Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

ROBUST HYBRID DESIGNS FOR REAL-TIME SIMULATION TRIALS

Russell C.H. Cheng
Owen D. Jones

Faculty of Mathematical Studies
University of Southampton
Southampton, SO17 1BJ, U.K.

ABSTRACT

Real time simulation trials involve people and are particularly subject to a number of natural constraints imposed by standard work patterns as well as to the vagaries of the availability of individuals and unscheduled upsets. They also typically involve many factors. Well thought-out simulation experimental design is therefore especially important if the resulting overall trial is to be efficient and robust. We propose hybrid experimental designs that combine the safety of matched runs with the efficiency of fractional factorial designs. This article describes real experiences in this area and the resulting approach and methodology that has evolved from these and which has proved effective in practice.

1 INTRODUCTION

This paper describes the design of real-time simulation trials and is based on work carried out for National Air Traffic Services in the UK. Details of the specific work involved and the trials themselves are subject to restrictions for reasons of confidentiality, however a number of issues arise of general interest. In particular, the design of simulation experiments in this context was of special interest because it involved real-time simulation studies and the presence of human operators and controllers. The individual runs that together made up the trial as a whole therefore had to be arranged to fit within the constraints of normal working hours. Also, the overall trial had to be sufficiently robust to be able to absorb interruptions and the occasional, often random, non-availability of individuals.

The two main examples, which we present below to illustrate general issues of interest, are based loosely on real trials. They are thus reasonably realistic; but we have not presented actual data or discussed the real issues that their analysis led to.

Here and in what follows we use the term *trial* to mean the overall simulation exercise as a whole, and a *run* to mean an individual simulation run, usually made at a given set of prescribed levels of a number of factors expected to be important in affecting the outcome of the simulation.

Indeed we will suppose that the objective of a trial is typically to study the influence of a number of factors on the behaviour of a complex system. The overall experiment has therefore to be efficiently planned to enable worthwhile information to be gathered from a relatively small number of runs.

Below we describe the methodology that has evolved from the air traffic work, and illustrate its use in the design of two hypothetical trials, which nevertheless are based on real studies. The discussion centres on the general design philosophy that was found to work well. Also of interest is the type of experimental design that has proved robust and flexible.

This last aspect was especially important to the organisers, as each trial was expensive and elaborate to set up both in time and financially. The organisers were therefore particularly keen to be assured at the outset that there was little chance of untoward eventualities seriously disrupting and possibly even ruining the usefulness of the trial as a whole.

The general philosophy of simulation experimental design is well enunciated by Barton (2002) which draws on many earlier discussions including Kleijnen (1987), Barton (1999) and Sanchez (2000). More general discussions, but still in a simulation context appear in Banks (1998) and Law and Kelton (2000).

The design of experiments in a general statistical context, but where the discussion has a focus that is especially appropriate to our simulation viewpoint, can be found in McLean and Anderson (1984), Montgomery (1997) and Wu and Hamada (2000).

We draw on the general ideas of the above references. Indeed we assume a basic knowledge of fractional factorial design in what follows. However our discussion has a much more specific focus. We shall discuss particular types of design that seem especially useful in the real-time simulation context. The background and rationale governing such designs are discussed in the next two Sections. We then describe two examples in particular detail to illustrate the design issues involved. In Sections 6 and 7 we suggest a regression viewpoint involving a stronger focus on estimation which avoids some of the practical difficulties evident in the ANOVA approach of the two examples.

2 BACKGROUND CONTEXT

The basic length of a run was one hour. This might appear arbitrary, but a little thought indicates that for real-time simulation work such a run-length will be a useful one to take in many contexts. It allows the individuals concerned to settle into a regular work pattern and be observed, but is not so long that fatigue becomes a serious issue. (Obviously trials studying fatigue itself might need to be handled differently, but even then a basic unit of measure of one hour might be perfectly justified.)

We see therefore that a typical trial cannot have more than an absolute maximum of 40 runs per week. A more realistic upper limit is nearer thirty or so runs each week, bearing in mind that debriefing and initial training sessions will often be required.

