Discussion paper

No. 2001-46

DETERMINISTIC VERSUS STOCHASTIC SENSITIVITY ANALYSIS IN INVESTMENT PROBLEMS: AN ENVIRONMENTAL CASE STUDY

By Willem J.H. van Groenendaal and Jack P.C. Kleijnen

July 2001

Deterministic versus Stochastic Sensitivity Analysis in Investment Problems: An Environmental Case Study

Willem J.H. Van Groenendaal[1]
Jack P.C. Kleijnen

Department of Information Systems/Center for Economic Research (CentER)
School of Management and Economics
Tilburg University
P.O. Box 90153
5000 LE Tilburg, The Netherlands
Tel. +31-13-4662423
Fax. +31-13-4663377
E-mail: W.J.H.vGroenendaal@kub.nl

---

[1] Corresponding author.

Deterministic versus Stochastic Sensitivity Analysis in Investment Problems: An Environmental Case Study

Abstract

Sensitivity analysis in investment problems is an important tool to determine which factors can jeopardize the future of the investment. Information on the probability distribution of those factors that affect the investment is mostly lacking. In those situations the analysts have two options: (i) apply a method that does not require knowledge of that distribution, or (ii) make assumptions about the distribution. In both approaches sensitivity analysis should result in practical information about the actual importance of potential factors. For approach (i) we apply statistical design of experiments (DOE) in combination with regression analysis or meta-modeling. For approach (ii) we investigate five types of relationships between the model output and each individual factor; Pearson's ρ, Spearman's rank correlation, and location, dispersion, and statistical dependence. We introduce two distribution types popular with practitioners: uniform and triangular. In an environmental case study both approaches identify the same factors as important.

Keywords: sensitivity analysis, experimental design, investment analysis, simulation

1   Introduction

In practice, investment decisions are often made using the Net Present Value (NPV) criterion; that is, a necessary condition to accept an investment proposal is that the NPV be non-negative. In this paper we address the problem of uncertainty in the model's "inputs" or "factors". To solve this problem, *risk analysis* was introduced by Hertz (1964) and Hillier (1963). That analysis assumes a known joint distribution function of the factor values, which is used to estimate the distribution of the output, NPV. To obtain this output distribution, analysts use either Monte Carlo (MC) or a statistical refinement called Latin Hypercube Sampling (LHS); both techniques are available in software such as @Risk and Crystal Ball (see Buede, 1998; and Evans and Olson, 1998). An investment proposal is accepted if $P(NPV \leq 0) \leq \alpha$, with α decided upon by the decision makers. In practice, however, most analyses are still deterministic, because either no information at all or only very limited information is available on the factor distribution. For both the deterministic and the stochastic problem formulations, a practical question is: Which factors can make a project go "wrong"; that is, which factors may cause $NPV < 0$ and $P(NPV < 0) > \alpha$ respectively. Decision makers ask for this type of information to support their decision making process; see Van Groenendaal (1998b).

One approach to obtain this information applies the statistical theory on design of experiments in combination with regression analysis or meta-modeling (further referred to as DOE). DOE shows which individual factors may jeopardize the results, and which factors interact. In earlier work we reported on the use of DOE in a large investment project; Van Groenendaal and Kleijnen (1997), and Van Groenendaal (1998a). These references, however, do not show how reliable this information is for decision making: does it reveal all important factors, in the correct order of importance? Therefore, we compare the deterministic approach (in which the stochastic character of factors is not modeled explicitly) with an approach that does account for the stochastic nature of factors.

In Risk Analysis a similar problem arises. Suppose we have perfect knowledge about the factor distribution, and have estimated that $P(NPV < 0) = \hat{\alpha} \leq \alpha$ holds. Then the next question is: "Which factors are important, and which are unimportant?". This question is asked in order to monitor the project after the project proposal is accepted. Van Groenendaal (1998b) argues that without information on factor importance the distribution function $P(NPV)$ is only of very

limited use to the decision makers, namely, only in case $P(NPV < 0) > \alpha$ holds and the project is rejected. It does not show which factors are important.

Kleijnen and Helton (1999a, b) analyze a problem that - from a mathematical viewpoint - is similar to our NPV problem. They too wish to detect what effects individual factors have on the output of their model (which concerns nuclear-waste disposal). They argue that simple linear regression models per factor can determine monotonic relationships only. Therefore they argue that additional methods are required. They propose a set of five meta-models and twelve statistical tests to determine factor importance in the case of stochastic inputs of deterministic simulation models, which is a setting similar to risk analysis. Their procedure starts with simple models: a first-order polynomial per factor. If this meta-model is rejected, a monotonic relationship per factor is assumed and tested. If this relation is rejected, they test location dependence, dispersion dependence, and statistical dependence between input and output respectively. They stop testing as soon as a statistically significant relationship is found that can also be explained by "domain" experts.

However, we claim that there is no reason to stop testing after the first statistically significant test. Suppose a number of factors are identified by the first test. If a factor is identified as important by subsequent tests and other factors are not, the relation between this factor and the output requires more thorough analysis. After all, the objective of the tests is to identify which factors to focus on in the sensitivity analysis, and during project implementation.

The goal of this paper is to investigate if DOE and Kleijnen-Helton's procedure may lead to different results. We use a case study concerning a simulation model developed for the NPV analysis of a biogas plant in China (ADB, 1996). We apply both the deterministic and the stochastic approach. For the stochastic approach we assume perfect knowledge about the joint factor distribution function. Next we perform a *robustness* analysis, i.e., we use different factor distributions, and analyze how this change affects the results of the Kleijnen-Helton procedure.

