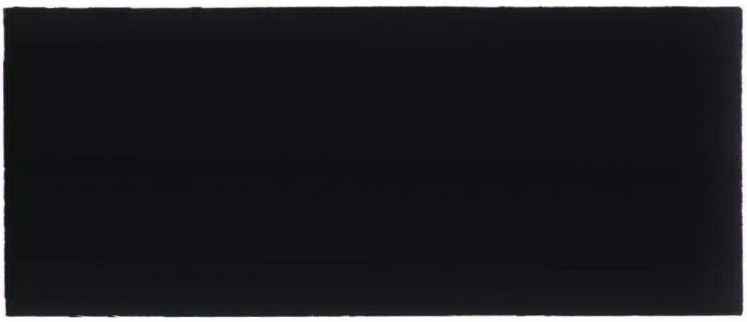


DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

SENSITIVITY ANALYSIS OF SIMULATION EXPERIMENTS: TUTORIAL ON REGRESSION ANALYSIS AND STATISTICAL DESIGN

Jack P.C. Kleijnen

FEW 474

SENSITIVITY ANALYSIS OF SIMULATION EXPERIMENTS: TUTORIAL ON 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 validation, what-if analysis, optimization, and so on. The analysis uses regression models and Least Squares algorithms. The design uses classical experimental designs such as 2^{k-p} factorials, which are more efficient than one at a time designs are. Moreover, classical designs make it possible to estimate interactions among inputs to the simulation. Simulation models may be optimized through Response Surface Methodology, which combines steepest ascent with regression analysis and experimental design. If there are very many inputs, then special techniques such as group screening and sequential bifurcation are useful. Several applications are discussed.

INTRODUCTION

Simulation is a mathematical technique that is applied in all those scientific disciplines that use mathematical modeling. These disciplines range from sociology to astronomy, as Karplus (1983) discussed so eloquently. Simulation is a very popular technique because of its flexibility, simplicity, and realism. By definition, however, simulation involves experimentation, namely experimentation with a model.

Experimentation requires an appropriate statistical *design and analysis*. For real systems, mathematical statistics has been applied since the 1930's, when Sir Ronald Fisher focussed on agricultural experiments. Since the 1950's George Box concentrated on chemical experimentation, which resulted in the well-known textbook by Box and Draper (1987). Tom Naylor organized a conference on the design of simulation experiments back in 1968; the proceedings are found in Naylor (1969). In 1974/75 my first book (Kleijnen, 1974/1975) covered both the 'tactical' and the 'strategic' issues of experiments with random and with deterministic simulation models. The term *tactical* was introduced into simulation by Conway (1963); it refers to the problems of runlength and variance reduction, which arise only in random simulations (for example, queuing simulations). *Strategic* questions are: which combinations of values for the input variables should be simulated, and how can the resulting output be analyzed? Obviously strategic issues arise in both random and deterministic simulations. Mathematical statistics can be applied to solve these questions, also in deterministic simulation; see Kleijnen (1987, 1990) and Sacks et al. (1989). This contribution focusses on the strategic issues in simulation experiments.

Strategic problems are also addressed under names like model validation, what-if analysis, goal seeking, and optimization. Table 1 summarizes some terminology; we shall return to this table.

PART 1: REGRESSION METAMODELS

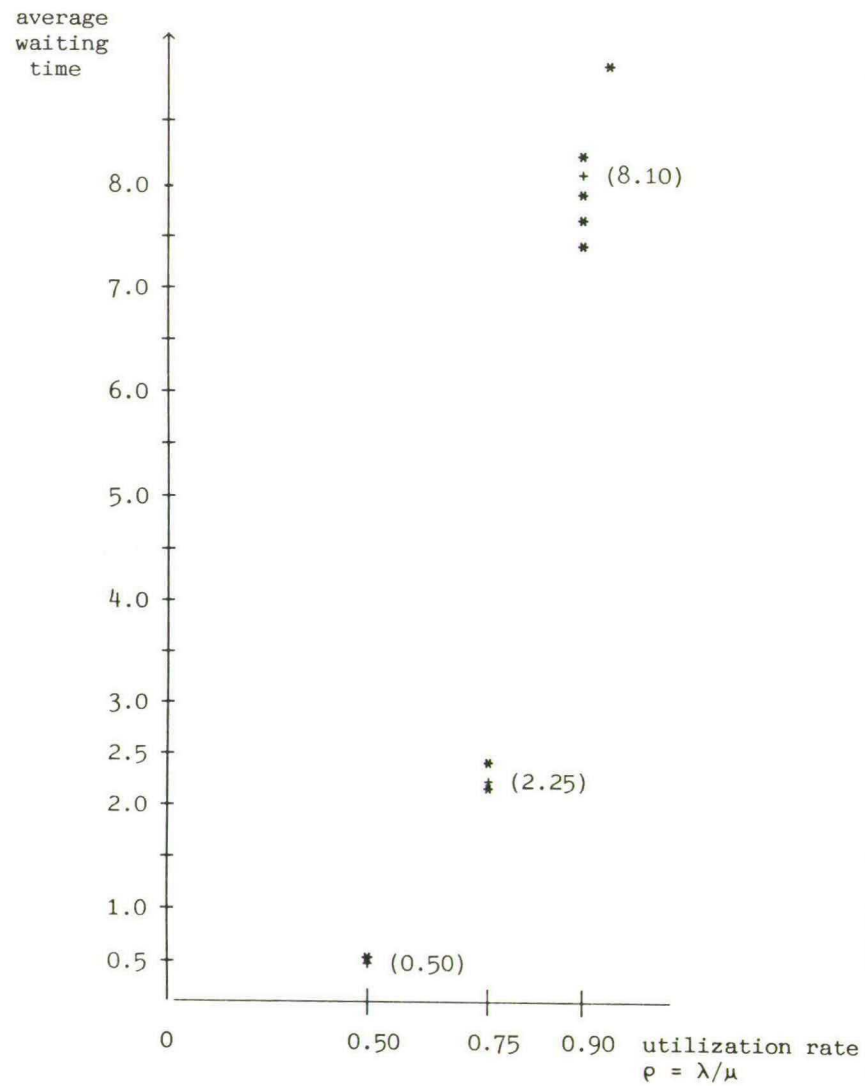
Figure 1 illustrates the type of graphs that is often used in the analysis of simulation results; it represents the relationship between the input and output of a simulation model. The visual analysis of the resulting 'cloud' of observed points should provide insight into the global effect an input change has on the output. Moreover, such a figure enables interpolation and extrapolation to get a quick estimate of the expected simulation response at an input value not yet observed. Note that in figure 1 the observations at low utilization rates show so little variation that they seem to coincide, given the scale of that figure.

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		

The visual analysis becomes impractical, when there are several input variables. Therefore we propose to replace the graphical analysis by an algebraic analysis: we use a regression model. Why is such an analysis useful? Most important is the resulting insight into the general behavior of the simulated system. A well-known disadvantage of simulation is its *ad hoc* character: the simulation responses are known only for the selected input combinations. Through regression analysis we can simultaneously examine all input/output combinations used in the simulation experiment, and gain more insight, while using a minimum of computer time. Computer

Figure 1: Relationship between average waiting time and utilization rate in simulated M/M/1
$\mu = 1.0$, $\lambda = 0.50, 0.75, 0.90$,
4 times 100 000 customers simulated per combination



time plays an important role in the search for optimal input combinations and in sensitivity analysis. We will show that the *linear* regression model is an adequate model to gain insight into simulation models. Because we make a (regression) model of the simulation model, we speak of (regression) *metamodels*. That simulation model is treated as a black box in our approach.

