Tilburg University

Techniques for sensitivity analysis of simulation models

Kleijnen, J.P.C.; van Ham, G.; Rotmans, J.

Publication date:
1990

Citation for published version (APA):
Kleijnen, J. P. C., van Ham, G., & Rotmans, J. (1990). *Techniques for sensitivity analysis of simulation models: A case study of the CO2 greenhouse effect*. (Research memorandum / Tilburg University, Department of Economics; Vol. 448). Unknown Publisher.

ERSITY

E

UNIVERSITEIT
BRABANT

POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

CIN01138
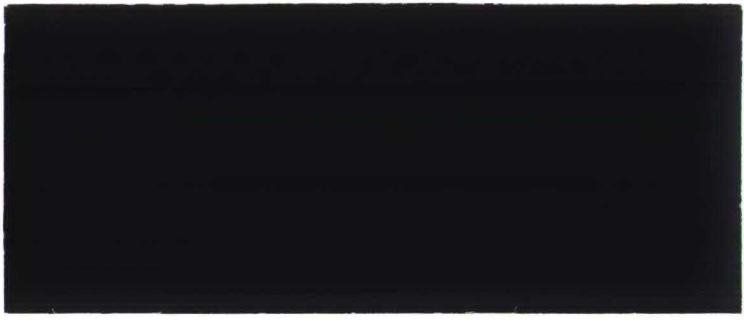
DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

TECHNIQUES FOR SENSITIVITY ANALYSIS OF SIMULATION MODELS: A CASE STUDY OF THE CO_2 GREENHOUSE EFFECT

Jack P.C. Kleijnen and Greet van Ham

FEW 448

ERRATUM

The cover page should be:

TECHNIQUES FOR SENSITIVITY ANALYSIS OF SIMULATION MODELS: A CASE STUDY OF THE CO_2 GREENHOUSE EFFECT

Jack P.C. Kleijnen, Greet van Ham and Jan Rotmans

FEW 448

Tilburg, July 27, 1990

Techniques for sensitivity analysis of simulation models:

a case study of the CO_2 greenhouse effect

Jack P.C. Kleijnen   Greet van Ham

Katholieke Universiteit Brabant

(Tilburg University)

5000 LE Tilburg

The Netherlands

Jan Rotmans

RIVM

(National Institute of Public Health and Environmental Protection)

The Netherlands

Techniques for sensitivity analysis of simulation models:

a case study of the CO_2 greenhouse effect

Jack P.C. Kleijnen   Greet van Ham

Katholieke Universiteit Brabant

(Tilburg University)

5000 LE Tilburg

The Netherlands

Jan Rotmans

RIVM

(National Institute of Public Health and Environmental Protection)

The Netherlands

Abstract

      Sensitivity analysis is needed for validation, what-if analysis, and optimization of complicated simulation models. One set of techniques for sensitivity analysis are least squares curve fitting, regression analysis, and statistical designs such as factorial designs. These techniques are applied to several modules of a large integrated assessment model for the greenhouse effect, developed in The Netherlands. The regression models turn out to be valid approximations to the simulation models. Some estimated effects are quite surprising for the simulation users.

Keywords: What-if analysis, validation, optimization, least squares, regression analysis, experimental design, factorial design, ecology, greenhouse effect, simulation, metamodel.

Introduction

Complicated simulation models have been constructed in many dis-
ciplines. All these models confront the analysts with the problem of sen-
sitivity analysis; that is, what are the effects of changing the parame-
ters and input variables of the simulation model? That question arisis in
'what if' analysis, validation, optimization, and so on. This article
introduces and illustrates the application of *simple* techniques that ori-
ginated in the discipline of mathematical statistics. These techniques are
least squares curve-fitting, regression analysis, and statistical designs
such as 2^{k-p} designs. The techniques are applied to an integrated assess-
ment model for the greenhouse effect. This model has been developed at a
large Dutch institute called National Institute of Public Health and Envi-
ronmental Protection (or RIVM in Dutch).

One of the major imminent ecological threats of the world is the
'enhanced greenhouse problem': the earth and the lower layers of its at-
mosphere have shown rising temperatures over the past hundred years. This
heating phenomenon is probably caused by an increase of greenhouse gases
(such as carbon dioxyde, methane, and ozon) that absorb the earth's heat
radiation, so the global average temperature rises. Mankind is largely
responsible for this increased 'greenhouse' gas concentration.
Temperatures are expected to rise, but with different amounts in different
regions of the earth (the tropics will be less affected probably). Higher
temperatures will cause thermal expansion of the oceans and melting of
artic ice, which raise the sea level. Many more processes, however, are
involved; see [IPCC, 1990]. One consequence of a higher sea level is the
need to raise the level of the dikes in the Netherlands. A survey of the
effects for society is given in [Gezondheidsraad, 1987].

