Tilburg University

The role of statistical methodology in simulation

Kleijnen, J.P.C.

Publication date:
1978

EW

faculteit der economische wetenschappen

RESEARCH MEMORANDUM

TILBURG UNIVERSITY

DEPARTMENT OF ECONOMICS

Postbus 90135 - 5000 LE Tilburg
Netherlands
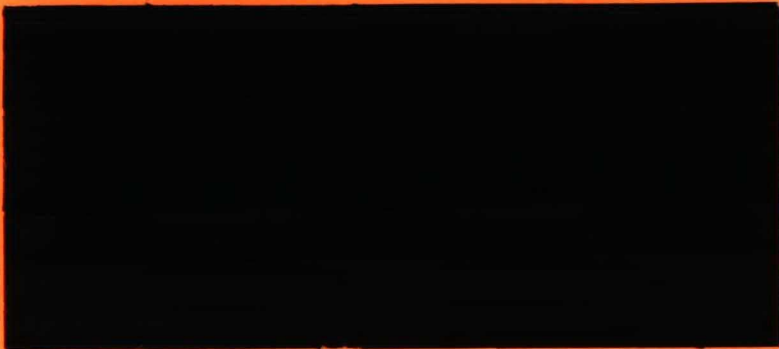
THE ROLE OF STATISTICAL METHODOLOGY IN
SIMULATION

Jack P.C. Kleijnen

R 35

T statistical methods

T simulation

THE ROLE OF STATISTICAL METHODOLOGY IN SIMULATION

Jack P.C. Kleijnen
Department of Business and Economics
Katholieke Hogeschool
Hogeschoollaan 225
5000 LE Tilburg
Netherlands

June 1978

THE ROLE OF STATISTICAL METHODOLOGY IN SIMULATION

by Jack P.C. KLEIJNEN
Department of Business and
Economics
Katholieke Hogeschool
5000 LE Tilburg
Netherlands

ABSTRACT

Statistical methods relevant to both digital and hybrid simulation are presented, using a minimum of formulas. A strategic issue is the ad hoc character of simulation. Statistical methods are surveyed which help to generalize and interpret simulation output data. Moreover statistical tools can show which system variants should be simulated, in order to obtain an understanding of the simulated system configurations. For Monte Carlo simulations some tactical problems are discussed: runlength and variance reduction.

More specifically, the ad hoc character of simulation is mitigated by a formal metamodel (auxiliary model), for which familiar regression analysis is used. The metamodel may include interactions among factors in the simulation experiment, and can be tested for its adequacy. Selecting the values of the input variables is the domain of experimental design. An example demonstrates that seven factors can be examined in only sixteen rather than $2^7=128$ runs. Situations with, say, a thousand factors require special screening designs. Requirements for "optimal" designs are briefly discussed.

In stochastic simulation two tactical problems exist: Variance reduction can be achieved through special techniques such as common random numbers, antithetic variates, control variates (regression sampling), importance sampling

(virtual measures). The variance of the simulation output is further affected by the simulation runlength, which leads to questions such as: are we really interested in steady-state behavior, how can we compute a confidence interval for the simulation output, how should we initialize the run, etc.

The survey includes 47 references, many just recently published, and provides references to a number of practical applications of statistical methods.

CONTENTS

Abstract
1. Strategic problems in simulation
2. Tactical problems in stochastic simulation
3. Formal metamodels: regression analysis
4. Experimental design
5. Summary of strategic problem
6. Tactical problems in stochastic simulation: variance reduction
7. Tactical problems: runlength
8. Miscellaneous statistical problems
9. Conclusion
Appendix: Survey of simulation types
Notes
References

1. STRATEGIC PROBLEMS IN SIMULATION

At this conference both digital and hybrid simulation are covered. We hope that the present paper is relevant to the two simulation areas, though we shall emphasize digital simulation, especially discrete-event simulation. In practice, simulation is a method very frequently used to study complex systems[1]. However, a major practical drawback of simulation is its ad hoc character: After spending much

mental effort and computer time to develop, program and run a computer simulation - strictly speaking - the results are valid only for the specific parameter values and structural relationships of the executed simulation program. Changing a parameter or relationship means that the simulation program has to be run again. Nevertheless, such changes are necessary to answer "what if" questions, or to find optimal system configurations, or - not to be forgotten - to establish the sensitivity of the conclusions to specific model assumptions.

Even after a great many simulation runs have been performed, it is difficult to obtain a general understanding of how the simulated system works: During the construction of the simulation model and its program much knowledge about the details of specific system components is acquired. However, insight into the behavior of the total system requires execution of the simulation program. These computer runs yield a mass of data but this mass may turn into a mess. Hence the output data should be summarized by a limited number of measures such as averages, peaks, and correlation coefficients (or spectra). In this way the various system configurations (system variants) are characterized by a few "statistics". Remain the problems of how to determine whether system variants show significantly different outputs, how to discern "patterns" in output changes as system configurations change, and so on.

In this paper we shall present a statistical methodology for generalizing the results of simulation experiments. Because of the survey character of our paper we shall avoid the use of mathematical-statistical formulas as much as possible. As we shall see, the accuracy of the resulting generalizations (metamodels) can be made explicit. Moreover, we shall present a systematic and efficient methodology for the exploration of the great many systems that can be simulated. In Kleijnen (1977) we further described how this methodology fits in the sequence of mental modeling - formal

modeling - computer programming - program running - system
understanding - optimization - implementation. Before we pro-
ceed to the formal presentation of our methodology, we in-
troduce a different problem area in simulation.

2. TACTICAL PROBLEMS IN STOCHASTIC SIMULATION

In Appendix 1 we survey several types of simulation
for those readers who do not feel familiar with terms such
as discrete-event, stochastic, difference-equation models,
etc. Besides the "strategic" problems in simulation models
of any type, there are some nagging "tactical" issues, if
the simulation is of the Monte Carlo type, i.e., when the
program contains stochastic variables generated by means
of pseudo-random numbers. Even if we concentrate on a single
system configuration with all parameters fixed, we have to
decide on the simulation run length. Once we terminate the
run, we wish to know the output's accuracy, specified by a
statistical confidence interval. The determination of this
accuracy may be complicated by serial correlations (auto-
correlations) among successive simulation observations so
that simple statistical methods are misleading. However,
as we shall see, the majority of practical, non-academic
simulation experiments can be analyzed without sophisticated
statistical analysis techniques. The reason is that most
practical studies do not concern long run, steady-state
behavior.

Even with modern high-speed digital computers,
Monte Carlo simulation may require runs of such lengths that
computing time becomes a bottleneck. We might then try to
apply special statistical techniques to reduce the variabili-
ty of the output, and hence the required runlength. This
is the area of Variance Reduction Techniques (VRT's), also
known as Monte Carlo Techniques, to be discussed in section
6.