Another way to look at the two approaches is to compare their use of scenarios (combinations of input factors) to generate information. Risk analysis uses a random selection of scenarios (likely as well as unlikely) to generate information. Besides the base case scenario, DOE uses extreme points of the experimental area to generate information, which can be interpreted as a non-random selection of scenarios. The result of DOE may be that some factors are overemphasized in decision making, whereas others are neglected. One of the goals of this paper

is to examine the validity of such a DOE analysis. Both methods, however, have the same goal: determine which factors are important.

The remainder of this paper is organized as follows. Section 2 discusses the two approaches in more detail. Section 3 reviews the NPV case study. Section 4 applies DOE. Section 5 applies the Kleijnen-Helton approach. Section 6 compares the results. Section 7 contains conclusions.

2 Tests for Sensitivity Analysis

The goal of sensitivity analysis is to determine which factors within the total set of factors in the model should be the focus of attention; that is, which factors have really important effects on the output. In an investment model the important factors are the factors that may jeopardize a positive *NPV*. A typical investment model has an evaluation period of more than ten years, and for many factors there is no historic information to estimate future factor values. Therefore practitioners often treat investment project analysis as a deterministic problem.

In DOE information on the effect of factor changes on the *NPV* is obtained by simulating (a subset) of the extreme points of the parameter space, and estimate a linear regression (meta)model to detect which factors are important. This approach also allows us to search for interactions between important factors. Although large investment problems are influenced by many factors, the nature of the problem allows us to analyze the effect of "composite" factors which act as a "funnel" (Van Groenendaal, 1998b). For example, total investment costs depend on many factors. At the highest aggregation level, however, we can restrict the analysis to the effect of *total* investment costs without bothering about the many factors that influence total investment. If this factor is important, the effect of the main factors within total investment (construction costs and material costs) can be analyzed next, etc. In many cases decision makers are actually interested in main categories only and not in details, because the latter become important only in the detailed engineering phase, which is mostly performed after the actual investment decision has been taken.

Kleijnen and Helton (1999a, b) try to find patterns in *scatter plots* of model output versus each factor separately. A first impression of how a stochastic input affects an output can be obtained through a scatter plot; an example is given in Figure 1. They distinguish five types of relationships, for which they apply a number of tests:

(i) Correlation analysis using *Pearson's* ρ for the pairs (X_k, Y) with output Y and factor X_k, $k =$

1, ... , K; K denotes the number of factors. If (X_k, Y) is bivariate normally distributed, then $E(Y|X_k)$ = $\beta_0 + \beta_1 X_k$ and $\beta_1 = \rho\sigma_y/\sigma_x$. Let N denote the sample size in the MC or LHS sample; for example, in Figure 1 $N = 500$, and in our case study $K = 14$. In Figure 1 we display only 1 of the 14 factors. Note that each sampled vector of K factors $(x_{1i}, ... , x_{k,i}, ... , x_{K,i})$, $i = 1, ...,$ N, is input to the deterministic simulation model, and gives a scalar value y_i for Y. The significance of the correlation is tested by Student's statistic.

(ii) Monotonic relations estimated through *Spearman*'s rank correlation, and tested through an approximate normal distribution (Conover, 1999, pp. 314-319). For this test the $x_{k,i}$ are replaced by $r(x_{k,i}) = 1$ for the smallest $x_{k,i}$, $r(x_{k,i}) = 2$ for the next smallest value of $x_{k,i}$, etc. The same is applied to y_i. Then the pairs *($r(x_{k,i})$, $r(y_i)$)*, $i = 1,...,N$, are formed and the monotonic relationship is tested.

(iii) *Location* of Y dependent on X_k, as follows.

a. *Common means*. Kleijnen and Helton divide the domain of x into ten classes. Assuming that the conditional distribution of Y on x_k is approximately normal, they apply the classic ANOVA F-statistic to test whether or not the conditional means $E(Y|X_k=x_k)$ depend on x.

b. *Common locations or Kruskal-Wallis test*. Assuming identical conditional distributions for Y, they apply the Kruskal-Wallis rank test for $E(r(Y|X_k=x_k))$, where $r(Y|X_k=x_k)$ denotes the rank of Y given $X_k = x_k$ (Conover, 1999, p. 288).

c. *Common medians*. To test whether the different classes of input X_k have different median values for Y, they apply the chi-square test for contingency tables; Conover (1999, pp. 218-224).

(iv) *Dispersion* of Y dependent on X_k, tested by the ANOVA F-statistic after jackknifing the variances, and by the chi-square contingency table statistic for interquartile ranges. The latter test is formulated by Kleijnen and Helton (1999a, b), based on the quantile test mentioned in Conover (1999, p. 223).

(v) Statistical *dependence* between the factor X_k and output Y, tested by a chi-square contingency statistic. For this we partition the domain of X_k and Y into ten classes.

For these five types of relationships Kleijnen and Helton (1999a, b) calculate the *critical value*, also known as the probability value or *p*-value: the smallest value of α at which the null-hypothesis would be rejected (Type I error) for the observed value of the test (Iman and Conover, 1983, p. 279).

5

Note, Kleijnen and Helton apply twelve tests in total, whereas we apply only eight tests. We neglect their standardized regression and standardized rank regression tests, because they lead to results very similar to the Pearson and Spearman correlation tests. They also apply two tests on partial correlation coefficients. However, we agree with Conover (1999, p. 327) that this concept is difficult to grasp and hard to interpret. For this reason we do not repeat these tests here.

3 An environmental Case Study: A Chinese Biogas Plant

The Chinese government sees large-scale biogas production as an opportunity to solve several problems simultaneously, namely: (i) the lack of energy in rural areas, (ii) the pollution of the environment by large breeding farms, and (iii) the lack of fertilizer for the agricultural sector. Large-scale biogas digesters produce a convenient form of energy (biogas), while recycling the manure of one or more breeding farms. The residuals of this production process can be used as fertilizer in the production of vegetables, and as an addition to fodder for other stock, such as, pigs, fish, and prawns. A number of factors affect the profitability of investing in large scale biogas. To analyze these factors we formalize the problem as follows.

