…RSITY
…OLIEKE UNIVERSITEIT BRABANT

POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

[redacted]

STATISTICS AND DETERMINISTIC SIMULATION MODELS: WHY NOT?

Jack P.C. Kleijnen

FEW 435

Kenmerk: 320.90.161

STATISTICS AND DETERMINISTIC SIMULATION MODELS:

WHY NOT?

Jack P.C. Kleijnen
Katholieke Universiteit Brabant
(Tilburg University)
5000 LE Tilburg
(Bitnet: Kleijnen@KUB.NL)

March 1990

STATISTICS AND DETERMINISTIC SIMULATION MODELS: WHY NOT?

Jack P.C. Kleijnen
Katholieke Universiteit Brabant
(Tilburg University)
5000 LE Tilburg
The Netherlands
(Bitnet: Kleijnen@KUB.NL)

ABSTRACT

First deterministic simulation models are compared with random simulation models and real-life experiments. In deterministic simulation no mathematical statistics is needed in the experimental design and in the Least Squares curve fitting to the resulting data. Further analysis becomes possible if certain statistical models are specified for the fitting errors. Kleijnen (1987) proposed normally identically and indepently distributed errors. Sacks et al. (1989 a, b) proposed dependent errors with a specific correlation structure. Needs for further research are indicated.

1. INTRODUCTION

The unique characteristic of a deterministic simulation model is that its responses y are fixed, given its set of input variables $\mathbf{X} = \{x_1, \ldots x_j, \ldots, x_k\}$. Hence we get:

$$\mathrm{var}\ (y|\mathbf{X}) = 0. \qquad (1)$$

This characteristic distinguishes these models from random simulation models and real-life experiments. *Random simulation models* use pseudomrandom numbers; so different seeds produces different y values, in general. In the analyses of these models we treat the pseudorandom numbers as if they were truly random numbers, which are distributed uniformly and independently. Hence we view the response y as a random variable with variance

$$\mathrm{var}(y|\mathbf{X}) = g(\mathbf{X}). \tag{2}$$

Some authors assume that $\mathrm{var}(y|\mathbf{X})$ is independent of $\mathbf{X}$, so $\mathrm{var}(y|\mathbf{X})$ reduces to a constant (say) σ^2. We, however, prefer to assume that the variance depends on the combination of input values or $\mathbf{x}_i = \{x_{i1},...,x_{ij},...,x_{ik}\}$. So if we select an $n \times k$ experimental design matrix $\mathbf{D}$, then each design "point" i has its own variance σ_i^2. We can estimate this variance through replication (that is, we feed the simulation model with different random numbers). See Kleijnen (1987).

Risk Analysis is an interesting combination of deterministic simulation models and Monte Carlo sampling. Dynamic models may be deterministic. Examples are financial models (which have become popular since spreadsheet software has become widespread) and ecological models. These models, however, depend on a number of inputs which are unknown. Therefore the user specifies a (prior) distribution of possible values: that user may select a distribution type (normal, uniform, etc.). The computer samples values from that distribution, and generates output values, which are summarized in an output distribution. Also see Iman and Helton (1988).

In *real life experiments* repeated observations of the same system generate different responses y, because the system operates in a noisy environment; that is, the environment is not fully controlled and it perturbes the system.

We emphasize that a *random* simulation model is not completely different from a deterministic simulation model: every time we feed the random model with the same input combination $\mathbf{x}_i$ *and* the same pseudorandom number seed, we get the same response. We "solved" this problem by assuming that the pseudorandom number generator produces truly random numbers, *and* that the seed selection does not affect the distribution of these numbers.

If we feed a given deterministic simulation model with the input combination $\mathbf{x}_i$, it always generates the same response (say) y_i. So a deterministic simulation model provides a perfectly controlled world. Is mathematical statistics still relevant in such an "ideal" world? In section 2 we shall first show that basic ideas of the statistical theory of experimental design (originated by Fisher) still applies. The resulting simulation observations can be analyzed through Least Squares curve fitting and "eyeballing" of the results. Further analysis, however, requires statistical assumptions. In section 3 we shall present the statistical model proposed in Kleijnen (1987): the errors are normally, independently and identically distributed. Sacks, Schiller and Welch (1989) replaced the independence assumption by a stochastic process assumption; henceforth we shall refer to that article as Sacks et al. (1989a). In section 4 we shall sketch future research needs.

2. DESIGN AND ANALYSIS WITHOUT STATISTICS

In deterministic simulation models we should still apply the statistical theory of experimental design to select the input combinations. For example, it is obviously not smart to change two inputs (say) x_1 and x_2 simultaneously in the experiment. And changing one factor at a time does not allow the detection of any interactions among inputs. We also point out that many Response Surface designs have been derived under the assumption that noise can be neglected so that only bias is to be minimized; see Kleijnen (1987, p. 314) and Sacks, Welch, Mitchell and Wynn (1989, p. 420); the latter reference is abbreviated to Sacks et al. (1989b) further on.

