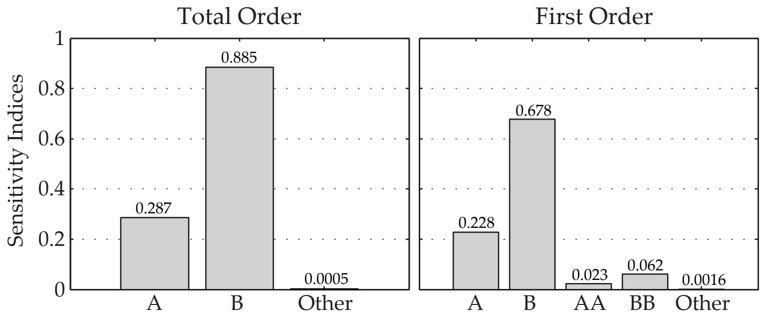



Figure 3.15 Sensitivity indices obtained with the Regression Based Sensitivity Analysis. Down-link margin.

formances was already evident from the graphs previously discussed. However, the RBSA provides us with a quantitative measure of the absolute importance of the design factors in the determination of the performances. The *Output RF Power* and the *Antenna Diameter* contribute for more than 50% of the variability of the subsystems mass while they influence almost all the variability of the *down-link margin*. The *Type of Antenna* (C) and the *Type of Transmitter* (E) affect the mass of the subsystems because of their power density with respect to the aperture diameter and the output power, according to the mathematical models discussed in Appendix A. These interactions are evident in the bars of Figure 3.14 named (BC) and (AE), respectively. The *Type of Solar Array* (D) contributes to a limited extent to the variability of the mass of the two subsystems. This is due to the relatively low difference between the values of the power density of the solar cells selected from the data base for the analysis, see Table A.3. Their contribution is mainly quadratic, correctly indicating that there is a minimum (in this case) of mass when the selected array is the one corresponding to the second level of the discrete design variable. The contribution of the *Type of Antenna* (C) to the *down-link margin* is very limited, and hidden in the *Other* bar of Figure 3.15. For a given diameter, in the given frequency range, the aperture and horn antenna have similar performances in terms of gain, which lead to similar performances in terms of *down-link margin*. On the other hand, the influence on the *mass* of the *Type of Antenna* is more significant, with the aperture antenna being lighter than the horn one for a given reference diameter.

The sum of the total-order sensitivity indices is larger than one, in both cases. This means that interactions or higher-order effects are present. Indeed these effects were detected and quantified in the computation of what we called first-order effect (the meaning of the term *first order* was already discussed). A quantitative indication of the importance of the factors for the determination of the performances, provides the engineering team with fundamental information to understand the effects of the design choices on the final design. In this case, for instance, one may easily conclude that the *Type of Solar Array* does not affect the performances much, thus it might be frozen to a particular type, based for instance, on the availability at the moment of implementation, or its cost, or based on experience on past missions. The *Type of Antenna* can be selected on the basis of its sole contribution to the mass (the aperture antenna minimizes the mass for a given *down-link margin* performance). This reduces *de facto* the dimensions of the design space allowing the engineering team to channel the effort on the more relevant design parameters. Very often the expert engineers, or the developers of the mathematical model themselves, are already able to predict in advance the effects of the design choices on the output. However, this does not have to be the case, especially in the presence of less expert engineers or team members, which were not directly involved in the development of the mathematical model.

In the following subsection, a formal validation of the performances of the RBSA is presented. The RBSA is compared to other methods for sensitivity analysis in the computation of the sensitivity indices of six test functions of increasing complexity.

No analytic solution is available for the problems proposed hereafter. The comparison is made between the results obtained with RBSA and the results provided by the software package Simlab (Simlab, 2011). In particular the sensitivity indices using the method of Sobol' with a large number of sample points is taken into account as comparative baseline.

### 3.2.5 Comparison of the methods for global sensitivity analysis

In this section some methods for global sensitivity analysis, including RBSA, are tested against six problems of increasing complexity, derived from Helton and Davis (2002). The purpose is to evaluate the performance of RBSA in determining the sensitivity indices of the various factors, comparing it with the method of Sobol', FAST, the method of Morris, and the SRCs. The comparison is based on the number of model evaluations, indicated with $N$, needed to obtain the sensitivity indices, and their accuracy. For a given model, a smaller number of evaluations indicates that the computational time needed to obtain the sensitivity indices is lower. It is useful to remind that the evaluation of the model is considered the computationally expensive part of the analysis. The analysis of the data to perform sensitivity analysis is relatively fast in all cases presented here.

