TILBURG UNIVERSITY
KATHOLIEKE UNIVERSITEIT BRABANT

POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

SENSITIVITY ANALYSIS OF SIMULATION EXPERIMENTS: REGRESSION ANALYSIS AND STATISTICAL DESIGN

Jack P.C. Kleijnen

FEW 440

SENSITIVITY ANALYSIS OF SIMULATION EXPERIMENTS:

REGRESSION ANALYSIS AND STATISTICAL DESIGN.

A contribution to the EUROPEAN SIMULATION MEETING

on

PROBLEM SOLVING BY SIMULATION,

organized by

International Association for Mathematics and
Computers in Simulation (IMACS)

ESZTERGOM, HUNGARY

23-30 August 1990

Jack P.C. Kleijnen
Katholieke Universiteit Brabant
(Tilburg University)
Postbox 80153
5000 LE Tilburg
The Netherlands

E-mail: kleijnen@kub.nl
Fax: 013-663072

SENSITIVITY ANALYSIS OF SIMULATION EXPERIMENTS: REGRESSION ANALYSIS AND STATISTICAL DESIGN

Jack P.C. Kleijnen
Katholieke Universiteit Brabant
(Tilburg University)
Tilburg, Netherlands

ABSTRACT

This tutorial gives a survey of strategic issues in the statistical design and analysis of experiments with deterministic and random simulation models. These issues concern what-if analysis, optimization, and so on. The analysis uses regression (meta)models and Least Squares. The design uses classical experimental designs such as 2^{k-p} factorials, which are efficient and effective. If there are very many inputs, then special techniques such as group screening and sequential bifurcation are useful. Applications are discussed.

INTRODUCTION

Simulation is a mathematical technique that is applied in all scientific disciplines that use mathematical modeling, ranging from sociology to astronomy; also see Karplus [1]. It is a very popular technique because of its flexibility, simplicity, and realism. By definition, simulation involves experimentation, namely with the model of a real system. Consequently it requires on appropriate <u>design and analysis</u>. For real systems mathematical statistics has been applied since the 1930's: Sir Ronald Fisher focussed on agricultural experiments in the 1930's; George Box concentrated on chemical experimentation, since the 1950's; see [2]. Tom Naylor organized a conference on the design of simulation experiments back in 1968; see [3]. In 1974/75 my first book [4] covered both the 'tactical' and 'strategic' issues of experiments with random and deterministic simulation models. The term tactical was introduced into simulation by Conway [5]; it refers to issues such as runlength and variance reduction, which arise only in random simulations such as queuing simulations. Strategic questions are: which combinations of input variables should be simulated, and how can the resulting output be analyzed? Obviously strategic issues arise in both random and deterministic simulations. Mathematical statatistics can be applied to solve these questions, also in deterministic simulation; see the recent publications [6]. [7], and [8]. This contribution focusses on these strategic issues in simulation experiments.

Strategic issues concern problems that are also addressed under names like model validation, what-if analysis, goal seeking, and optimization; see table 1, reproduced from my recent book [6, p. 136]. We shall return to this table.

REGRESSION METAMODELS

Before the systems analyst starts experimenting with the simulation model, he (or she) has accumulated prior knowledge about the system to be simulated: he may have observed the real system, tried different models, debugged the final simulation model, and so on. This tentative knowledge isformalized in a regression or Analysis of Variance (ANOVA) model. ANOVA models are elementary in the statistical theory of experimental design: Sums of Squares (SS's) are compared through the F test to detect significant main effects and interactions; see below. The simplest ANOVA models can be easily translated into regression models; see [6, pp. 263-293]. Because regression analysis is more familiar than ANOVA is, we shall use regression terminology henceforth.

Table 1: Terminology

Computer program	Simulation model	Regression model	User view
Output	Response	Dependent variable y	Result
Input	Parameter	Independent variable x	Environment
	Variable		
	Enumeration	Continuous	Validation Risk Analysis
	Function	Discrete	Controllable
	Scenario	Binary	Optimization
			Goal output (control) Satisfy (what-if)
	Behavioral relationship		

So prior knowledge is formalized in a tentative regression model. In other words, this model is tested later on to check its validity as we shall see. The regression model specifies which inputs seem important, which interactions among these inputs seem important, and which scaling seems appropriate. We shall discuss these items next.

