Estimating net present value variability for deterministic models
van Groenendaal, W.J.H.

Publication date:
1995

Citation for published version (APA):
Abstract: For decision makers the variability in the net present value (NPV) of an investment project is an indication of the project’s risk. So-called risk analysis is one way to estimate this variability. However, risk analysis requires knowledge about the stochastic character of the inputs. For large, long-term investment projects (such as energy infrastructures) modelling their stochastic character is often difficult, if not impossible. The analysis of the variability is then restricted to deterministic sensitivity analysis, such as one factor at a time and scenario analysis. However, these deterministic analyses do not account for the total variability in the NPV. It will be shown that the use of experimental design, taken from statistical theory, in combination with regression metamodelling is a better approach to estimating the variability of the NPV.

Keywords: Risk, Sensitivity analysis, Experimental design, Metamodels, Regression.

1 Introduction
For the assessment of the feasibility of large investment projects, such as gas transmission systems, large non-linear models are often built. In general, only one solution for the investment problem is designed [9, 14]. An important criterion for evaluating an investment project is its net present value (NPV). However, if for the most likely or base case scenario the project is profitable (NPV > 0), information on the robustness of the base case result is needed to support the decision makers in their assessment of a project’s risk. To support such risk assessment, risk analysis is often proposed; that is, the probability distribution of the NPV is estimated. To obtain this distribution, information is required on the distribution of the input variables and model parameters (further referred to as factors). However, modelling the joint distribution of even a limited number of input variables and parameters can be rather problematic, if not impossible, for large projects. Especially for investments in large energy infrastructures this is true, due to the versatile nature of the energy market. As a result of these problems, the analysis is often restricted to one factor at a
time sensitivity analysis or the analysis of a limited number of scenarios, where we use the term sensitivity analysis to indicate deterministic approaches to the problem [7]. Note that the one factor at a time approach is the most popular one in applied work; for case studies see [2, 12, 16].

Furthermore, both deterministic methods result in limited and often misleading information about the variability of the \( \text{NPV} \); especially the one factor at time approach, which is crude and is known to be naive and inefficient from a statistical point of view [8, pp. 171-2]. More reliable information about the variability of the \( \text{NPV} \) can be obtained through the application of experimental designs that are taken from statistical theory.

In this paper we outline a procedure to improve traditional sensitivity analysis through the application of experimental design theory in combination with regression metamodelling. We will show that the use of experimental designs leads to more adequate information about the variability of the \( \text{NPV} \). and that metamodelling can be used to evaluate the variability. The procedure developed requires no knowledge about the distribution of the input variables or the parameters of the simulation model. This procedure has the advantage that it is easy to apply and leads to better results than traditional sensitivity analysis does.

We demonstrate the approach by a case study, namely the investment in gas transmission on the Indonesian island of Java. This project requires an overall investment of U.S.$ 1.9 billion in 1990 prices, and will be realized in three separate investment phases. Phase I is utilizing the gas reserves in the vicinity of Java, to develop the market. Phase II is the connection to gas reserves in East Kalimantan more than 1,000 miles away, and Phase III is the installation of compressor stations to boost the systems capacity. For the base case scenario the \( \text{NPV} \) for the project is 898.5 billion Indonesian rupiah in 1990 prices (1 U.S.$ = 1770 Rupiah), so the project is feasible. Ten factors are apriori thought to influence the \( \text{NPV} \) (see the first column of Table 1); they present other projects or government policies (diversifiable risk) as well as indicators for economic development (non-diversifiable risk). Our results for this project give the Indonesian government clear insight in the variability of the \( \text{NPV} \) and thus the risks of investing in gas transmission on Java.

This paper is organized as follows. In Section 2 a three-step procedure for the determination of the variability in the \( \text{NPV} \) is outlined. In Section 3 the procedure is applied to investments in gas transmission on Java; significant main effects and factor interactions are estimated. In Section 4 the robustness of the estimation results is investigated. Section 5 contains conclusions.

2 Determination of NPV variability through deterministic simulation

The first step in the estimation of the \( \text{NPV} \) variability is the determination of the effects of positive and negative change in the factor’s value. Normally, decision makers are more interested in information on what can go wrong than in information on ‘pleasant surprises’. For our analysis of variability this implies that we are especially interested in factor changes that have a negative effect on the \( \text{NPV} \). In our case study the project is financially sound: for the base case scenario the \( \text{NPV} \) is 898.5 billion rupiah. The next thing decision makers want to know is: does the project go wrong, when the actual factor values deviate from the base case values? In other words, deviations from the base case values that will improve the result, are less interesting (a positive \( \text{NPV} \) is sufficient to accept the project). The decision makers want to know under what circumstances the decision to invest will be incorrect, and they wish to judge the likelihood of these unfavorable circumstances. To determine factor changes that cause negative effects, the first step is applying one factor at a time sensitivity analysis.

Note that restricting the analysis to negative effects only, implies that we reduce the experimental area for our design. This affects the results that can be obtained through metamodelling, as we will show later.
Note further that this first step is required only when it is not known in which direction factors affect the $NPV$, which in general will be the case for large investment projects. For example, faster economic growth will, among others, lead to more sales, but will also lead to an earlier construction of the second and third phase of the project. More sales will have a positive effect on the $NPV$, but earlier investment in a connection to gas resources further away will affect the $NPV$ negatively. The direction of the overall effect of increased economic growth is therefore not known in advance.