The investment is in a large scale bio-digester with an annual *rated or design production capacity* of

$$\boldsymbol{Q}^r = F(I,\ L,\ \boldsymbol{M},\ \boldsymbol{E},\ \boldsymbol{Z})$$

where $\boldsymbol{Q}^r = (Q_1,\ Q_2,\ Q_3,\ Q_4)^T$ is a vector of system outputs, with Q_1 biogas, Q_2 liquid sludge, Q_3 fertilizer, and Q_4 regenerated fodder; F is a non-specified production function (with multiple inputs and multiple outputs); I is the investment amount; L denotes labor, $\boldsymbol{M}$ is the vector of other inputs (desulfurizer and water); $\boldsymbol{E}$ is the vector of energy inputs (electricity, coal, and diesel oil); and $\boldsymbol{Z} = (Z_1,\ Z_2,\ Z_3)^T$ are the three raw materials used in production, namely cow dung (Z_1), chicken dung (Z_2), and industrial waste (Z_3), all expressed in metric tons. The investment amount I follows a fixed scheme, and is zero in most years. (We suppressed the time index in all equations that are not dynamic.) The time period studied is $0,\ \dots,\ J$ with J the length of the evaluation period; after J periods the salvage sum is assumed zero.

The design capacity $\boldsymbol{Q}^r$ is achieved under good management practice, but such practice is often lacking. Therefore we introduce the *actual production capacity*

$$Q = {}^T Q^r$$

where $= (\ _1, \ _2, \ _3, \ _4)^T$ is a vector of efficiency rates (mostly, but not necessarily, smaller than 1), which can be improved through better management.

The annual *sales value* of the output is

$$S = \boldsymbol{P_Q}^T \boldsymbol{Q}$$

with $\boldsymbol{P_Q} = (P_{Q_1}, P_{Q_2}, P_{Q_3}, P_{Q_4})^T$ the vector of output prices. The price of biogas P_{Q_1} is strongly *correlated* with the prices of other energy sources. The prices for liquid sludge (P_{Q_2}), fertilizer (P_{Q_3}), and regenerated fodder (P_{Q_4}) depend on the composition of raw materials $\boldsymbol{Z} = (Z_1, Z_2, Z_3)^T$. A higher usage of industrial waste decreases the value of the end products Q_2, Q_3, and Q_4.

The annual *operating costs* are

$$TC = P_L L + \boldsymbol{P_M^T M} + \boldsymbol{P_E^T E} + \boldsymbol{P_Z^T Z}$$

with P_L the price of labor, $\boldsymbol{P_M}$ the vector of prices of intermediary inputs, $\boldsymbol{P_E}$ the vector of prices for energy inputs, and $\boldsymbol{P_Z}$ the vector of raw material prices.

The annual *net benefits* of the investment are

$$NB = S + A - TC$$

where A represents the avoided indemnities and damages to the environment that result from the investment; avoided indemnities are a benefit because without the investment they would have to be paid to the government.

The *net present value* at time t_0 ("now") is

$$NPV_{t_0} = -\sum_{j=0}^{J} \frac{P_{I,t_0+j} I_{t_0+j}}{(1+r)^j} + \sum_{j=1}^{J} \frac{NB_{t_0+j}}{(1+r)^j} \tag{1}$$

where P_I is the price of the investment. This *NPV* is used to evaluate the investment.

If all variables are at their *base case* value (see below), the *NPV* in (1) turns out to be 2.56 million Yuan, so the project is justified financially. There are, however, a number of factors that may affect the *NPV*. For this study we consider the following eight factors; for convenience we also give their base values.

1. The shares of the different inputs in the total Z, for which the vector of base values is $(0.808, 0.114, 0.078)^T$.

2. The total amount of annual input $\sum_{i=1}^{3} Z_i$; base value is 31,000 metric ton.

3. The total investment costs $P_I^T I$; base: 4,961,000 Yuan and a building time of one year.

4. Environmental benefits A; 564,900 Yuan per year.

5. The prices of labor P_L; 4,200 Yuan per year, and the intermediary inputs water and desulferizer $P_M^T = (P_W, P_D)$; (0.48;2,034) Yuan per unit.

6. The price of biogas P_{Q_1}; 0.8 Yuan/m$_g^3$, and the prices of the other energy inputs electricity, diesel oil, and coal $P_E^T = (P_{electricity}, P_{diesel\,oil}, P_{coal})$; (0.375, 1780, 285) Yuan per unit.

7. The prices of the post-processing output liquid sludge (Q_2), fertilizer (Q_3), and fodder (Q_4) ($P_{Q_2}, P_{Q_3}, P_{Q_4}$); (1.627, 813.7, 537.0) Yuan per unit.

8. The efficiency $_1$ of the biogas installation; its base value is 1.029 m$_{gas}^3$/m$_{digester}^3$.

3.1 Factor Uncertainty

We set magnitudes for the possible changes in the base values listed above, as follows. For the factors 1, 2, 5, and 7 we set the maximum changes at ± 20%. For factor 3 the change is ± 25%, based on our previous experience. Factor 4 contains 209,900 Yuan per year of avoided damages, but these are highly uncertain. So we set the change of avoided damages at ± 50%. Given the current law, the indemnities are assumed fixed. For factor 6 we vary the price of biogas by ± 25%, whereas we vary the other energy prices by ± 20%. (The difference in price variation is due to differences in quality between biogas and other fuels.) We vary the efficiency of factor 8 by ± 17%, a value obtained from operating other digesters.