Moreover there is a typical natural length to a trial. Almost any sort of real-time trial will require physical resources, both space and hardware with their attendant costs. The financial implication of this is that a trial will typically be at most one or two weeks in length. For this reason we shall confine our discussion to trials where no more than 30 to 60 runs each of length one hour are to be made.

In the air traffic control context and in other complex systems, there will be many factors that will influence the outcome which itself is almost certainly multidimensional in its own right. The natural experimental formats to use are therefore factorial experiments. We shall only consider factors with two, three or four levels.

The main steps in the design process are set out in Table 1.

Table 1: Trial Design

Step	Action
1.	Identify Factors and Levels
2.	Produce Approximate Initial Design
3.	Adjust design to meet special conditions
4.	Test Robustness of Design

3 DESIGN RATIONALE

The maximum number of runs possible places a clear restriction on how many factors can sensibly be studied. Table 2 gives the number of main effects and first order interactions corresponding to different numbers of factors. If therefore the maximum number of runs available is 40 then one cannot have more than 8-10 factors if all main effects and first order interactions are to be estimated.

Table 2: Number of Main Effects and First Order Interactions

Number of Factors	Main Effects	First Order Interactions
5	5	10
6	6	15
7	7	21
8	8	28
9	9	36
10	10	45
11	11	55
12	12	66

The number of factors can be increased if it is possible to eliminate a sufficient number of interactions as being negligibly small. Care has obviously to be exercised in selecting a design that ensures all interactions that are likely to be non-negligible can be estimated.

We shall assume in what follows that our main interest is in the estimation of main effects and some or all first order interactions. We allow the possibility that some interactions can, *a priori*, be assumed to be zero. We shall assume that interactions of order higher than the first can be neglected, or at least treated as part of the unexplained 'error', in the ANOVA sense.

Even with this last assumption, it will still be necessary to employ a *fractional* factorial design, relying on correct design to avoid the usual difficulties of aliasing and confounding. The selection of suitable designs is the main focus of the remainder of the paper.

A particular format of special interest is where the trial is intended to compare *two* systems of operation. Typically one system, which we call the *base* system, is the existing one. The other is a new alternative, which we shall refer to as the *new* system. The obvious way of handling this is to treat the system being studied as being a factor, where there are two levels corresponding to whether it is the base or the new system that is being used in the run. We shall call this the *system factor*.

There are two things of note both of which lead us into wishing to treat the system factor in a special way rather than as being an ordinary factor.

Firstly, the new system is usually being considered as a candidate replacement for the existing old system. However rather less is known about it than the base system. It is therefore desirable or even necessary to explore its likely behaviour more fully than that of the base, which is usually already better understood. This makes it desirable to make more runs using the new, than the base system.

The second thing is to do with the use of fractional factorial designs. These designs make use of assumptions about interaction effects to reduce the total number of runs from that of a full factorial whilst still allowing all non negligible effects and interactions of interest to be estimated. However there is a danger that, if the assumptions turn out to be unjustified, then this can render suspect the subsequent analysis in the fractional factorial trial.

Use of matched pairs of runs is a much more robust procedure, but is conservative in that it provides a less full exploration of the effect of different factor combinations than that provided by a fractional factorial design, when the latter works properly.

We therefore consider *hybrid* designs. These are where the total number of runs is divided into three. One set of runs uses the base system. The other two use the new system. In one set the runs are *matched runs* with the same set of combinations of levels of factors, other than the system factor, used as for the set of runs using the base system. The other set of runs using the new system has combinations of other factor level settings chosen to form a fractional factorial design when taken with the first set of runs where the base system is used.

The hybrid design is a trade-off. It retains the robustness of matched pairs but gives some of the efficiency of fractional factorials in exploring the factor space.