We distinguish between variables and parameters. Following Zeigler (1976), we define a parameter as a quantity that cannot be observed in the real system, whereas a variable can be observed directly. Examples of variables are customer arrival times and number of servers; examples of parameters are the arrival rate λ of the Poisson arrival process, and the service rate μ. When the simulation program is run, the parameters are known inputs.

REGRESSION METAMODELS

Before systems analysts start experimenting with a simulation model, they have accumulated p*rior* knowledge about the system to be simulated: they may have observed the real system, tried different models, debugged the final simulation model, and so on. This tentative knowledge can be formalized in a regression model or an Analysis of Variance (ANOVA) model. ANOVA models are discussed in elementary statistical theory of experimental design. The simplest ANOVA models, however, can be easily translated into regression models, as Kleijnen (1987 pp. 263-293) shows. Because regression analysis is more popular than ANOVA is, we shall use regression terminology henceforth.

So prior knowledge is formalized in a *tentative regression model*. (Because this model is only tentative, it must be tested later on to check its validity.) 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 possible 'inputs' are parameters and variables, but also 'behavioral relationships'. Examples of such relationships are priority rules in queuing models, and consumption curves in economic models. So 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 variables. For example, x = -1 if the priority rule is First Come First Served, and x = +1 if the rule is Last Come First Served. Note that 'inputs' are called 'factors' in experimental design terminology. The tentative regression model becomes

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k-1} \sum_{g=j+1}^{k} \beta_{jg} x_j x_g + \\ + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \varepsilon, \qquad (1)$$

where the symbols have the following meanings. The symbol y denotes the simulation response. β_0 represents the overall mean response. β_j gives the main or first-order effect of factor j (j = 1,..., k). β_{jg} denotes the two-factor interaction between the factors j and g (g = j+1,..., k); that interaction implies that the effect of factor j depends on the 'level' of factor g; the response curves do not run parallel. β_{jj} represents the quadratic effect of factor j; it means that the response function shows curvature; such effects assume quantitative factors. Finally ε denotes 'fitting errors' or noise. Under certain restrictive mathematical conditions the response function in Eq. (1) is a *Taylor series* expansion of the simulation model $y(x_1,...,x_k)$. Unfortunately, these conditions do not hold in simulation. Therefore we propose to *start* with an initial model, for example, a model like Eq. (1) or even a simpler model with only main effects β_j besides the overall mean β_0. The purpose of that model is to guide the design of the simulation experiment and to interpret the resulting simulation data.

Note that Eq. (1) excludes interactions among three or more factors. Such interactions are popular in ANOVA, but they are hard to interpret. A regression model without such interactions often suffices, as we have observed in practice.

The regression variables x in Eq. (1) may be *transformations* of the original simulation parameters and variables; for example, $x_1 = \log(z_1)$ where z_1 denotes the original simulation input. We shall return to such transformations. *Scaling* is also important: if the lowest value of z_1 corresponds with $x_1 = -1$ and the highest value of z_1 corresponds with $x_1 = +1$, then β_1 measures the relative importance of factor 1 when that

factor ranges over the whole experimental area. In optimization, however, the response curve is explored only locally if *Response Surface Methodology* (RSM) is used, as we shall see. The *local* regression model may be a first-order model:

$$y = \gamma_0 + \sum_{j=1}^{k} \gamma_j z_j + \varepsilon. \qquad (2)$$

The midpoint of the local experiment is $(\bar{z}_1, \ldots, \bar{z}_j, \ldots, \bar{z}_k)$ with

$$\bar{z}_j = \frac{\sum_{i=1}^{n} z_{ij}}{n} \qquad (j=1,\ldots,k), \qquad (3)$$

where z_{ij} denotes the value of factor j in factor combination i, and n denotes the total number of factor combinations explored locally. The importance of factor j at the midpoint of the local experiment, is then measured by $\gamma_j \bar{z}_j$. For further details we refer to Bettonvil and Kleijnen (1990).

In all experiments analysts use models such as Eq. (1), explicitly or implicitly. For example, if they change one factor at a time, then they assume that all interactions are 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.

As an example we model a delicatessen store with one queue and s employees; the service rate is μ, and the customer arrival rate is λ. Possible criteria, which summarize the results of the simulation, could be the utilization rate of the servers, and the mean waiting time of customers. In this example, however, we concentrate on a single output, namely average waiting time $\bar{w}$. So we summarize the time series formed by the individual waiting times $w_1, w_2, \ldots$ through a single number, namely the average $\bar{w}$. We formulate a regression metamodel; for example $y = \beta x^2 + \varepsilon$, where y represents the average waiting time $\bar{w}$, and x represents the utilization or traffic rate $\rho = \lambda/(\mu s)$. This yields a figure that resembles figure 1.

LEAST SQUARES

Suppose that in the queuing example we study six combinations of the three simulation inputs (arrival rate λ, service rate μ, and number of servers s). We select these input values such that the utilization factor ρ lies between 0.5 and 0.7. We simulate each input combination for (say) 2000 customers, and compute the average waiting time $\bar{w}$. This we repeat 20 times; each replicated simulation run starts in the empty state (no customers waiting), and uses a new stream of pseudorandom numbers.

A naive metamodel for this queuing simulation would be

$$\bar{\bar{w}} = \gamma_0 + \gamma_1 \lambda + \gamma_2 \mu + \gamma_3 s + \varepsilon. \tag{4}$$

The simulation inputs correspond with the unscaled regression variables in Eq. (2), as follows: $z_1 = \lambda, z_2 = \mu$, $z_3 = s$, and a dummy variable $z_0 = 1$. This yields the scaled variables x. For example, since λ ranges between (say) 1 and 4, we get

$$x_1 = \frac{\lambda - 2.5}{1.5}. \tag{5}$$

Of course, there is a dummy variable $x_0 = 1$ that corresponds with β_0 in Eq. (1). For simplicity's sake we suppose that each combination is replicated an equal number of times (say) m; in the example m is 20. Then the Least Squares algorithm can be applied to the *average* responses; in the example we have

$$\bar{\bar{w}}_i = \Sigma_{r=1}^{m} w_{ir}/m \text{ with } i = 1,\ldots,n, \tag{6}$$

where n denotes the number of combinations; in the example n = 6. So the response variable in Eqs. (1) and (2) is $y_1 = \bar{\bar{w}}_1$. Queuing theory suggests the use of a logarithmic scale:

$$\log \bar{\bar{w}} = \gamma_0 + \gamma_1 \log \lambda + \gamma_2 \log \mu + \gamma_3 \log s + \varepsilon, \tag{7}$$

which implies $y = \log \bar{\bar{w}}$, $z_1 = \log \lambda$, etc. If $\gamma_1 = -\gamma_2 = -\gamma_3$ (which we can easily test), then (7) implies

$$\bar{\bar{w}} = \gamma_0^* \left[\frac{\lambda}{\mu s}\right]^{\gamma_1} \varepsilon^* = \gamma_0^* \rho^{\gamma_1} \varepsilon^*, \tag{8}$$

where $\gamma_0 = \log \gamma_0^*$ and $\varepsilon = \log \varepsilon^*$. Earlier we suggested to approximate the input/output behavior in figure 1 by the model

$$\bar{\bar{w}} = \gamma \rho^2 + \varepsilon. \tag{9}$$

Equation (9) implies $x_1 = \rho^2$; in other words, a metamodel that is quadratic in its inputs can still be linear in the regression parameters. Eqs. (4) through (9) give alternative metamodels for a simple queuing simulation. The specification of regression models depends on knowledge of the simulated system, not on knowledge of mathematical statistics! Mathematical statistics can help to validate the regression model, as we shall see.