To gain quantitative insight into the greenhouse problem and de-
velop long-term strategies for coping with climatic changes, RIVM deve-
loped the Integrated Model for the Assessment of the Greenhouse Effect
or *IMAGE*. This model is a deterministic simulation (but most of the sensi-

tivity techniques applied to this model can also be used in random simula-
tion models). The state of the dynamic biospheric system is computed per
half year, up to the year 2100, while starting at the year 1900. The model
is composed of modules, which treat specific parts of the greenhouse pro-
blem. Modules get inputs from other modules. Also see figure 1 and the
references [Rotmans, 1990; Rotmans, et al. 1990; den Elzen & Rotmans,
1988].

The sensitivity analysis techniques are applied to several mo-
dules. This paper concentrates on the carbon-cycle module; the dike rai-
sing modules are briefly discussed; see the shaded modules in figure 1.

Note that there are alternative techniques for sensitivity analy-
sis. *Latin Hypercube* sampling is a Monte Carlo method, which is discussed
at length in [Iman and Helton, 1988] and criticized in [Easterling, 1986];
also see [Kleijnen, 1987, p. 143-145]. This technique was applied to seve-
ral *IMAGE* modules in [Lammerts, 1989]; it gave results similar to the
results of this paper. More sophisticated techniques do not treat the
simulation model as a black box; they use analytical differential
analysis; see [McRae, 1989].

This article is organized as follows. First the need for sensiti-
vity analysis is discussed, and the greenhouse case study is introduced.
Then metamodels, which explain the input/output behavior of the underlying
simulation model, are explained. The coefficients of the metamodel are
estimated by least squares regression analysis. The resulting metamodel
can be validated. Closely related to the metamodel specification is the
selection of an efficient experimental design. All techniques are demon-
strated by their application to several modules of the greenhouse simula-
tion model.

Metamodeling through regression analysis

A simulation model maps its inputs into one or more outputs; hence
a simulation model is a mathematical function (say) s( ). The inputs are
parameters, input variables, and behavioral relationships (or submodules);

Figure 1: The modules of *IMAGE*;

the shaded modules are submitted to sensitivity analyses
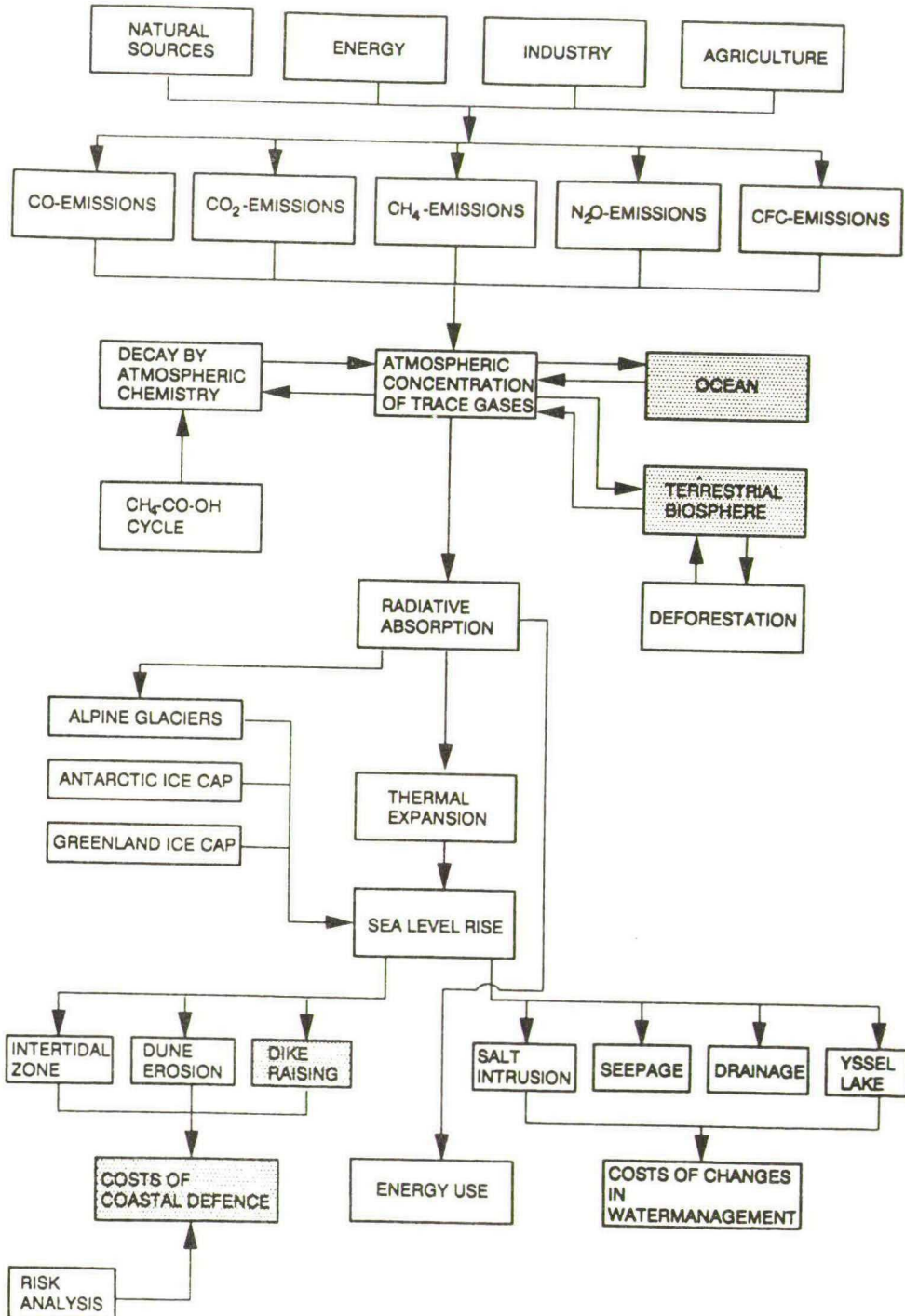
see [Kleijnen, 1987, p. 136]. These inputs are called *factors* in the statistical design of experiments. They may be represented by z_j with $j = s1, \ldots, k$ and $k \geq 1$. In the greenhouse model all factors are quantitative, but the techniques also apply to qualitative factors. The case study concentrates on a single output variable (say) y, namely the global average atmospheric CO_2 concentration in the year 2100. If there are several outputs, the technique can be applied per output. This yields