The simplest designs are those where each factor is held at two levels. A factor, A say, with four levels can often be dealt with by using two pseudofactors, A1 and A2 say, each at two levels. The four levels of A can then be identified with the four combinations of A1 and A2 levels. The analysis is done using A1 and A2, with a reinterpretation of the results at the end in terms of A.

There is one aspect involving the design of an experiment that should be borne in mind. The usual way that fractional factorial designs are presented in standard texts is through tabled appendices including detailed aliasing properties. This is very convenient for standard designs like those of Resolution III and IV whose properties can be succinctly summarized in terms of aliasing characteristics of *all* interactions of a given order. However such tabulations tend to be less convenient to use in the grey area where we are limited by the number of trials into having to consider designs which may not have the full resolution that we should like. Thus for example it may be that we expect some but not all first order interactions to be important, but we are not able to use a design that allows *all* first order interactions to be estimated. However we may be fairly sure that certain interactions will not be present; perhaps because we know that the interactions of a certain factor are unlikely. Then we can estimate any interaction that is aliased only with these interactions, and it may be that such a design exists. However the process of identifying such a design is often not straightforward and requires either expert knowledge or else an assiduous search of tabulations in the hope of finding one with the right aliasing structure for our purpose.

An alternative approach is to write down a parsimonious model, including those effects and interactions that might be present, but which explicitly *excludes* those that we think will be negligible. We can then construct a design usually without having to be dictated by laudable but potentially awkward design principles. The only requirement is *estimability*: does the design that we have constructed allow us to estimate all the effects and interactions included in our model?

This is easily tested by checking if the least squares estimator is computable or not. Explicitly the key question is: If X is the design matrix, then is X^TX invertible?

There is of course widespread recognition that both ANOVA and regression are both manifestations of the linear model, however the development of each approach tends to follow rather different paths, with usually relatively little discussion of the equivalence of the two.

The main emphasis in the literature on ANOVA is on the construction of the ANOVA table and the construction of sums of squares. There is relatively less emphasis on the design matrix as such. Estimability is viewed very much from the approach of aliasing and confounding.

In contrast (multiple linear) regression analysis tends to focus much more on the estimation of coefficients in the regression model, and so there is a much more direct focus on estimability, and on the properties of the design matrix itself.

We feel that there is much to commend an approach to design of simulation experiments using the regression viewpoint.

In the next two sections we consider two explicit examples. The first, 'ZIP' trial, example is an illustration of a 2^k design where $k = 11$.

The second example is what is called a *mixed-level* design. It only has five factors. However though four of the factors are 2-level, the fifth is a 3-level factor. The full design is thus $2^4.3$.

4 EXAMPLE OF A 2^K HYBRID DESIGN

Our first example is of a 2k hybrid design. It is a typical example of experimental design in that, despite the tabulation of very many fractional factorial 2k designs, we nevertheless cannot simply take one 'off the shelf', but have to adjust and manipulate such a known design to get what we want. Thus in producing the final design we will make use of:

(i) Pseudofactors
(ii) Examples of Resolution III and Resolution IV designs
(iii) The foldover principle
(iv) Hybrid Construction

We shall suppose that a special concern of the trial is to investigate an augmented method of system operation, which we shall call 'with ZIP'. Thus a major requirement was to make at least some of the runs in matched pairs with one run in each pair conducted 'with ZIP', and the other 'without ZIP'. The runs were to investigate the effect of the following factors.

- General Factors
  - At 2 levels:

  - System (ZIP, the main factor under investigation)
  - Weather Factor1 (D)
  - Weather Factor2 (W)
- At 4 levels:
  - Controller Position (C)
  - Traffic Sample (T)

- ZIP only Factors
  - At 2 levels:
    - H-Factor (H)
    - A-Factor (A)
    - S-Factor (S)
    - P-Factor (P)

4.1 Constraint Requirement

For organisational reasons the 'with ZIP' runs had to be carried out in pairs. This imposed the constraint that, in each pair of runs, all the ZIP only factors had to be held at the same level throughout.