Note that factor 1 actually comprises two factors, factors 1a and 1b: the share of chicken dung (say) α_2 and the share of industrial waste α_3 in the total annual input (the sum of all shares $(\alpha_1, \alpha_2, \alpha_3)$ equals 1). We vary α_2 and α_3 in the same way; that is, if α_2 is at its maximum

(minimum) than so is α_3; hence in DOE the two components are treated as a single factor. In the same way factor 6 comprises four factors, and factor 7 three factors. In the Kleijnen-Helton analysis these factors will be treated as different factors, but they will be made correlated. So in total there are fourteen factors for the stochastic approach (eight in DOE).

We have no other information besides the ranges of the factor values. This lack of more specific information may be quantified through *uniform* marginal factor distributions, with the range as support for these distributions. (Such uniform distributions are called non-informative prior distributions in Bayesian analysis.) For the Kleijnen-Helton analysis this leads to the following stochastic structure.

1. The input shares α_2 and α_3 are uniformly and independently distributed, with $\alpha_2 \sim U(0.091; 0.137)$ and $\alpha_3 \sim U(0.063; 0.094)$.

2. The amount of total input Z is uniformly distributed over the range of $\pm 20\%$ around the base case value.

3. The investment costs $P_I^T I$ are uniformly distributed $\pm 25\%$.

4. Environmental benefits A are uniformly distributed on $U(460,000; 670,000)$.

5. The prices P_M of the intermediary inputs are uniformly distributed $\pm 20\%$, with correlation coefficients of one, so they act as a single factor as they did in the DOE approach.

6. The variation in the prices of energy inputs $(P_{electricity}, P_{diesel\ oil}, P_{coal})$ are correlated with the price of biogas (P_{Q_1}). The price of biogas is uniformly distributed $\pm 25\%$, the other energy prices are uniformly distributed $\pm 20\%$, and the correlation coefficients between all individual prices are assumed to be 0.8. The four energy prices $(P_{Q_1}, P_{electricity}, P_{diesel\ oil}, P_{coal})$ are further called factors 6a through 6d.

7. The prices $(P_{Q_2}, P_{Q_3}, P_{Q_4})$ of the post-processing output liquid sludge (Q_2), fertilizer (Q_3), and fodder(Q_4), called factors 7a, 7b, and 7c, are uniformly distributed $\pm 20\%$. Since liquid sludge and fertilizer are partly substitutes, their correlation coefficient is set at 0.8, whereas the correlation coefficient with fodder is set at 0.6.

8. The efficiency of biogas production, $_1$, is uniformly distributed $U(0.829; 1.171)$.

4 Deterministic Sensitivity Analysis through DOE

For the deterministic investment model we denote the eight factors by X_i ($i = 1, 2, \ldots, 8$);

for these X_i we consider only three values: -1, 0, and 1, where -1 denotes the low value of the range, 0 denotes the base case value, and 1 denotes the high value of the range. In a so-called star design all factors, except one, are kept zero (the star design is a one-factor-at-a-time design). For the specific star design given below, we added 10% to (subtracted 10% of) the high (low) value of the range.

To analyze the effects of the eight factors, we select an unreplicated central composite design (CCD) including a 2_V^{8-2} design (Montgomery, 1991). The star design comprises the 16 axial points, two for each factor i, $(0, ... , a_i, ... ,0) \in \mathbb{R}^8$ and $(0, ... , -a_i, ... ,0) \in \mathbb{R}^8$ plus the central point $(0, ... , 0, ... ,0) \in \mathbb{R}^8$. This design has 81 data points: 64 points of the 2_V^{8-2} design, and the 17 points of the star design. This CCD gives unbiased estimators of the main effects β_i, the two-factor interactions β_{ij}, and the quadratic effects β_{ii} in

$$\underline{Y} = \beta_0 + \sum_{i=1}^{8} \beta_i X_i + \sum_{i=1}^{7} \sum_{j=i+1}^{8} \beta_{ij} X_i X_j + \sum_{i=1}^{8} \beta_{ii} X_i^2 + \underline{\ } , \tag{2}$$

where stochastic variables are underlined.

The result of our analysis is given in Table 1: All main effects turn out to be significant and there are ten significant two-factor interactions, and no significant quadratic effects. R_{adj}^2 is high: 0.98.

Because the CCD uses extreme points of the experimental area, it is not reasonable to assume that the error or residue term $\underline{\ }$ in (2) will be normally distributed. To test normality of the residues we applied Wald's statistic on skewness and kurtosis, and a combined test (Greene, 1993, pp. 309-311). All three statistics are $\ _2^2$ distributed. The statistics turn out to be highly significant, so the assumption of normality of the residues has to be rejected. Therefore, we cannot use the F-test on model reduction that is, $H_0: R\beta = 0$ (Kleijnen, 1987, pp. 155-57). To test for model reduction, we first used the limiting distribution of Wald's statistic

$$W = (R\hat{\beta})^T [R\hat{\sigma}^2 (X^T X)^{-1} R^T] R\hat{\beta} \tag{3}$$

which converges to a chi-square distribution with degrees of freedom equal to the rank of the matrix R (Greene, 1993, pp. 300-301). In (2) there are $(1 + 8 + 8x7/2 +8=)$ 45 coefficients, of which 26 are assumed to be zero. This model reduction is accepted: the value of W is only 6.68

($\chi^2_{26;\,0.05}$ = 38.9).

However, this Wald statistic on model reduction assumes homoscedasticity. Because we simulate extreme points, this assumption may not hold. So next we tested the model reduction assuming heteroscedasticity. Let ($e_1^2, ..., e_i^2, ..., e_m^2$) be the vector of squared estimated residues, with m the number of observations (in our case study $m = 81$), and let $\hat{\Omega}$ denote an estimated covariance matrix with ($e_1^2, ..., e_i^2, ..., e_m^2$) on the main diagonal and zeroes elsewhere. Wald's statistic for the heteroscedastic model is

$$W = (R\hat{\beta})^T [R(X^TX)^{-1}(X^T \hat{\Omega} X)(X^TX)^{-1}R^T]^{-1} R\hat{\beta} \tag{4}$$

which has the same chi-square limiting distribution as before. The value of W turns out to be 17.9, so the model reduction is again accepted.