$$y = s(z_1, \ldots, z_j, \ldots, z_k). \tag{1}$$

A mathematical function may be approximated by a Taylor series, under certain mathematical conditions. Suppose the initial approximation is

$$\hat{y} = \gamma_0 + \sum_{j=1}^{k} \gamma_j z_j + \sum_{j=1}^{k-1} \sum_{h=j+1}^{k} \gamma_{jh} z_j z_h, \tag{2}$$

where γ_0 is called the overall or grand mean; γ_j is the first-order or main effect of factor j; and γ_{jh} is the interaction between the factors j and h, that is, the effect of factor j depends on the level of factor h.

Note that the variables in approximation (2) may be functions of the variables in the simulation model (1); for example, $\log(z_j)$ or $1/z_j$. The approximation remains linear in the parameters γ; so linear regression analysis still applies; see [Kleijnen, 1987, pp. 160-161].

The approximation in (2) is called a *metamodel*: it is a model of the input/output behavior of the underlying simulation model. The Taylor series argument may be one inspiration for the specification of a metamodel. Because the mathematical conditions of the Taylor series do not hold in complicated simulation models, the validity of the metamodel must be checked. In other words, the metamodel is only an approximation. Before that model can be validated, it must be calibrated, that is, its coefficients or parameters γ must be estimated. Moreover there is a scaling problem. These issues are discussed now.

For simplicity's sake the interactions in the metamodel (2) are ignored temporarily. If the input variable z_j increases by one unit, then the output changes by γ_j units. We assume, however, that sensitivity analysis is meant to determine the effects of changes of the inputs over the *whole* experimental area, in order to detect the importance of those inputs. (Next those important factors are further investigated to validate and optimize the simulation model. If only optimization were the goal, then local marginal effects would suffice.) So the importance of factor j is measured by the difference between the outputs at the lowest and the highest value of that factor. Denoting those two extreme values by L_j and H_j respectively (so the 'experimental area' is a k-dimensional rectangle), the original variables z_j yield the *standardized variables* x_j, which range between -1 and +1:

$$z_j = a_j x_j + b_j \text{ with } a_j = \frac{H_j - L_j}{2} \text{ and } b_j = \frac{H_j + L_j}{2} . \qquad (3)$$

The simple transformation (3) together with the original metamodel (2) yields the standardized metamodel:

$$\hat{y} = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k-1} \sum_{h=j+1}^{k} \beta_{jh} x_j x_h . \qquad (4)$$

It is simple to prove that β_j reflects the *importance* of factor j: $\beta_j = \gamma_j (H_j - L_j)/2$, ignoring interactions. See [Bettonvil and Kleijnen, 1990].