A possible way of handling the factors, C and T with 4 levels each, is to use a pair of 2-level pseudofactors to model each. Thus C is treated as (CA, CB) with CA and CB each taking two levels. The four possible combinations of CA and CB levels then represent the four original C levels.

Using the hybrid principle we might consider a 48 run trial, with 32 combinations of 'with ZIP' runs and 16 'without ZIP' runs.

We consider possible trial designs based on $2^{(k-p)}$ fractional factorial designs, where k is the number of factors (including pseudofactors) and $1/p$ is the fraction of a full ($2^{**}k$) factorial design used.

One possible way of meeting the Constraint Requirement is to take a $2^4 = 16$ run design for the 'without ZIP' runs. The factors here are then TA, TB, D, W, CA, CB making $k = 6$ and p = 2. One can have a $2^{(6-2)}$ design of resolution IV using the generators 5=123 and 6=234.

To handle the 'with ZIP' runs, we note that here the factors are TA, TB, D, W, CA, CB, augmented by H, A, S and P. This makes k = 10. But we have to satisfy the Constraint Requirement on H, A, S and P. We can do this simply by using a pair of replicated 16 run designs each exactly the same as the 'without ZIP' 16 run design. Each 16 run design is a $2^{(10-6)}$ design using 5=123, 6=234, as before, and taking 7=134, 8=124, 9=1234, (10)=12. This $2^{(10-6)}$ design is of resolution III.

Thus, as far as the unconstrained factors TA, TB, D, W, CA, CB are concerned, these are handled in exactly the same way in each of the 16 run sets, 'with ZIP' or 'without ZIP'.

The advantage of this design is that we can estimate 'pure error' from the two 'with ZIP' replications. Its disadvantage is that the 'with ZIP' design has resolution III only.

We could instead obtain a resolution IV design for the 'with ZIP' set of runs that satisfies the Constraint Requirement. We start with a 9-factor, 16 run, $2^{(9-5)}$ of resolution III. Such a design exists; it is precisely our previous $2^{(10-6)}$ design but simply omitting the (10)*th* factor. Thus we generate the fifth to ninth factor level settings using: 5=123, 6=234, 7=134, 8=124, 9=1234. We choose our nine factors and their order as follows. Let H, A, S be the first three factors, and the fourth be one of the unconstrained 2 level factors, D, say.

Let the fifth, sixth, seventh, eighth, and ninth factors be the remaining unconstrained factors: TA, TB, W, CA, CB. The point of this arrangement is that the runs will be in pairs where all the first three factors will be at the same level. Thus each pair can be one of the paired runs.

Now we use the technique of 'foldover'. First add an extra 2 level factor, in our case the remaining constrained factor P, taking its level as -1 throughout all 16 runs. We now simply copy the entire 16 run design, except that the level of each factor is interchanged for its other level (i.e. a +1 becomes a -1, and a -1 becomes a +1). This creates a $2^{(10-5)}$ design of resolution IV. By construction the Constraint Requirement is met for all four constrained factors.

For the 'without ZIP' set of runs we simply take the initial 9-factor, 16 run, $2^{(9-5)}$ design of resolution III and ignore the first three (constrained) factors used. The remaining design with the six factors D, TA, TB, W, CA, CB is a $2^{(6-2)}$ design. It is only of resolution III however.

We do not know whether it is possible to produce resolution IV designs for both the 'with ZIP' and the 'without ZIP' set of experiments that both satisfies the Constraint Requirement and also that produces a subset of 16 runs within the full set of 32 runs of the 'with ZIP' set matched to the 16 of the 'without ZIP' set. We suspect not, but would be happy to stand corrected.

On balance we think the first design, using replications, is preferable.

5 A MIXED LEVEL DESIGN

The second example is one involving a mixed level design. The trial involved the following factors and was to be conducted over two weeks.

- At 2 levels:
  - System (S, The main factor under investigation)
  - Weather Factor1 (D)
  - Weather Factor2 (W)
  - Controller Factor1 (A)
- At 3 levels:
  - Controller Factor2 (R)

5.1 Constraint Requirement

The factor A concerned the availability of controllers. Only one level of A could be used in a given week.