The standard deviation (or standard error) of the output y_i can be estimated from the m replications, provided $m > 1$:

$$\hat{\sigma}_i = \left[\frac{\sum_{r=1}^{m} (y_{ir} - y_i)^2}{(m-1)\, m}\right]^{\frac{1}{2}}, \tag{10}$$

where the factor m is needed because y_i is the average of m responses. The m replications use identical independent combinations of inputs but different pseudorandom number seeds. In deterministic simulation, replication would make no sense; the same output would result.

We can now apply the Ordinary Least Squares (OLS) technique to estimate the vector of regression parameters β:

$$\hat{\beta} = (X'X)^{-1}X'y . \tag{11}$$

y is a vector with n elements, **X** is an $n \times q$ matrix, where q denotes the number of regression parameters; in Eq. (4) q is four. The variance-covariance matrix of **y**, denoted by **cov** (y), is a diagonal matrix with the elements σ_i^2 ; see (10). We can apply not only OLS, but also Estimated Weighted Least Squares (EWLS), as Kleijnen (1987) discusses in detail. If different input combinations use the same pseudorandom seed, then **cov** (y) is not diagonal anymore. Its elements can be estimated:

$$\hat{\text{cov}}\,(y_i, y_{i'}) = \frac{\sum_{r=1}^{m} (y_{ir} - y_i)(y_{i'r} - y_{i'})}{(m-1)}$$

$$(i, i' = 1,\ldots,n). \qquad (12)$$

For $i = i'$ this equation reduces to Eq. (10). Common seeds in regression metamodelling are further discussed in Kleijnen (1991). The OLS estimator $\hat{\beta}$ remains unbiased, even if the distribution of the simulation responses is not Gaussian. Normality usually holds if the simulation responses are averages, such as average queuing time.

In *deterministic* simulation we can also apply linear regression models. Then we assume that ε, the deviations between the simulation output and the metamodel prediction, is normally distributed, again with zero mean but now with constant variance σ^2. The simplest model assumes that these deviations are independent, so OLS suffices; see Kleijnen (1990). A more sophisticated model for ε is discussed by Sacks et al. (1989).

VALIDATION OF REGRESSION METAMODELS

Suppose we have specified and 'calibrated' a metamodel; that is, we have estimated the values of the regression parameters β. We can then use this metamodel to *predict* the output for *new* values of the parameters and input variables of the simulation model: inter- and extrapolation. So we select a *new* combination of the q independent variables $\mathbf{x}'_{(n+1)} = (1, x_{(n+11)}, \ldots, x_{(n+1)q-1})$, and through the metamodel we predict the simulation result:

$$\hat{y}_{n+1} = \hat{\beta}_0 + \sum_{h=1}^{q-1} \hat{\beta}_h x_{(n+1)h} = \mathbf{x}'_{n+1}\hat{\beta} . \qquad (13)$$

To test whether the predictor (13) is accurate (so the metamodel is valid), we actually run the simulation program with the new input combination $\mathbf{x}_{n+1}$. In *deterministic* simulation we 'eyeball' the relative prediction error $\hat{y}_{n+1}/y_{n+1}$: is this error 'acceptable'? (We shall discuss random simulation later.)

A disadvantage of this method is that the simulation program must be executed for one or more extra input combinations. An alternative is *cross-validation*, which runs as follows.

(i) We eliminate one combination, say i, from the old data. So we eliminate row i of $\mathbf{X}$ and element y_i of $\mathbf{y}$. The remaining data are denoted by $\mathbf{X}_{-i}$ and $\mathbf{y}_{-i}$.

(ii) We estimate the regression parameters β from the remaining data. So Eq. (11) is replaced by

$$\hat{\beta}_{-i} = (\mathbf{X}'_{-i}\mathbf{X}_{-i})^{-1}\mathbf{X}'_{-i}\mathbf{y}_{-i}. \qquad (14)$$

We assume that the number of combinations n exceeds the number of regression parameters q; otherwise the (n-1) × q matrix $\mathbf{X}_{-i}$ would certainly be singular.

(iii) To predict the result for the eliminated combination we replace Eq. (13) by

$$\hat{y}_i = \hat{\beta}_{0(-i)} + \sum_{h=1}^{q-1} \hat{\beta}_{h(-i)}\, x_{ih} = \mathbf{x}'_i\hat{\beta}_{-i} \qquad (15)$$

(iv) We calculate the relative prediction errors $\hat{y}_i/y_i$.

(v) We repeat steps (i) through (iv) for all n values of i.

So without any extra simulation effort, we have n prediction errors at our disposal. These errors can be 'eyeballed' again. We shall give an application later on.

For *random* simulation models with Gaussian responses we can adapt the cross-validation approach, as explained in Kleijnen (1987, 1991). But there is a more powerful validation test, namely *Rao (1959)'s lack-of-fit F test*. This test compares two variance estimators. One estimator uses the Mean Squared Residuals (MSR):

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n} (y_i - \tilde{y}_i)^2}{n-q} = (\mathbf{y} - \tilde{\mathbf{y}})'(\mathbf{y} - \tilde{\mathbf{y}})/(n-q), = \tilde{\mathbf{e}}'\,\tilde{\mathbf{e}}/(n-q), \qquad (16)$$

where $\tilde{\mathbf{y}} = \mathbf{X}\,\tilde{\beta}$ and $\tilde{\beta}$ denotes the Estimated Generalized Least Squares (EGLS) estimator

$$\tilde{\beta} = (X' [\hat{cov}\ (y)]^{-1} X)^{-1} X' [\hat{cov}\ (y)]^{-1} y; \qquad (17)$$

obviously in Eq. (16) the estimated fitting errors are $\tilde{e}$. The MSE is unbiased only if the model is correct; otherwise it overestimates the variance. The other estimator, based on replication, was defined in Eq. (12). Obviously, the latter estimator does not depend on the regression model. The model is rejected if the ratio of these two estimators exceeds F_{v_1,v_2} with $v_1 = n - q$ and $v_2 = n(m-1)$:

$$F_{n-q,m-n+q} = \frac{m-n+q}{n-q} \frac{1}{(m-1)} \tilde{e}' [\hat{cov}\ (y)]^{-1} \tilde{e} \qquad (18)$$

USE OF METAMODELS

Only after the metamodel has been validated, can it be used to explain the behavior of the simulation model. To test if an estimated regression parameter differs significantly from zero, the 1-α confidence interval for the OLS estimator is computed. For random simulation this interval is

$$\hat{\beta}_h \pm t_m^{\alpha/2} \sqrt{\hat{var}\ (\hat{\beta}_h)} \qquad (h=1,\ldots,q), \qquad (19)$$

where $\hat{var}\ (\hat{\beta}_h)$ is the h^{th} diagonal element of

$$\hat{cov}\ (\hat{\beta}) = (X'X)^{-1} X'\ \hat{cov}\ (y)\ X\ (X'X)^{-1}\ , \qquad (20)$$

where $\hat{cov}\ (y)$ is given by Eq. (12). For deterministic simulation we assume that cov $(y) = \sigma^2 I$ where σ is estimated by (16) provided we replace the EGLS estimator $\tilde{\beta}$ in (16) by the OLS estimator $\hat{\beta}$. So for deterministic simulation we get

$$\hat{cov}\ (\hat{\beta}) = (X'X)^{-1}\ \hat{\sigma}^2 \qquad (21)$$

and

$$\sigma^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-q}. \qquad (22)$$

The degrees of freedom (m) in Eq. (19) must be replaced by n-q (deterministic simulation has no replications: m = 1).

If $\hat{\beta}_h$ differs from zero significantly, then x_h is important. This is the information we seek in *what-if* analysis. Suppose x_h represents a parameter or a variable that can be influenced by the user. Then to (say) maximize the system output, the user increases that parameter or variable if the sign of the estimated regression coefficient $\hat{\beta}_h$ is positive. The user may even try to optimize the system, as we shall see later.