In the *analysis*, however, we may use mathematical curve fitting and "eye balling" instead of statistical analysis (that analysis includes regression analysis and Analysis of Variance or ANOVA). So we may apply the mathematical criterion of Least Squares to fit a metamodel to the simulation data; those data consist of $\{\mathbf{x}_i, y_i\}$ with i=1,...,n. To validate the resulting "calibrated" metamodel we can use mathematical criteria such as the multiple correlation coefficient R^2. We can also predict the response for a new input combination $\mathbf{x}_{n+1}$, and "eyeball" the relative prediction error $\hat{y}_{n+1}/y_{n+1}$ where $\hat{y}_{n+1}$ is the value predicted by the metamodel and y_{n+1} is the response of the simulation model for the new input combination. A refinement is "cross validation": we delete combination i, fit the metamodel to the remaining simulation data $\{\mathbf{X}_{-i}, y_{-i}\}$; predict the deleted response through $\hat{y}_{-i}$; and "eyeball" the relative prediction error $\hat{y}_{-i}/y_i$; this procedure we repeat for i = 1,...,n. An example is Kleijnen and Standridge (1988) who discuss a deterministic simulation of Flexible Manufacturing Systems (FMS).

Once we have validated the model as a whole, we can study the individual Least Squares coefficients $\hat{\beta}_j$ (j=1,...,k). In order to determine which factors are most important, we may sort the coefficients $\hat{\beta}_j$, provided we have standardized the factors; see Bettonvil and Kleijnen (1989). If, however, we wish to identify coeffcients that are so small that they actually reflect "noise", then we need mathematical statistics; see the next section.

We can also use the regression model to *predict* the simulation responses at input values *not* contained in the simulation data $\{\mathbf{X}, y\}$. (In the validation we predicted responses for the "old" input X.) Usually we interpolate, not extrapolate, because the simulation data are based on an experimental design that includes extremal values: the "experimental area" covers the "area of interest". The predictors are

$$\hat{y}_2 = X_2 \hat{\underline{\beta}}, \tag{3}$$

where X_2 corresponds with the set of new input values (and $\hat{\underline{\beta}}$ is still the Least Squares estimator). These predictors are computed faster than the simulation responses $y_2 = h(X_2)$, where h(.) denotes the simulation model. For example, in a study of the Rotterdam container harbor we answered *ad hoc* management questions through the metamodel; in the beginning we were not completely sure that this approach was adequate, so we checked the predictions by running the expensive (random) simulation model overnight; see Kleijnen, Van den Burg and Van der Ham (1978).

To quantify the *uncertainty* of these predictions we must specify a statistical model for the fitting errors **e**: (3) yields

$$\mathrm{var}(\hat{y}_2) = X_2 \ \mathrm{cov}(\hat{\underline{\beta}}) \ X_2' \ , \tag{4}$$

where

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

In the next section we shall discuss statistical models for **e**, which also specify **cov** (**e**) in (5).

Note that Sachs et al. (1989a) introduced a more complicated predictor that has the nice property that at the observed input combinations the predictor equals the observed response: $\hat{y}(X) = y(X)$. In random simulation we do not expect such an equality, since the observed responses have zero probability of being observed again (when new seeds are employed). In random simulation we test whether $E[\hat{y}(X)] = E[(y(X)]$; see Kleijnen (1990).

3. STATISTICAL ERROR MODELS

Kleijnen (1987, pp. 163-164) discusses why a statistical error model may be appropriate in deterministic simulation: "Since infinitely many combinations of simulation parameters.... are possible, there are infinitely many errors e. The population of these errors has a specific variance, denoted by σ^2.... Now we *sample* the simulation parameters.... We may perform this sampling randomly or more or less systematically.... in the metamodelling of deterministic simulation we may model the independent variables as random variables. Consequently, the regression parameter estimator $\hat{\beta}$ being a function of **x** ... becomes random, and so does $\hat{\mathbf{y}}$ ... so that $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$ is random too ..."

Note that "random design" were elaborately discussed in *Technometrics*, back in 1959. In these designs the input combinations are sampled by flipping a coin so $P(x_{ij}=1) = 0.5$ and $P(x_{ij}= -1) = 0.5$; see Kleijnen (1987, pp. 321-323).

Sacks et al. (1989a, pp. 41-47) model the fitting errors "as a realization of a stochastic process in which the covariance structure of [e] relates to the smoothness of the response". Sacks et al. (1989a, p. 42) further assume that **e** is Gaussian with zero expectation. They assume that if two design points are further apart along one of the k axes, then the covariance of the two fitting errors decreases exponentially. Their procedure is computationally complex; it uses a supercomputer.