Table 1 showed that 'inputs' are not only parameters and variables but may also be 'behavioral relationships'; that is, a module of the simulation model may be replaced by a different module. In the regression model such a qualitative change is represented by one or more binary (0,1) variables. Note that 'inputs' are called 'factors' in experimental design terminology. 'Interaction' means that the effect of a factor depends on the values (or 'levels') of another factor:

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \sum_{g=1}^{k} \beta_{jg} x_j x_g +$$

$$+ \sum_{j=1}^{k} \sum_{g=1}^{k} \sum_{h=1}^{k} \beta_{jgh} x_j x_g x_h + \ldots + e, \tag{1}$$

where y is the simulation response; β_0 is the overall or grand mean; β_j is the main or first-order effect of factor j (j = 1,..., k); β_{jg} is the two-factor interaction between the factors j and g (g = 1,..., k; g ≠ j); β_{jj} is the quadratic effect of factor j; β_{jgh} is the three-factor interaction among the factors j, g, and h (h = 1,..., k; h ≠ g ≠ j); and so on; e denotes 'fitting 'errors' or noise. Under certain strict mathematical conditions the 'response curve' in Eq. (1) is a Taylor series expansion of the simulation model $y(x_1,\ldots,x_k)$. Unfortunately these conditions do not hold in simulation. Therefore we propose to start with an initial model that excludes interactions among three or more factors: such high-order interactions are popular in ANOVA but they are hard to interpret. The purpose of the regression model is to guide the design of the simulation experiment and to interpret the resulting simulation data; a regression model without high-order interactions suffices, as we observed repeatedly in practice.

The regression variables x in Eq. (1) may be <u>transformations</u> of the original simulation parameters and variables; for example, $x_1 = \log(z_1)$ where z_1 denotes the original simulation input. <u>Scaling</u> is also important: if the lowest value of z_1 corresponds with $x_1 = -1$ and its highest value corresponds with $x_1 = +1$, then β_1 measures the relative importance of factor 1 when that factor ranges over the experimental area. In optimization we explore the response curve only locally if we use <u>Response Surface Methodology</u> (RSM). Then the local regression model is a first-order model:

$$y = \gamma_0 + \Sigma \gamma_j z_j + e, \tag{2}$$

where the importance of factor j at $\bar{z}_j$, the midpoint of the local experiment, is measured by $\gamma_j \bar{z}_j$; obviously $\bar{z}_j = \Sigma_{j=1}^{n} z_{ij}/n = (L_j+H_j)/2$ where $L_j \le z_{ij} \le H_j$ with local experimental area $[L_1, H_1] \times \dots \times [L_k, H_k]$; z_{ij} denotes the value of factor j in simulation run or observation i. See [9] and [2].

In any experiment the analyst uses a model such as Eq. (1), explicitly or implicitly. For example, if he changes one factor at a time, then (implicitly) he assumes all interactions (β_{jg}, β_{jgh},...) to be zero. Of course it is better to make the regression model explicit and to find a design that fits that model, as we shall see next. But first note that we call the regression model a <u>metamodel</u> because it models the input/output behavior of the underlying simulation model.

EXPERIMENTAL DESIGN

Based on a tentative regression (meta)model we select an experimental design. The design matrix $\mathbf{D} = (d_{ij})$ specifies the n combinations of the k factors that are to be simulated. (In multi-stage experimentation such as RSM that set of n combinations is followed by a next set.) Classical statistical theory gives designs that are 'efficient' and 'effective'. <u>Efficiency</u> means that the number of combinations or simulation runs is 'small'. Suppose there are Q effects in the regression model. The number of runs should satisfy the condition $n \ge Q$; for example, we need $k + 1$ runs if there are no interactions at all. So we may do one base run (say) $\mathbf{x}'_0 = (-1, -1, \dots, -1)$; and then we change one factor at a time: $\mathbf{x}'_1 = (+1, -1, \dots, -1)$, $\mathbf{x}'_2 = (-1, +1, -1, \dots, -1), \dots, \mathbf{x}_k = (-1, \dots, -1, +1)$; see table 2 for k = 3. However, to estimate the effects $\underline{\beta}' = (\beta_0, \beta_1, \dots, \beta_k)$ we fit a curve to the simulation data (**X**, y) where $\mathbf{X} = (\mathbf{1}, \mathbf{D})$ in the first-order model; **1** denotes a vector of n ones. The classical fitting criterion is <u>Least Squares</u>. This criterion yields the estimator