We can also give predictions for new input combinations; that is, we can interpolate and extrapolate from old simulation data. This is important if a simulation run costs much computer time.

To illustrate the use of metamodels we summarize Kleijnen and Standridge (1988). They apply regression analysis to a deterministic simulation model of a Flexible Manufacturing System (FMS). The goal is to select x_i, the number of machines of type i with i = 1,...,4, such that a given production volume is realized. The inputs to the simulation may vary over the following 'experimental domain': $x_1 \in (5,6)$, $x_2 \in (1,2)$, $x_3 \in (2,3)$, and $x_4 \in (0,1,2)$; that is, 24 combinations are of interest. Actually they simulate only eight combinations and estimate the first-order metamodel. If one combination is deleted from the data (cross-validation), then the OLS estimates of β change; especially the values of the non-significant effects change since they represent noise. Note that upon deletion of a combination, the degrees of freedom are $v = n - q = 7 - 5$. The effects of deleting a combination can be shown more compactly by concentrating on the responses; that is, the criterion becomes prediction instead of explanation. This yields the relative prediction errors $\hat{y}_i/y_i$ in cross-validation. The prediction errors $\hat{y}/y$ and the instabilities of $\hat{\beta}$ are so large that the additive metamodel is rejected, and a new model is investigated. This model still uses the old data of the simulation experiment, $\mathbf{X}$ and $\mathbf{y}$. The $\hat{\beta}$ suggested that the factors 1 and 3 are unimportant. Therefore a regression model in the remaining factors is formulated, but now interaction is included:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_2 x_2 + \hat{\beta}_4 x_4 + \hat{\beta}_{24} x_2 x_4 . \qquad (23)$$

In this model all estimated effects remain significant upon run deletion. The relative prediction errors become much smaller. The new metamodel

suggests that type 2 and 4 machines are the bottle-necks in the FMS. The negative sign of the interaction $\hat{\beta}_{24}$ means that there is a trade-off between these two machine types.

PART 2: DESIGN OF EXPERIMENTS

Simulation practitioners change the values of parameters and input variables, and the model structure, in order to analyse the effects of such changes. In most experiments the examination of all factor combinations would take too much computer time. Therefore the goal of experimental design is to gain insight into the system's behavior while observing relatively few factor combinations.

CLASSICAL EXPERIMENTAL DESIGNS

Figure 2 and table 2 illustrate the following experimental designs that are often used in practice.

(i) Change a single factor at a time: 'one factor at a time designs'.

(ii) Examine all possible combinations of factor levels: 'full factorial designs'.

(iii) Examine only part of all possible combinations: 'incomplete factorial designs', in particular 'fractional designs'.

We shall show that designs of type (iii) can give the same information as types (i) and (ii) do; yet type (iii) requires fewer observations. In contrast to type (i), type (iii) can give information on interactions.

Except for the section on RSM, we assume that each factor is observed at only two 'levels' or values. Therefore the independent variables x satisfy

$$x_{ij} = \begin{cases} -1 \text{ if factor j is } \textit{off} \text{ in combination i} \\ (i = 1,2,\ldots n) \quad (j = 1,2,\ldots,k) \\ \\ +1 \text{ if factor j is } \textit{on} \text{ in combination i.} \end{cases} \tag{24}$$

Figure 2: Experimental design for three factors

Legenda:
the numbers (1),...,(8) correspond with the combinations in table 7

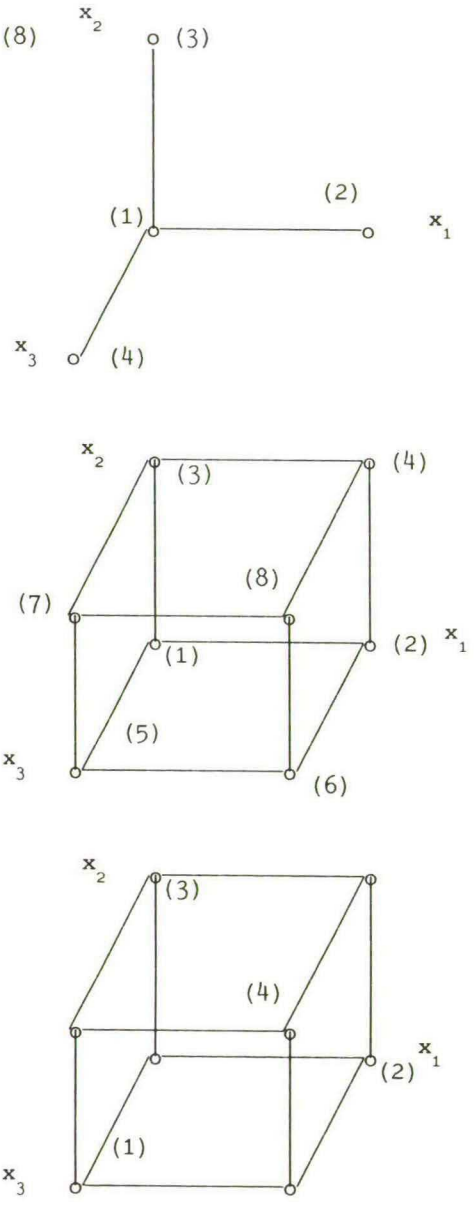


Table 2: Experimental design for three factors

(i) one factor at a time

combination	x_1	x_2	x_3
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	-1	-1	1

(ii) full factorial design

combination	x_1	x_2	x_3
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

(iii) fractional design: 2^{3-1} factors

combination	x_1	x_2	x_3
1	-1	-1	1
2	1	-1	-1
3	-1	1	-1
4	1	1	1

For quantitative factors, 'off' may mean that the factor has a low value. Equation (24) also holds for qualitative factors, which have only nominal values. For example, if factor 1 denotes the priority rule and if in combination 3 the priority rule is First Come First Served (FCFS), we may say that $x_{31} = 1$ holds; if in combination 4 the priority rule is Last Come First Served (LCFS), then $x_{41} = -1$ holds (also see Part 1 above Eq. 1). For qualitative factors with more than two levels we refer to Kleijnen (1987, pp. 275-278).

In the example of figure 2 there are three factors, each with two levels. The underlying simulation model may be a queuing system with the following factors and levels: priority rule is either FCFS or LCFS, number of servers is 1 or 3, and utilization degree is 0.5 or 0.8. Table 2 gives the matrices $\mathbf{X} = (x_{ij})$ for the three designs; we do not show the dummy variable x_0 because it is always +1. Consider a regression *metamodel* or *response surface* with main effects β_j, two-factor interactions β_{jg}, and noise ε that is normally distributed with zero mean and constant variance σ^2 (also see Eq. 1):

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{j=1}^{k-1} \sum_{g=j+1}^{k} \beta_{jg} x_{ij} x_{ig} + \varepsilon_i ;$$

$$\varepsilon_i \sim N(0,\sigma^2) \qquad (i=1,\ldots,n). \tag{25}$$

Table 2(i) shows that changing *one factor at a time* yields four combinations ($n = k + 1 = 4$), namely the base variant ($i = 1$) and three more variants. This design assumes that there are no interactions. Intuitively we estimate the main effects of the factors by

$$\hat{\alpha}_j = y_i - y_1 \quad \text{with } i = j + 1 \ (j = 1,\ldots,k). \tag{26}$$