The main purpose of this comparison is to demonstrate that the regression-based sensitivity analysis approach is able to successfully provide quantitative sensitivity indices (as the Sobol' and the FAST approach) with a low number of model evaluations (as most of the screening methods). The method of Morris is executed with increasing levels ($P$) and increasing number of samples-per-level ($R$) for the same purpose. The values of the SRC are reported from the original study of Helton and Davis (2002) for comparison with the results obtained with the other methods. The RBSA method is executed with models of increasing order and with an increasing number of sample points until a satisfactory level of $R^2_{adj}$ is obtained, as discussed in Section 3.2.3. The sensitivity indices obtained with the method of Sobol', FAST and RBSA are only reported in terms of total-order sensitivity indices, $S_{Ti}$. The methods are executed on each problem with an increasing number of sample points to determine the minimum number of model evaluations that allows to stabilize the value of the sensitivity indices. By *stable* it is intended that the sensitivity indices do not change significantly for increasing sample size, *i.e.,* they are constant to the second meaningful decimal digit. To obtain the sensitivity indices with the methods of Sobol', FAST, and Morris, the *Simlab* software suite was used (Simlab, 2011). In the comparison presented in this section, we consider the converged values from the method of Sobol' and FAST to be the correct results for the sensitivity indices. With RBSA, we try to obtain the same results, in a computationally cheaper way.

The first test problem considered is linear with only three uniformly distributed variables (Problem 1, Helton and Davis (2002)):

$$f(\mathbf{x}) = \sum_{i=1}^{3} x_i, \qquad \mathbf{x} = [x_1, x_2, x_3] \tag{3.36}$$

with $x_i : U(\bar{x}_i - \sigma_i, \bar{x}_i + \sigma_i), \bar{x}_i = 3^{i-1}, \sigma_i = 0.5\bar{x}_i$ for $i = 1, 2, 3$.

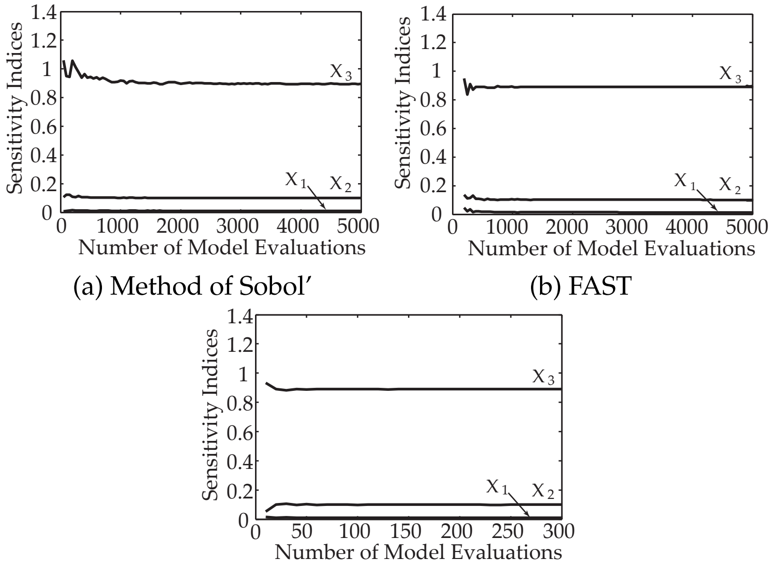
The results of the comparison are summarized in Table 3.9. Considering the low complexity of the problem, the method of Sobol' and FAST converge to a stable value of the sensitivity indices with a relatively large sample size, $1000$ model evaluations, while the RBSA provides satisfactory results already after $20$ model evaluations.

This is demonstrated with the graphs in Figure 3.16. They show the trend of the sensitivity indices computed with the method of Sobol' (a), FAST (b), and RBSA (c), as a function of the number of model evaluations. The first two methods provide a definite distinction between

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] |
|---|---|---|---|---|---|---|---|
| | $N = 1000$ | $N = 1000$ | $N = 8$ | | | $N = 100$ | $N = 20$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ |
| $x_1$ | 0.011 | 0.014 | 3 | 9 | 0 | 0.105 | 0.013 |
| $x_2$ | 0.099 | 0.101 | 2 | 3 | 0 | 0.316 | 0.097 |
| $x_3$ | 0.892 | 0.890 | 1 | 1 | 0 | 0.946 | 0.890 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Linear regression model with 2-factors interaction terms. $R^2_{adj} = 1.00$

Table 3.9 Comparison of sensitivity analysis methods. Problem 1 (see Eq. (3.36)).