The one factor at a time approach has as an advantage that this approach is easy to understand because it treats every factor separately. However, changing \textit{one factor at a time} is also its main weakness, because it does not allow us to check for interactions among the factors. To obtain information on both main effects (as the one factor at a time approach does) and interactions (combined effects), other experimental designs are necessary. Popular designs are full factorial designs with every factor at two levels only, which require $2^k$ simulation runs for $k$ factors; so ten factors require $2^{10} = 1024$ simulation runs. Such designs enable us to estimate all possible interactions, not only between two factors [8, pp. 172-4]. In other words, a full factorial design - even with a moderate number of factors - requires too many simulation runs; and it is certainly not a practical solution in case there are many factors. The large number of simulation runs was one of the reasons for proposing risk analysis [10]. Furthermore, since the simulation model is large, each simulation run requires substantial computer time. This is not feasible in a commercial setting for project evaluation. Moreover, except for the pathological case where interactions among all factors exist, a full factorial design is not necessary, as we shall see. In most cases, accounting for two factor interactions is sufficient.

Fortunately, there are experimental designs that allow us to estimate all main effects and a limited number of interactions, but require much less than $2^k$ simulation runs. So experimental design theory can help to meet the information needs of the decision makers. This alternative has - to the best of our knowledge - been neglected in project appraisal (none of the many guidelines for project evaluation used by international lending agencies mentions this approach).

What we need is an approach that allows us to estimate \textit{all main effects, and possibly some interactions}. Suppose for the time being, that we have $k$ factors and there are no interactions at all. Then the \textit{full factorial} design needs $2^k$ observations to estimate only $k+1$ effects, namely $k$ main effects plus the overall mean. In principle $k+1$ observations suffice to estimate $k+1$ effects [8, p. 175]. In other words, for the $k+1$ effects it suffices to simulate a fraction of the $2^k$ observations. These designs are called fractional factorial designs; such as, $2^{k-p}$ designs and Plackett-Burman designs. Obviously, if only main effects are important, the minimum number of simulation runs is obtained by choosing the maximum $p$ such that $2^{k-p} \geq k+1$ still holds. In our case we have $k+1 = 11$, which calls for a $2^{10.6} (= 16)$ design.

Fractional factorial designs have a number of runs equal to a power of two. So especially when the number of factors becomes large, the number of simulation runs can still become too large for our purpose. A class of designs that allows a more gradual increase in the number of simulation runs is the \textit{Plackett-Burman} design type. These designs require a number of runs equal to a multiple of four. Hence, for ten factors the Plackett-Burman design with twelve runs can be used [5, pp. 329-36]. The transposed Plackett-Burman design matrix (say) $X^T \in \mathbb{R}^{10 \times 12}$ is given in Table 1. Every column in that table represents a simulation run. A plus sign (+) stands for the base case value of the corresponding factor, and a minus sign (-) for the value that has a negative influence on the base case result. It is easy to see that the columns of the Plackett-Burman design matrix are orthogonal; thus $(X^T X)^{-1} = 12^{-1} I$, where $I \in \mathbb{R}^{10 \times 10}$ denotes the identity matrix. Furthermore, the design satisfies one linear constraint: the sum of the first eleven rows of $X$ equals minus row twelve. The augmented matrix $(e : X)$, with $e = (1, 1, \ldots, 1) \in \mathbb{R}^{12}$ has the same
properties as the matrix $\mathbf{X}$. We shall return to these mathematical properties. Note that the rows of the design matrix (or columns in Table 1) can be interpreted as scenarios, some of which may make economic sense.

Table 1: Plackett-Burman design for ten factors (Source: [5, p. 332])

<table>
<thead>
<tr>
<th>combination factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 investment costs</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 construction time</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3 reserves West Java</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4 real GVA</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5 energy prices</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6 relative gas/oil price</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>7 purchase prices</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8 coal prices</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9 other costs</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10 discount rate</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The third (and final) step in our procedure is analyzing the data obtained in the second step, by fitting a regression metamodel; that is, we approximate the I/O behavior of the (complicated) simulation model through a regression model [8, pp. 149-50]. If we assume that only main effects and two-factor interactions are important, the metamodel has the form

$$
\mathbf{y}_i = \beta_0 + \sum_{j=1}^{10} \beta_j \mathbf{x}_j^i + \sum_{j=1}^{10} \sum_{k=j+1}^{10} \beta_{jk} \mathbf{x}_j^i \mathbf{x}_k^i + \epsilon_i, \quad i = 1, 2, \ldots, 12
$$

(1)

where $\mathbf{y}_i$ denotes $\mathbf{NPV}_i$ (that is, the result of the i-th simulation run), $\mathbf{x}_j^i$ denotes the standardized value of factor $j$ in simulation run $i$, and $\epsilon_i$ denotes the approximation error, which is assumed to be 'white noise', that is, $\epsilon_i$ is additive normally independently distributed noise with $E(\epsilon_i) = 0$ and $V(\epsilon_i) = \sigma^2$. The 'white noise' assumption allows us to use simple statistical analysis and tests (such as the F and t tests, as we shall see). Alternatively, we can drop this assumption, and use the Ordinary Least Squares (OLS) curve fitting, which yields a less powerful analysis (such as $R^2$, no t values).