The base level of factor S was 0, the new level corresponded to level 1. A hybrid design with twice the number of level S1 runs as S0 runs was proposed.

A full trial would require $2^4.3 = 48$ combinations, with 6 main effects and 15 first order interactions. Connor and Young (1961) discuss the construction of mixed two and three level designs where all main effects and two-factor interations can be estimated. The article is reproduced in McLean and Anderson (1984 Appendix 1).

The 3-level R factor places a severe constraint on construction of a design. The best that seems possible is a ¾ fractional design with 36 runs. This requires three sets T_1, T_2, T_3 of factor combinations for the four 2-level factors. Each has to be combined with the full T' set of factor combinations (ie the full replicate) of the 3-level factor. The overall design is $\{T_1T', T_2T', T_3T'\}$. In this design all main effects and first-order interactions are estimable.

This design requires modification to meet the Constraint Requirement. As things stand each set T_i requires both levels of each two-level factor to be used. If we take A to be the first factor in the Table 3 then we can arrange the table in the form of Table 4.

In Table 4 all the runs of T'_1 have factor A at setting zero, whilst those of T'_2 have factor A at setting 1.

Table 3: 2-level Factor Combinations of the $2^4.3^1$ Design

T_1	T_2	T_3
0000	0001	0011
1100	0110	0100
0111	1010	1000
1011	1101	1111

Table 4: 2-Level Factor Combinations of the 2^43^1 Design

Run	T'_1	T'_2
1	0100	1100
2	0000	1000
3	0111	1111
4	0011	1011
5	0001	1010
6	0110	1101

Note that, in both T'_1 and T'_2, Runs 1 and 2 and also Runs 3 and 4 form matched pairs with regard to the *second* factor. Thus if the second factor is taken to be S, the main system factor, then this modified design $\{T'_1T', T'_2T'\}$ already automatically incorporates 12 pairs of matched runs.

The design is readily extended to include additional matched pairs. For example additional runs, 0010 and 1001, matched to Runs 6 in T'_1 and T'_2 can be added. Combined with T' this gives an additional six runs.

Alternatively the design can be extended to include additional runs with the system factor setting at level S2. For example additional runs, 0101 and 1110, can be added, which combined with T' gives six additional runs. This latter scheme has the advantage of yielding runs matched to the Runs 5 in T'_1 and T'_2 thereby increasing the number of matched pairs to a total of 18.

Wu and Hamada (2000, Chapter 7) discuss the construction of mixed level designs. They give a number of *Orthogonal Array* (OA) designs. If only main effects are of interest then the OA(12, $2^4.3$) design is a very useful parsimonious design that could be used as a basis for this particular example. This design provides a very efficient core of runs from which fuller designs can be formed.

An alternative approach (see Montgomery, 1997) is to use two 2-level pseudofactors, R_1 and R_2 say, to model the 3-level factor R. Thus the combination $R_1R_2 = 00$ can represent $R = 0$, *both* combinations $R_1R_2 = 10$ and $R_1R_2 = 01$ can represent $R = 1$ and $R_1R_2 = 11$ can represent $R = 2$. Thus we are modelling $2^4.3$ as being an embedded subset of 2^6.

A little care is needed in using this technique as the interaction between R_1 and R_2 is being treated as a main effect of R, thus a fractional design of 2^6 has to be selected to ensure that the effects associated with all levels of R can be estimated. The quarter replicate 2^{6-1} using the generator 6=12345 has resolution VI, and this allows all the main effects to be estimated.

6 REGRESSION APPROACH AND ESTIMABILITY

As remarked in the introduction, the approach of selecting, from known designs, one with all the properties that are required by the experimental setup, can be one fraught with pitfalls. It often requires perusing extensive tabulations, especially to check the properties of aliasing, to ensure that all effects and interactions that might be important in the trial are not confounded.