$$\hat{\underline{\beta}} = (\mathbf{X}'\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{y}. \tag{3}$$

Now consider the classical 2^{3-1} design of Table 2. The corresponding **X** is orthogonal, so (3) reduces to the scalar expression

$$\hat{\beta}_{j'} = \Sigma_{i=1}^{n} x_{ij'} y_i/n \qquad (j' = 0,1,\ldots,k). \tag{4}$$

Table 2. Two designs for three factors.

	One at a time			2^{3-1} Design		
Run	d_1	d_2	d_3	d_1	d_2	d_3
1	-	-	-	-	-	+
2	+	-	-	+	-	-
3	-	+	-	-	+	-
4	-	-	+	+	+	+

How can we choose between the two designs of table 2? Classical theory assumes that the fitting errors e are <u>white noise</u>: e is normally, and independently distributed with zero mean and constant variance (say) σ^2. Then (3) yields the variance-covariance matrix

$$\mathbf{cov}(\hat{\underline{\beta}}) = (\mathbf{X}'\ \mathbf{X})^{-1}\ \sigma^2\ . \tag{5}$$

An orthogonal matrix **X** is <u>optimal</u>: it minimizes var $(\hat{\beta}_j)$, the elements on the main diagonal of Eq. (5); see [6, p.335]. There are straightforward procedures for deriving design matrices, if $n = 2^{k-p}$ with $(p=0,1,\ldots)$; for other n values results are tabulated; see [2] and [6].

So the classical designs are efficient under the white noise assumption (recent research uses alternative assumptions; see Sachs et al. [7]). Moreover, these designs are <u>effective</u>: they permit the estimation of interactions. If we allow for two-factor interactions, then the number of effects becomes $Q = 1 + k + k(k-1)/2$. If k is small, we may simulate $n \geq Q$ combinations; for example, if $k = 5$ then a 2^{5-1} design is suitable. (If k is large, then we may hope that some factors will turn out to give nonsignificant main effects; we may assume that factors without main effects have no two-factor interactions either; there are designs that yield unbiased estimators for main effects with $n = 2k < 1 + k + k(k-1)/2$; see [6, pp. 303 - 309], [9].) If the factors are quantitative, then a second-order regression model has k quadratic effects too. Then n must increase, and more than two levels per factor must be simulated: RSM designs; see [2] and [6].

SCREENING

For didactic reasons we discuss 'screening' designs _after_ classical experimental designs. In practice the simulation model has a great many factors that may be important; of course the analyst assumes that only a few factors are really important: parsimony. So in the beginning of a simulation study it is necessary to search for the few really important factors among the many conceivably important factors: _screening_.

Classical textbooks do not discuss screening situations, because in real-life experiments it is impossible to control (say) a hundred factors. In simulation, however, we perfectly control all inputs and we indeed use models with many inputs. One approach is _group screening_, introduced in the early 1960's by Watson, Jacoby and Harrison, Li, and Patel. This technique aggregates the many individual factors into a few group factors. Some simulation applications can be found in [6, p.327]. Recently Bettonvil [10] further developed group screening into _sequential bifurcation_, a very efficient technique that accounts for white noise and interactions. He applied this technique to an ecological model with nearly 300 factors.

REGRESSION ANALYSIS: TECHNICALITIES

Eq. (3) gave the Ordinary Least Squares (OLS) estimator $\hat{\underline{\beta}}$. In deterministic simulation that estimator may suffice, although Sachs et al. [7] give a better estimator if the white noise assumption is dropped (and replaced by a stationary covariance assumption). In random simulation the classical assumptions seldom hold. If the response variances differ with the inputs (as the response means do), then Weighted Least Squares (WLS) is better. If common random numbers drive the various factor combinations, then Generalized Least Squares (GLS) is best. See [6, pp. 161-175].