Equation (25) implies that β_j is the differential effect of factor j when the other factors are fixed at the origin ($x_{ig} = 0$). Since Eq. (25) means that the response surface has no curvature, the differential effect $(\partial/\partial x_j)$ equals the difference quotient $\Delta y / \Delta x_j$. If $\Delta x_j = 1$ then $\beta_j = \Delta y$. The estimator

$$\hat{\beta}_j = \hat{\alpha}_j / 2 \tag{27}$$

is an unbiased estimator of β_j. Equations (25) and (26) imply that the variance of the estimator $\hat{\alpha}$ is $2\sigma^2$; so $V\{\hat{\beta}_j\} = \sigma^2/2$. The intuitive estimator of β_0 is the overall average:

$$\hat{\beta}_0 = \Sigma_{i=1}^{n} y_i/n. \tag{28}$$

A *full factorial* design consists of all possible combinations of factor levels. In case of two levels per factor, this design results in 2^k combinations. To estimate the main effect intuitively, we adapt Eq. (26): we take the difference between the average result for $x_j = -1$ and $x_j = 1$ respectively, so table 2(ii) yields

$$\begin{aligned}
\hat{\alpha}_1 &= (y_2 + y_4 + y_6 + y_8)/4 - (y_1 + y_3 + y_5 + y_7)/4 \\
\hat{\alpha}_2 &= (y_3 + y_4 + y_7 + y_8)/4 - (y_1 + y_2 + y_5 + y_6)/4 \\
\hat{\alpha}_3 &= (y_5 + y_6 + y_7 + y_8)/4 - (y_1 + y_2 + y_3 + y_4)/4.
\end{aligned} \tag{29}$$

In the general case of k factors, each observed at two levels, the intuitive estimators become

$$\hat{\alpha}_j = \frac{\Sigma_{i=1}^{n} x_{ij} y_i}{(n/2)} \qquad \text{with } n = 2^k. \tag{30}$$

We now prove that this *intuitive* estimator is identical to the *OLS* estimator. The matrix **X** for the 'additive' model consists of the dummy column (with $n = 2^k$ elements equal to +1) and the k columns of the experimental design in table 2(ii); so in the example **X** is an 8×4 matrix. This **X** is *orthogonal*: $\mathbf{X'X} = n\mathbf{I}$. Hence the OLS estimator becomes

$$\begin{cases}
\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'y} = \frac{1}{n}\mathbf{X'y} \quad \text{or} \quad \hat{\beta}_j = \frac{1}{n}\Sigma_{i=1}^{n} x_{ij} y_i \\
\qquad (j=0,\ldots,k), \\
\operatorname{cov}(\hat{\beta}) = \sigma^2(\mathbf{X'X})^{-1} = \frac{\sigma^2}{n}\mathbf{I} \quad \text{or} \quad \operatorname{var}(\hat{\beta}_j) = \frac{\sigma^2}{n} \quad (n=2^k).
\end{cases} \tag{31}$$

So Eq. (30) specifies the intuitive estimator $\hat{\alpha}_j$ for a 2^k design; Eq. (31) proves that the intuitive estimator equals the OLS estimator $\hat{\beta}_j$ multiplied

by two; Eq. (27) relates $\hat{\alpha}_j$ and $\hat{\beta}_j$. Hence the OLS estimator and the intuitive estimator are identical indeed. Eq. (31) proves that the k+1 estimators have a constant variance and are independent.

To see which design is more *efficient*, we compare the variances of $\hat{\beta}_j^{(i)}$ and $\hat{\beta}_j^{(ii)}$, which denote the estimators for designs (i) and (ii) respectively, but we account for the different numbers of observations, k + 1 and 2^k respectively:

$$\frac{\text{var}\{\hat{\beta}_j^{(i)}\}\times(k+1)}{\text{var}\{\hat{\beta}_j^{(ii)}\}\times 2^k} = \frac{(\sigma^2/2)\times(k+1)}{(\sigma^2/n)\times n} = \frac{k+1}{2} . \qquad (32)$$

So type (ii) designs are more efficient. This can be explained intuitively: in type (i) designs the variance remains the same (namely $\sigma^2/2$) for all values of k, whereas in type (ii) designs the variance decreases as $n = 2^k$ increases; in type (ii) designs each observation i is used when estimating an effect β_j. It can be proved that an orthogonal matrix **X** minimizes the variances of the $\hat{\beta}_j$'s if **cov** (**y**) = **cov**(ε) = σ^2**I**.

Type (ii) designs are not only more efficient; they are also more *effective* in that they enable the estimation of *interactions*. To estimate the interaction between x_j and x_g we multiply the vectors of these two independent variables elementwise; this yields table 3, which follows from the design in table 2(ii) and the model in Eq. (25). Each column in table 3, except for the x_0 column, has an equal number of plus signs and minus signs. All vectors of **X** are orthogonal. So analogous to Eq. (31), the OLS estimator of the two-factor interaction β_{jg} becomes:

$$\hat{\beta}_{jg} = \frac{1}{n} \Sigma_{i=1}^{n} (x_{ij} x_{ig}) y_i \quad (n = 2^k). \qquad (33)$$

The estimators of main effects and interactions remain independent with constant variance σ^2/n.

Upon including two-factor interactions, the number of effects q becomes

$$q = 1 + k + k(k-1)/2. \qquad (34)$$

Table 3: Independent variables **X** for the 2^3 experimental design including factor interactions

combination	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3
1	1	-1	-1	-1	1	1	1
2	1	1	-1	-1	-1	-1	1
3	1	-1	1	-1	-1	1	-1
4	1	1	1	-1	1	-1	-1
5	1	-1	-1	1	1	-1	-1
6	1	1	-1	1	-1	1	-1
7	1	-1	1	1	-1	-1	1
8	1	1	1	1	1	1	1

So the inequality $q < n = 2^k$ holds for $k > 2$. For example, in table 3 we have: $q = 7 < n = 8$; and for $k = 7$ we have $q = 29 << n = 128$. But, if $q < n$ holds, then we could estimate regression models with more parameters, in particular models with interactions among three or more factors. These higher-order interactions, however, are difficult to interpret, so we suggest not to add high-order interactions to the metamodel. Ignoring high-order interactions leads to the experimental designs of type (iii) in table 2.

If there are no interactions, then the k+1 effects can be estimated from only a *fraction* of the 2^k observations, as is illustrated by table 2(iii). The intuitive estimators are analogous to Eq. (29):

$$\begin{aligned} \hat{\alpha}_1 &= (y_2 + y_4)/2 - (y_1 + y_3)/2 \\ \hat{\alpha}_2 &= (y_3 + y_4)/2 - (y_1 + y_2)/2 \\ \hat{\alpha}_3 &= (y_1 + y_4)/2 - (y_2 + y_3)/2. \end{aligned} \tag{35}$$

The new matrix **X** remains orthogonal; so the OLS formulas in Eq. (31) still apply, but now with $n = 2^{k-p}$.

Designs of types (ii) and (iii) are equally efficient, if the number of observations is taken into account:

$$\frac{\text{var}\{\hat{\beta}_j^{(iii)}\}\times 2^{k-p}}{\text{var}\{\hat{\beta}_j^{(ii)}\}\times 2^{k}} = \frac{(\sigma^2/2^{k-p})\times 2^{k-p}}{(\sigma^2/2^{k})\times 2^{k}} = 1 \ . \qquad (36)$$

In practice, however, fractional designs are often preferred, because their number of combinations is considerably smaller, so less computer time is needed for the simulation experiment. Table 2 showed n = 8 and 4 for the 2^3 and the 2^{3-1} experiment respectively; for k = 7 the full factorial design needs 128 combinations whereas a fractional design may need only 8 combinations, as we shall see in table 5. On the other hand, the full factorial design has a smaller variance because of the larger number of observations. In comparison with the one at a time design, the fractional design becomes more efficient, the more factors there are, as Eq. (32) proves. Both types may require an equal number of observations (namely k+1). In the example of table 7 n is 4 and $V\{\hat{\alpha}_j\}$ is $2\sigma^2$ and σ^2 in the type (i) and type (iii) designs respectively.