3. FORMAL METAMODELS: REGRESSION ANALYSIS

A real-life system can be modeled by means of a
simulation model. The relationship between the inputs and
outputs of the simulation program (model) can in turn be mo-
deled by a metamodel or auxiliary model, to meet the strate-
gic demands mentioned in section 1. Several types of meta-
models are surveyed in Kleijnen (1977): common sense graphi-
cal approach, tables with two or three factors, explicit
formal metamodels such as the Meisel & Collins (1973) piece-
wise linear approximations. In this paper, however, we con-
centrate on linear regression metamodels. The common-sense
graphical approach is related to the regression approach, as
follows.We may change one factor,say x;observe the resulting
output y; repeat this procedure a number of times; plot the
(x, y) combinations; fit a curve by hand; and conclude
whether x has an important effect on y. The regression
approach formalizes this hand-fitting by applying the least
squares algorithm. It extends the procedure into multiple
dimensions. It further systematizes the various steps, in-
cluding tests for the importance (significance) of factors
together with their interactions, and tests for the adequa-
cy of the fitted regression metamodel. Linear regression
analysis has the great advantage of being a familiar tech-
nique for most scientists: Regression models have been
extensively applied to interpret and generalize experimental
results in agriculture, chemistry, engineering, psychology,
etc., where these models are also known as Analysis of
Variance (ANOVA).Regression metamodels in simulation have
been advocated by a few other authors. For instance, Week
& Fryer (1977) estimated "main" effects and "interactions"
in their job shop simulation study (effects to be defined
more precisely below). In their study regression models
are further utilized to answer "inverse" questions, i.e.,
which input values yield fixed, desired output values; see
also Kleijnen & Rens (1978), Kleijnen et al. (1978), Koons &

Perlic (1977) and Sherden (1976). The need for some kind
of metamodel to assist the more detailed model (simulation,
mathematical programming model) has been emphasized by several
more authors: Blanning (1974, 1975), Geoffrion (1976), Pegels
(1976, p 205).

Note that the variables in a model may be partitioned
into decision and environmental variables (factors). Decision
(control) variables are under the control of management. En-
vironmental variables are not under management control but
do influence the outputs (responses, criteria). In the long
run some environmental variables may become controllable,
for instance, arrival rates may be influenced by sales pro-
motion. The criterion variables are either satisfied or
optimized by a correct selection of the decision variables.
The sensitivity of this choice to the assumed environmental
factors, must also be investigated.

Let x denote a factor influencing the outputs of
the real-world system. The factor may be qualitative or quan-
titative, continuous or discrete. In Kleijnen (1975, p. 300)
it is shown how we can represent a qualitative factor by
several dummy variables assuming only the values zero or one.
We shall concentrate on qualitative factors (besides quanti-
tative factors) which are studied for only two "levels" or
"values". Then this qualitative factor can be represented
by a single dummy variable x with the values -1 and +1; see
Table 1 below. The response (output) of the real-world system
is a time-series. We shall concentrate on a single response
variable; for multiple outputs we apply our procedure to each
variable separately. In order to compare system configurations,
we characterize a whole time path by either a single measure
or a few measures: average, standard deviation, correlation
coefficients (spectra), slope of a fitted linear trend, peak,
etc. Let y_R denote such a measure, characterizing a time path
of the real-world system. Hence the response variable y_R is
a function of the factors x:

$$y_R = f_1(x_1, x_2, \ldots, x_m) \tag{3.1}$$

The real system is approximated by a <u>simulation model</u>, where $\underline{y}$ is a function of k factors x_j (j=1,...,k), plus a vector of random numbers $\vec{\underline{r}}$. (Vectors and matrices are denoted by $\rightarrow$; stochastic variables are underlined.) Hence the simulation output $\underline{y}$ can be represented as

$$\underline{y} = f_2(x_1, x_2, \ldots, x_k, \vec{\underline{r}}) \tag{3.2}$$

where k is much smaller than the unknown m in eq. (3.1), and $\vec{\underline{r}}$ symbolizes the joint effect of all factors x in eq. (3.1) not explicitly represented in eq. (3.2). The simulation model is specified by a computer program denoted by the function f_2. This model may be approximated in turn by a <u>metamodel</u> (within a specific experimental area E; see below). We propose a metamodel that is <u>linear</u> in its <u>parameters</u> β. This linearity does not mean that the metamodel is linear in its variables x; see eq. (3.5) below. Before we proceed to various specific metamodels we observe that the metamodel approach also applies when no sampling is used so that $\vec{r}$ in eq. (3.2) vanishes; see Appendix 1 for the various simulation types.

The simplest metamodel to express the effects of the k factors would be:

$$\underline{y}_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \underline{e}_i \qquad (i=1,\ldots,N) \tag{3.3}$$

where in simulation run i (observation i) factor j has the value x_{ij} (j=1,...,k) and $\underline{e}_i$ represents the noise (disturbance, error) in the metamodel which is assumed to have zero expectation. Such a simple metamodel implies that a <u>change</u> in x_j has a constant effect on the expected response, $\&(\underline{y})$:

$$\frac{\partial[\&(\underline{y})]}{\partial x_j} = \beta_j \qquad (j=1,\ldots,k) \tag{3.4}$$

(a) $\beta_{12} = 0$: no interaction

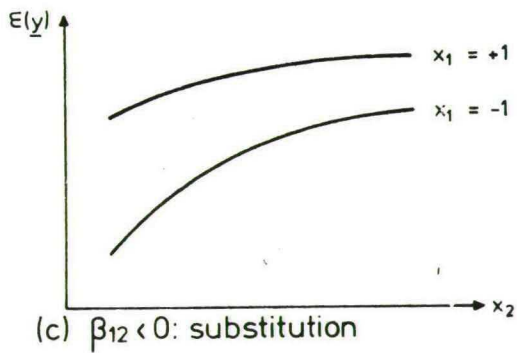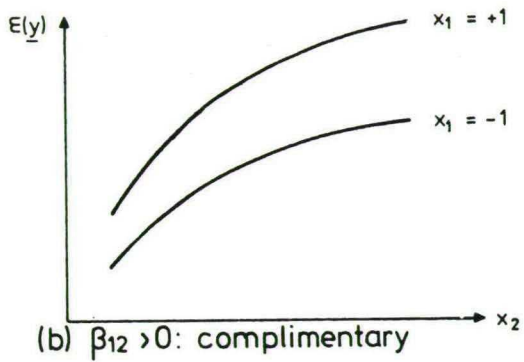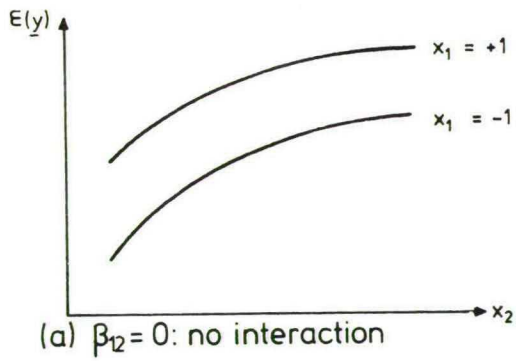(b) $\beta_{12} > 0$: complimentary

(c) $\beta_{12} < 0$: substitution

FIG. 1 Interactions

A more general metamodel postulates that the effect of factor
j also depends on the values of the other factors j' (j'$\neq$ j).
This can be formalized as in eq. (3.5) where for illustration
purposes we take k = 3:

$$\underline{y}_i = \beta_0 + (\beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}) +$$

$$+ (\beta_{12} x_{i1} x_{i2} + \beta_{13} x_{i1} x_{i3} + \beta_{23} x_{i2} x_{i3}) + \underline{e}_i$$

$$(i=1,\ldots,N) \quad (3.5)$$

Here the parameters (coefficients) β_{12}, β_{13} and β_{23} denote
interactions between the factors 1 and 2, 1 and 3, and 2 and
3 respectively. A graphical illustration of interaction in
the case of two factors (k=2), is shown in FIG. 1. In case
(a) of FIG. 1 the curves are parallel, i.e., the effect of
x_2 on $\&(\underline{y})$ does not depend on the level of x_1. In case (b)
the interaction coefficient β_{12} is positive. Hence the two
factors are complimentary, i.e., as x_2 increases the increase
in $\&(\underline{y})$ is stimulated when the increase of x_2 is accompanied
by an increase in x_1. In case (c) the marginal output of x_2
is much smaller when more of x_1 is available which can be
substituted for x_2.[2] The need to consider interactions among
factors when analyzing simulation results, has been empha-
sized by Koons & Perlic (1977) in their case-study of a steel
plant.