The choice of the experimental area affects this analysis. From a statistical point of view, we should use the total experimental area, that is, use the widest range in factor levels. This will give the best variance estimator for the metamodel. However, the metamodel is only an approximation to our simulation results and the fit will be better when the experimental area is smaller. We regard the latter argument more important than the first. So there is also a statistical motivation for the analysis of negative effects only.

If there are no interactions, the OLS estimates for the coefficients $\beta_0$ and $\mathbf{\beta}_M = (\beta_{1,2}, \ldots, \beta_{9,10})^T$ in (1) are: $\hat{\beta}_0 = 12^{-1} \sum_{i=1}^{12} \mathbf{y}_i$, because $e^T \mathbf{X} = 0$ since every column in $\mathbf{X}$ contains an equal number of +1 and -1 entries; and $\hat{\mathbf{\beta}}_M = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = 12^{-1} \mathbf{I} \mathbf{X}^T \mathbf{y}$. However, if the assumption of no interactions is false, these estimates are biased, as we show now.

Denote the vector of two-factor interactions by $\mathbf{\beta}_A = (\beta_{1,2}, \ldots, \beta_{1,10}, \beta_{2,3}, \ldots, \beta_{9,10})^T \in \mathbb{R}^{45}$. 
Denote the matrix of independent variables associated with by 
\( \mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_9) \in \mathbb{R}^{12 \times 45} \) 
with 
\( \mathbf{V}_j = (x_{j1}, x_{j2}, \ldots, x_{j12}). \) 
\( x_{j1} \) denotes the 12-dimensional vector with elements 
\( x_j x_{j2}, \) 
j = 1, 2, \ldots, 12. Then the expected value of \( \mathbf{\hat{p}}_M \) is

\[
E(\mathbf{\hat{p}}_M) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \mathbf{p}_M + \mathbf{V} \mathbf{p}_A) = \mathbf{p}_M + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V} \mathbf{p}_A
\]

(2)

The matrix \( 12^{-1} \mathbf{X}^T \mathbf{V} \) is called the alias or bias matrix; this matrix shows how the main effects are confounded with the interactions. (For a comprehensive treatment of the subject we refer to \cite{11}.)

The estimator for the main effects \( \mathbf{\hat{p}}_M \) would be unbiased if \( \mathbf{X}^T \mathbf{V} = \mathbf{0} \) in (2). Unfortunately \( \mathbf{X}^T \mathbf{V} = \mathbf{0} \) does not hold for the Plackett-Burman design; fortunately this equality can be achieved by applying the Box-Wilson foldover theorem; see \cite[p. 303]{6}. A foldover is obtained by adding \( -\mathbf{X} \) to the original design matrix \( \mathbf{X} \), so twenty-four instead of twelve simulation runs are executed. Through this addition of \( -\mathbf{X} \) to \( \mathbf{X} \) the OLS estimator for \( \mathbf{\hat{p}}_M \) is no longer biased by two-factors interactions \cite[pp. 413-4]{5}. A resolution IV design is a design in which no main effect is confounded with any other main effect or any two-factor interaction, but in which the two-factor interactions are confounded with each other \cite[Chapter 25]{6}. (Obviously no unbiased estimators of all main effects and two-factor interactions are possible: there are \( 1 + 10 + (10 \times 9/2) = 56 \) effects and only 24 runs.) Applying the foldover technique changes (1) into

\[
\mathbf{\hat{p}} = (\mathbf{X}^T \mathbf{X})^{-1}\mathbf{X}^T \mathbf{V} \mathbf{p}_A
\]

(3)

The OLS estimator \( \mathbf{\hat{p}} \) follows from (3):

\[
\mathbf{\hat{p}} = (\mathbf{Z}^T \mathbf{Z})_Y, \text{ where } \mathbf{Z}^T \mathbf{Z} = \begin{pmatrix}
24 & 0 & 0 \\
0 & 24I & 0 \\
0 & 0 & 2V^T V
\end{pmatrix}.
\]

(4)

This is easy to prove: remember that every column in \( \mathbf{X} \) and \( \mathbf{V} \) contains as many -1 as +1 elements, and that \( \mathbf{X} \) is an orthogonal matrix.

The maximum number of two-factor interactions we can estimate with the foldover is equal to the rank of \( \mathbf{V} : \mathbf{V}^T \), which is equal to the rank of \( \mathbf{V} \). Because \( \mathbf{V} \) consists of combinations of columns of \( \mathbf{X} \), the rank of \( \mathbf{V} \) can not exceed the row rank of the Plackett-Burman design matrix for twelve factors, which is only eleven (since the addition of the first eleven rows of this design matrix is equal to minus row twelve). If we calculate the rank of \( \mathbf{V} \) it is actually eleven. This rank implies that we can estimate up to eleven two-factor interactions.