(a) Method of Sobol′ (b) FAST

(c) RBSA

Figure 3.16 Total-order sensitivity indices as a function of the sample size. Problem 1 (see Eq. (3.36)).

the effects of the three factors already with few sample points but the values of the sensitivity indices are stable only after many more model evaluations. This effect will be more evident in the presence of more complex problems. Much less model evaluations are needed by the method of Morris to obtain a qualitative measure of sensitivity, *i.e.,* the ranking of the factors according to their importance in the determination of $f(\mathbf{x})$. RBSA provides very precise quantitative sensitivity indices with a reduced computational effort when compared to the method of Sobol′ and FAST. Indeed, only 20 model evaluations are required to obtain in practice the same results as the method of Sobol′ and FAST. The SRCs provide a correct ranking of the relevance of the factors, but the sensitivity indices are much different from these provided by the other methods. Indeed, $x_2$ results to be much more important than it actually is.

The second test problem is again a linear one but with a larger number, *i.e.,* 22, of uniformly distributed variables (Problem 2, Helton and Davis (2002)):

$$f(\mathbf{x}) = \sum_{i=1}^{22} c_i \left(x_i - 0.5\right), \qquad \mathbf{x} = [x_1, x_2, \cdots, x_{22}] \tag{3.37}$$

with $x_i : U(0,1)$ and $c_i = (i-11)^2$ for $i = 1, 2, \cdots, 22$.

The large number of variables of Problem 2 causes the method of Sobol′ and FAST to con-

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] |
|---|---|---|---|---|---|---|---|
| | $N = 10,000$ | $N = 24,000$ | $N = 46$ | | | $N = 100$ | $N = 600$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ |
| $x_1$ | 0.149 | 0.192 | 2 | 100 | 0 | 0.381 | 0.152 |
| $x_2$ | 0.0979 | 0.115 | 3 | 81 | 0 | 0.308 | 0.100 |
| $x_3$ | 0.0619 | 0.0660 | 4 | 64 | 0 | 0.243 | 0.0633 |
| $x_4$ | 0.0369 | 0.0372 | 5 | 49 | 0 | 0.186 | 0.0369 |
| $x_5$ | 0.0199 | 0.0212 | 6 | 36 | 0 | 0.136 | 0.0198 |
| $x_6$ | 0.0093 | 0.0130 | 7 | 25 | 0 | 0.0951 | 0.0096 |
| $x_7$ | 0.0038 | 0.0081 | 8 | 16 | 0 | 0.0608 | 0.0039 |
| $x_8$ | 0.0012 | 0.0023 | 9 | 9 | 0 | 0.0342 | 0.0012 |
| $x_9$ | 0.0002 | 0.0013 | 10 | 4 | 0 | 0.0152 | 0.0002 |
| $x_{10}$ | 0 | 0.0011 | 11 | 1 | 0 | 0.0038 | 0 |
| $x_{11}$ | 0 | 0.0011 | 12 | 0 | 0 | 0 | 0 |
| $x_{12}$ | 0 | 0.0011 | 11 | 1 | 0 | 0.0038 | 0 |
| $x_{13}$ | 0.0002 | 0.0012 | 10 | 4 | 0 | 0.0152 | 0.0002 |
| $x_{14}$ | 0.0012 | 0.0026 | 9 | 9 | 0 | 0.0342 | 0.0012 |
| $x_{15}$ | 0.0038 | 0.0059 | 8 | 16 | 0 | 0.0609 | 0.0039 |
| $x_{16}$ | 0.0093 | 0.0076 | 7 | 25 | 0 | 0.0951 | 0.0096 |
| $x_{17}$ | 0.0199 | 0.0160 | 6 | 36 | 0 | 0.136 | 0.0197 |
| $x_{18}$ | 0.0367 | 0.0390 | 5 | 49 | 0 | 0.186 | 0.0371 |
| $x_{19}$ | 0.0619 | 0.0520 | 4 | 64 | 0 | 0.243 | 0.0627 |
| $x_{20}$ | 0.0980 | 0.116 | 3 | 81 | 0 | 0.307 | 0.100 |
| $x_{21}$ | 0.149 | 0.174 | 2 | 100 | 0 | 0.380 | 0.153 |
| $x_{22}$ | 0.218 | 0.232 | 1 | 121 | 0 | 0.460 | 0.224 |