      If all factors are quantitive, continuous variables,
then we add "purely quadratic" effects β_{jj} to eq. (3.5). This
yields

$$\underline{y}_i = \beta_0 + \sum_{j=1}^{3} \beta_j x_{ij} + \sum_{j<j'}^{2\ 3} \beta_{jj'} x_{ij} x_{ij'} +$$

$$+ \sum_{1}^{3} \beta_{jj} x_{ij}^2 + \underline{e}_i \qquad (3.6)$$

which represents the Taylor series expansion of eq. (3.2),
cut off after the second-degree terms. For an application
of eq. (3.6) we refer to Koons & Perlic (1977, p.7). We
feel that in practice it is rare that all factors are quan-
titative, so that we shall concentrate on the metamodel with
k main effects β_j, $k(k-1)/2$ two-factor interactions $\beta_{jj'}$,
and the general (overall) mean β_0. In symbols:

$$\underline{y}_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{j < j'}^{k-1} \sum^{k} \beta_{jj'} x_{ij} x_{ij'} + \underline{e}_i$$

$$(i=1,\ldots,N) \qquad\qquad (3.7)$$

We start by assuming a metamodel, such as eq. (3.7),
but next we test statistically whether this assumption was
realistic! Two statistical tests can be used:
(1) Generate some new observations $\underline{y}$ using the simulation
model. Use the familiar Student t-statistic to compare these
observations $\underline{y}$ to the predicted value $\hat{\underline{y}}$ based on the re-
gression metamodel which was estimated from the old obser-
vations.[3)]
(2) A so-called lack-of-fit F-statistic can be computed
which compares the "mean residual sum of squares" to the
"pure error".
For details on both approaches we refer to Kleijnen et
al. (1978).

If the assumed metamodel turns out to be unreaso-
nable, we have several alternatives:
(1) Make the metamodel more complicated by adding terms such
as three-factor interactions. If $\underline{y}'$ is a shorthand notation
for $\underline{y}$ in eq. (3.5), then we may expand (3.5) to

$$\underline{y}_i = \underline{y}'_i + \beta_{123} x_{i1} x_{i2} x_{i3} \qquad\qquad (3.8)$$

The intuitive interpretation of three-factor interactions
is difficult. Moreover these interactions mean additional
parameters. Hence we prefer the next alternative.

(2) Look for transformations of x. For instance, if $\underline{y}$ de-
notes waiting time, and x_1 and x_2 denote mean arrival and
service rate, then the transformation $x' = x_1/x_2$ implies that
the original metamodel

$$\underline{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \underline{e} \qquad (3.9)$$

is replaced by a new metamodel:

$$\underline{y} = \gamma_0 + \gamma_1 x' + \underline{e}' \qquad (3.10)$$

We strongly recommended to look for such transformations,
from the very beginning of the study! The transformation
can be based on the relevant theory, for instance, queuing
analysis in the above example. A popular transformation in
econometrics is $x' = \log x$, $\underline{y}' = \log \underline{y}$, so that the para-
meters β represent elasticity coefficients. A more com-
plicated example is provided by Yen & Pierskalla (1977). A
simple transformation is also utilized in the case study by
Kleijnen et al. (1978).

(3) Reduce the experimental area E. This option limits the
generality of our conclusions. However, if the only purpose
of the metamodel is to find the optimum x-values, then a
small area E can be used, a metamodel fitted, and the direc-
tion of better x-values determined. See Montgomery & Betten-
court (1977) for details on this so-called Response Surface
Methodology (RSM). For a bibliography we also refer to
Kleijnen (1975).

Note that, after we have used the metamodel to meet
the demands of sensitivity analysis, optimization, and so
on, we can return to the original simulation model to study
the system behavior in detail, e.g., to study its dynamics.

The parameters β in the above equations can be estimated and tested for their significance, using the familiar technique of regression analysis, applying either Ordinary or Generalized Least Squares; see e.g. Draper & Smith (1966). Least Squares is also summarized in Kleijnen et al. (1978).

5. EXPERIMENTAL DESIGN

As we mentioned in section 1 parameters, variables, and structural relationships in the simulation model are changed in order to perform sensitivity analysis, answer what-if questions, do optimalization studies, etc. "Infinitely" many system configurations (system variants or briefly systems) may be of potential interest. Even with modern fast computers simulating all system variants is out of the question. The problem of selecting a limited number of system variants for actual simulation evaluation can be solved by means of statistical methods known as experimental design methodology. Experimental design theory has been developed since the 1920's and has been widely applied to experiments in agriculture, chemistry, etc. Unfortunately, its application to the management and social sciences is still in its infancy. The reason is that in sociotechnical systems the scientific design of experiments is difficult and expensive (disruption of the organization). However, in a simulation model of such a social system, the experimental factors are completely under the scientist's control, so that experimental design methods become relevant. Let us consider some simple examples.

If we study the effects of just a few factors, say three, then we may start by letting each factor assume only two "values" or "levels" denoted by $x = +1$ and $x = -1$ respectively, and evaluating the responses at all combinations, i.e., at $2^3 = 8$ combinations. (Remember that one factor combination specifies one system variant to be simulated.) But,

if we are willing to assume that the "first order" model of
eq. (3.3) is adequate, then we need to estimate only four
effects, namely β_0, β_1, β_2 and β_3. Before reading on, the
reader is challenged to specify his own selection of the
x-combinations! Next he may compare his selection to Table 1.
This table is constructed using a "trick" developed in ex-
perimental design theory: The last column x_3 is obtained by
multiplying the corresponding elements in the x_1 and x_2
columns. Such tricks certainly become necessary as the num-
ber of factors increses, since in that case the number of
combinations N grows dramatically. For instance, for seven
factors we have $N = 2^7 = 128$. Table 2 displays a design for
seven factors[4]: If we assume a first order model, then only
the first eight combinations need to be evaluated, in order to
estimate β_0, β_1, ...,β_7. If we leave open the possibility that
such a metamodel is inadequate, then the next eight combi-
nations should also be evaluated. The first eight combinations
form a so-called 2^{7-4} fractional factorial design of "resolu-
tion III". The sixteen combinations together form a 2^{7-3}
fractional factorial of "resolution IV", and can give us an
idea of the importance of interactions besides main effects;
see Kleijnen (1975) for more details. Observe that a design
matrix such as Table 1 also specifies the cross-products
x_1x_2, x_1x_3 and x_2x_3 in eq. (3.5).