The structure of \( \mathbf{Z}^T \mathbf{Z} \) implies that we can compute \( \mathbf{\hat{p}}_0 \) independently of \( \mathbf{\hat{p}}_M \) and \( \mathbf{\hat{p}}_A \), namely

\[
\mathbf{\hat{p}}_0 = \sum_{i=1}^{24} y_i / 24. \quad \text{Similarly we get } \mathbf{\hat{p}}_M = 24^{-1} \mathbf{X}^T \mathbf{Y}. \quad \text{Next we select eleven independent}
\]
columns from $\mathbf{V}$ to form the matrix $\mathbf{V}_j$. The rest of the columns of $\mathbf{V}$ are combined in the matrix (say) $\mathbf{V}_A$. The resulting alias matrix $\left(V_i^T V_j\right)^{-1} V_i^T V_A$ for the eleven interactions can be formed in the same way as we did for $\hat{\mathbf{b}}_p$ in (2); also see Kleijnen [5, pp. 415-6]. The alias structure for the Plackett-Burman design and for its foldover is rather complicated [13, p. 223].

Note that if we want estimates of more than eleven interactions, we have to further augment our design (beyond the foldover), and perform new simulation runs. Next we will apply this three-step procedure to estimate the variability in $\mathbf{NPV}$.

3 Applying the Three-Step Procedure to the Case Study

The results of the first step (one factor at a time analysis) are given in Table 2, which allows us to determine those changes that have a negative effect on the $\mathbf{NPV}$. The factor values in Table 2 are chosen such that the experimental area is as large as possible; that is, the boarders of the experimental area are the extreme values for the variables considered.

Note that if the values associated with factor changes are chosen correctly, the magnitudes of the estimated $\hat{\beta}_j$'s in (1) are comparable and show the importance of the effects [1].

Note further that all entries in Table 2 are positive, so no single effect causes the project to go wrong. For decision makers this information might be misleading.

Most results are as one would expect: higher (lower) costs lead to lower (higher) $\mathbf{NPV}$ values (factors 1, 2, and 9). An increase in the price of coal leads to an increase in the demand for natural gas by the power sector, and thus to a higher $\mathbf{NPV}$.

The change in initial energy prices (factor 5) has the largest effect on $\mathbf{NPV}$. This needs some explanation. We priced natural gas at fuel oil parity, and the change in initial energy prices implies that relative gas-fuel oil prices remain the same. Therefore the change in volume sold is small for the manufacturing sector. However, given the cost curves for the different technologies of the power sector, the changes in price levels of coal and natural gas cause changes in the technology mix for the power sector. The volume effect of this change is considerable. Factor 5 shows the importance of a correct (economically efficient) pricing of energy.

Changing the price of natural gas relative to the price of oil (factor 6) has only a minor effect on the $\mathbf{NPV}$. However, lowering the relative gas-oil price has a positive effect. The cause is that lowering the relative price leads to conversion of existing production processes to natural gas, but does not affect the demand by new investments. In the base case natural gas is priced at fuel oil parity (154 rupiah per m³). At this price nearly all new investments will use gas, whereas most existing facilities will not convert to gas; also see [16]. Figure 1 shows the gas-oil price sensitivity pattern for existing production processes. If the price is lowered to 152 rupiah, some food processing switches to natural gas. Next textile and cardboard changes; the last big increase in demand is between 146 and 144 rupiah per cubic meter, when the paper and pulp production switches to natural gas. (The bar at 0 rupiah per cubic meter in Figure 1 indicates the maximum demand for the period displayed.) All changes in demand discussed are within the -10% relative price decrease; the increase in demand volume offsets the decrease in price, which explains the increase in $\mathbf{NPV}$. (Apparently the price elasticity of demand is larger than -1, in this price range.)
Table 2: One-factor-at-a-time sensitivity analysis for \( NPV \)

<table>
<thead>
<tr>
<th>factor j</th>
<th>value</th>
<th>( NPV )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 investment costs</td>
<td>( \pm 25% )</td>
<td>659.1</td>
</tr>
<tr>
<td>2 construction time</td>
<td>one year extra</td>
<td>822.0</td>
</tr>
<tr>
<td>3 gas reserves in West Java</td>
<td>( \pm 20% )</td>
<td>1,141.5</td>
</tr>
<tr>
<td>4 real GVA</td>
<td>see Appendix 6.1</td>
<td>992.7</td>
</tr>
<tr>
<td>5 energy prices</td>
<td>( \pm 5% )</td>
<td>1,254.1</td>
</tr>
<tr>
<td>6 relative gas-oil price</td>
<td>( \pm 10% )</td>
<td>737.9</td>
</tr>
<tr>
<td>7 purchase prices</td>
<td>( \pm 5% )</td>
<td>472.6</td>
</tr>
<tr>
<td>8 coal price</td>
<td>( \pm 3% )</td>
<td>956.2</td>
</tr>
<tr>
<td>9 other costs</td>
<td>( \pm 25% )</td>
<td>829.9</td>
</tr>
<tr>
<td>10 discount rate</td>
<td>10% ( \pm 2% )</td>
<td>577.4</td>
</tr>
</tbody>
</table>

| base case                                     |                  | 898.5      |

1) \( + \) and \(- \) correspond with a positive and negative change in the input respectively.

Note that analysts often simulate more than two extra levels in one factor at a time sensitivity analysis. In general, intermediate simulation runs will add no extra information, unless there are very strong non-linearities; also see [3].