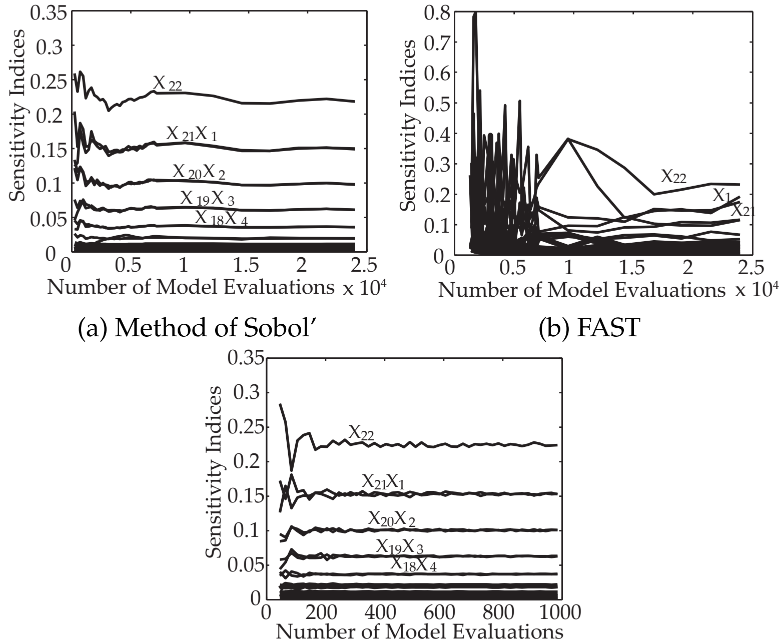
[a]Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)

[b]Linear regression model. $R^2_{adj} = 1.00$

Table 3.10 Comparison of sensitivity analysis methods. Problem 2 (see Eq. (3.37)).

verge to a stable value for the sensitivity indices only after $10,000$ and $24,000$ model evaluations, respectively. However, a clear distinction between the factors is already in place after $5,000$ samples in the case of the method of Sobol′, see Figure 3.17(a). In Figure 3.17(b) instead, the apparently chaotic behavior of FAST indicates that this method particularly suffers with low sample size when the number of variables is large, even with a linear model like Problem 2. The method of Morris provides excellent results in ranking the factors with a very low number of simulations. This is due to the fact that Problem 2 is linear, and a precise estimation of the variability of the data using the elementary effect is already possible with two sample points per variable. The RBSA method requires more model evaluations than the method of Morris (regression model with linear terms) and still much less with respect to the method of Sobol′ and FAST. The RBSA provides very precise (compared to the method of Sobol′, for instance) quantitative sensitivity indices, see Table 3.2.5, already after 600 model evaluations.

In the case of linear functions the method of Morris provides results that remain constant if the number of levels $P$ or the number of samples-per-level $R$ increases. This is the natural consequence of the fact that the elementary effects, *i.e.*, the incremental ratios, of a linear function are constant in the entire interval of variation of the variables. This is not valid in general, for instance when curvature is present, as will be demonstrated later in this section. In both Problem 1 and 2 the method of Morris is able to correctly rank the factors and to correctly indicate the absence of interactions or non-linear effects (since the value of $\sigma$ is zero for all factors). Also the SRCs provide a correct indication on the relative importance of the factors but they do not provide any information on the presence (or not) of higher-order effects. The method of Sobol′ and FAST provide quantitative sensitivity indices at the expenses of a large computational effort. The RBSA method provides very precise quantitative sensitivity indices,




(a) Method of Sobol′ (b) FAST

(c) RBSA

Figure 3.17 Total-order sensitivity indices as a function of the sample size. Problem 2 (see Eq. (3.37)).

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] | RBSA[c] |
|---|---|---|---|---|---|---|---|---|
| | $N = 1,500$ | $N = 2,000$ | $N = 330$ | | | $N = 100$ | $N = 200$ | $N = 250$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ | $S_{Ti}$ |
| $x_1$ | 0.543 | 0.539 | 1 | 1 | 0 | 0.740 | 0.537 | 0.539 |
| $x_2$ | 0.461 | 0.475 | 2 | 0.79 | 0.055 | 0.587 | 0.452 | 0.460 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Quadratic regression model. $R^2_{adj} = 0.989$
[c] Fourth-order regression model. $R^2_{adj} = 1.00$

Table 3.11 Comparison of sensitivity analysis methods. Problem 3a (see Eq. (3.38)).

even in problems with a large number of variables, as Problem 2, at a computational cost that is much lower when compared to the method of Sobol′ and FAST.

The third test problem is monotonic, non-linear, with 2 uniformly distributed variables (Problem 3, Helton and Davis (2002)):

$$f(\mathbf{x}) = x_1 + x_2^4, \quad \mathbf{x} = [x_1, x_2] \tag{3.38}$$

with $x_i : U(0,1)$ for $i = 1,2$ (Problem 3a), $x_i : U(0,3)$ for $i = 1,2$ (Problem 3b), or $x_i : U(0,5)$ for $i = 1,2$ (Problem 3c).