Further reduction leads to significant W-values for both tests, which indicates that assuming homoscedasticity is permitted here. In summary, Wald's test on model reduction indicates that out of the 45 effects in (2) 19 (the constant, 8 main effects, 10 two-factor interactions, and no quadratic effect) are significant, see Table 1.

TABLE 1 ABOUT HERE

Table 1 further shows that the grand mean $\hat{\beta}_0$ = 2,558,301 is almost equal to the base case value (namely NPV = 2,557,937 Yuan). Equation (2) in combination with Table 1 implies that at the center of the design ($X_i = 0$, $i = 1, ... , 8$) the metamodel correctly "predicts" the simulation outcome. Further, all main effects have the signs expected by experts. Their absolute values indicate their relative importance (because we standardized: $-1 \le X_i \le 1$), assuming the experimental area (the combination of factor ranges) is chosen correctly (see Kleijnen and Van Groenendaal, 1992, pp. 177-178).

5 Stochastic Sensitivity Analysis through the Kleijnen-Helton Approach

To implement the simulation of the NPV model with stochastic inputs, we use the spreadsheet software Excel, combined with Crystal Ball's risk analysis. We apply LHS with $N = 500$

simulation runs. The Lilliefors test for normality (Conover, 1999, p 443) of the 500 observations does reject the assumption of normality, in contrast with what most theorists expect.

Figure 1 shows a scatter plot of the amount invested (X_3) and the NPV (Y). The Kleijnen-Helton analysis tries to find what type(s) of relationships can explain this scatter plot (and others); that is, what type of relationships are there between investment and NPV? So we test the five relationships of § 2, using the eight statistical tests mentioned. The results are as follows (also see Table 2).

(i) The *Pearson* Correlation coefficient, indicating a linear relationship between input and output, is significant at the 5% level for twelve of the fourteen factors. The *p*-values and the Student test results in Column 1 of Table 2 may deviate slightly, because the *p*-values are based on an approximation (Press at al., p. 631).

(ii) The *Spearman* Rank Correlation coefficient, indicating a monotonic relationship, gives a pattern similar to the Pearson correlation; only the order of factors 7a and 1a has changed.

(iii) The *location* of Y depends on X_k. This hypothesis is tested through the following three tests. *Common means*. The ANOVA F-test is significant for nine of the fourteen variables; namely the factors 6a through 7c. The *Kruskal-Wallis* rank test gives the same nine factors as the test on common means does, but in a slightly different order. Finally, the test for *Common medians* gives significant results for the same nine factors as the other two tests on location, but again in a different order.

(iv) *Dispersion of Y depends on X_k.* The test on *common variances* indicates that only two factors have significant effects, namely 8 and 7c. Only one factor is significant, namely factor 6a, using the test on *common interquartiles*. These two tests use related, yet different variability measures. They give quite different conclusions, e.g., factors 8 and 7c may affect the variances, likewise, factor 6a may affect the interquantiles.

(v) *Statistical independence*. Eight factors give significant dependence between X and Y. (We also partitioned the domain for Y and X_k into five instead of ten classes, but this did not change our conclusions.)

TABLE 2 ABOUT HERE

Comparing the results of the various tests shows that all eight tests lead to significant results at

the 5% level. (Increasing the level to 10% or 25% does not alter this conclusion noticeably; not displayed in Table 2.) To see if the tests result in roughly the same factor ranking (identify the same factors as important), we calculate the *top-down correlation*, defined by Iman and Conover (1987) as follows.

The top-down correlation uses the Savage scores $s(h)$ of input X_k: if X_k is ranked (say) h by a test, then $s(h) = \sum_{j=h}^{N} 1/j$. We have 14 variables, so the maximum score is $s(1) = \sum_{j=1}^{14} 1/j \cong 3.25$ and the minimum score is $s(14) = 1/14$. We can form the 14 by 8 matrix $S = (S_1, \dots, S_8)$ of scores with $S_i = (s_{i,1}, \dots, s_{i,14})^T$ the vector of scores of the fourteen input variables X_k according to test i. For this S we then calculate the classical Pearson correlation coefficients $\hat{\rho}_{i,j}$ for the pairs (S_i, S_j), $i = 1, \dots, 7$ and $j = i + 1, \dots, 8$; see Table 3 above the diagonal. Again we test the significance of these correlations by *p*-values using $P(|\rho| > |\hat{\rho}_{i,j}|) = erfc(|\hat{\rho}_{i,j}| \sqrt{14-2}/\sqrt{2})$, where *erfc* is the complementary error function defined in Press et al. (1992, p. 631). These *p*-values are displayed below the diagonal in Table 3. For example, the correlation between the ranking according to tests 2 and 3 is 0.991 (which is in cell {2, 3}), and the corresponding *p*-value is 0.0004 (in cell {3, 2}). Table 3 shows that the correlations between the significant test results are high and significant; i.e., the tests give roughly the same important factor lists.

In the Kleijnen-Helton analysis we used *uniform* distributions. To examine the effects of these distributions, we repeat the analysis with *symmetric triangular* distributions. For every factor the base value is taken as the midpoint of the triangular distribution, and the low and high value as the minimum and maximum. We do not report the full analysis, but the main conclusions only:

i)   In almost all cases the significant factors are the same as in Table 2; that is, the tests show the same pattern of statistically significant results.

ii) The tests indicate the same factors as important, although there is a difference in ranking. In Table 2, factor 8 seems more important than factors 3 and 2, but this is the other way around for triangular distributions. We test the results for uniform versus triangular distribution, using the Iman-Conover top-down correlation; see Table 4, upper part. The correlations for six tests are significant (but all less than 1).