      The above phase of the investigation may be preceded
by a pilot or scréening phase. In the latter stage a great
many, say a hundred or a thousand factors may be conceived of,
but hopefully only relatively few are really important. De-
tecting these important factors can be based on special ex-
perimental designs: group-screening, random designs, etc.
Random designs (randomly selected factor combinations) were
applied in a water-resource simulation; Maass et al. (1962).
Group-screening [5] was utilized in the simulation of computer
systems - see Schatzoff & Tillman (1975) - but applications
of this class of designs are extremely rare. Nevertheless we
imagine that it is quite common to have simulation models

Table 1
Experimental Design for Three Factors

Combination	x_1	x_2	x_3 $(= x_1 x_2)$
1	+1	+1	+1
2	-1	+1	-1
3	+1	-1	-1
4	-1	-1	+1

Table 2
Experimental Design for Seven Factors

Combination	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+
9	+	+	+	-	-	-	+
10	-	+	+	+	+	-	-
11	+	-	+	+	-	+	-
12	-	-	+	-	+	+	+
13	+	+	-	-	+	+	-
14	-	+	-	+	-	+	+
15	+	-	-	+	+	-	+
16	-	-	-	-	-	-	-

with a great many parameters and variables, which could
benefit from group-screening designs. Several types of screening
designs are evaluated in Kleijnen (1975 b).

Let us briefly evaluate the designs derived in the
statistical design literature. The traditional, "common sense"
approach is to change one factor at a time: ceteris paribus
approach. We proposed to change several factors simultaneous-
ly; see Tables 1 and 2. Such factorial designs (full or
fractional) are more efficient, i.e., they yield more accurate
estimators of the factor effects,and they provide estimates
of possible interactions among factors. Some problems, however,
remain: Specific designs such as 2^{k-p} designs, and the con-
comittant regression analysis yield "optimal" results only
under certain statistical assumptions such as constant va-
riances. How robust are these optimality properties and
what are the alternatives? Ad hoc optimal designs specified
by computer, generalized least squares, robust estimation
procedures, etc. are surveyed in Kleijnen et al. (1977).
Designs such as in Tables 1 and 2 may be evaluated against
the following requirements:

(1) A small number of runs N: Obviously, to estimate q para-
meters it is necessary that $N \geq q$. However, N may be much
smaller than 2^k, for instance, in Table 2 N = 16 whereas
$2^7 = 128$.

(2) Maximum statistical accuracy, given the number of runs:
If the classical statistical assumptions hold, then the
accuracy requirement is satisfied by choosing an orthogonal
design; otherwise the selection of the design matrix poses
a problem not yet solved. Note that Tables 1 and 2 do yield
orthogonal columns.

(3) Providing a measure for the adequacy of the fitted meta-
model: If N > q then a lack-of-fit F-statistic exists. If
besides the N observations we have one or more runs, not
used in the estimation of the parameters, then "validation"
of the model is possible, using a t-test.

(4) Desirable "confounding" (bias, alias pattern): If not
all factor effects can be estimated from N runs only, the
main effects should be biased by high order effects (say,

three-factor interactions), not by other main effects. The designs derived by experimental design theory immediately show how effects are biased by other effects.

(5) Flexibility of the design: Unfortunately, in many standard designs the number of runs N is restricted to a power of two (2^{k-p} designs) or a multiple of four. Fortunately, it remains possible to start with only a small number of runs, to test the results, and to proceed to a larger design that yields more detailed estimates, so-called sequentialized designs; see Kleijnen (1975, pp. 344-345, 367-370).

(6) Numerical inaccuracy caused by an ill-conditioned matrix $\overset{*}{X}$: An orthogonal matrix $\overset{*}{X}$ eliminates such problems. When using normalized variables (between -1 and +1) we should not forget to translate the estimated effects back into the original effects; see Kleijnen et al. (1978).

Note that, if the assumed metamodel turns out to be completely misleading, then the "optimal" properties of the experimental design break down. For instance, if the interaction between the factors 1 and 2 in Table 1 is actually important then we cannot estimate the main effect of factor 3 since it is completely confounded with that interaction. To reduce the possibility of such events, preliminary experimentation and analysis is necessary; see also Kleijnen (1975, pp. 391-393) and Kleijnen et al. (1978).

In conclusion, the literature on experimental design is overwhelming, and still growing! As a sample we mention the recent textbook Daniel (1976). In Kleijnen (1975, pp. 287-450) we have given a selection from the vast literature, tailored to the needs of the simulation practitioner: The focus is on simple designs such as 2^{k-p} designs, excluding sophisticated designs such as "partially balanced incomplete block designs". Excluded are techniques not needed in simulation, e.g., randomization and blocking are needed in case of incomplete control over the experiments as in areas outside simulation.

5. SUMMARY OF STRATEGIC PROBLEM

Formal metamodels are a useful technique for generalizing and interpreting simulation output. An important aspect of this interpretation is the concept of factor interaction. Efficient exploration of the simulation space requires an experimental design. Work on statistical designs is abundant but unfamiliar to the majority of simulation practitioners. Our experience is that the necessary statistical techniques can be learned without too many problems. Observe, however, that these techniques alone cannot solve the problem for the scientist! The models and hypotheses to be evaluated by the statistical techniques, have to be provided by management or by other, non-statistical specialists. The use of the techniques leaves much freedom: freedom of interpretation, choice of significance levels α, etc. Hence an automated application of techniques is impossible. Moreover, all statistical techniques are based on certain statistical assumptions such as constant variances, which are not satisfied in practice. It remains a challenge to develop more general and robust techniques. In the mean time, the practioner must use his judgement in the selection and use of his statistical tools. Nevertheless we feel that these tools result in a more efficient exploration of the experimental area, and in a better idea of both the limitations and the generalizations of the simulation experiment. In this way, one important drawback of simulation is mitigated, namely its ad hoc character. For an elaborated case study using a variety of statistical techniques, we refer to Kleijnen et al. (1978).

6. TACTICAL PROBLEMS IN STOCHASTIC SIMULATION: VARIANCE REDUCTION

In the following sections we shall focus on problems arising when simulating one specific system configu-

ration, i.e., a <u>single</u> factor combination. The problems we shall discuss arise in <u>stochastic</u> simulation models only: runlength determination in relation to estimation of the variances of the simulation response, and reduction of that variance through special statistical techniques, so-called <u>Variance Reduction Techniques (VRT's)</u>.

Though we devoted a doctoral dissertation to the <u>issue of VRT's</u> - see Kleijnen (1975, pp. 105-285) - over the years we have grown very pessimistic as to the practicality of such techniques. Note that computer time can be saved (and variance reduced through additional runs) by other devices than VRT's, e.g., more efficient random number generation, faster sampling procedures, better software. In Kleijnen (1975, pp. 105-285) we discussed six VRT's in detail including some applications, and provided references to many more techniques. Four VRT's will be briefly discussed in the present paper.

(1) <u>Common random numbers</u>

A system configuration may be simulated using the same random number seed as in the other system variants, so that systems are compared "under the same circumstances". This is the only VRT often applied by practitioners. Kleijnen (1975) discussed the practical complication of synchronizing random number streams per type of stochastic process. A complication overlooked by most practitioners, is that the <u>analysis</u> of the simulation results gets more difficult when the outputs become dependent, e.g., ordinary least squares assumes independent responses y. The following VRT is (nearly) as simple as the use of the same random numbers, but does not complicate the analysis.