Note that to search for the optimum combination of the input factors, Response Surface Methodology (RSM) combines a first-order metamodel with the steepest ascent technique. That search should not use the original or standardized model but a centered model:

$$\hat{y} = \delta_0 + \sum_{j=1}^{k} \delta_j (z_j - \bar{z}_j) + \sum_{j=1}^{k-1} \sum_{h=j+1}^{k} \delta_{jh} (z_j - \bar{z}_j)(z_h - \bar{z}_h), \qquad (5)$$

where $\bar{z}_j = \sum_{i=1}^{n} z_{ij}/n$ assuming that n combinations of input factors are simulated. See [Bettonvil and Kleijnen, 1990].

Calibration means that the parameters of the model are quantified. So the metamodel's parameters β in (4) are estimated. Therefore the meta-model is fitted to the simulation data. Let q denote the number of parameters in the metamodel; for example, in (4) q equals $1 + k + k(k-1)/2$. To get estimated parameter values $\hat{\beta}$, n combinations of the factor values are simulated. That set of simulated combinations yields the n × q matrix of independent variables X corresponding to the metamodel of (4):

$$
X = \begin{bmatrix} 1, & x_{11}, & \cdots & x_{1k}, & x_{11}\ x_{12}, & \cdots & ,x_{1,k-1}\ x_{1k} \\ 1, & x_{i1}, & \cdots & x_{ik}, & x_{i1}\ x_{i2}, & \cdots & ,x_{i,k-1}\ x_{ik} \\ 1, & x_{n1}, & \cdots & x_{nk}, & x_{n1}\ x_{n2}, & \cdots, & x_{n,k-1}\ x_{nk} \end{bmatrix} . \tag{6}
$$

Example: Suppose there are three factors (k=3), which in combination i have the values +1, -1 and -1 respectively. (Remember that standardization means that factor 1 is at its highest level H_1, factor 2 is at its lowest level L_2 in this combination, and so on; see equation 3.) Then the interaction variable $x_1 x_2$ has the value (+1) (-1) = -1 in this combination, and so on. Obviously β_0 corresponds to the 'variable' that is +1 in all combinations. So row i of X equals $x_i' = (+1,\ +1,\ -1,\ -1,\ -1,\ -1,\ +1)$.

The output of combination i is y_i; see (1). Fitting the metamodel to the simulation data, using the *least squares* citerion, yields the estimated parameters $\hat{\beta}$:

$$
\hat{\underline{\beta}} = (X'X)^{-1}\ X'\ y . \tag{7}
$$

The least squares criterion yields unique estimates only if X is non-singular so that the inverse of (X'X) exists. A necessary condition is $n \geq q$ (the number of simulated factor combinations is not smaller than the number of parameters in the metamodel). This condition, however, is not sufficient. For example, if the factors 1 and 2 are changed simultaneously in the n combinations, then their two columns are identical ($x_{i1} = x_{i2}$ for i

= 1,...,n) and X is singular. Obviously X is not singular if all its co-
lumns are *orthogonal* ($\Sigma^n_{i=1} x_{ij} x_{ih} = 0$, etc.). Under certain statistical
assumptions, an orthogonal matrix X is optimal; see the next section.

The calibrated metamodel can now be *validated*. One aspect is how
well this metamodel fits the simulation data. One overall criterion is R^2:

$$R^2 = 1 - \frac{\Sigma^n_{i=1} (\hat{y}_i - y_i)^2}{\Sigma^n_{i=1} (y_i - \bar{y})^2} \; , \tag{8}$$

where $\bar{y} = \Sigma^n_{i=1} y_i/n$. A 'perfect' fit means that $y_i = \hat{y}_i$ for all i, so the
upper limit for R^2 is 1. Unfortunately, a lower threshold for R^2 is hard
to give. Therefore we propose to compute the relative errors $(y_i - \hat{y}_i)/y_i$,
which can be 'eyeballed' by the user.

Validating a model, however, usually means that the model is used
to forecast the output; next that forecast is confronted with the true
output. Therefore *cross-validation* should be used. Delete one combination
(x'_i, y_i) from the old data set (X,y); denote the remaining set by
(X_{-i}, y_{-i}). Reestimate the metamodel's parameters $\underline{\beta}$, analogous to (7):

$$\hat{\underline{\beta}}_{-i} = (X'_{-i} X_{-i})^{-1} X'_{-i} y_{-i} \tag{9}$$

Predict the output of combination i, not using the data of combination i:

$$\hat{y}_i = x'_i \hat{\underline{\beta}}_{-i} \; . \tag{10}$$

Compute the forecast errors

$$e_{-i} = \hat{y}_{-i} - y_i \; . \tag{11}$$

The user may again evaluate the relative errors e_{-i}/y_i. This procedure is
repeated for all i (i=1,...,n).