In each experiment (uniform and triangular) all factors have the same distribution. To analyze the effect of *different distribution* types, we use uniform distributions for all factors except for factor 4; for this factor we use the triangular distribution with the midpoint identical to the maximum

value (asymmetry). The effects of this change are moderate; see Table 4 center part.

Finally, we use symmetrically triangular distributions for all factors except for factor 4; for this factor we use again the asymmetric triangular. The same six tests as under ii) give significant results and identify the same factors as being important, but again in a slightly different order. The Iman-Conover top-down correlations are significant - see Table 4, lower part -, indicating that changes in the distributions are important, but do not necessarily lead to completely different results.

6 Comparing DOE and Kleijnen-Helton approach

Both the deterministic and the stochastic approaches can rank factors in order of "importance". The question is: do both methods lead to the same ranking? In our case study, DOE indicated that all factors are important (§ 4). One way to rank the factors according to DOE is by looking at the absolute values of the main effect. For example, factor 3 would then be the most important factor followed by factor 6; see Table 1. However, interaction effects are then neglected, and these can be substantial. We therefore calculate the effect of factor i as the absolute value of its main effect plus the most favorable outcome of the significant interactions between factor i and the other factors. For example, to calculate the effect of factor 5 we assume $X_1 = -1$, $X_2 = 1$, $X_5 = 1$ and all other factors are zero. The effect of factor 5 is calculated as $|-264,147 - 43,461 - 46,371| = 353,979$. This combination of main effects and two-factor interactions leads to the following order of importance: factors 6, 8, 2, 3, 4, 1, 7, 5. The contributions of the first four factors ranges between 1.235 and 1.353 million Yuan, which indicates that these factors are roughly of the same importance. The effect of the next most important factor (factor 4) is only half that size (0.645 million Yuan), and the least important factor (factor 5) is about 25% of that of the most important factors (0.354 million Yuan).

Returning to the Kleijnen-Helton analysis, we now use the significant test results in Table 2. Then clearly the five most important factors are 6, 8, 3, 2, and 4 (which are significant in six of the eight tests). So the tests give almost the same order of importance as DOE does; however, they do not differentiate among factors. DOE clearly indicates that the factors 6, 8, 3, and 2 are more important than factor 4.

Is one approach superior? Certainly not. Apparently DOE can be used to indicate important

factors when no information on the factor distributions is available. An advantage of DOE is that it directly quantifies the magnitude of the effects, which is valuable information for decision makers. The Kleijnen-Helton approach does indicate the same factors as important, but with no information on the magnitude. It does indicate the order through the *p*-values. This may, however, be hard to interpret by non-statisticians. The Kleijnen-Helton approach does, however, result in detailed information on the nature of the relationship between the inputs and the output.

Furthermore, the limited size of the case study and the fact that several sets of factors can be identified that are strongly related (for example, factor 6, energy prices) favor DOE. When the number of factors increases, DOE requires more runs. Actually, we applied a resolution V design, but when the number of factors further increases this may not be feasible. Kleijnen and Helton (1999) analyzed 75 factors. DOE for 75 factors would require much more work and a resolution V design would not be feasible. DOE does on the other hand offer screening methods to deal with very large numbers of factors; for example Bettonvil and Kleijnen (1997) study 281 factors (also see Campolongo et al. (2000)).

Investment analysis uses the $NPV < 0$ as its criterion. Therefore, it is worthwhile to look at the number of times $NPV < 0$ occurs in both approaches. In DOE 9 of the 81 (11.1%) data points are negative. In the stochastic simulation with uniform distributions, $P(NPV \leq 0)$ is less than 1%. To see if this low probability is by accident, we simulated the problem several more times; in all cases $P(NPV \leq 0)$ was less than 2%. The same holds for the simulation with the symmetric triangular distributions. These results may tempt to conclude that the chance that the project goes wrong, is very small. However, we should be cautious. Because we have no information on the exact form of the factors' probability distributions, the estimation of the tail of the NPV distribution is likely to be sensitive to specification errors. To check this, we simulated the model with asymmetric triangular distributions with midpoints identical to the extreme value that has a negative effect on the NPV. In this case $P(NPV \leq 0)$ is 10.26 %, which is similar to DOE.

7 Conclusion

In practice, NPV calculations are made for the base case scenario, using the NPV formula displayed in (1). Analysts and clients are aware of the fact that other scenarios may materialize during the life span of the investment. The aim of sensitivity analysis in investment analysis is to

determine which factors are important and need to be made more precise during the investment analysis and to be monitored more carefully during the construction phase.

Mostly factors are unknown or stochastic by nature. We presented a case study where the only information on the stochastic nature of the factors was the range over which factors vary and the most likely (base case) value. The further analysis depends on whether or not it is reasonable (or necessary) to assume knowledge on the joint probability function of the factors. We compared two approaches: (i) DOE (design of experiments in combination with regression analysis), which assumes no knowledge on the joint probability distribution of the factors except for their ranges, and (ii) an approach developed by Kleijnen and Helton, which assumes the joint probability distribution is known. The two approaches were applied to a case study, namely a model of an investment problem in a large scale biogas plant in China.

In case the analyst is not prepared to make assumptions about the factor distributions, DOE can be applied to identify important factors. Efficient sensitivity analysis is possible if the simulation model implies an I/O transformation that can be approximated by a first- or second-order polynomial in the factors, such as the second-order polynomial in eq. (2). DOE implies that extreme or unlikely scenarios are investigated, namely 'corners' in the space of factor values. We used an un-replicated central composite design to obtain data on the changes of NPV due to factor changes. Estimation of eq. (2) was not straightforward, because the assumption of a normally distributed error term will in general not be met. This was corroborated by Wald's statistics on skewness, kurtosis, and a combined test of the residuals. We therefore used Wald's statistics for model reduction in case of homo- and heteroscedasticity.