To analyze the effects of the size of the experimental area, we use the data of Table 2 to estimate the main effects for the one factor at a time approach. We associate the vector \( (1,0,\ldots,0) \in \mathbb{R}^{11} \) with the base case. With a positive and negative factor change we associate the values +2 and -2 respectively (instead of the more familiar values +1 and -1). The estimates are then scaled in such a way that they become comparable with the results for the negative area only (on which we will concentrate in the next step of our procedure). Without this scaling the estimates cannot be compared. The design matrix then becomes:

\[
X = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2
\end{pmatrix} \in \mathbb{R}^{21 \times 10}.
\]

(5)

For the estimation of \( \beta_0 \) this matrix has to be augmented with the vector \( (1,1,\ldots,1) \in \mathbb{R}^{21} \).
Table 3: Estimation results for a one factor at a time design

<table>
<thead>
<tr>
<th></th>
<th>full experimental area</th>
<th>negative area only</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_0$</td>
<td>871.8</td>
<td>898.5</td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>119.7</td>
<td>119.7</td>
</tr>
<tr>
<td>$\hat{\beta}_2$</td>
<td>45.2</td>
<td>38.3</td>
</tr>
<tr>
<td>$\hat{\beta}_3$</td>
<td>210.9</td>
<td>300.4</td>
</tr>
<tr>
<td>$\hat{\beta}_4$</td>
<td>49.9</td>
<td>52.6</td>
</tr>
<tr>
<td>$\hat{\beta}_5$</td>
<td>242.2</td>
<td>306.6</td>
</tr>
<tr>
<td>$\hat{\beta}_6$</td>
<td>51.0</td>
<td>80.3</td>
</tr>
<tr>
<td>$\hat{\beta}_7$</td>
<td>213.0</td>
<td>213.0</td>
</tr>
<tr>
<td>$\hat{\beta}_8$</td>
<td>41.5</td>
<td>34.3</td>
</tr>
<tr>
<td>$\hat{\beta}_9$</td>
<td>34.3</td>
<td>54.2</td>
</tr>
<tr>
<td>$\hat{\beta}_{10}$</td>
<td>212.1</td>
<td>160.6</td>
</tr>
</tbody>
</table>

$R^2_{adj}$ 1)

D.F. 2)

1) $R^2_{adj} = \left( R^2 - \frac{k}{m-1} \frac{m-1}{m-k-1} \right)$, with m the number of observations, and k the number of factors (excluding $x_0 \equiv 1$).

2) Not defined for $m = k$.

3) D.F.: degrees of freedom $m - k - 1$.

Figure 1: Demand-price sensitivity of manufacturing sector
Table 4: \( NPV \) results for the Plackett-Burman design (\( X \)) and its foldover (\( -X \))

<table>
<thead>
<tr>
<th>comb.</th>
<th>factor ( 1 ) ( 2 ) ( 3 ) ( 4 ) ( 5 ) ( 6 ) ( 7 ) ( 8 ) ( 9 ) ( 10 )</th>
<th>( X )</th>
<th>( -X )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+ + + + + + - - - -</td>
<td>130.6</td>
<td>-370.8</td>
</tr>
<tr>
<td>2</td>
<td>+ + + + + + - - - -</td>
<td>-1,017.8</td>
<td>125.4</td>
</tr>
<tr>
<td>3</td>
<td>+ + + + + + - - - -</td>
<td>-189.4</td>
<td>-554.2</td>
</tr>
<tr>
<td>4</td>
<td>+ + + + + + - - - -</td>
<td>-271.9</td>
<td>20.5</td>
</tr>
<tr>
<td>5</td>
<td>+ + + + + + - - - -</td>
<td>191.3</td>
<td>-572.2</td>
</tr>
<tr>
<td>6</td>
<td>+ + + + + + - - - -</td>
<td>208.0</td>
<td>-672.2</td>
</tr>
<tr>
<td>7</td>
<td>+ + + + + + - - - -</td>
<td>-154.5</td>
<td>-90.7</td>
</tr>
<tr>
<td>8</td>
<td>+ + + + + + - - - -</td>
<td>446.9</td>
<td>-842.7</td>
</tr>
<tr>
<td>9</td>
<td>+ + + + + + - - - -</td>
<td>-609.9</td>
<td>220.2</td>
</tr>
<tr>
<td>10</td>
<td>+ + + + + + - - - -</td>
<td>294.9</td>
<td>-588.2</td>
</tr>
<tr>
<td>11</td>
<td>+ + + + + + - - - -</td>
<td>-879.8</td>
<td>179.8</td>
</tr>
<tr>
<td>12</td>
<td>+ + + + + + - - - -</td>
<td>-730.4</td>
<td>898.5</td>
</tr>
</tbody>
</table>

The OLS estimates of the eleven coefficients are given in Table 3 under the heading 'full experimental area'. Because of the structure of \( X \), there are only two values for the standard deviations: \( \hat{\sigma}_j = 22.0 \) and \( \hat{\sigma}_j = 35.7 \) \( (j = 1, 2, \ldots, 10) \). With \( 21-11 = 10 \) degrees of freedom, the estimates for \( \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_3, \hat{\beta}_5, \hat{\beta}_7, \) and \( \hat{\beta}_{10} \) in the \( NPV \) model are significant at the level \( \alpha = 0.01 \) \( (t_{10,0.01} = 2.76) \). \( \hat{\beta}_4 \) and \( \hat{\beta}_6 \) are only just significant at the \( \alpha = 0.10 \) level \( (t_{10,0.10} = 1.37) \). All other estimates are insignificant.