The computation of the errors e_i is possible without applying the least squares criterion n times (to n-1 combinations). First no data are eliminated; see (7). Next the so-called 'hat' matrix H is computed:

$$H = X(X'X)^{-1} X'$$
(12)

with diagonal elements h_{ii}. Then

$$e_{-i} = e_i/(1-h_{ii}) .$$
(13)

Many modern regression analysis packages give those 'leave one out residuals'. See [Kleijnen, 1987, p. 178] and [Atkinson, 1985, p. 13].

The mathematical analysis can be refined if a *statistical* (sub)model is added for the *fitting errors* e. Kleijnen (1987, p. 164) assumes that these errors are normally and independently distributed with common variance (say) σ^2. Then the least squares algorithm yields the Best Linear Unbiased Estimators (BLUE); that is, the estimators have minimum variances and correct expected values. Those variances equal the main-diagonal elements of the variance-covariance matrix of $\hat{\underline{\beta}}$:

$$cov (\hat{\underline{\beta}}) = (X' X)^{-1} \sigma^2 .$$
(14)

The parameter σ^2 can be estimated through the Mean Squared Error:

$$\hat{\sigma}^2 = \Sigma_{i=1}^{n} (\hat{y}_i - y_i)^2/(n-q) .$$
(15)

The estimated variances (or standard errors) yield the Student t statistic with n - q degrees of freedom:

$$t_{n-q} = (\hat{\beta}_j - \beta_j)/\hat{\sigma}_j \qquad (j=1,\ldots, q) ,$$
(16)

where $\hat{\sigma}_j$ denotes $\{\hat{var}(\hat{\beta}_j)\}^{\frac{1}{2}}$ and β_j is the j th element of $\hat{\underline{\beta}}$; so β_1 in (16) is identical to β_0 in (4), β_2 in (16) is β_1 in (4),...., β_q in (16) is $\beta_{k-1,k}$ in (4). The *significance* of $\hat{\beta}_j$ can be tested statistically, using the t table for a given significance level or type-I error (say) α; for

example, for $\alpha = 0.05$ and $v = 12$ the critical t value is 2.18 in a two-sided test, which considers the absolute value of t_{n-q}.

Note that a more sophisticated model for the fitting errors is used by Sachs et al. (1989). They assume that the errors are not independent but form a stationary process with a specific correlation function. Also see [Kleijnen, 1990].

Statistical design of experiments

The metamodel determines the experimental design;, for example, a model with interactions such as (4) cannot be calibrated through a design that changes one factor at a time; see [Kleijnen, 1987, pp. 266-267]. If purely quadratic terms $\beta_{jj} x_j^2$ are added to (4), then the variable x_j cannot be observed at only two levels (-1 and +1).

Given the metamodel, there is more than one design to calibrate that model. A necessary condition for the design is that the resulting matrix of independent variables X is non-singular; see (7). Consider, for example, a first-order model: in (4) the double summation term vanishes. That model has $q = 1 + k$ effects; so a necessary condition is that the number of combinations satisfies: $n \geq k + 1$. For $k = 3$ table 1 gives two designs that give a non-singular X; X was defined in (6).

Table 1: Two designs for a first order model with k = 3

Combination	One factor at a time design			2^{3-1} design		
	x_1	x_2	x_3	x_1	x_2	x_3
1	-1	-1	-1	-1	-1	+1
2	+1	-1	-1	+1	-1	-1
3	-1	+1	-1	-1	+1	-1
4	-1	-1	+1	+1	+1	+1

The 2^{3-1} design is *balanced*: each column of **X** has an equal number of plus and minus signs, and each pair of columns has an equal number of the four combinations $(-, -)$, $(-, +)$, $(+, -)$, $(+, +)$. If the classical statistical model for the errors is assumed, then the covariance matrix of $\hat{\beta}$ is given by (14). An orthogonal **X** *minimizes* the variances of $\hat{\beta}_j$; see [Kleijnen, 1987, p.335].

If the metamodel includes interactions, then the number of effects increases considerably. To keep the number of combinations relatively small, the user may specify which interactions may be important; the remaining interactions are assumed to be negligible. Examples will be presented later.