Several scenarios result in negative NPV values, indicating that some factors or rather factor combinations can jeopardize the investment. Eight main effect and ten two-factor interactions were identified as important; this is valuable information for the analyst *and* the client.

In case the analyst is prepared or obliged (by the client) to make assumptions on the joint probability distribution of the factors, we can apply a stochastic approach to identify important factors developed by Kleijnen and Helton. They analyze scatter plots of individual factor values versus model output. In this setting every combination of draws from the marginal probability distributions is a scenario, which may or may not be a likely one. Kleijnen and Helton identified five types of relationships between output and factors: linear, monotonic, location of output depends on factor, dispersion of output depends on factor, and statistical dependence of output

16

and factor. They give twelve tests to see what relationships are present; we applied eight of their tests.

For the factor distributions we introduced two assumptions popular with practitioners: uniform and triangular marginal distributions. The factor ranges used in DOE, are also used to support the uniform marginal distributions. Triangular marginal distributions, with the mode at the base case value, are often used when analysts and clients feel that the base case value is more likely than the extreme values.

For both types of probability distributions the Kleijnen-Helton approach indicates that three types of relationships are significant in the case study: linear, the location of the output depends on particular factors, and there is statistical dependence between the output and factors.

To investigate the effect of the assumed factor distributions, we also introduced a mix of thirteen uniform and one triangular distribution. The same factors are found to be important.

In all stochastic analyses the probability of a negative NPV is small (less than 2%). However, this result needs to be interpreted with care. A simulation with asymmetric triangular marginal distributions shows that the 2% becomes 10%.

In the case study, DOE and the Kleijnen-Helton approach identify the same factors, in almost the same order. DOE, however, indicates possible interactions between factors, whereas the Kleijnen-Helton approach analyzes the relation between individual inputs and outputs only. In case of investment projects, however, information on interactions is valuable.

A disadvantage of DOE is that it takes many more runs when the number of factors increases. The Kleijnen-Helton approach is more robust in this respect.

References

ADB (1996) Rural energy study in the People's Republic of China, Final Report, Asian Development Bank TA 2100, April 1996.

Bettonvil, B. and J.P.C. Kleijnen (1997) Searching for important factors in simulation models with many factors: sequential bifurcation. European Journal of Operations Research, 96(1): 180-194.

Buede, D (August 1998) Decision analysis; software survey: aiding insight IV. *OR/MS TODAY*, pp. 56-64.

Campolongo, F., J.P.C. Kleijnen, and T. Andres (2000) Screening Methods. In A. Saltelli, K. Chan, and E.M. Scott (eds.) *Sensitivity Analysis*, John Wiley & Sons, Chichester, England, pp. 65-89.

Conover, W.J. (1999), *Practical Non-parametric Statistics; Third Edition*. Wiley, New York.

Evans, J.R. and D.L. Olson (1998) *Introduction to simulation and risk analysis*. Prentice-Hall, Upper saddle River, New Jersey.

Greene, W.H. (1993), *Econometric Analysis*, Macmillan Publishing Company, New York.

Hertz, D.B. (1964) Risk analysis in capital investment. *Harvard Business Review*, Jan.-Feb., 95-106.

Hillier, F.S. (1963) The derivation of probabilistic information for the evaluation of risky investments. *Management Science*, Vol 9, April, 443-57.

Iman, R.L. and W.J. Conover (1983) Modern Business Statistics; Second Edition. Wiley, New York.

Iman, R.L. and W.J. Conover (1987), A measure of top-down correlation. *Technometrics*, 29, no. 3, pp. 351-357.

Kleijnen J.P.C. (1987) *Statistical tools for simulation practitioners*, Marcel Dekker, New York, U.S.A.

Kleijnen J.P.C. and W.J.H. Van Groenendaal (1992) *Simulation: a statistical perspective*, John Wiley & Sons, Chichester, England.

Kleijnen, J.P.C., and J. Helton (1999a) Statistical analyses of scatter plots to identify important factors in large-scale simulations, 1: review and comparison of techniques. *Reliability Engineering and Systems Safety*, 65, 2, pp. 147-185.

Kleijnen, J.P.C., and J. Helton (1999b) Statistical analyses of scatter plots to identify important factors in large-scale simulations, 2: robustness of techniques. Reliability Engineering and Systems Safety, 65, 2, pp. 187-197

Montgomery, D.C. (1991), *Design and Analysis of Experiments*, *3rd Edition*. Wiley, New York.

Press, W.H., S.A. Teukolsky, W.T. Vetterlings, and B.P. Flannery (1992), *Numerical Recipes: the Art of Scientific Computing*. 2nd edition. Cambridge University Press, Cambridge, England.

Van Groenendaal, W.J.H. (1998a) Estimating NPV variability for deterministic models, *European Journal of Operations Research*, Vol. 107(1):202-213.

Van Groenendaal, W.J.H. (1998b) *The Economic Appraisal of Natural Gas Projects*, Oxford University Press, Oxford.

Van Groenendaal, W.J.H. and J.P.C. Kleijnen (1997) On the assessment of economic risk: factorial design versus Monte Carlo methods. *Journal of Reliability Engineering and Systems Safety*, 57, pp. 91-102.