An alternative is to focus much more specifically on the estimation issue. It is then much more natural to use the regression modelling approach, and to write down the statistical model with all the effects and interactions that are deemed important represented by *parameters*. The design can then be progressively built up to ensure that all the parameters of the model can be estimated. One can of course utilise good design principles and methods, like orthogonality, and balanced designs, and efficient fractional designs. However this can be done in a flexible way. If use of a good design principle cannot, for whatever reason, be easily adhered to, then some slippage from the ideal can however be tolerated.

The key requirement of a *workable* design is *estimability*.

This can be directly checked. Barton (2002) points out that this can easily be done by compiling some artificial data from the model, and then running this simulated data through a statistical package to see if it encounters any difficulties. In fact there is no need even to produce any simulated data, for the following reason.

The linear model has the form

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

where **y, θ** and **ε** are respectively the vectors of observations, parameters and errors. The key however is the design matrix **X**. The parameters **θ** are estimable if and only if the corresponding sum of squares matrix, $\mathbf{X}^T\mathbf{X}$, is invertible. *This is the only check that is needed.*

Though it is a well known fact, many text books tend only to make passing reference to the equivalence of ANOVA and regression techniques. An exception is Wu and Hamada (2000) which contains some simple examples illustrating and highlighting the estimation aspects of ANOVA.

We illustrate this approach for the orthogonal array design, OA(12, 2^4.3), given in Wu and Hamada (2000).

Table 5: OA(12, 2^4.3) Design (Columns a, b, c, d, e)

μ	a	b	c	d	e_1	e_1	ab	ac	ad	ae_1	ae_2
1	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	0	0	0	0
1	1	0	1	1	0	0	0	1	1	0	0
1	1	1	1	0	0	0	1	1	0	0	0
1	0	0	1	1	1	0	0	0	0	0	0
1	0	1	1	0	1	0	0	0	0	0	0
1	1	0	0	1	1	0	0	0	1	1	0
1	1	1	0	0	1	0	1	0	0	1	0
1	0	0	1	0	0	1	0	0	0	0	0
1	0	1	0	1	0	1	0	0	0	0	0
1	1	0	0	0	0	1	0	0	0	0	1
1	1	1	1	1	0	1	1	1	1	0	1

The basic design of OA(12, 2^4.3) is given in Table 5 and is defined by **a**, **b**, **c**, and **d** (level 2 factors) and the two $\mathbf{e}_1$ and $\mathbf{e}_2$ columns corresponding to the two degrees of freedom of the 3-level factor. Here we have used bold-faced symbols to denote the columns of the matrix. If we add the column of the general mean **μ** to these columns then we see that the regression model, $\mathbf{y} = \mathbf{X}\,\boldsymbol{\theta} + \boldsymbol{\varepsilon}$ in this case has the design matrix:

$$\mathbf{X}_0 = [\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{d}\ \mathbf{e}_1\ \mathbf{e}_2\,]$$

and that the sum of squares matrix $\mathbf{X}_0^T\mathbf{X}_0$ is non-singular. Thus all the main effect parameters are estimable, if there are no first order interactions.

Suppose now that we are interested in the first order interactions AB, AC, AD, AE, AE^2 associated with factor A. We can do this by adding the corresponding interaction columns **ab**, **ac**, **ad**, $\mathbf{ae}_1$ and $\mathbf{ae}_2$ to $\mathbf{X}_0$ We find that if we exclude just $\mathbf{ae}_2$, and take the design matrix to be

$$\mathbf{X}_1 = [\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{d}\ \mathbf{e}_1\ \mathbf{e}_2\ \mathbf{ab}\ \mathbf{ac}\ \mathbf{ad}\ \mathbf{ae}_1\,],$$

then its sum of squares matrix $\mathbf{X}_1^T\mathbf{X}_1$ is still invertible, indicating that these interactions can also be estimated in the design. However, if we include $\mathbf{ae}_2$, then

$$\mathbf{X}_2 = [\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{d}\ \mathbf{e}_1\ \mathbf{e}_2\ \mathbf{ab}\ \mathbf{ac}\ \mathbf{ad}\ \mathbf{ae}_1\ \mathbf{ae}_2\,]$$

has corresponding $\mathbf{X}_2^T\mathbf{X}_2$ that is *not* invertible. Thus not all the main effects and all first order interactions associated with A can be estimated.