(2) <u>Antithetic variates</u>

Suppose the first run of a specific system configuration is generated from the random number stream r_0, r_1, $r_2, \ldots$ and yields the result y_1. Then the "antithetic" run is generated from the complements $1-r_0$, $1-r_1$, $1-r_2, \ldots$ and yields y_2. The idea is, that when y_1 happens to undershoot

its expected value, then y_2 is expected to overshoot that value.(Example: In run 1 most r's happen to be small, so that service times are short, and waiting times are short. In run 2 most r's are large, etc.) Statistically speaking, $\underline{y}_1$ and $\underline{y}_2$ are conjectured to be <u>negatively</u> correlated, so that the variances of their average decreases. The statistical analysis remains simple since it can be based on the n/2 averages of the antithetic pairs (y_1, y_2) (y_3, y_4), ..., (y_{n-1}, y_n). Kleijnen (1975a) discussed the surprising fact that it is not necessarily optimal to combine antithetic variates and common random numbers when comparing <u>two</u> system variants. Recently Schruben & Marjolin (1977) investigated the joint application of these two VRT's when investigating N system variants in an experimental design. They found that applying either common random numbers only, or a specific combination[6] of common and antithetic random numbers, reduced the estimated variance by 80% compared to independent random number streams.

(3) <u>Control variates or regression sampling</u>

During a simulation run we may keep track of the average value $\hat{\mu}$ of the <u>input</u> variable, say, interarrival time. Note that the expected value μ is known, since we sample the input from a known distribution function. If we wish to estimate, say, average waiting time η for a specific average input value μ, then we may correct our estimate via the regression model

$$\underline{y}_i = \beta_0 + \beta_1 \cdot \underline{x}_i + \underline{u}_i \qquad (i=1,\ldots,n) \qquad (6.1)$$

where $\underline{y}_i$ is the average waiting time of run i, $\underline{x}_i$ is the average interarrival time ($\underline{x} \equiv \hat{\underline{\mu}}$) of run i, $\underline{u}$ is an error term, and the β's are regression coefficients. Hence

$$\hat{\underline{y}}_\mu = \hat{\underline{\beta}}_0 + \hat{\underline{\beta}}_1 \cdot \mu$$

$$= \bar{\underline{y}} + \hat{\underline{\beta}}_1 (\mu - \bar{\underline{x}}) \qquad (6.2)$$

where $\bar{y}$ is the "crude" average response of the n simulation responses y_i, which is corrected by the "control variable" x. We found a variance reduction factor of 3.84 in a case study, a telephone exchange simulation; however, the estimator (6.2) does involve some statistical complications as shown in Hopmans & Kleijnen (1977).

(4) Importance sampling and virtual measures

There is one class of systems that might benefit very much from variance reduction, viz., systems where we are interested in "rare events" such as "excessive" waiting times, inventory stockouts, etc. During most of the simulation run nothing of interest happens. A VRT especially relevant for such systems seems importance sampling (IS), and a closely related technique known as "virtual measures". These techniques sample more frequently that part of the time path during which rare events tend to occur more frequently (and correct for that oversampling). In a recent case study we applied this idea to a telephone-exchange simulation studying blocking probabilities (all lines occupied). Unfortunately, the practical results were very disappointing; see Hopmans & Kleijnen (1978).

7. TACTICAL PROBLEMS: RUNLENGTH

Under the heading "runlength" we shall discuss a set of related questions such as:
- How long to continue the simulation run?
- How to start the run (initialization)?
- How often to replicate (repeat) the run with different random number seeds?
- How accurate is the estimated response (confidence intervals)?

An important remark to start with, is that in practical, as opposed to academic, simulations these questions can often be answered using only elementary statistical techniques, such as t-statistics.

In practice simulation models are usually
terminating, i.e., the simulation run is stopped when a
specific event occurs. Simple examples are:

(1) In studying maintenance policies the simulation run may
end when the equipment (say, a computer) breaks down. A new
run starts with a "perfect" piece of equipment.

(2) A queuing (waiting-line) system such as a bank or hospi-
tal clinique, is closed at 5 P.M. (critical event). The new
run corresponds with a new day, starting in the "empty"
state.

(3) A corporate simulation model can be utilized to examine a
policy's effect on profit over the next three months (planning
horizon). The simulation run of three months is repeated
for different policies (what-if), starting each run from
the most recent "situation" (system state). Note that cor-
porate models are often non-stochastic.

(4) Queuing systems that never close down are, e.g., a te-
lephone exchange and a highway crossing. Such systems may
be simulated to see whether the system configuration can
handle peak traffic. As soon as the rush hour is over (cri-
tical event, defined fuzzily), the simulation run is ter-
minated. A next run starts from a pre-rush-hour situation.
The same or different random numbers (see section 6 on
variance reduction) may be utilized in that next run, to
study the sensitivity to the starting conditions (and
to the random number seed). Example 4 seems to be the most
problematic example, and may merit additional research.

In the above examples there is no interest in
steady-state responses! Steady-state (stationary, equilibri-
um, long-run) behavior means that the distribution function
(probability law) does not change over time. In the above
examples, however, start-up and end effects are part of the
relevant output. Each simulation run yields a single obser-
vation on the output, say, the average waiting time or the
total profit. (The relevance of transient behavior is also em-
phasized in Fox (1978) and Lam & Pedersen (1977).) If the simulation

is stochastic, more accurate estimates can be obtained by
repeating (replicating) the run with different random num-
bers (possibly antithetic, see section 6). The statistical
analysis is straightforward for terminating systems. For ex-
ample, if y_i denotes average waiting time in run $i (i=1,...,n)$,
then its standard deviation(standard error) is estimated by

$$\underline{s}_y = \{ \sum_{i=1}^{n} (\underline{y}_i - \bar{\underline{y}})^2 / (n-1) \}^{\frac{1}{2}} \tag{7.1}$$

A confidence interval for the expected value η can be based
on the Student t-statistic:

$$P\{\eta \leq \bar{\underline{y}} + t_{n-1}^{\alpha} \cdot \underline{s}_y / \sqrt{n}\} = 1-\alpha \tag{7.2}$$

For an application we refer to Kleijnen (1978) where Chapter
IX concerns a simulation experiment with an IBM management
game, used to study the financial benefits of accurate in-
formation.
          If the confidence interval in (7.2) turns out to
be too long, we may increase the accuracy of the average
simulation output $\bar{\underline{y}}$ by generating additional runs. The total
number of runs for a fixed length c of the confidence inter-
val, should be

$$\underline{n} = \{t_{n-1}^{\alpha}/c\}^2 \cdot \underline{s}_y^2 \tag{7.3}$$

For additional comments[7] on such a sequential approach we
refer to Kleijnen (1975).

          In the simulation literature most attention is fo-
cussed on steady-state behavior. Such behavior is primarily
of academic interest: Simulation is used by many academics
in the study of analytic models such as queuing models; see
the examples in Ignall et al. (1978). Transient behavior of

such models is difficult to analyse - Kotiah(1978), Liittschwayer &
Ames(1975) - so that most academic studies concentrate on steady-state(li-
miting) behavior.Simulations with the (practical!) aim of
assisting such theoretical studies, are confronted with
serious problems: Should the simulation be continued or
should replicated runs be used? Replicated runs yield inde-
pendent observations (on, say, steady-state average waiting
time) but each run creates an initialization problem
(transient behavior). A single prolonged run consists of
many dependent individual responses: autocorrelation or
serial correlation problem. For instance, if customer i
has to wait "very" long (longer than average) then the next
customer probably has to wait longer too (positive corre-
lation). Elementary statistical techniques assuming indepen-
dence, are misleading in that case. The variance of the
average of the continued run $\bar{w}$ based on m autocorrelated
individual observations w is given by:

$$\text{var}(\bar{w}) = \{1+2 \sum_{j=1}^{m} (1 - \frac{j}{m})\rho_j\}\sigma_w^2/m \qquad (7.4)$$

where ρ_j is the autocorrelation between w_t and w_{t+j}, and
σ_w^2 is the variance of an individual observation w. If the
w's were independent ($\rho_j = 0$), then the variance of their
average would reduce to the familiar expression σ_w^2/m. So
the autocorrelation ρ inflates the variance of the average,
and should not be ignored. Several approaches are possible.