Once the regression model is calibrated (that is, the parameters β are estimated), the metamodel's validity must be tested. For deterministic simulation models we propose cross validation: delete factor combination i $(\mathbf{x}_i', y_i)$; reestimate $\underline{\beta}$ from the ramaining simulation data $(\mathbf{X}_{-i}, \mathbf{y}_{-i})$; predict the deleted response y_i through the reestimated regression model $(\hat{y}_i = \hat{\underline{\beta}}'_{-i}\mathbf{x}_i)$; "eyeball" the relative prediction errors $\hat{y}_i/y_i$: are these errors acceptable to the user?

In random simulation we prefer Rao's adjusted lack-of-fit F-test: the estimated response variances and covariances are compared with the residuals $(\hat{\mathbf{y}} - \mathbf{y})$; see [11].

SOME APPLICATIONS

Applications of our approach are getting numerous. For example, a simple - but realistic - case study concerns a Flexible Manufacturing System (FMS). Input to the FMS simulation is the 'machine mix', that is, the number of machines of type i with i = 1,...,4. Intuitively selected combinations of these four inputs give inferior results when compared with a classical design. The throughput predicted by the simulation is analyzed through two different regression models. These models are validated. A regression model in only two inputs but including their interaction, gives valid predictions and sound explanations [12].

Another application concerns a decision support system (DSS) for production planning, developed for a Dutch company. To evaluate this DSS, a simulation model is built. The DSS has 15 control variables that are to be optimized. The effects of these 15 variables are investigated, using a sequence of classical designs. Originally, 34 response variables were distinguished. These 34 variables, however, can be reduced to one criterion variable, namely productive machine hours, that is to be maximized, and one commercial variable measuring lead times, that must satisfy a certain side-condition. For this optimization problem the Steepest Ascent technique is applied to the experimental design outcomes. See [13].

A final case study concerns a set of deterministic ecological simulation models that require sensitivity analysis to support the Dutch government's decision making. First results for a model of the 'greenhouse' effect are given in [14]; additional results are given in [10].

CONCLUSIONS

Experimental design and regression analysis are statistical techniques that have been widely applied in the design and analysis of data obtained by <u>real life</u> experimentation and observation. In simulation, these techniques are gaining popularity: a number of case studies have been published. The techniques need certain adaptations to account for the peculiarities of deterministic and random simulations.

REFERENCES

[1] Karplus, W.J., The spectrum of mathematical models. <u>Perspectives in Computing</u>, <u>3</u>, no. 2, May 1983, pp. 4-13.

[2] Box, G.E.P. and N.R. Draper, <u>Empirical Model-Building And Response Surfaces</u>. John Wiley & Sons, New York, 1987.

[3] Naylor, T.H., editor, <u>The Design of Computer Simulation Experiments</u>, Duke University Press, Durham, N.C. 1969.

[4] Kleijnen, J.P.C., <u>Statistical Techniques in Simulation</u>. Marcel Dekker, Inc., New York, 1974/1975. (Russian translation, Moscow, 1978.)

[5] Conway, R.W., Some tactical problems in digital simulation. Management Science, 10, no. 1, Oct. 1963, p. 47-61.

[6] Kleijnen, J.P.C., Statistical Tools for Simulation Practitioners, Marcel Dekker, Inc., New York, 1987.

[7] Sachs, J., W.J. Welch, T.J. Mitchell, and H.P. Wynn, Design and analysis of computer experiments, Statistical Science, 4, no. 4, 1989, pp. 409-435.

[8] Kleijnen, J.P.C., Statistics and deterministic simulation: Why not? Katholieke Universiteit Brabant (Tilburg University), Tilburg, 1990.

[9] Bettonvil, B. and J.P.C. Kleijnen, Measurement scales and resolution IV designs. American Journal of Mathematical and Management Sciences (accepted).