The one factor at a time approach uses 21 observations to estimate 11 parameters. However, eleven simulation runs would have been sufficient. Because we are more interested in negative effects than in windfalls, we also estimated the main effects based on negative changes only. The estimates of \( \hat{\beta} \) based on these data are given in Table 3 under the heading 'negative area only'.

The latter estimates show a slightly different ranking of factors. For factor 10 (the discount rate) a difference in magnitude is obvious, because of its non-linear effect in the calculation of \( NPV \). The differences between the two estimates for the factors 3 and 5 indicate that the metamodel is inappropriate; that is, higher order effects or interactions between factors might be important.

Remark: The literature on risk analysis has suggested other methods to detect what factors are important. To rank factors, Hull \[4, pp. 120-1\] introduces the term range coefficient \( (NPV_{H} - NPV_{L})^2 / \max (NPV_{H} - NPV_{L})^2 \) , with \( NPV_{H} \) and \( NPV_{L} \) the maximum \( H \) and minimum \( L \) value for factor \( j \) respectively. It is easy to see that \( (NPV_{H} - NPV_{L})^2 = 4\hat{\beta}_j \) (\( \hat{\beta}_j \) the OLS estimate of the \( j \)-th main effect). Because taking the square of a non-negative variable is a monotone transformation, the range coefficient will lead to the same ranking as our approach does. If we apply Hull's criterion to eliminate factors at a level of 0.10, we would declare the factors 1, 3, 5, 7, and 10 to be important. So the results of both methods are the same.
The results of the twenty-four simulation runs for $NPV$ using the foldover of the Plackett-Burman design are shown in Table 4. Combinations 1 to 12 are identical to the twelve columns of Table 1. Remember that our analysis focusses on those conditions that will jeopardize the positive advise for the investment project that follows from the positive $NPV$ value in the base case. Hence all results are expected to be worse than the base case result. This explains the many (fourteen) negative entries in the columns denoted by $X$ and $-X$ in Table 4. Each single minus value of the ten factors lowers the $NPV$; so only if the interactions of the factors have positive influences, the negative effects of the factors are mitigated. In case the interaction of factors strengthens the main effects, the results will be even more negative.

For example, in combination 12 all factors are at their 'minus' level; hence the $NPV$ is negative (namely, -730.4). In the foldover part ($-X$) the corresponding combination has all factors at their 'plus' or base case level. Hence this $NPV$ corresponds to the base case result and is positive (898.5). However, all factors at their 'minus' level does not give the most negative result. The effect of a larger discount rate (factor 10) can mitigate the effect of negative cash flows. This occurs for combinations 2 and 11 in the column denoted by $X$, and for combination 8 in the column denoted by $-X$.

Estimation of main effects and important two-factor interactions

We start our analysis with the OLS estimation of the main effects for both the original Plackett-Burman design and the foldover (using the $NPV$ data of Table 4). When we simply eyeball the estimation results in Table 5, it looks as if the foldover does not have much effect on the estimated main effects. The main effects that seem important when using the data of the Plackett-Burman design only, are in order of importance $\beta_5$, $\beta_3$, $\beta_7$, and $\beta_1$ (besides the overall mean $\beta_0$). When using the foldover, these effects remain the most important ones; the t-values for these five coefficients are all significant at $\alpha = 0.05$ ($t_{13-24-11} = 2.85$). If we compare Tables 3 and 5 we see that, with the exception of $\beta_{10}$, the same main factors are important.

If we delete those regressors whose estimated main effects do not significantly differ from zero at $\alpha = 0.05$, then the $R^2_{adj}$ increases from 0.87 to 0.91 (not displayed in the table). Note that because $X^T X = 24 I$ is a diagonal matrix, the estimates of the main effects do not change when we delete or add regressors.

Once we know which main effects matter, we want to determine whether or not there are significant two-factor interactions. So we estimate a model with only the significant main effects plus the first eleven independent interactions, which are $(\hat{\beta}_{1,2}, \hat{\beta}_{1,3}, ..., \hat{\beta}_{1,10}, \hat{\beta}_{2,3}, \hat{\beta}_{2,6}) \in R^{11}$. This raises the $R^2_{adj}$ from 0.90 to 0.97. The estimates for the interactions and $R^2_{adj}$ are also displayed in Table 5 (again, because $X$ is an orthogonal matrix, the estimates for the main effects remain the same). The increases in $R^2_{adj}$ suggest that there are significant interactions; indeed several interactions have the same magnitudes as some main effects have.

A more formal test on interactions is the F-test for the hypothesis $H_0: \beta_{1,2} = \beta_{1,3} = ... = \beta_{2,6} = 0$ (see Kleijnen [6, pp. 156-7]). The value of the F-statistic is 4.80, which is significant at the 5% level, since $F_{11}^{0.05} = 3.31$ (the degrees of freedom are explained as follows: 11 interactions are hypothesized to be 0, and $24 - (5 + 11) = 8$ degrees of freedom.
Table 5: Estimates for the Plackett-Burman and its foldover