This test problem is executed with increasing dimensions of the ranges of variability, giving rise to three sub-problems to analyze. The results are presented in Tables 3.11 to 3.13. The RBSA method is implemented with the regression models from linear- to fourth-order. As expected, the second-order model presents some lack-of-fit that increases as the variability range gets larger. This indicates that higher-order effects of the variables are present in the model. However the sensitivity indices are very close to those computed with the method of Sobol′ and FAST, and to those computed with the fourth-order model. The fourth-order regression model, with the same number of model evaluations of the second-order one, allows

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] | RBSA[c] |
|---|---|---|---|---|---|---|---|---|
| | $N = 1,000$ | $N = 1,000$ | $N = 330$ | | | $N = 100$ | $N = 200$ | $N = 250$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ | $S_{Ti}$ |
| $x_1$ | 0.001 | 0.0021 | 2 | 3 | 0 | N/A | 0.0018 | 0.0015 |
| $x_2$ | 0.999 | 0.996 | 1 | 64 | 4.613 | N/A | 0.976 | 0.9985 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Quadratic regression model. $R^2_{adj} = 0.978$
[c] Fourth-order regression model. $R^2_{adj} = 1.00$

Table 3.12 Comparison of sensitivity analysis methods. Problem 3b (see Eq. (3.38)).

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] | RBSA[c] |
|---|---|---|---|---|---|---|---|---|
| | $N = 1,000$ | $N = 1,000$ | $N = 330$ | | | $N = 100$ | $N = 200$ | $N = 250$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ | $S_{Ti}$ |
| $x_1$ | 0 | 0.0013 | 2 | 5 | 0 | N/A | 0.0004 | 0.0001 |
| $x_2$ | 1 | 0.999 | 1 | 501.9 | 34.83 | N/A | 0.9756 | 0.999 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Quadratic regression model. $R^2_{adj} = 0.976$
[c] Fourth-order regression model. $R^2_{adj} = 1.00$

Table 3.13 Comparison of sensitivity analysis methods. Problem 3c (see Eq. (3.38)).

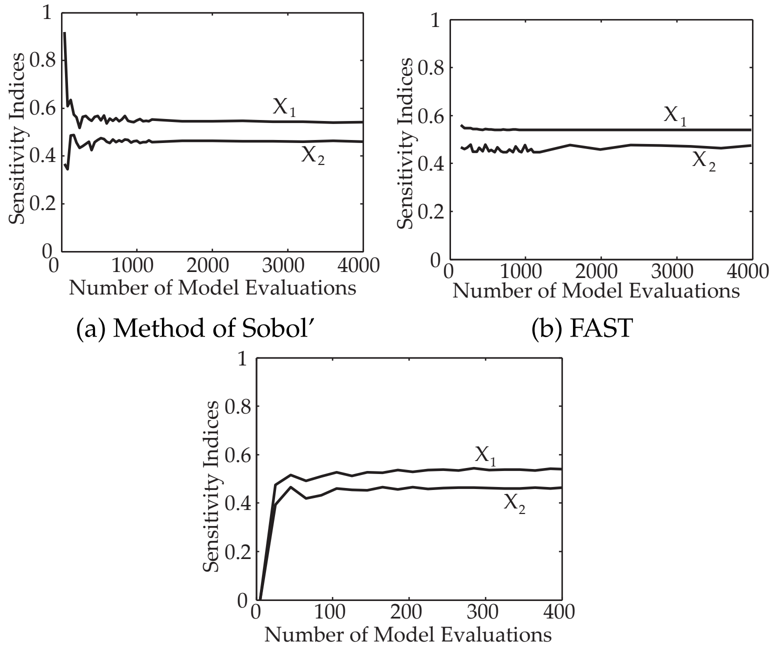
to account for all the variability of the samples providing more precise values of the sensitivity indices. In case of Problem 3a the SRCs provide only an insight in the factors importance, and in case of Problem 3b and Problem 3c the values were not available in Helton and Davis (2002).

The method of Sobol′ and FAST provide a clear distinction between the factors based on their importance, already with a low number of function evaluations. However, stable values of the sensitivity indices are only obtained after $1,500$ and $2,000$ model evaluation respectively for Problem 3a, and with $1,000$ samples for Problem 3b and 3c. The RBSA provides accurate sensitivity indices already with a sample size of $250$, as shown by the trends in Figure 3.18.

In the case of Problem 3, and in all other non-linear problems, the method of Morris provides results that are not constant as the number of levels (P) and the number of sample-points per level (R) changes. This is due to the fact that the non-linear behavior is better approximated with a large number of intervals (levels) in which linear elementary effects are computed and averaged out. However, due to the nature of this method, increasing the number of levels (P) alone is not sufficient for improving the results of the sensitivity analysis. The number of model evaluations should also be increased to have higher chances of obtaining at least a few points for every level.