(1) <u>Repeated runs</u>

        As mentioned above, replicated runs result in in-
dependent observations so that the analysis becomes simple.
However, each run is confronted with the initialization
problem. References on this approach and its alternatives
will be given below.

(2) <u>Prolonged run with (nearly) independent subruns of fixed length</u>

Assume that the serial correlation among the individual observations decreases as the observations are farther apart. Divide the total run (usually after removing the initial phase as in approach 1 above), into subruns of fixed length. Though the first "few" observations of a subrun still depend on the last "few" observations of the preceding subrun, the subrun <u>averages</u> will be independent, practically speaking, provided the subrun length is "long enough".Therefore a subrun length may be selected intuitively; the autocorrelation among subrun averages be tested (through the estimated autocorrelation coefficient of lag 1 or through von Neumann's ratio); if the autocorrelation is too high (empirical threshold) then the subrun length is increased, etc. A recent paper · on this approach is Law & Carson (1977). In practice, approach 2 is often followed, but with the subrun length being selected purely intuitively. Relying on intuition alone seems dangerous: Analytical results for simple queuing systems demonstrate that for heavy traffic individual observations remain correlated over surprisingly long lags. On the other side, overestimating the autocorrelation means that the subruns are too long, so that too few subruns remain.[8]

(3) <u>Prolonged run without subrun distinction, but with estimated individual autocorrelations</u>

Eq. (7.4) displayed the effects of the autocorrelation coefficients ρ_j ($j=1,2,...$) among the individual observations. In the sixties Fishman and Kiviat, at that time with the RAND Corporation, published several reports in which the variance of the average simulation output was based on the estimation of those coefficients ρ_j (or their Fourier transformations: spectral analysis). For a recent discussion we refer to Clark (1977). In practice this approach has never been popular: cumbersome estimation of the ρ's; difficult selection of m (the number of ρ's to be incorporated).

(4) <u>Prolonged run with truly independent subruns based on</u>
<u>the renewal property</u>

Both Iglehart and Fishman have pioneered the appli-
cation of the renewal or regenerative property shown by many si-
mulated systems. Consider a queuing system that has become
empty. Then the next history (timepath) is independent of
the past history! Consequently, the total run can be divided
into subruns, each subrun starting as soon as the system
has become empty. In contrast to the subruns of fixed length
(approach 2) the new subrun definition creates subruns of
stochastic lengths: when does the system return to the
empty state? These subruns are exactly independent! The
estimation of the total run's average involves some statis-
tical complications: ratio estimators for point estimation,
jackknifing for confidence intervals, etc. Estimating
quantiles[9] such as the 90%-point, involves some more pro-
blems; see Seila (1978) and also Coppus et al. (1977). Per-
centiles (e.g., the probability of queue-sizes exceeding
the waiting room capacity) are studied by Fishman & Moore
(1977). However, it is our experience (with graduate students in
management science and econometrics) that these statistical
complications are easily overcome. In general, any Markov
system shows the renewal property. However, practical
problems remain: The renewal state may occur so infrequently
that too few subruns result; see Hopmans & Kleijnen (1978)
for an example. Approximative renewal states may then be
formulated; see Gunther (1975). A recent textbook on the
renewal approach is Crane & Lemoine (1978); see also
Fishman (1978). Since we feel that this is an important
technique for the exact analysis of steady-state (academic)
simulations, we mention some more applications: Lavenberg
& Slutz (1975) and Schwetman & Bruell (1976). A case study
(a simple time sharing system) comparing this approach to
replicated runs (approach 1) is provided by Sargent (1977).

The initialization problem was mentioned several
times in the present section. Remember that this problem
exists primarily in steady-state, academic simulations. An
example of initialization in a practical simulation for
planning purposes, is provided by Jain (1975, p. 85): "at
the start of a simulation the model represents the actual
state of the machine shop." Note further that in the renewal
approach to steady-state simulations, there is no start-up
problem: observations can be collected immediately when the
simulation starts in the renewal state (e.g. the empty
state). In the other three approaches the transient phase
does pose a problem: Usually initial observations are thrown
away though this is not necessarily optimal. Recently
Wilson & Pritsker (1977) investigated a variety of heuristics,
and concluded that it seems best to start the simulation
run in the most likely stationary state, and to retain all
observations (transient and steady-state).

8. MISCELLANEOUS STATISTICAL PROBLEMS

In the above sections we concentrated on those
statistical problems we thought to be most relevant in
the analysis and design of simulation experiments. In the
present section we briefly discuss some remaining issues.

(1) Multivariate responses
In practice a number of criteria and measures are of interest,
e.g., in a queuing situation we may be interested in both
waiting time and server utilization, maesured by their
means and their 90% quantiles, etc. Though sophisticated
statistical tools exist (e.g. Multivariate Analysis of Va-
riance or MANOVA), it is practical to use univariate tech-
niques and to account for the multivariate character by the
choice of an appropriate error rate (Bonferroni inequality);
see Kleijnen (1975) for more details.

(2) <u>Multiple comparison and ranking procedures</u>

In the first few sections we discussed situations where a number of factors define a great many system variants of potential interest. Relevant techniques are regression analysis and experimental design. A different situation exists, when there are only a few system variants, say, 10, or in general k.   These systems may correspond to different queuing disciplines, etc. Multiple comparison procedures (MCP) are suited to situations with k systems (or populations in statistical jargon) and a <u>fixed</u> number of simulation runs (observations) per system. MCP give exact statistical results (controlled α errors) when comparing $k(k-1)/2$ systems with each other, when selecting a subset containing the "best" system (say, highest mean response), etc. Multiple ranking or selection procedures (MRP) have been developed for situations where the number of runs is not fixed, but has to be determined such that the best system can be selected with a prespecified probability of correct selection. Both MCP and MRP are discussed at length in Kleijnen (1975); a recent publication is Dudewicz (1977). At present simulation practitioners have shown little interest in these procedures; the only references we are aware of are Lin (1975) and Vicéns & Schaake (1972).

(3) <u>Statistical input: random numbers, etc.</u>

We have investigated the statistical analysis of the simulation <u>output</u>. On the input side we have the traditional problems of random number generation (multiplicative and shift back generators) and sampling from distributions (including multivariate distributions). References can be found in Kleijnen (1975).

(4) <u>Model validation</u>

Checking whether the model's output conforms with the real world observations can be based on a variety of statistical techniques: t-tests, goodness-of-fit tests, regression analysis, etc. However, this issue involves many more aspects than just statistics; see, e.g., Zeigler (1976).