Table 1: Regression meta-model based on an unreplicated central composite design; only estimates significant at $\alpha = 0.05$ are displayed

Coefficient	Estimate	Coefficient	Estimate
$\hat{\beta}_0$	2,558,301	$\hat{\beta}_{1,2}$	61,967
$\hat{\beta}_1$	309,817	$\hat{\beta}_{1,5}$	43,461
$\hat{\beta}_2$	912,900	$\hat{\beta}_{1,6}$	34,736
$\hat{\beta}_3$	-1,235,250	$\hat{\beta}_{1,7}$	34,178
$\hat{\beta}_4$	644,872	$\hat{\beta}_{1,8}$	26,877
$\hat{\beta}_5$	-264,147	$\hat{\beta}_{2,5}$	-46,371
$\hat{\beta}_6$	1,084,915	$\hat{\beta}_{2,6}$	216,987
$\hat{\beta}_7$	310,418	$\hat{\beta}_{2,7}$	62,100
$\hat{\beta}_8$	869,229	$\hat{\beta}_{2,8}$	173,846
		$\hat{\beta}_{6,8}$	217,307
$R^2_{adj} = 0.98$			

Table 2: Ranking of factors according to their p-values

Linear relationship		Monotonic relationship		Location of Y depends on X_k						Dispersion of Y depends on X_k				Statistical independence	
Pearson correlation		Spearman correlation		common means		Kruskal-Wallis		common medians		common variance		common interquartiles		statistical independence	
ranked factors	p-value	ranked factors	p-value	ranked factors	p-value	ranked factors	p-value	ranked factors	p-value	ranked factors	p-value	ranked factors	p-value	ranked factors	p-value
6a[1]	0.0000	**6a**	0.0000	**6a**	0.0000	**6a**	0.0000	**6a**	0.0000	**8**	0.0216	**6a**	0.0034	**6a**	0.0000
6c	0.0000	**6c**	0.0000	**6c**	0.0000	**6c**	0.0000	**8**	0.0000	**7a**	0.0464	4	0.0605	**8**	0.0000
8	0.0000	**8**	0.0000	**8**	0.0000	**8**	0.0000	**3**	0.0000	1b	0.0609	6c	0.0943	**6b**	0.0000
6b	0.0000	**6b**	0.0000	**6b**	0.0000	**3**	0.0000	**6c**	0.0000	6b	0.1656	8	0.3314	**6c**	0.0000
3	0.0000	**3**	0.0000	**3**	0.0000	**6b**	0.0000	**6d**	0.0000	7c	0.2660	1a	0.3838	**3**	0.0000
6d	0.0000	**6d**	0.0000	**6d**	0.0000	**6d**	0.0000	**6b**	0.0000	4	0.2713	6d	0.4410	**6d**	0.0000
2	0.0000	**2**	0.0000	**2**	0.0000	**2**	0.0000	**2**	0.0000	6a	0.2878	7c	0.4866	**2**	0.0004
4	0.0000	**4**	0.0000	**4**	0.0000	**4**	0.0000	**4**	0.0000	6d	0.3948	7b	0.5341	4	0.2404
7c	0.0000	**7c**	0.0000	**7c**	0.0056	**7c**	0.0080	**7c**	0.0343	6c	0.5555	7a	0.5503	5	0.4236
7a	0.0018	**1a**	0.0027	7b	0.0669	7b	0.0870	1a	0.2044	1a	0.5665	3	0.5831	7c	0.4729
1a	0.0042	**7a**	0.0037	7a	0.1375	5	0.1150	1b	0.3071	3	0.6371	6b	0.7319	1a	0.5105
5	0.0111	**5**	0.0047	5	0.2183	1b	0.1404	7b	0.3703	2	0.7836	1b	0.7479	7a	0.5232
7b	0.0544	**7b**	0.0430	1b	0.2619	7a	0.1960	5	0.3977	5	0.9119	2	0.9114	1b	0.5232
1b	0.0552	1b	0.0596	1a	0.3100	1a	0.2374	7a	0.8514	7b	0.9828	5	0.9401	7b	0.6115

[1] Bold numbers indicate insignificant test results at the 5% level.

Table 3: Uniform versus triangular top-down correlations and their *p*-values

		top-down correlation coefficients $\hat{\rho}_{i,j}$, $i = 1, \dots ,7$ and $j = i + 1, \dots , 8$							
		1	2	3	4	5	6	7	8
estimated *p*-values	1	-	0.999	0.993	0.993	0.919	*0.115*	0.712	0.944
	2	0.0003	-	0.991	0.991	0.922	*0.098*	0.718	0.945
	3	0.0003	0.0004	-	0.992	0.917	*0.098*	0.698	0.937
	4	0.0003	0.0004	0.0003	-	0.938	*0.051*	0.694	0.925
	5	0.0009	0.0009	0.0009	0.0007	-	*0.166*	0.673	0.942
	6	0.6789	0.7243	0.7229	0.8535	0.5500	-	*0.065*	*0.233*
	7	0.0103	0.0096	0.0118	0.0123	0.0153	0.8160	-	0.633
	8	0.0007	0.0007	0.0007	0.0008	0.0007	0.4012	0.0224	-

Remark: The cells above the diagonal contain the top-down correlations and the cells below the diagonal the corresponding *p*-value.

.

Table 4: Top-down correlations for different factor distributions

	Pearson correlation	Spearman Correlation	common means	Kruskal-Wallis	common medians	common variance	common interquartiles	statistical independence
	all uniform versus all symmetric triangular							
$\hat{\rho}$	0.7405	0.7373	0.7324	0.7610	0.7332	0.4747	0.0201	0.7376
p-value	0.0076	0.0079	0.0083	0.0061	0.0082	0.0870	0.9422	0.0078
	all uniform versus all uniform except factor 4 which is non-symmetric triangular							
$\hat{\rho}$	0.7495	0.7676	0.7088	0.7599	0.7878	-0.0543	-0.0554	0.7627
p-value	0.0069	0.0056	0.0106	0.0045	0.0045	0.8447	0.8418	0.0060
	all triangular versus all triangular except factor 4 which is non-symmetric triangular							
$\hat{\rho}$	0.9212	0.9543	0.9607	0.9706	0.9608	0.2428	0.3968	0.9518
p-value	0.0009	0.0006	0.0005	0.0005	0.0005	0.3813	0.1525	0.0006

Figure 1: Scatterplot of NPV versus Investment Costs