7 SUMMARY AND CONCLUSIONS

In summary, we have pointed out how the design of real-time simulation trials is constrained by a certain inherent structure dictated by time and space dependent resource constraints. Nevertheless it is important to ensure that the experimental design is robust against chance variations outside the control of the experimenter, and the main approach suggested is the use of *hybrid* designs.

The example of section 4 illustrates how such a design can be constructed starting with a standard design, but where this design has to be modified using a number of design techniques including: (i) fractional factorial design, (ii) foldover (iii) pseudofactors.

The example of section 5 illustrates how mixed level designs are handled in a similar way, and shows how they are just as amenable to hybridization.

The two examples reinforce the view that practical design cannot be easily made automatic. However they do suggest that a more general and flexible *regression* approach can be formulated based on the estimability test of the parameters of the regression model.

We have not consider the issue of robustness. However the regression approach indicates that a simple way of handling this is to construct simulated data sets which incorporate realistic scenarios involving chance variations of the selected design and even missing data. The estimability test can then be extended into a full analysis of these artificial data sets to investigate robustness issues. It is hoped to discuss this idea more fully elsewhere.

REFERENCES

Banks, J. et al. 1998. *Handbook of simulation: principles, methodology, advances, applications and practice.* ed. J. Banks. New York: John Wiley and Sons.

Barton, R. R. 1999. *Graphical methods for the design of experiments.* New York: Springer-Verlag.

Barton, R. R. 2002. Designing Simulation Experiments. In *Proceedings of the 2002 Winter Simulation Conference*

E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds. 45-51. Piscataway, New Jersey: IEEE.

Connor, W. S. and Young, S. Fractional Factorial Designs for Experiments with Factors at Two and Three Levels. *National Bureau of Standards Applied Mathematics Series 58.*

Kleijnen, J. P. C. 1987*. Statistical tools for simulation practitioners.* New York: Marcel Dekker.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*, 3rd ed. New York: McGraw-Hill.

McLean, R. A. and Anderson, V. L. 1984. *Applied Factorial and Fractional Designs.* New York: Marcel Dekker.

Montgomery, D. C. 1997. *The design and analysis of experiments,* 4th Ed. New York: John Wiley and Sons.

Sanchez, S. M. 2000. Robust design: seeking the best of all possible worlds. In *Proceedings of the 2000 Winter Simulation Conference, ed. J. A. Joines, R. R. Barton, K. Kang and P. A. Fishwick, 69-76.* Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Wu, C. F. J. and Hamada, M. 2000. *Experiments Planning, Analysis, and Parameter Design Optimization.* New York: Wiley.

AUTHOR BIOGRAPHIES

RUSSELL C. H. CHENG is Professor, Head of Operational Research, and Deputy Dean of the Faculty of Mathematical Studies at the University of Southampton. He has an M.A. and the Diploma in Mathematical Statistics from Cambridge University, England. He obtained his Ph.D. from Bath University. He is a former Chairman of the U.K. Simulation Society, a Fellow of the Royal Statistical Society, Member of the Operational Research Society. His research interests include: variance reduction methods and parametric estimation methods. He was a Joint Editor of the *IMA Journal of Management Mathematics*. His email and web addresses are <R.C.H.Cheng@maths.soton.ac.uk> and <www.maths.soton.ac.uk/staff/Cheng>.

OWEN D. JONES is a Lecturer in Operational Research at the University of Southampton. He is a Fellow of the Cambridge Commonwealth Trust and a member of the Australian Mathematical Society. His research interests are in stochastic analysis, in particular for self similar processes. His email and web addresses are <O.D.Jones@maths.soton.ac.uk> and <www.maths.soton.ac.uk/staff/ODJones>.