The metamodel may be expanded with *purely quadratic* effects: $\sum_{j=1}^{k} \beta_{jj} x_j^2$ is added to (4). These quadratic effects quantify the curvature of the response surface. Then more than two values per factor must be simulated (otherwise all columns for x_j^2 are identical to the column for x_0). A classical design is the *central composite* design: each factor is observed not only at -1 and +1 but also at the 'center point' 0 and at two other values, for example, -2 and +2; together five values. The 2^{k-p} designs that is used to estimate main effects and interactions, is augmented with an observation at the center plus the following 2k combinations:

$$
\begin{array}{llll}
(\ 0, & 0, & \ldots, & 0) \\
(-2, & 0, & \ldots, & 0) \\
(+2, & 0, & \ldots, & 0) \\
(\ 0, & -2, & \ldots, & 0) \\
(\ 0 & +2, & \ldots, & 0) \\
& \vdots & & \\
(\ 0, & 0, & \ldots, & -2) \\
(\ 0, & 0, & \ldots, & +2) \ .
\end{array}
\qquad (17)
$$

Next applications of metamodeling and experimental design will be presented. First the results for a relatively simple module of IMAGE are discussed; then results and technical details for a more complicated module are presented.

Dike raising in IMAGE

One module of IMAGE estimates the magnitude of the necessary dike raise and the resulting costs; see the lower part of figure 1. Eleven factors are examined (k=11); for example, the "unit dike raising cost", which is the cost of increasing the dike level by one meter. Finding a valid meta-model takes several iterations; altogether nine different models (and their concomitant designs) are tried. In an early iteration the metamodel helped to detect a serious error in the underlying simulation model: the original module needed to be split into two modules; the first submodule yields the dike raise necessary to keep the flooding probability under a fixed safety level; the second submodule takes that raise as input and yields the costs as output. So metamodeling may serve *verification* of the simulation model. Moreover, metamodeling may show in which area the simulation model is valid; for factor combinations outside that experimental area the simulation is not a correct model.

To obtain a valid metamodel for the dike raising costs module, the ranges of the original input variables must be decreased. This makes sense mathematically: a Taylor approximation is better in a smaller area. The final metamodel yields relative forecast errors smaller than 10%, which is acceptable for the IMAGE analysts. Nine of the eleven factors are significant, and so is one interaction. The most important factor is the "unit dike raising cost", as the analysts expected. The order of importance of the other factors was surprising, and gives more insight into the simulation model; for details see [Van Ham et al., 1990].

The carbon-dioxide cycle in IMAGE: ocean module

There are two modules for the CO_2 cycle in IMAGE: one for the oceans and one for the terrestrial biosphere; also see the upper part of figure 1. This section covers the first module; the next section discusses the second module.

The oceans show three CO_2 processes, described in [Goudriaan & Ketner, 1984]. For these processes ten factors are investigated; for example, thickness of ocean layers is factor 5. For each factor a range is specified by the analysts; for example, factor 5 varies between 3,000 and 4,000 meters; factor 2 (diffusion coefficient) ranges between 3,716 and 5,984 cm^2/second. These variables are standardized, as described by (3). The analysts list eleven specific interactions that might be important; the remaining 34 interactions are neglected. To verify the design the reader should know that the following interactions may be important: 1 3, 1 5, 2 3, 2 5, 3 6, 3 7, 5 6, 5 7, 5 8, 6 8 and 7 8, where 1 3 stands for $\beta_{1\ 3}$, and so on. So the metamodel is given by (4) with k = 10 and only eleven specific interactions β_{jh}.

A 'full factorial' design requires 2^{10} combinations, which takes too much computer time. The number of effects in the metamodel is: q = 1 + 10 + 11 = 22. Hence a classical 2^{k-p} design with enough combinations requires: n = $2^{10-p} \geq 22$ or p ≤ 5 (least squares applied to the whole data set requires n ≥ q, whereas cross validation requires n > q). There are many 2^{10-5} designs. Accounting for the eleven specific interactions, the following design is selected; details are given in [Kleijnen, 1987, pp. 295-300]. Write down all $2^{k-p} = 2^5$ combinations of the factors 1,2,4,9, and 10. Write down element i of the column for factor 3 as the product of the elements i in the columns for the factors 9 and 10; that is, $x_{i3} = x_{i9} x_{i10}$ for i = 1,..., n and n = $2^{k-p} = 2^5$ or in short-hand: 3 = 9 10. This is called a 'generator' of the design. The 2^{10-5} design is fully specified by its p = 5 generators:

$$3 = 9\ 10 \quad 5 = 4\ 10 \quad 6 = 1\ 9 \quad 7 = 2\ 9 \quad 8 = 1\ 2\ 4, \tag{18}$$

where 8 = 1 2 4 means $x_{i8} = x_{i1} x_{i2} x_{i4}$. The generator 3 = 9 10 means that the main effect of factor 3 is confounded or aliased with the interaction between the factors 9 and 10; that is, $\hat{\beta}_3 = \hat{\beta}_{9\ 10}$ and $E(\hat{\beta}_3) = \beta_3 + \beta_{9\ 10}$. If indeed the interaction $\beta_{9\ 10}$ is negligible, then this confounding is acceptable. Analogously, the generator 8 = 1 2 4 means: $E(\hat{\beta}_8) = \beta_8 + \beta_{1\ 2\ 4}$ where $\beta_{1\ 2\ 4}$ is a 'three factor' interaction, which was not yet

defined in this paper; such high-order interactions, however, are assumed negligible in metamodeling.