[10] Bettonvil, B., Detection of Important Factors by Sequential Bifurcation. Doctoral dissertation, Katholieke Universiteit Brabant (Tilburg University). Tilburg, November 1990.

[11] Kleijnen, J.P.C., Regression metamodels for simulation with common random numbers: comparison of techniques. Katholieke Universiteit Brabant (Tilburg University), Tilburg, January 1990.

[12] Kleijnen, J.P.C. and C.R. Standridge, Experimental design and regression analysis in simulation: an FMS case study. European Journal of Operational Research, 33, 1988, pp. 257-261.

[13] Kleijnen, J.P.C., Simulation and Optimization in Production Planning: a Case Study, Katholieke Universiteit Brabant (Tilburg University), Tilburg, 1988.

[14] Rotmans, J. and O.J. Vrieze, Metamodelling, a sensitivity analysis on the greenhouse effect. European Journal of Operational Research (accepted).

IN 1989 REEDS VERSCHENEN

368 Ed Nijssen, Will Reijnders
"Macht als strategisch en tactisch marketinginstrument binnen de distributieketen"

369 Raymond Gradus
Optimal dynamic taxation with respect to firms

370 Theo Nijman
The optimal choice of controls and pre-experimental observations

371 Robert P. Gilles, Pieter H.M. Ruys
Relational constraints in coalition formation

372 F.A. van der Duyn Schouten, S.G. Vanneste
Analysis and computation of (n,N)-strategies for maintenance of a two-component system

373 Drs. R. Hamers, Drs. P. Verstappen
Het company ranking model: a means for evaluating the competition

374 Rommert J. Casimir
Infogame Final Report

375 Christian B. Mulder
Efficient and inefficient institutional arrangements between governments and trade unions; an explanation of high unemployment, corporatism and union bashing

376 Marno Verbeek
On the estimation of a fixed effects model with selective non-response

377 J. Engwerda
Admissible target paths in economic models

378 Jack P.C. Kleijnen and Nabil Adams
Pseudorandom number generation on supercomputers

379 J.P.C. Blanc
The power-series algorithm applied to the shortest-queue model

380 Prof. Dr. Robert Bannink
Management's information needs and the definition of costs, with special regard to the cost of interest

381 Bert Bettonvil
Sequential bifurcation: the design of a factor screening method

382 Bert Bettonvil
Sequential bifurcation for observations with random errors

383 Harold Houba and Hans Kremers
Correction of the material balance equation in dynamic input-output models

384 T.M. Doup, A.H. van den Elzen, A.J.J. Talman
Homotopy interpretation of price adjustment processes

385 Drs. R.T. Frambach, Prof. Dr. W.H.J. de Freytas
Technologische ontwikkeling en marketing. Een oriënterende beschouwing

386 A.L.P.M. Hendrikx, R.M.J. Heuts, L.G. Hoving
Comparison of automatic monitoring systems in automatic forecasting

387 Drs. J.G.L.M. Willems
Enkele opmerkingen over het inversificerend gedrag van multinationale ondernemingen

388 Jack P.C. Kleijnen and Ben Annink
Pseudorandom number generators revisited

389 Dr. G.W.J. Hendrikse
Speltheorie en strategisch management

390 Dr. A.W.A. Boot en Dr. M.F.C.M. Wijn
Liquiditeit, insolventie en vermogensstructuur

391 Antoon van den Elzen, Gerard van der Laan
Price adjustment in a two-country model

392 Martin F.C.M. Wijn, Emanuel J. Bijnen
Prediction of failure in industry
An analysis of income statements

393 Dr. S.C.W. Eijffinger and Drs. A.P.D. Gruijters
On the short term objectives of daily intervention by the Deutsche Bundesbank and the Federal Reserve System in the U.S. Dollar - Deutsche Mark exchange market

394 Dr. S.C.W. Eijffinger and Drs. A.P.D. Gruijters
On the effectiveness of daily interventions by the Deutsche Bundesbank and the Federal Reserve System in the U.S. Dollar - Deutsche Mark exchange market

395 A.E.M. Meijer and J.W.A. Vingerhoets
Structural adjustment and diversification in mineral exporting developing countries