9. CONCLUSION

Simulation means experimentation, albeit experimentation using a mathematical model instead of the real world. Any experiment requires a sound design. Without such a design even the most sophisticated analysis fails, e.g., if the factors 1 and 2 are changed simultaneously, their separate effects cannot be estimated. Scientific designs such as 2^{k-p} factorials further make it possible to explore the simulation space much more efficiently. The statistical analysis, given the design, should extract as much information from the experiment as is possible, e.g., estimate interactions. Such an analysis can be done systematically by means of a formal metamodel, i.e., a regression model. Moreover, such an analysis shows the limitations of the conclusions. For instance, if the simulation run is too short, so much stochastic noise may be present that instead of an expensive simulation model, a toss of the coin had better been used.

APPENDIX 1: SURVEY OF SIMULATION TYPES

We distinguish the following types of simulation; see also Kleijnen (1976):

(1) Discrete-event models, e.g., queuing systems modeled by a language such as GPSS. Most of the time these models are inherently stochastic, i.e., without the probabilistic character of certain variables the queuing problems would disappear (a scheduling problem would remain).

(2) Time-slicing models of discrete-event systems. It might be convenient to model, say, a highway interchange, by dividing the time-axis into time slices of equal length, and checking which events, if any, occurred.

(3) Difference-equation models, e.g., national econometric and corporate models. Decisions and data for such systems

are of a periodic nature: quarterly, yearly, etc. Usually
these models are deterministic.

(4) Difference-equation models for differential equation
systems: Forrester's Industrial or System Dynamics using
DYNAMO, is a well-known example. Nearly always these models
are deterministic.

(5) Digital computer models (by means of difference equa-
tions) of continuous phenomena in science and technology:
the spontaneous model is formulated in differential equations
based on laws of nature known from physics etc. For various
reasons a digital computer may be preferred over an analogue/
hybrid computer in the solution of the model. Languages
such as CSMP and CSSM are used on the digital computer.
Note that it is rare that the opposite occurs: an ana-
logue computer (differential equation) used to solve diffe-
rence equation problems of types 1 through 4.

NOTES

1) In Operations Research, for instance, several surveys
   have shown that simulation is a most popular technique
   (together with linear programming and statistical tech-
   niques such as regression analysis); see Ledbetter &
   Cox (1977) for a recent survey and for additional re-
   ferences.

2) Actually the curves in FIG. 1 are no straight lines
   so that they represent more general formulations than
   eq. (3.5).

3) These "new" observations might correspond to the "center"
   of the design ($x_j = 0$ for all factors), in order to
   check whether pure quadratic effects are zero.

4) In the first eight combinations we use $4 = 12$, $5 = 13$,
   $6 = 23$ and $7 = 123$. In the next eight combinations we switch
   signs: $x_{i'j} = -x_{ij}$ ($i'=i+1$) ($i=1,\ldots,8$).

5) Suppose the individual factors are x_1, x_2,... and we know the signs of the β's, e.g., $\beta_1 \geq 0$ and $\beta_2 \geq 0$. Then x_1 and x_2 can be combined into a single group-factor z_1 with main effect $\gamma_1 = \beta_1 + \beta_2 (\geq 0)$. Hence z_1 is +1 (and -1 respectively) if all its component factors are +1 (and $x_1 = x_2 = -1$ respectively). Execute only two runs (instead of $2^2 = 4$ runs):

      Run 1: $x_1 = -1$ $x_2 = -1$ → $z_1 = -1$

      Run 2: $x_1 = +1$ $x_2 = +1$ → $z_2 = +1$

If responses do not differ significantly, we can conclude that neither x_1 nor x_2 are important, and eliminate x_1 and x_2 from further experimentation. In general, k individual factors x can be combined into g group-factors z which can be tested in a 2^{g-p} fractional factorial design; see Kleijnen (1975b).

6) Split the N design points in two orthogonal blocks, and run one block with common random numbers, and the other block with the antithetic numbers.

7) Notice that n in (7.2) is deterministic, whereas $\underline{n}$ is stochastic in (7.3). Nevertheless (7.3) gives satisfactory results.

8) When the number of subruns is M, then the variance of the estimated variance is $2\sigma^4/M$ so that the confidence interval for the mean η becomes less stable (but has the same expected length).

9) For the definition of quantiles and percentiles consider $P(x < y) = z$. If z $(0 \leq z \leq 1)$ is fixed to, say, 90% then we have to estimate the "quantile" y. If we fix y, then we have to estimate the "percentile" z.

REFERENCES

BLANNING, R.W., The sources and uses of sensitivity information. INTERFACES, 4, no. 4, August 1974, pp. 32-38. (See also 5, no. 3, May 1975, pp. 24-25.)

CLARK, G.M., AN IMPROVED PROCEDURE FOR ESTIMATING THE
VARIANCE OF A TIME-SERIES AVERAGE CALCULATED FROM SAMPLE
AUTOCOVARIANCES OF SIMULATION OUTPUT DATA. Department of
Industrial and Systems Engineering, The Ohio State Univer-
sity, Columbus, 1977.

COPPUS, G.,M. VAN DONGEN and J.P.C. KLEIJNEN, Quantile
estimation in regenerative simulation: a case study. PER-
FORMANCE EVALUATION REVIEW, 5, no. 3, Summer 1977, pp. 5-15.
(Reprinted in SIMULETTER, 8, no. 2, Jan. 1977, pp. 38-47.)

CRANE, A. and J. LEMOIGNE, AN INTRODUCTION TO THE REGENE-
RATIVE METHOD FOR SIMULATION ANALYSIS. Springer-Verlag,
Berlin, 1978.

DANIEL, C., APPLICATIONS OF STATISTICS TO INDUSTRIAL EX-
PERIMENTATION. John Wiley & Sons, Inc., New York, 1976.

DRAPER, N.R. and H. SMITH, APPLIED REGRESSION ANALYSIS. John
Wiley & Sons, Inc., New York, 1966.

DUDEWICZ, E.J., New procedures for selection among (simula-
ted) alternatives. PROCEEDINGS 1977 WINTER SIMULATION CON-
FERENCE, edited by H.J. HIGHLAND, R.G. SARGENT and J.W.
SCHMIDT, A.C.M., New York, 1977, pp. 59-62.

FISHMAN, G.S., Achieving specific accuracy in simulation
output analysis. COMMUNICATIONS ACM, 20, no. 5, May 1977,
pp. 310-315.

FISHMAN, G.S. and L.R. MOORE, ESTIMATING THE MEAN OF A COR-
RELATED BINARY SEQUENCE WITH AN APPLICATION TO DISCRETE
EVENT SIMULATION. Technical Report No. 77-2, Curriculum on
Operations Research and Systems Analysis, University of
North Carolina, Chapel Hill, April 1977.

FOX, B., Estimation and simulation. MANAGEMENT SCIENCE, 24, no. 8, April 1978, pp. 860-861.

GEOFFRION, A.M., The purpose of mathematical programming is insight, not numbers. INTERFACES, 7, no. 1, Nov. 1976, pp. 81-92.

GUNTHER, F.L., THE ALMOST REGENERATIVE METHOD FOR STOCHASTIC SYSTEM SIMULATIONS. Research Report no. 75-21, Operations Research Center, University of California, Berkeley, Dec. 1975.

HOPMANS, A.C.M. and J.P.C. KLEIJNEN, REGRESSION ESTIMATORS IN SIMULATION. Report FEW-70, Department of Economics, Katholieke Hogeschool, Tilburg (Netherlands), Dec. 1977.