<table>
<thead>
<tr>
<th></th>
<th>Plackett-Burman ($\chi$)</th>
<th>foldover ($\chi$: $-\chi$)</th>
<th>two factor interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_{10}$</td>
<td>-54.4</td>
<td>17.5</td>
<td>$\hat{\beta}_{26}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{9}$</td>
<td>-42.5</td>
<td>-10.0</td>
<td>$\hat{\beta}_{23}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{8}$</td>
<td>-71.9</td>
<td>-16.2</td>
<td>$\hat{\beta}_{20}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{7}$</td>
<td>220.1</td>
<td>206.8</td>
<td>$\hat{\beta}_{19}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{6}$</td>
<td>-20.4</td>
<td>-26.8</td>
<td>$\hat{\beta}_{18}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{5}$</td>
<td>272.1</td>
<td>293.2</td>
<td>$\hat{\beta}_{17}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{4}$</td>
<td>-20.8</td>
<td>-9.1</td>
<td>$\hat{\beta}_{16}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{3}$</td>
<td>243.5</td>
<td>237.3</td>
<td>$\hat{\beta}_{15}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{2}$</td>
<td>-91.9</td>
<td>-7.5</td>
<td>$\hat{\beta}_{14}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{1}$</td>
<td>137.7</td>
<td>146.7</td>
<td>$\hat{\beta}_{13}$</td>
</tr>
<tr>
<td>$\hat{\beta}_{0}$</td>
<td>-215.2</td>
<td>-201.2</td>
<td>$\hat{\beta}_{12}$</td>
</tr>
</tbody>
</table>

| $R^2_{adj}$ | 0.87 | 0.87 | 0.97 |
| D.F.        | 1    | 13   | 8    |

are left for the full model). So the $H_0$ hypothesis is rejected for the $NPV$ model, which suggest important interaction effects.

Note that accepting $H_0$ only means zero weighted sums of two-factor interaction effects that can be obtained from the alias matrix; see [13, p. 223]. Failure to reject $H_0$ does not prove that there are no interactions!

There are $\sum_{j=1}^{11} \binom{45}{2}$ possible models with two-factor interactions. So determining which individual interactions are significant, is problematic without making further assumptions (or executing more simulation runs). A first new assumption is that there are only interactions between factors with significant main effects; this assumption is also often made in Risk Analysis to limit the number of factors. For the $NPV$ model this assumption refers to the interactions between investment costs and West Java’s reserves ($\hat{\beta}_{13}$), investment costs and all energy prices ($\hat{\beta}_{15}$), investment costs and the purchase prices ($\hat{\beta}_{17}$), West Java’s reserves and all energy prices ($\hat{\beta}_{3,5}$), West Java’s reserves and purchase prices ($\hat{\beta}_{3,7}$), and all energy prices in combination with a change in purchase prices ($\hat{\beta}_{8,7}$). Estimation of this model, and testing it against the model without interactions shows that there are indeed some significant interactions. However, the $F$-statistic is 3.14, which is barely significant at the 5% level ($F_{13,0.05} = 2.92$). We do not find this result very satisfactory.

Taking into account those interactions that are related to significant main effects, is a reasonable approach when no other information is available; the simulation model is then treated as a black box. However, in our case (and in most other applications) this approach is not necessary. We can derive some clues from our simulation model itself and from the intermediate simulation results that lead to the $NPV$ values in Table 4. That table gives only the $NPV$, not any more details on the simulation runs. In the simulation model, economic growth plays a role in nearly all its submodels. However, Table 5 shows that economic growth (factor 4) has no significant effect on the $NPV$. This seems odd, and is also contrary to what economic theory tells us. Studying the
detailed simulation results of the twenty-four simulation runs shows that economic growth does strengthen the effect of a change in the West Java reserves (factor 3), but it also strengthens the effects of some of the changes in prices (factors 5, 6, 7, and 8). We therefore restrict our search for interactions to these six variables (factors 3-8). After some analyses we find that four main effects (see above) and four two-factor interactions give the best results:

\[ N\hat{PV} = -201.2 + 146.7x_1 + 237.3x_3 + 293.2x_5 + 206.8x_7 + \\
+70.8x_3x_4 + 102.8x_5x_8 + 106.3x_4x_6 - 38.1x_7x_8 \]  

\( R^2 \text{adj} \) is 0.98. The hypothesis \( H_0: \beta_{3,4} = \beta_{3,8} = \beta_{4,6} = \beta_{7,8} = 0 \) yields \( F_{15} = 15.09 \) and is significant, even at the 0.5% level (\( F_{15,0.005} = 5.80 \)). Equation (6) shows that economic growth \((x_4)\) plays a role in the determination of the \( N\hat{PV} \) in combination with West Java’s reserves \((x_3)\), which seems logical since faster growth means earlier depletion. Economic growth also affects the \( N\hat{PV} \) in combination with the relative gas-oil price \((x_6)\). There are also significant interactions between the depletion of West Java’s reserves and the price of coal \((x_8)\), and between the purchasing price of natural gas and the price of coal \((x_{1,8})\). The fact that the constant \( \beta_0 \) differs from zero means that there is some non-linearity in the simulation results. However, the magnitude is moderate and requires no further research. Because we have chosen the range for every factor as wide as possible the magnitude of the coefficients indicates their relative importance.

Note that for the base case (all \( x_i = 1 \)) the value for \( N\hat{PV} = 924.6 \), which is a good forecast for the base case simulation result (898.5). Factors are usually not at their extreme values; fractions of -1 or +1 can be introduced in to predict the result.