The combinations of the 2^{10-5} design are simulated, and the outputs are compared with the predictions of the calibrated metamodel. This results in relative errors exceeding 10% in eight combinations, which is considered unacceptable. Shrinking the ranges of the original variables does not help. Next the metamodel is expanded with purely quadratic effects. The central composite design of (17) requires $1 + 2k$ extra combinations. To save computer time, only five of the ten factors are investigated, namely those five factors that are significant in the previous metamodel. (Because that metamodel is not valid, it is dangerous to use it for the selection of factors; the resulting new metamodel, however, will be validated again.)

The quadratic model is used for the factors 3,4,5,7 and 10. Only four (not ten) interactions between these five factors are conjectured to be important. So the number of effects excluding purely quadratic effects, is: $1 + 5 + 4 = 10$. So the 2^{k-p} design, which is part of the central composite design, must satisfy: $n = 2^{5-p} \geq 10$ or $p \leq 1$. So one generator is selected; namely $4 = 3\ 7\ 10$. These sixteen combinations are augmented with eleven more combinations, following (17). This experiment yields a calibrated metamodel, which is cross-validated. Six more combinations are selected randomly, simulated, and compared with the predicted outcomes. Finally, a 'base' combination is examined; this combination is not the center combination $(0,0,\ldots,0)$ of (17), but is close to it; it is a combination intuitively specified by the analysts. All validation results are acceptable: the errors are smaller than 10%. The individual effects of this accepted metamodel are discussed next.

Because a statistical model for the fitting errors e is assumed, (16) gives the Student t statistic for effects, where the degrees of freedom is $n-q = (16+11) - (10+5) = 12$. For $\alpha = 0.05$ the critical t value is $t_{12}^{\alpha/2} = 2.18$. Table 2 shows significant effects only.

Note that if no such statistical model were assumed, then the last column should be ignored. If the design were orthogonal, then 'significance' and importance would coincide: (14) through (16) show that $t_{n-q} = \hat{\beta}_j/(\hat{\sigma}/\sqrt{n})$. So if effects are sorted in order of magnitude ($|\beta_j|$), they are sorted in order of significance ($|t_{n-q}|$). The central composite design, however, is not orthogonal: quadratic effects and the overall mean are not orthogonal.

Summarizing, originally ten factors are investigated for the ocean module. Because the metamodel without quadratic effects cannot be accepted, a model including such effects is specified. That model, however, is restricted to five factors. The latter model can be accepted, and yields only four important factors. These factors have significant main effects, one significant quadratic effect, and three significant interactions.

Table 2: Significant effects of ocean module

Effect	Estimate	t Statistic
β_0	1074.66	154.08
β_5	-244.95	-139.35
β_4	158.37	98.10
β_3	51.77	32.07
β_{55}	27.79	14.05
β_7	18.78	11.63
β_{45}	- 17.65	- 8.93
β_{47}	- 12.26	- 6.20
β_{35}	- 8.35	- 4.22

The carbon-dioxide cycle in IMAGE: terrestrial biosphere module

The terrestrial biosphere module is described in [Goudriaan & Ketner, 1984]. The analysis of this module is presented, because the mo-

dule contains many input variables: $k = 62$. These 62 variables are described in [Van Ham et al., 1990]. There are designs that yield estimators of main effects without being biased by possible interactions; moreover, these designs yield estimators of confounded interactions: so-called *resolution-IV* designs (see Kleijnen, 1987, p. 301). Such designs require at least $2k$ combinations (so $k + 1 \ll n \ll 1 + k + k(k-1)/2$). For $k = 62$ a resolution IV 2^{k-p} design satisfies: $2^{62-p} \geq 124$ or $p = 55$. So 55 generators must be selected. Each selection yields a specific confounding pattern of estimated effects. The analysts give 26 interactions that might be important. Based on that list, 55 generators are selected; see [Van Ham et al.,1990]. First one estimated effect turns out to have the wrong sign: the effect is significantly positive whereas the analysts expect a negative effect. Next the ranges of the input variables are decreased, and the results become acceptable: the relative forecast errors are small; all (26 confounded) interactions are non-significant; all significant (unbiased) main effects have the correct sign. There are only 13 significant main effects (significance is measured by the t statistic with 128 - (1+62+26) = 39 degrees of freedom). Note that for validation purposes an experiment with twelve randomly selected extra combinations is executed; its relative forecast errors vary between -5.5% and -0.10%.