The fractional factorial designs of the 2^{k-p} type imply that the number of combinations is a power of two: n = 2, 4, 8, 16, 32, and so on. There are so-called *Plackett and Burman designs* for first-order metamodels, where n is only a multiple of four. An example is shown in table 4, where k = 11 and n = 12. The Plackett-Burman designs have been tabulated; these tables are reproduced in Kleijnen (1975, pp. 332-333).

Table 4: Plackett-Burman Design for eleven factors

Combination	1	2	3	4	5	6	7	8	9	10	11
1	+	−	+	−	−	−	+	+	+	−	+
2	+	+	−	+	−	−	−	+	+	+	−
3	−	+	+	−	+	−	−	−	+	+	+
4	+	−	+	+	−	+	−	−	−	+	+
5	+	+	−	+	+	−	+	−	−	−	+
6	+	+	+	−	+	+	−	+	−	−	−
7	−	+	+	+	−	+	+	−	+	−	−
8	−	−	+	+	+	−	+	+	−	+	−
9	−	−	−	+	+	+	−	+	+	−	+
10	+	−	−	−	+	+	+	−	+	+	+
11	−	+	−	−	−	+	+	+	−	+	+
12	−	−	−	−	−	−	−	−	−	−	−

If the number of effects k+1 is not a multiple of four, then n is rounded upwards to the next multiple of four; for example, if k is 4,5,6 or 7 then n becomes 8; see table 4. For k = 4 we eliminate the columns x_5, x_6 and x_7; for k = 5 we do not use the columns x_6 and x_7; for k = 6 we eliminate the column x_7.

If there are *interactions*, then fractional factorials may give false conclusions. For example, in table 2(iii) we have $x_3 = x_1 x_2$. It can be proved that, for example, $\hat{\alpha}_3$ is Eq. (35) does not estimate α_3 but $\alpha_3 + \alpha_{12}$. Hence if the interaction α_{12} is important, $\hat{\alpha}_3$ is a biased estimator of α_3. Therefore we should validate the metamodel that guides the choice of the design. That validation was discussed in part 1.

The classical theory on experimental designs assumes that the *noise* ϵ is normally and independently distributed with constant variance. We can estimate this constant variance from the Mean Squared Residuals, as we saw in Eq. (22). In practice, the variances may differ very much. We can then estimate the heterogeneous variances σ_i^2, provided we simulate combination i of the k factors $m \geq 2$ times, as we saw in Eq. (10). If common pseudorandom seeds are used, then correlations can be estimated through Eq. (12).

The literature presents the metamodel and the experimental design in the form of *standardized* variables x; that is, x is either +1 or -1. We may also formulate the model in terms of the original variables z_j that are centered around their averages $\bar{z}_j$ (defined in Eq. 3):

$$y = \delta_0 + \Sigma_{j=1}^{k} \delta_j (z_j - \bar{z}_j) + \Sigma_{j=1}^{k-1} \Sigma_{h=j+1}^{k} \delta_{jh} (z_j - \bar{z}_j)(z_h - \bar{z}_h) + \epsilon. \qquad (37)$$

We assume a balanced experimental design; that is, there are as many 'low' values L_j as there are 'high' values H_j for factor j. Such a design implies

$$\bar{z}_j = \frac{1}{n} \Sigma_{i=1}^{n} z_{ij} = \frac{1}{n} \left(\frac{n}{2} L_j + \frac{n}{2} H_j \right) = (L_j + H_j)/2. \qquad (38)$$

The marginal or main effect of factor j is then δ_j. The total effect over the range $L_j \leq z_j \leq H_j$ is $\delta_j (H_j - L_j) = 2\beta_j = \alpha_j$. Factor 1 may have a larger

marginal effect than factor 2 has; yet the importance of factor 2 may be higher, given the experimental domain. In sensitivity analysis the interest is in α_j or β_j; then the standardized variables x should be used. Also see Bettonvil and Kleijnen (1991).

A case study that illustrates the use of fractional designs and regression analysis is presented in Kleijnen, van Ham, and Rotmans (1990). The study concerns a large deterministic ecological simulation model that consists of many modules. This model requires sensitivity analysis to support the Dutch government's decision making.

SCREENING

In the beginning of a simulation study there may be very many factors. Obviously not all these factors are important; unfortunately, we do not know *which* factors are really important. If there is a very large number of factors, then the designs discussed so far require too many combinations and hence too much computer time: they all require $n > k$. The following method may provide a solution.

Table 5: 2^{7-4} Experimental design

combination	x_1	x_2	x_3	$x_4 = x_1 x_2$ *)	$x_5 = x_1 x_3$	$x_6 = x_2 x_3$	$x_7 = x_1 x_2 x_2$
1	-1	-1	-1	1	1	1	-1
2	1	-1	-1	-1	-1	1	1
3	-1	1	-1	-1	1	-1	1
4	1	1	-1	1	-1	-1	-1
5	-1	-1	1	1	-1	-1	1
6	1	-1	1	-1	1	-1	-1
7	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1

*) $x_4 = x_1 x_2$ stands for $x_{i4} = x_{i1} x_{i2}$, and so on.

We partition the k factors into G groups. A *group factor* w_g (g = 1,2,...,G) has the value 1 (and -1 respectively) if all its component factors x have the value 1 (and -1 respectively); see table 6 with k = 100 and G = 2 (we shall discuss the last column of this table in a moment). For each group we test whether the group has a significant effect (say) $\hat{\gamma}_g$. If the group effect is not significant, we eliminate all factors of that group, in the next phase. It can be proved, under not too stringent assumptions, that if the group is not significant, then none of the individual factors is important. Group screening quickly reduces the number of factors in the pilot phase of a simulation project.

Suppose that only x_1 and x_2 are important; that is, only β_1 and β_2 are not zero, there are no interactions, and all other factors are unimportant (of course, in practice we do not have this prior knowledge). This yields the last column of table 6. The OLS estimator $\hat{\gamma}_1$ for the main effect of group 1 follows from the general formulas for OLS, which were given in Eq. (31). So

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^{4} w_{i1} y_i}{4} = \frac{-y_1 - y_2 + y_3 + y_4}{4} . \qquad (39)$$

Table 6: Group screening

combination i	group factor w_1	w_2	individual factors $x_1 \ldots x_{50}$	$x_{51} \ldots x_{100}$	expected result y
1	-1	-1	-1... -1	-1... -1	$-\beta_1 - \beta_2$
2	-1	1	-1... -1	1... 1	$-\beta_1 - \beta_2$
3	1	-1	1... 1	-1... -1	$\beta_1 + \beta_2$
4	1	1	1... 1	1... 1	$\beta_1 + \beta_2$

For group 2 we get:

$$\hat{\gamma}_2 = \frac{\sum_{i=1}^{4} w_{i2} y_i}{4} = \frac{-y_1 + y_2 - y_3 + y_4}{4}. \qquad (40)$$

It is simple to prove that the expected value of $\hat{\gamma}_1$ is $\beta_1 + \beta_2$ and that of $\hat{\gamma}_2$ is zero.

Next we test if the estimates $\hat{\gamma}_1$ and $\hat{\gamma}_2$ differ significantly from zero, as we saw in Eq. (19). Suppose we test each group effect at level α. The probability of finding a significant $\hat{\gamma}_2$ is α. The larger $\beta_1 + \beta_2$ is, the higher the probability of a significant γ_1 is. Obviously this approach fails if β_1 and β_2 happen to have opposite signs, and they are of the same absolute magnitude: $\beta_1 + \beta_2 = 0$. (Therefore we code the variables x such that they have non-negative effects, if they have any effects at all; for example, in multi-server queuing systems we take x = + 1 if traffic rate is high, service rate is low, number of servers is low, and so on.) So after only four runs we probably eliminate the 50 individual factors that form group 2. In the next stage of the investigation (not shown in the table) we further examine group 1; for example, we apply group screening to the factors 1,2,...,50. We continue, until so few individual factors remain that we can apply the classical designs such as 2^{k-p} designs. For further details we refer to Kleijnen (1987, pp. 320-328).