The three graphs in Figure 3.19 present the trends of the values of $\mu^*$ and $\sigma$ for increasing $P$ (going from Figure 3.19(a) to Figure 3.19(c)) and for increasing $R$. In each graph of Figure 3.19 the $\Diamond$ symbols indicate the effects of factor $x_1$, while the $\circ$ symbols indicate the effects factor $x_2$. The numbers from $1$ to $15$ indicate an increasing value of the parameter $R$. For ease of visualization, only the results obtained with six different values of $R$ are reported ($1 \Rightarrow R = 10$, $2 \Rightarrow R = 110$, $3 \Rightarrow R = 210$, $4 \Rightarrow R = 310$, $5 \Rightarrow R = 410$, $15 \Rightarrow R = 10,000$). Already with a value of $R = 110$, $\mu^*$ and $\sigma$ do not change significantly for increasing $R$. In particular, $R$ affects $\sigma$ the most. Remember that the method of Morris is qualitative, usually used as a screening method, and with $R = 110$ the relative importance of the parameters is already well defined (at least with this test problem having only 2 variables). The number of levels (P) mostly affects the value of $\mu^*$ that shows a decreasing trend for increasing $P$.

In Figure 3.20 the results obtained using the method of Morris on Problem 3b and 3c are presented. The approach is the same followed for Problem 3a, but in this case only the results for $P = 6$ are reported. In all cases regarding Problem 3, the method of Morris is able to




(a) Method of Sobol′ (b) FAST

(c) RBSA

Figure 3.18 Total-order sensitivity indices as a function of the sample size. Problem 3(a) (see Eq. (3.38)).

correctly rank the factors and to correctly indicate a higher-order effect for the factor $x_2$. As expected, in all cases the values of $\mu^*$ and $\sigma$ of the factor $x_1$, $\Diamond$ symbol, do not change for increasing $R$, since $x_1$ is only responsible for a linear effect.

The fourth problem is monotonic, non-linear, with 6 uniformly distributed variables (Problem 4, Helton and Davis (2002)):

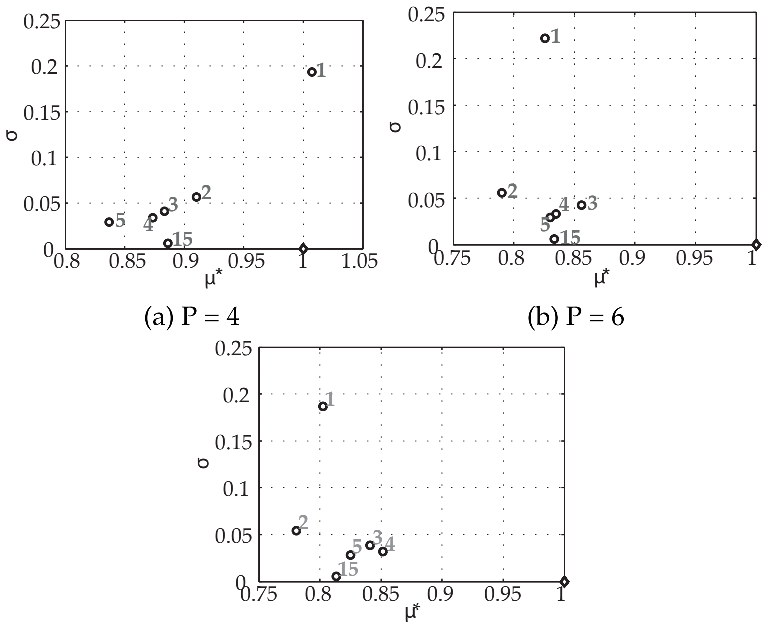
$$f(\mathbf{x}) = \exp\left(\sum_{i=1}^{6} b_i x_i\right) - \prod_{i=1}^{6} \frac{(e^{b_i} - 1)}{b_i}, \qquad \mathbf{x} = [x_1, x_2, \ldots, x_6], \tag{3.39}$$

with $x_i : U(0,1)$ for $i = 1, 2, \ldots, 6$ and $b_1 = 1.5, b_2 = b_3 = \cdots = b_6 = 0.9$.

The results of the comparison are shown in Table 3.14. The method of Sobol′ and FAST converge to stable values for the sensitivity indices after $7,500$ model evaluations, as shown in Figure 3.21. The RBSA is able to account for almost all the variability of Problem 4 with a cubic regression model ($R^2_{adj} = 0.99$), and already with $500$ sample points the estimation of the sensitivity indices is very precise, see Table 3.14.