So Kleijnen (1987) proposes the same *marginal* distribution as Sacks et al. (1989a) assume. Kleijnen implicitly assumes independent errors, whereas Sacks et al. postulate a stationary stochastic process with a particular covariance function. For simplicity's sake we may prefer to stick to the model with independent errors that underlies Ordinary Least Squares (OLS(. Sacks et al.'s model seems more realistic, if the response surface is smooth. For, suppose the error is positive for some $\mathbf{x}_i$ in the k-dimensional space. Suppose further that we wish to interpolate for $\mathbf{x}_i + \varepsilon$, which is a point "close" to the point $\mathbf{x}_i$. Then $\hat{y}(\mathbf{x}_i+\varepsilon)$, the response predicted by OLS, tends to underestimate the true response $y(\mathbf{x}_i+\varepsilon)$. Sacks et al., procedure does not have that unattractive characteristic; unfortunately they must assume a *specific* covariance function; moreover, their computations for that function are formidable.

A less fundamental discussion of statistical errors in deterministic simulation can be found in Olivi (1984) and Olivi and Pike (1981). They distinguish two groups of independent variables, namely controlled variables that are supposed to be of major importance, and uncontrolled variables of minor importance. They sample the minor variables, which results in experimental error. Similarly Owen, Koehler and Sharifzadeh (1989) consider "the response as a function of the most important inputs, possibly with some noise due to the other inputs".

4. FUTURE RESEARCH

In the tradition of Popper a scientific model should be refutable. Therefore we should develop tests that allows us to *reject* the hypothesis that the fitting errors have a specific distribution; in this case we assume either independent normal errors or normal errors with a specific covariance structure. Also see Kleijnen (1987, pp. 178-179) and Sachs et al. (1989b. p. 417).

The stochastic process specification introduced by Sachs et al. (1989a) looks promising. *Practitioners* will probably apply this approach, once the conceptual and computational details have been worked out. The "classical" regression analysis and experimental design (based on independent errors) have already been applied to many simulation experiments, as is documented in Kleijnen (1987, p. 241).

In practice, simulation models often have *many inputs*. For example, Bettonvil (1990) investigates a deterministic ecological model with 281 inputs. Het assumes "white" noise: normally identically and independently distributed errors. Sachs et al. (1989a,b) limited their approach to small problems with, for example, six inputs. The computational burden of problems with many inputs seems formidable.

We emphasize that metamodels have two goals: prediction and explanation. For *prediction* purposes the metamodel is a black box; the only question is: does the black box predict "well"? *Explanation* means that the user gets insight into the behavior of the underlying simulation model. For example, Kleijnen and Standridge (1988, p. 261) report that the final metamodel (after the original metamodel was rejected) explained the behavior of the underlying Flexible Manufacturing System: "Statistical techniques... reduce the drawbacks of an empirical technique like simulation, i.e.,... the regression metamodel... helped the authors to better understand how an FMS works!". Sachs et al. (1989a,b) concentrate on prediction.

In summary, the statistical analysis of deterministic simulation data is controversial. Sachs et al. (1989b, p.435) state: "In earlier drafts we did attempt to discuss these philosophical matters [Bayesian and frequentist views] more fully, but we gave up due to differences among the authors!" Note that we did not discuss the Bayesian viewpoint at all; Sachs et al. (1989b) do.

REFERENCES

Bettonvil, B., Sequential Bifurcation: Screening For Important Factors. Ph. D. dissertation, Katholieke Universiteit Brabant (Tilburg University). Tilburg, November 1990.

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

Box, G.E.P. and N.R. Draper, Empirical Model-Building with Response Surfaces. John Wiley & Sons, New York, 1987.

Iman, R.L. and J.C. Helton, An investigation of uncertainty and sensitivity analysis techniques for computer codes. Risk Analysis, 8, 1988, pp. 71-90,

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

Kleijnen, J.P.C., Statistical Tool for Simulation Practioners, Marcal Dekker, Inc., New York, 1987.

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.

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

Olivi, L. and J.D. Pike, Response surface methodology for the analysis of reactor safety: methodology, development and inplementation. ISPRA, 21020 ISPRA (Varese), Italy, Sept. 1981.

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

Owen, A., J. Koehler and S. Sharifzadeh, Comment. (Comment on Sachs et al., 1989b.) Statistical Science, 4, no. 4, 1989, pp. 429-430.

Sachs, J., S.B. Miller, and W.J. Welch, Design for computer experiments. Technometrics, 31, no. Feb. 1989, pp. 41-47.

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.

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

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

414 C.B. Mulder en A.B.T.M. van Schaik
Een nieuwe kijk op structuurwerkloosheid

415 Drs. Ch. Caanen
De hefboomwerking en de vermogens- en voorraadaftrek

416 Guido W. Imbens
Duration models with time-varying coefficients

417 Guido W. Imbens
Efficient estimation of choice-based sample models with the method of moments

418 Harry H. Tigelaar
On monotone linear operators on linear spaces of square matrices

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