Conclusion

Any simulation model requires sensitivity analysis. That model can be treated as a black box, if the techniques of regression analysis and experimental design are applied. The regression model is a metamodel of the simulation model, and guides the experimental design. The design leads to efficient and effective experimentation.

The case study demonstrates that application of these statistical techniques requires knowledge of the underlying simulation model and real world system. For example, potentially important factors and their ranges must be given by the analysts. Some statistical expertise is needed to select the generators for the design.

The case study was succesful. The metamodels give acceptable fore-cast errors. The significance of certain effects surprises the analysts. For example, quadratic effects in the ocean module were not expected; and the major importance of the 'biotic growth' factor in the terrestrial module is also surprising. Another surprise is the 'bug' in the dike rai-sing module; the metamodel shows that this module must be split into two submodules.

The sensitivity analysis of IMAGE took quite some time and effort. Fortunately this investment in metamodeling is judged to be profitable. The conclusions of this analysis will guide the development of an interac-tive version of IMAGE.

Summarizing, regression metamodels and experimental designs are useful in the sensitivity analysis of simulation models, as the case study demonstrates. Details on the techniques can be found in [Kleijnen, 1987]; for sensitivity analysis of simulation models with hundreds of factors, special screening designs are presented in [Bettonvil, 1990].

References

Atkinson, A.C., 1985. *Plots, Transformations, and Regression*. Clarendon Press, Oxford.

Bettonvil, B., 1990. *Detection of Important Factors by Sequential Bifurca-tion*, Tilburg University Press, Tilburg, Netherlands.

Bettonvil, B. and J.P.C. Kleijnen, 1990. "Measurement Scales and Resolu-tion IV Designs: a Note", *American Journal of Mathematical and Management Sciences* (to appear).

Easterling, R.G., 1986." Letter to the editor." *Technometrics*, 28: 91-92.

Den Elzen, M.G.J. and J. Rotmans, 1988. *Simulatiemodel voor een Aantal Maatschappelijke Gevolgen van het Broeikaseffect voor Nederland*, (Simulation model for a number of societal consequences of the

greenhouse effect for the Netherlands), RIVM report nr /58471008, Bilthoven, Netherlands.

Gezondheidsraad, 1987. *2e Deeladvies inzake de CO_2-problematiek: Weten-schappelijke Inzichten en Maatschappelijke Gevolgen*, (2nd Report on the CO_2 problem: scientific insights and societal consequences), Staatsdrukkerij, The Hague, Netherlands.

Goudriaan J. and P. Ketner, 1984. "A Simulation Study for the Global Carbon Cycle, Including Man's Impact on the Biosphere", *Climatic Change* 6: 167-192.

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

Intergovernmental Panel on Climate Change (IPCC), 1990. *"Scientific Assessment of Climate Change"*. Report of Working Group I, WMO/UNEP.

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

Kleijnen, J.P.C., 1990. *Statistics and Deterministic Simulation: Why Not?* Katholieke Universiteit Brabant (Tilburg University), Tilburg, Netherlands.

Lammerts, I., 1989. *Onzekerheidsanalyse toegepast op het broeikassimula-tiemodel IMAGE* RIVM, Bilthoven, Netherlands.

McRae, G., 1989. *Designing Air Quality Control Strategies*. Chemical Engineering and Engineering and Public Policy, Carnegie-Mellow University, Pittsburgh (Pennsylvania).

Rotmans, J., 1990. *Modelling the Greenhouse Effect*, Kluwer, Dordrecht (The Netherlands).

Rotmans, J., H. de Boois and R.J. Swart, 1990. "IMAGE: An Integrated Model for the Assessment of the Greenhouse Effect: the Dutch Approach", *Climatic Change*, 16: 331-356.

Sachs, J., W.J. Welch, T.J. Mitchell and H.P. Wynn, 1989. "Design and analysis of computer experiments." *Statistical Science*, 4: 409-435.

Va Ham, G., J. Rotmans and J.P.C. Kleijnen, 1990. *Een Gevoeligheidsanalyse via Metamodellen en Experimentele Proefopzetten, toegepast op een Simulatiemodel voor het Broeikaseffect*, RIVM report, nr. 758471010, Bilthoven, Netherlands.

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 go-
vernments 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

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