups they have minimum absolute correlation 0. A random design certainly yields biased estimators of individual effects β_j, given any realization of the sampled design. Group-screening is based on the aggregation principle. The assumption of known signs (or directions) of the individual effects does not seem very restrictive, since we may rely on the low probability of "pathological" behavior (important factors occur within the same group and have effects of equal magnitudes with opposite signs) and we may examine some factors individually. Two-factor interactions do not bias main effects, if we examine the G group factors in a resolution IV design (with number of runs $n > 2G$).

REFERENCES

Bettonvil, B. (1987). Sequential bifurcation for factor screening. Catholic University Brabant, Tilburg, July 1987.

Bettonvil, B. and J.P.C. Kleijnen (1987). Measurement scales and resolution IV designs. Catholic University Brabant, Tilburg, February 1987.

Booth, K.H.V. and D.R. Cox (1962). Some systematic supersaturated designs. Technometrics, 4: 489-495.

Cyert, R.M., and J.G. March (1963). A Behavioral Theory of the Firm. Prentice-Hall, Englewood Cliffs (N.J.).

De Hoogh, M.P.A.J. (1982). De ontwikkeling van een metamodel voor de duur van de bouw van de stormvloedkering in de Oosterschelde. (The development of a metamodel for the duration of constructing a storm-surge barrier in the Eastern Scheldt), Thesis, Department of Mathematics and Informatics, Delft University of Technology, Delft (Neth.).

Ghosh, S. (1987). Influential nonnegligible parameters under the search linear model. Communications in Statistics, Theory and Methods, 16, no. 4: 1013-1025.

Kleijnen, J.P.C. (1974/1975). Statistical Techniques in Simulation. Volumes I and II. Marcel Dekker, Inc., New York. (Russian translation: Publishing House "Statistics", Moscow, 1978).

Kleijnen, J.P.C. (1975). Screening designs for poly-factor experimentation. Technometrics, 17, no. 4: 487-493.

Kleijnen, J.P.C. (1983). Cross validation using the t statistic. European Journal of Operational Research, 13: 133-141.

Kleijnen, J.P.C. (1987). Statistical Tools for Simulation Practitioners. Marcel Dekker, Inc., New York.

Kleijnen, J.P.C., A.J. Van den Burg and R.T. Van der Ham (1979). Generalization of simulation results: practicality of statistical methods. European Journal of Operational Research, 3: 50-64.

Maass, A., et al. (1962). Design of Water-Resource Systems. Harvard University Press, Cambridge (Massachusetts).

Mauro, C.A. (1984). On the performance of two-stage group screening experiments. Technometrics, 26, no. 3: 255-264.

Mauro, C.A. and K.C. Burns (1984). A comparison of random balance and two-stage-group screening designs: a case study. Communications in Statistics, Theory and Methods, 13, no. 21: 2625-2648.

Mauro, C.A. and D.E. Smith (1982). The performance of two-stage group screening in factor screening experiments. Technometrics, 24, no. 4: 325-330.

Mihram, G.A. (1972). Simulation: Statistical Foundations and Methodology. Academic Press, New York.

Morris, M.D. and T.J. Mitchell (1983). Two-level multifactor designs for detecting the presence of interactions. Technometrics, 25, no. 4: 345-355.

Nolan, R.L. and R. Mastroberti (1972). Productivity estimates of the strategic airlift system by the use of simulation. Naval Research Logistics Quarterly. 19: 737-752.

Olivi, L., editor (1984). Response Surface Methodology Handbook for Nuclear Reactor Safety. EUR 9600, Commission of the European Communities, Joint Research Centre, Ispra (Italy).

Ottieno, J.A.M. and M.S. Patel (1984). Two-stage group-screening designs with unequal a-prior probabilities. Communications in Statistics, Theory and Periods, 13, no. 6: 761-779.

Rooda, J.E. and W. Van der Schilden (1982). Simulation of maritime transport and distribution by sea-going barges; an application of multiple regression analysis and factor screening. Bulk Solids Handling, 2, no. 4: 813-824.

Rosenkranz, F. and R. Bürgisser (1976). Automatisches Planen und Auswerten von Simulationsexperimenten mit einer Unternehmungs-Simulationssprache. Angewandte Informatik, 5: 216-222.

Samuels, S.M. (1978). The exact solution to the two-stage group-testing problem. Technometrics, 20, no. 4: 497-500.

Satterthwaite, F.E. (1959). Random balance experimentation. Technometrics, 1: 111-137, 184-192.

Schatzoff, M. and C.C. Tillman (1975). Design of experiments in simulator validation. IBM Journal of Research and Development. 19, no. 3: 252-262.

Shubik, M. (1975). The Uses and Methods of Gaming. Elsevier Scientific Publishing Company, Inc., New York.

Srivastava, J.N. (1976). Smaller sized factor screening designs through the use of search linear models. 9th International Biometric Conference, Boston.

Received by Special Editor 28 November, 1986; Revised May, 1987.

Refereed by J. N. Srivastava, Colorado State University, Fort Collins, Colorado, U. S. A.