Equation contains the information on the variability in the \( N\hat{PV} \) and on the most important factors that can cause this variation. This is the kind of information decision makers require. They can learn the following from equation:

- West Java gas reserves are of crucial importance to the project. These reserves are needed for the development of the market for natural gas in West Java; smaller reserves strongly affect the project’s feasibility. Many of the two-factor interactions are related to West Java’s reserves. These interactions imply that investment in exploration and exploitation in West Java will be profitable.
- The project is sensitive to energy prices. A correct pricing of energy products is therefore a prerequisite for the success of the project. Most responsive to relative price changes is the power sector. Changing the relative coal-gas price affects the economic viability of the project, mainly through the power sectors’ fuel-mix. The importance of the power sector is supported by our analysis of an overall change in energy prices, which strongly affects the \( N\hat{PV} \). In this case too, the effect on the amount of gas sold results from the fuel-mix for the power sector.
- The demand for natural gas by the non-bulk consumers of the manufacturing sector is quite robust with respect to relative gas-oil price changes. The explanation is that these are high valued applications.
- The \( N\hat{PV} \) is strongly affected by purchase price increases (which is equivalent to a reduction of the transmission margin).
- Changes in investment costs have only a moderate influence on the \( N\hat{PV} \).
- It is further remarkable that economic growth is important, only in combination with West Java’s reserves and energy prices. This importance through combined effects (interaction) clearly shows that the widely used practice in risk analysis of deleting those factors from the analysis whose main effect is insignificant, can easily lead to errors.
4 Sensitivity Analysis of the Variability

The validity of approximation can be tested through cross-validation. Cross-validation means that we eliminate factor input combinations (say) \( x_i \) (the columns of \( X \) and \( X^\top \)) and the corresponding simulation result \( y_i \) one by one, and re-estimate the regression model. With the resulting regression model we can predict the simulation result for the combination eliminated. We denote this prediction by \( \hat{y}_{-i} \). If the metamodel (5) is robust, the prediction \( \hat{y}_{-i} \) will be close to \( y_i \). For our Plackett-Burman foldover we have twenty-four data points \( (i = 1, 2, \ldots, 24) \), so we can make 24 predictions \( \hat{y}_{-i} \). (For more details on cross-validation see [7, pp. 156-7].)

To test the robustness of deterministic models Kleijnen and Van Groenendaal [8] suggest to compute the relative prediction error \( \frac{\hat{y}_{-i} - y_i}{y_i} \). However, in our model \( y \) denotes \( NPV \), which can change from positive to negative. This switching of signs implies that the relative prediction error is not a good measure. For example, assume that we predict \( NPV = 1 \) by \( NPV \hat{=}-10 \), and \( NPV = 1,000 \) by \( NPV \hat{=}-1,500 \). The first prediction error is -10, and the second one is 1.5. According to the criterion \( \frac{\hat{y}_{-i} - y_i}{y_i} \), the first error would be regarded more serious than the second one, whereas the absolute difference \( |\hat{y}_{-i} - y_i| \) for the first prediction is only 11 compared to 500 for the second one. It is obvious that the second prediction error is worse than the first. To indicate the quality of the predictions \( \hat{y}_{-i} \) obtained through cross-validation, we find a scatter plot more appropriate. If our model is perfect, the scatter plot is a straight line. We quantify this performance by the correlation coefficient between \( \hat{y}_{-i} \) and \( y_i \). This coefficient is 0.987. Figure 2 shows that the regression model performs satisfactorily.

![Figure 2: Scatter plot of NPV regression predictions and simulation realizations](image)

5 Conclusions

For many long-term investment projects no information on the probability distribution of the input
variables and model parameters is available. Hence, risk analysis (or any other probabilistic approach) is virtually impossible. However, decision makers need information on the variability of the project’s \( NPV \) for the assessment of its risk. Traditionally one factor at a time sensitivity analysis is used or a limited number of scenarios are analyzed. However, these two deterministic approaches are insufficient.

To obtain information on the variability of a project’s \( NPV \), we developed a three-step procedure based on experimental design theory and regression metamodelling. It was shown that metamodels can be used to gain insight into the variability, i.e., use a regression model to approximate the input-output behavior of the simulation model. Experimental design theory can be used to design simulation experiments that, through systematic factor variation, result in the data required for fitting the metamodel. Furthermore, using experimental design theory allows us to estimate interactions between factors, which is not possible in the traditional approaches (nor in risk analysis).

One could argue that our interpretation of the interactions is not supported by a complete statistical analysis, since we did not systematically check all factor-interactions. However, because the simulation model is not a black box analysts who understand their problem will normally be able to qualitatively derive where interactions can be expected; that is what we did here. The metamodel obtained for the variability of the \( NPV \) is statistically sound and is supported by our knowledge of the problem at hand.

The resulting information on the variability in the \( NPV \) tells decision makers which factors are important. Because we have carefully chosen the range within which the factors vary, the magnitude of an estimate \( \hat{\beta}_j \) is also an indication for its importance. We restricted our experimental area to those factor changes that have a negative influence on the \( NPV \), because decision makers are more interested in what can cause the project to go wrong then in pleasant surprises. The smaller experimental area will also improve the fit of the metamodel.

To test the robustness of the metamodel, we used cross-validation. This analysis showed that the parameter estimates are stable.
References