396 R. Gradus
About Tobin's marginal and average q
A Note

397 Jacob C. Engwerda
On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$

398 Paul C. van Batenburg and J. Kriens
Bayesian discovery sampling: a simple model of Bayesian inference in auditing

399 Hans Kremers and Dolf Talman
Solving the nonlinear complementarity problem

400 Raymond Gradus
Optimal dynamic taxation, savings and investment

401 W.H. Haemers
Regular two-graphs and extensions of partial geometries

402 Jack P.C. Kleijnen, Ben Annink
Supercomputers, Monte Carlo simulation and regression analysis

403 Ruud T. Frambach, Ed J. Nijssen, William H.J. Freytas
Technologie, Strategisch management en marketing

404 Theo Nijman
A natural approach to optimal forecasting in case of preliminary observations

405 Harry Barkema
An empirical test of Holmström's principal-agent model that tax and signally hypotheses explicitly into account

406 Drs. W.J. van Braband
De begrotingsvoorbereiding bij het Rijk

407 Marco Wilke
Societal bargaining and stability

408 Willem van Groenendaal and Aart de Zeeuw
Control, coordination and conflict on international commodity markets

409 Prof. Dr. W. de Freytas, Drs. L. Arts
Tourism to Curacao: a new deal based on visitors' experiences

410 Drs. C.H. Veld
The use of the implied standard deviation as a predictor of future stock price variability: a review of empirical tests

411 Drs. J.C. Caanen en Dr. E.N. Kertzman
Inflatieneutrale belastingheffing van ondernemingen

412 Prof. Dr. B.B. van der Genugten
A weak law of large numbers for m-dependent random variables with unbounded m

413 R.M.J. Heuts, H.P. Seidel, W.J. Selen
A comparison of two lot sizing-sequencing heuristics for the process industry

IN 1990 REEDS VERSCHENEN

419 Bertrand Melenberg, Rob Alessie
A method to construct moments in the multi-good life cycle consumption model

420 J. Kriens
On the differentiability of the set of efficient (μ,σ^2) combinations in the Markowitz portfolio selection method

421 Steffen Jørgensen, Peter M. Kort
Optimal dynamic investment policies under concave-convex adjustment costs

422 J.P.C. Blanc
Cyclic polling systems: limited service versus Bernoulli schedules

423 M.H.C. Paardekooper
Parallel normreducing transformations for the algebraic eigenvalue problem

424 Hans Gremmen
On the political (ir)relevance of classical customs union theory

425 Ed Nijssen
Marketingstrategie in Machtsperspectief

426 Jack P.C. Kleijnen
Regression Metamodels for Simulation with Common Random Numbers: Comparison of Techniques

427 Harry H. Tigelaar
The correlation structure of stationary bilinear processes

428 Drs. C.H. Veld en Drs. A.H.F. Verboven
De waardering van aandelenwarrants en langlopende call-opties

429 Theo van de Klundert en Anton B. van Schaik
Liquidity Constraints and the Keynesian Corridor

430 Gert Nieuwenhuis
Central limit theorems for sequences with m(n)-dependent main part

431 Hans J. Gremmen
Macro-Economic Implications of Profit Optimizing Investment Behaviour

432 J.M. Schumacher
System-Theoretic Trends in Econometrics

433 Peter M. Kort, Paul M.J.J. van Loon, Mikulás Luptacik
Optimal Dynamic Environmental Policies of a Profit Maximizing Firm

434 Raymond Gradus
Optimal Dynamic Profit Taxation: The Derivation of Feedback Stackelberg Equilibria

435 Jack P.C. Kleijnen
Statistics and Deterministic Simulation Models: Why Not?

436 M.J.G. van Eijs, R.J.M. Heuts, J.P.C. Kleijnen
Analysis and comparison of two strategies for multi-item inventory systems with joint replenishment costs

437 Jan A. Weststrate
Waiting times in a two-queue model with exhaustive and Bernoulli service

438 Alfons Daems
Typologie van non-profit organisaties

439 Drs. C.H. Veld en Drs. J. Grazell
Motieven voor de uitgifte van converteerbare obligatieleningen en warrantobligatieleningen