Recently several alternative techniques for screening have been developed:

(i) Sequential bifurcation by Bettonvil (1990) is a more efficient variation on group screening. He applied this technique to ecological models with nearly 300 factors.

(ii) Search linear models by Gosh (1987) are based on Srivastava (1975) 's designs.

(iii) The frequency domain technique by Schruben and Cogliano (1987) uses spectral analysis. It is criticized by Sargent and Som (1988).

(iv) Perturbation analysis by Ho et al. (1984) does not treat the simulation model as a black box to which experimental design theory is applied. Instead they apply mathematical analysis to derive marginal effects.

Figure 3: Steepest ascent

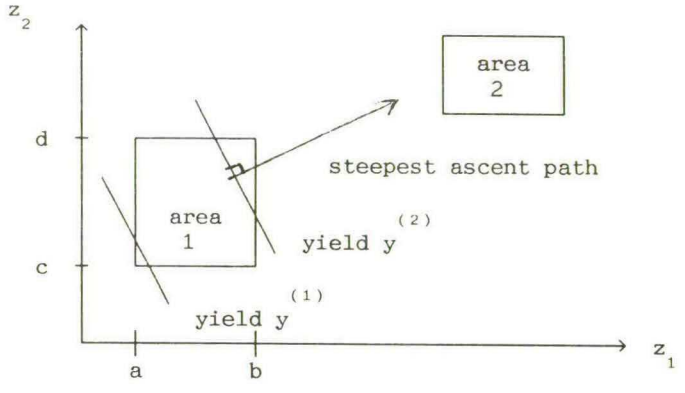
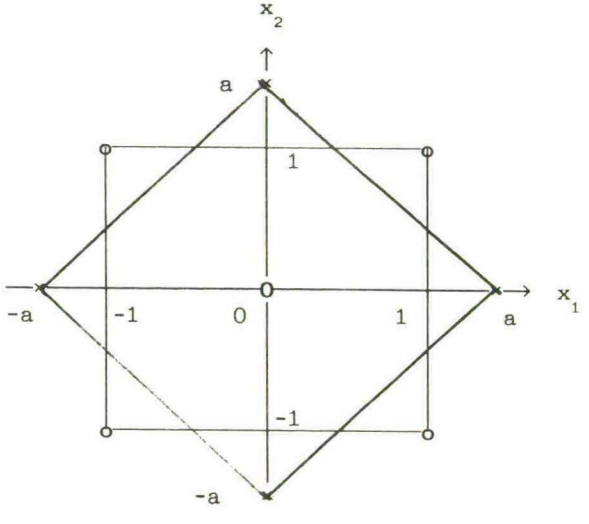


Figure 4: Experimental design for two factors with interaction and quadratic effects

Legenda: o 2^k design, × star design, 0 central point



OPTIMIZATION: RESPONSE SURFACE METHODOLOGY

The goal of Response Surface Methodology (RSM) is to find optimum values for the quantitative input variables that can be controlled by the user. The method is heuristic; so it is not certain that its answer is optimal indeed; for example, a *local* optimum may be found. RSM uses the *steepest ascent* algorithm, which determines in which direction the factors should be changed in order to reach the optimum; it does not give the step size along this path. We now sketch how RSM works; for illustrative purposes we take a problem with only two controllable variables, as is illustrated by figure 3.

(i) Initially, we vary the controllable variables z_1 and z_2 over a small range only. In figure 3 this area is the square with $a \leq z_1 \leq b$ and $c \leq z_2 \leq d$.

(ii) Because this area is small relative to the total experimental area, we apply the local first-order approximation

$$\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 z_1 + \hat{\gamma}_2 z_2, \qquad (41)$$

where the three regression parameters are estimated from the four observations at the four corners of area 1.

(iii) Next we change z_1 and z_2 according to the ratio $\hat{\gamma}_1/\hat{\gamma}_2$. It can be proved that the steepest ascent path is perpendicular to the line of constant yields ($y^{(1)}$ and $y^{(2)}$ respectively in figure 3). We know the direction $\hat{\gamma}_1/\hat{\gamma}_2$ from Eq. (41), but not the step size. We intuitively choose a step size, which takes us to area 2 in figure 3.

(iv) We repeat (i), (ii), and (iii) until the local first-order model is not adequate anymore. The literature often uses the significance of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ as a criterion; however, we can also test the validity of the model in the way explained in Part 1. An inadequate first-order model implies that we are close to an optimum: a plane such as Eq. (41) can model a hill *side*, but not a hill *top*. This leads to the next step.

(v) Near the optimum we use a second-order approximation:

$$\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 z_1 + \hat{\gamma}_2 z_2 + \hat{\gamma}_{12} z_1 z_2 + \hat{\gamma}_{11} z_1^2 + \hat{\gamma}_{22} z_2^2 . \qquad (42)$$

Comparing the first - and second - order approximations of Eqs. (41) and (42) shows that now we must estimate more parameters. So we have to examine more factor combinations: $n \geq 1+k+k(k-1)/2+k$. There are special RSM designs. Figure 4 shows such a design for $k = 2$ scaled variables; the five parameters of Eq. (42) are estimated from nine observations. Moreover because we want to estimate the quadratic effects γ_{11} and γ_{22}, the factors must have more than two levels (with two levels we would have: $x_0 = x_1^2 = x_2^2$). Actually there are five levels: -a,-1,0,1, and a with $a \neq 1$ and $a \neq 0$.

In the general case with k factors, the *star design* of figure 4 consists of 2k combinations. In each of these combinations, all factors are zero, except for one factor (the star design is a one factor at a time design). So the star design is specified by the following 2k rows where each row has k-1 zero's:

-a, 0, ..., 0
+a, 0, ..., 0
0, -a, ..., 0
0, +a, ..., 0
.
.
.
0, ..., 0, -a
0, ..., 0, +a

In the general case, the 2^k design of figure 4 becomes a 2^{k-p} design with p so small that all two-factor interactions can be estimated (such a 2^{k-p} design is called a 'resolution IV' design or 2^{k-p}_{IV} design).

(vi) Finally we differentiate the estimated local second-order model for the controllable variables, and solve the resulting system of linear equations for these variables:

$$\frac{\partial \hat{y}}{\partial z_1} = \hat{\gamma}_1 + \hat{\gamma}_{12} z_2 + 2\hat{\gamma}_{11} z_1 = 0$$

$$\frac{\partial \hat{y}}{\partial z_2} = \hat{\gamma}_2 + \hat{\gamma}_{12} z_1 + 2\hat{\gamma}_{22} z_2 = 0. \qquad (43)$$

The resulting values $(\hat{z}_1^*, \hat{z}_2^*)$ are the estimated optimum values.

It is possible that we have not estimated an optimum but either a saddle point or a ridge (an infinite number of optimum solutions). Furthermore, we may get stuck on a local optimum. For more details we refer to Kleijnen (1987, pp. 202-206, 312-316).

A case study is given by Kleijnen (1988). The study concerns the optimization of a simulated Production Planning System (PPS) with 14 controllable variables. A local first-order model is estimated from the 16 combinations of a 2^{14-10} design. RSM has scaling effects; heuristically a scale is selected such that changing x_j from -1 to +1 means a 20% increase of variable j (j = 1,...,14). Furthermore, not one but two response variables are of interest, namely the number of productive hours y (which excludes idle times and switchover times among machines) and the lead time it takes to deliver products, z. While y is maximized, z is kept below a given limit. Kleijnen (1988) selects a step size such that one controllable variable is doubled while the other inputs are less than doubled. A second local experiment is executed next. This experiment is again determined by a 2^{14-10} design. Unfortunately the improvements of y in the second experiment are rather small. For further details we refer to Kleijnen (1988).