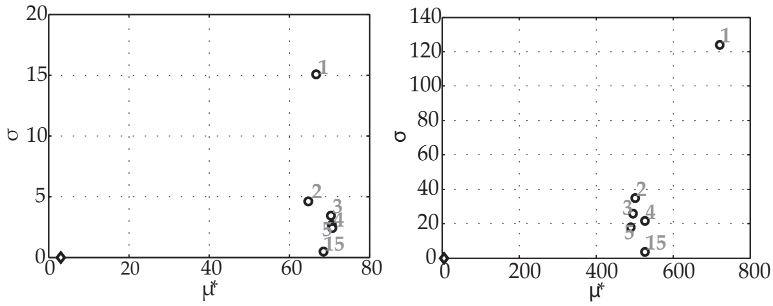
As anticipated in the previous discussion, the method of Morris does not provide a stable ranking of the factors as the number of levels (P) and the number of sample points per level (R) changes. In this case, however, the factor $x_1$ is identified as the most relevant one already with $R = 10$, thus with a sample size of 70, as shown in Figure 3.22(a), where the numbering has the same interpretation as discussed before. The relative ranking of the factors $x_2$ to $x_6$ keeps changing with increasing $R$. For this reason it was decided to report only the results for $R = 410$ in Table 3.14. As shown in Figure 3.22(b) the relative ranking of the factors $x_2$ to $x_6$ still changes from $R = 410$ (the black symbols) to $R = 10,000$ (the group of symbols with the value of $\sigma$ lower than $0.5$), but the absolute value of $\sigma$ and $\mu^*$ does not change significantly.

In the case of non-linear monotonic problems the method of Sobol′ and FAST provide as accurate results as in the linear case, even if they show poor performance when the sample size is relatively low. The method of Morris has shown one potential weakness that arises




(a) P = 4
(b) P = 6
(c) P = 10

Figure 3.19 Effect of the number of levels on the sensitivity analysis computed with the method of Morris, Problem 3a. ◊ factor $x_1$, ∘ factor $x_2$. $1 \Rightarrow R = 10$, $2 \Rightarrow R = 110$, $3 \Rightarrow R = 210$, $4 \Rightarrow R = 310$, $5 \Rightarrow R = 410$, $15 \Rightarrow R = 10000$.




(a) Problem 3b, P = 6
(b) Problem 3c, P = 6

Figure 3.20 Sensitivity Analysis using the method of Morris. ◊ factor $x_1$, ∘ factor $x_2$. $1 \Rightarrow R = 10$, $2 \Rightarrow R = 110$, $3 \Rightarrow R = 210$, $4 \Rightarrow R = 310$, $5 \Rightarrow R = 410$, $15 \Rightarrow R = 10,000$.

when non-linear problems are taken into account: the results are very sensitive to the number of levels $P$ and the number of sample points per level $R$. The SRCs perform well, even with non-linear monotonic problems. Linear approximations of the non-linear monotonic models provide a good indication of the general trends of the output, but this cannot be considered true in general. The RBSA demonstrates excellent performance also with this class of problems. Indeed, in both cases of Problem 3 and Problem 4 it provided very precise quantitative sensitivity indices at a relatively low computational cost.

The fifth problem is non-monotonic with 8 uniformly distributed variables (Problem 5, Helton and Davis (2002)):

$$f(\mathbf{x}) = \prod_{i=1}^{8} \frac{|4x_i - 2| + a_i}{1 + a_i} \qquad \mathbf{x} = [x_1, x_2, \ldots, x_8]\,, \tag{3.40}$$

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] |
|---|---|---|---|---|---|---|---|
| | $N = 7500$ | $N = 7500$ | $N = 2870$ | | | $N = 100$ | $N = 500$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ |
| $x_1$ | 0.399 | 0.391 | 1 | 36.97 | 1.22 | 0.522 | 0.392 |
| $x_2$ | 0.166 | 0.161 | 6 | 22.37 | 0.82 | 0.295 | 0.153 |
| $x_3$ | 0.153 | 0.165 | 4 | 22.61 | 0.84 | 0.297 | 0.157 |
| $x_4$ | 0.164 | 0.156 | 3 | 22.70 | 0.90 | 0.344 | 0.159 |
| $x_5$ | 0.158 | 0.173 | 5 | 22.46 | 0.81 | 0.351 | 0.157 |
| $x_6$ | 0.155 | 0.170 | 2 | 23.41 | 0.92 | 0.284 | 0.156 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Cubic regression model. $R^2_{adj} = 0.99$

Table 3.14 Comparison of sensitivity analysis methods. Problem 4 (see Eq. (3.39)).