HOPMANS, A.C.M. and J.P.C. KLEIJNEN, IMPORTANCE SAMPLING IN SYSTEMS SIMULATION: A PRACTICAL FAILURE? Report Department of Economics, Katholieke Hogeschool, Tilburg (Netherlands), June 1978.

IGNALL, E.J., P. KOLESAR and W.E. WALKER, Using simulation to develop and validate analytic models: some case studies. OPERATIONS RESEARCH, 26, no. 2, March-April 1978, pp. 237-253.

JAIN, S.K., A simulation-based scheduling and management information system for a machine shop. INTERFACES, 6, no. 1, part 2, Nov. 1975, pp. 81-96.

KLEIJNEN, J.P.C., STATISTICAL TECHNIQUES IN SIMULATION. (In two parts.) Marcel Dekker, Inc., New York, 1974/1975.

KLEIJNEN, J.P.C., Antithetic variates, common random numbers and optimum computer time allocation. MANAGEMENT SCIENCE, APPLICATION SERIES, 21, no. 10, June 1975 (a), pp. 1176-1185.

KLEIJNEN, J.P.C., Screening designs for poly-factor experimentation. TECHNOMETRICS, 17, no. 4, Nov. 1975 (b), pp. 487-493.

KLEIJNEN, J.P.C., Discrete simulation: types, applications, and problems. In: SIMULATION OF SYSTEMS, edited by L. DEKKER, North-Holland Publishing Company, Amsterdam, 1976.

KLEIJNEN, J.P.C., GENERALIZING SIMULATION RESULTS THROUGH METAMODELS. Working Paper 77.070, Department of Business and Economics, Katholieke Hogeschool, Tilburg (Netherlands), Dec. 1977.

KLEIJNEN, J.P.C., COMPUTERS AND PROFITS: QUANTIFYING FINANCIAL BENEFITS OF INFORMATION.(provisional title; to be published) 1978.

KLEIJNEN, J.P.C. and P.J. RENS, IMPACT revicited: a critical analysis of IBM's inventory package "IMPACT". PRODUCTION AND INVENTORY MANAGEMENT, 19, no. 1, first quarter, 1978, pp. 71-90.

KLEIJNEN, J.P.C., A.J. VAN DEN BURG en R.T. VAN DER HAM, Generalization of simulation results: practicality of statistical methods. EUROPEAN JOURNAL OPERATIONAL RESEARCH, 1978 (to appear).

KOONS, G.F. and B. PERLIC, A STUDY OF ROLLING-MILL PRODUCTIVITY UTILIZING A STATISTICALLY DESIGNED SIMULATION EXPERIMENT. Research Laboratory, United States Steel Corporation, Monroeville (Penn.), (1977?).

KOTIAH, T.C.T., Approximate transient analysis of some queuing systems. OPERATIONS RESEARCH, 26, no. 2, March-April 1978, pp. 333-346.

LAM, C.F. and J. PEDERSEN, Continuous simulation of a complex queuing system. SIMULATION, $\underline{29}$, no. 2, August 1977, pp. 42-48.

LAVENBERG, S.S. and D.R. SLUTZ, Regenerative simulation of a queuing model of an automated tape library. IBM JOURNAL OF RESEARCH AND DEVELOPMENT, $\underline{19}$, Sept. 1975, pp. 463-475.

LAW, A.M. and J.S. CARSON, A SEQUENTIAL PROCEDURE FOR DETER-MINING THE LENGTH OF A STEADY-STATE SIMULATION. WP 77-12, Department of Industrial Engineering, University of Wisconsin, Madison, April 1977.

LIITTSCHWAGER, J.M. and W.F. AMES, On transient queues: practice and pedagogy. In: COMPUTER SCIENCE AND STATISTICS, EIGHTH ANNUAL SYMPOSIUM ON THE INTERFACE, edited by J.W. FRANE, Health Sciences Computing Facility, University of California, Los Angeles, 1975.

LIN, W.T., MULTIPLE OBJECTIVE BUDGETING MODELS: A SIMULATION. Working Paper 12-01-75, Department of Accounting, Graduate School of Business Administration, University of Southern California, Los Angeles, 1975.

LEDBETTER, W.N., and J.F. COX, Operations research in production management: an investigation of past and present utilization. PRODUCTION AND INVENTORY MANAGEMENT, $\underline{18}$, no. 3, third quarter, 1977, pp. 84-92.

MAASS, A., et al.,DESIGN OF WATER-RESOURCE SYSTEMS. Harvard University Press, Cambridge (Massachusets), 1962.

MEISEL, W.S. and D.C. COLLINS, Repro-modeling: an approach to efficient model utilization and interpretation. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, vol. SMC-3, no. 4, July 1973, pp. 349-358.

MONTGOMERY, D.C. and V.M. BETTENCOURT, Multiple response surface methods in computer simulation. SIMULATION, 29, no. 4, Oct. 1977, pp. 113-121.

PEGELS, C.C., SYSTEMS ANALYSIS FOR PRODUCTION OPERATIONS. Gordon and Breach Science Publishers, New York, 1976.

SARGENT, R.G., Statistical analysis of simulation output data.   SIMULETTER, 8, no. 3, April 1977, pp. 21-31.

SCHATZOFF, M. and C.C. TILLMAN, Design of experiments in simulation validation. IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 19, no. 3, May 1975, p. 252-262.

SCHRUBEN, L.W. and B.H. MARJOLIN, PSEUDO-RANDOM NUMBER ASSIGNMENT IN STATISTICALLY DESIGNED SIMULATION AND DISTRIBUTION SAMPLING EXPERIMENTS. School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, Nov. 1977.

SCHWETMAN, H.D. and S.C. BRUELL, When to stop a simulation run: a case study. SIMULETTER, 7, no. 4, July 1976, pp. 131-137.

SEILA, A.F., ON THE PERFORMANCE OF TWO METHODS FOR QUANTILE ESTIMATION IN REGENERATIVE PROCESSES. Bell Telephone Laboratories, Inc., Holmdel (N.J. 07733), (1978?)

SHERDEN, W.A., Origin of simultaneity in corporate models. In: WINTER SIMULATION CONFERENCE, edited by H.J. HIGHLAND, T.J. SCHRIBER and R.G. SARGENT, ACM, New York, 1976.

VICENS, G.J. and J.C. SCHAAKE, SIMULATION CRITERIA FOR SELECTING WATER RESOURCE SYSTEMS ALTERNATIVES. Report no. 154, Ralph M. Parsons Laboratory, Department of Civil

Engineering, M.I.T., Cambridge (Mass.), Sept. 1972.

WEEKS, J.K. and J.S. FRYER, A methodology for assigning minimum cost due-dates. MANAGEMENT SCIENCE, 23, no. 8, April 1977, pp. 872-881.

WILSON, J.R. and A.A.B. PRITSKER, A PROCEDURE FOR EVALUATING STARTUP POLICIES IN SIMULATION EXPERIMENTS. Purdue University, Lafayette (Indiana),1977.

YEN, H.C. and W. PIERSKALLA, A SIMULATION OF A CENTRALIZED BLOOD BANKING SYSTEM. Michael Reese Hospital, Office of Operations Research, Chicago (Ill. 60616), 1977.

ZEIGLER, B.P., THEORY OF MODELLING AND SIMULATION. John Wiley & Sons, Inc., New York, 1976.