CONCLUSIONS

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

introduction to these techniques is given by Kleijnen and van Groenendaal (1991); a monograph on the subject is Kleijnen (1987).

REFERENCES

Bettonvil B., *Detection of Important Factors by Sequential Bifurcation*, Tilburg University Press, Tilburg, 1990.

Bettonvil B. and J.P.C. Kleijnen, Measurement Scales and Resolution IV Designs, *American Journal of Mathematical and Management Sciences*, 1991.

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

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

Friedman L.W. and H.H. Friedman, Validating the Simulation Metamodel; Some Practical Approaches, *Simulation*, 45(3), 1985, p. 144-146.

Ghosh S., Influential nonnegligible parameters under the search linear model. *Communications in statistics, theory and methods*, 16, no. 4, 1987, pp. 1013-1025.

Ho Y.C., R. Suri, X.R. Cao, G.W. Diehl, J.W. Dille and M. Zazanis, Optimization of large multiclass (non-product-form) queueing networks using perturbation analysis, *Large Scale Systems* 7 (1984) pp. 165-180.

Karplus W.J., The spectrum of mathematical models. *Perspectives in Computing*, 3, no. 2, May 1983, pp. 4-13.

Kleijnen J.P.C., *Statistical Techniques in Simulation*. Marcel Dekker, Inc., New York, 1974/1975. (Russian translation, Moscow, 1978.)

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

Kleijnen J.P.C., Simulation and Optimization in Production Planning: A Case Study, *Research Memorandum* FEW 308, Tilburg University, 1988.

Kleijnen J.P.C., Statistics and Deterministic Simulation Models: Why not?, *Proceedings Winter Simulation Conference*, edited by O. Balci, 1990, pp. 344-346.

Kleijnen J.P.C., Regression Metamodels for Simulation with Common Seeds: Comparison of Techniques, *Management Science* (accepted), 1991.

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. and W. van Groenendaal, *Simulation and Statistics: an Introduction*, Wiley, Chichester, 1991.

Kleijnen J.P.C., G. van Ham and J. Rotmans, Techniques for Sensitivity Analysis of Simulation Models: a Case Study of the CO2 Greenhouse Effect, Tilburg University, 1990.

Naylor T.H., editor, *The Design of Computer Simulation Experiments*, Duke University Press, Durham, N.C. 1969.

Rao C.R., Some problems involving linear hypotheses in multivariate regression, *Biometrika*, 46, 1959, pp. 49-58.

Sacks 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.

Sargent R.G. and T.K. Som, *A different view of frequency domain experiments*. Working paper no. 88-01, Syracuse University, Syracuse (New York 13244), Sept. 1988.

Schruben L.W. and V.J. Cogliano, An experimental procedure for simulation response surface model identification. Communications ACM, 30, no. 8, August 1987, pp. 716-730.

Srivastava J.N., 'Designs for searching nonnegligible effects' in: J.N. Srivastava (ed.). A survey of statistical design and linear models, North Holland Publishing Company, Amsterdam, 1975, 507-519.

Zeigler B., *Theory of Modelling and Simulation*, Wiley Interscience, New York, 1976.

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

440 Jack P.C. Kleijnen
Sensitivity analysis of simulation experiments: regression analysis and statistical design

441 C.H. Veld en A.H.F. Verboven
De waardering van conversierechten van Nederlandse converteerbare obligaties

442 Drs. C.H. Veld en Drs. P.J.W. Duffhues
Verslaggevingsaspecten van aandelenwarrants

443 Jack P.C. Kleijnen and Ben Annink
Vector computers, Monte Carlo simulation, and regression analysis: an introduction

444 Alfons Daems
"Non-market failures": Imperfecties in de budgetsector

445 J.P.C. Blanc
The power-series algorithm applied to cyclic polling systems

446 L.W.G. Strijbosch and R.M.J. Heuts
Modelling (s,Q) inventory systems: parametric versus non-parametric approximations for the lead time demand distribution

447 Jack P.C. Kleijnen
Supercomputers for Monte Carlo simulation: cross-validation versus Rao's test in multivariate regression

448 Jack P.C. Kleijnen, Greet van Ham and Jan Rotmans
Techniques for sensitivity analysis of simulation models: a case study of the CO_2 greenhouse effect

449 Harrie A.A. Verbon and Marijn J.M. Verhoeven
Decision-making on pension schemes: expectation-formation under demographic change

450 Drs. W. Reijnders en Drs. P. Verstappen
Logistiek management marketinginstrument van de jaren negentig

451 Alfons J. Daems
Budgeting the non-profit organization
An agency theoretic approach

452 W.H. Haemers, D.G. Higman, S.A. Hobart
Strongly regular graphs induced by polarities of symmetric designs

453 M.J.G. van Eijs
Two notes on the joint replenishment problem under constant demand

454 B.B. van der Genugten
Iterated WLS using residuals for improved efficiency in the linear model with completely unknown heteroskedasticity

455 F.A. van der Duyn Schouten and S.G. Vanneste
Two Simple Control Policies for a Multicomponent Maintenance System

456 Geert J. Almekinders and Sylvester C.W. Eijffinger
Objectives and effectiveness of foreign exchange market intervention
A survey of the empirical literature

457 Saskia Oortwijn, Peter Borm, Hans Keiding and Stef Tijs
Extensions of the τ-value to NTU-games

458 Willem H. Haemers, Christopher Parker, Vera Pless and
Vladimir D. Tonchev
A design and a code invariant under the simple group Co_3

459 J.P.C. Blanc
Performance evaluation of polling systems by means of the power-series algorithm

460 Leo W.G. Strijbosch, Arno G.M. van Doorne, Willem J. Selen
A simplified MOLP algorithm: The MOLP-S procedure

461 Arie Kapteyn and Aart de Zeeuw
Changing incentives for economic research in The Netherlands

462 W. Spanjers
Equilibrium with co-ordination and exchange institutions: A comment

463 Sylvester Eijffinger and Adrian van Rixtel
The Japanese financial system and monetary policy: A descriptive review

464 Hans Kremers and Dolf Talman
A new algorithm for the linear complementarity problem allowing for an arbitrary starting point

465 René van den Brink, Robert P. Gilles
A social power index for hierarchically structured populations of economic agents

IN 1991 REEDS VERSCHENEN

466 Prof.Dr. Th.C.M.J. van de Klundert - Prof.Dr. A.B.T.M. van Schaik
Economische groei in Nederland in een internationaal perspectief

467 Dr. Sylvester C.W. Eijffinger
The convergence of monetary policy - Germany and France as an example

468 E. Nijssen
Strategisch gedrag, planning en prestatie. Een inductieve studie binnen de computerbranche

469 Anne van den Nouweland, Peter Borm, Guillermo Owen and Stef Tijs
Cost allocation and communication

470 Drs. J. Grazell en Drs. C.H. Veld
Motieven voor de uitgifte van converteerbare obligatieleningen en warrant-obligatieleningen: een agency-theoretische benadering

471 P.C. van Batenburg, J. Kriens, W.M. Lammerts van Bueren and R.H. Veenstra
Audit Assurance Model and Bayesian Discovery Sampling

472 Marcel Kerkhofs
Identification and Estimation of Household Production Models

473 Robert P. Gilles, Guillermo Owen, René van den Brink
Games with Permission Structures: The Conjunctive Approach