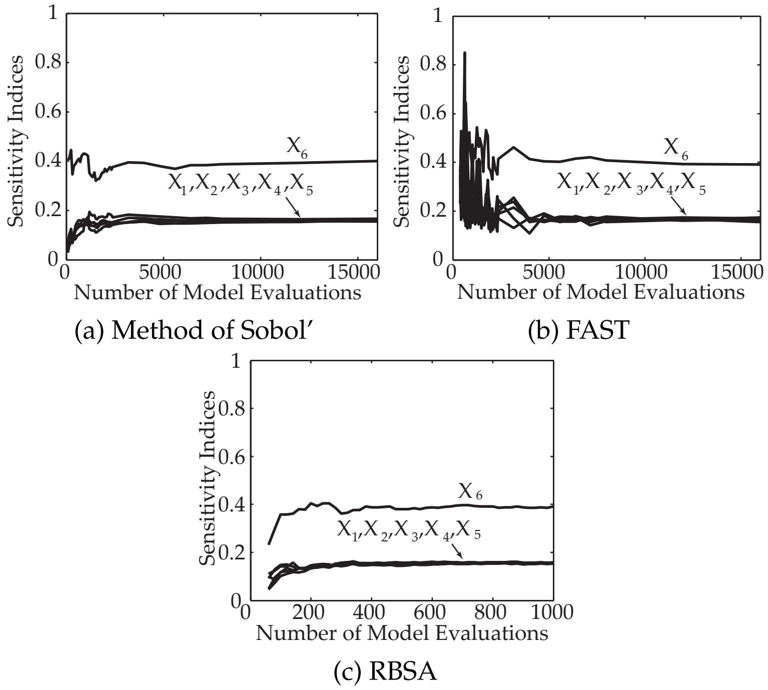
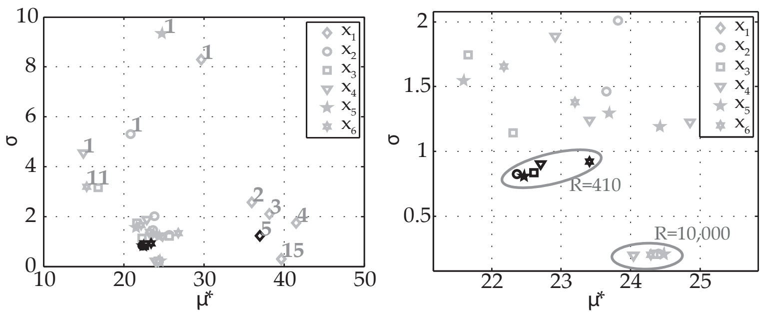



Figure 3.21 Total-order sensitivity indices as a function of the sample size. Problem 4 (see Eq. (3.39)).

with $x_i : U(0,1)$ for $i = 1, 2, \ldots, 8$ and $a_1 = 0, a_2 = 1, a_3 = 4.5, a_4 = 9, a_5 = a_6 = a_7 = a_8 = 99$.

The results of the comparison are presented in Table 3.15. The first aspect worth mentioning is that the SRCs are not able to distinguish any of the variables effects. This is probably an expected result since the model of Problem 5 presents an absolute value, which causes the linear model to be deceived. The method of Sobol′ and FAST provide stable results after $3,000$ and $5,000$ model evaluations, respectively. As already anticipated in the brief description of the methods, and as demonstrated in this test case, they do not suffer the highly non-linear behavior of the problem under investigation in the design region of interest. The polynomial regression models of the RBSA cannot perfectly cope with a functional like the absolute value, by definition. However, with a fifth-order model and $1,000$ sample points, see Figure 3.23, the RBSA can already account for almost $94\%$ of the variability of the data, providing a good quantitative distinction between the effects of the factors, and quantitative sensitivity indices that are close to the actual ones.




(a) Method of Morris (b) Zoom

Figure 3.22 Sensitivity Analysis using the method of Morris, with $P = 6$. Problem 4 (see Eq. (3.39)). $1 \Rightarrow R = 10$, $2 \Rightarrow R = 110$, $3 \Rightarrow R = 210$, $4 \Rightarrow R = 310$, $5 \Rightarrow R = 410$, $15 \Rightarrow R = 10,000$

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] |
|---|---|---|---|---|---|---|---|
| | $N = 3,000$ | $N = 5,000$ | $N = 90,000$ | | | $N = 100$ | $N = 1,000$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ |
| $x_1$ | 0.792 | 0.794 | 1 | 0.0060 | 0.0260 | $\sim 0$ | 0.704 |
| $x_2$ | 0.244 | 0.239 | 2 | 0.0048 | 0.0151 | $\sim 0$ | 0.175 |
| $x_3$ | 0.0338 | 0.0355 | 4 | 0.0014 | 0.0059 | $\sim 0$ | 0.0214 |
| $x_4$ | 0.0104 | 0.0114 | 3 | 0.0028 | 0.0033 | $\sim 0$ | 0.0120 |
| $x_5$ | 0.0001 | 0.0006 | 7 | 0.0003 | 0.0003 | $\sim 0$ | 0.0075 |
| $x_6$ | 0.0001 | 0.0006 | 6 | 0.0004 | 0.0003 | $\sim 0$ | 0.0110 |
| $x_7$ | 0.0001 | 0.0006 | 5 | 0.0007 | 0.0003 | $\sim 0$ | 0.0118 |
| $x_8$ | 0.0001 | 0.0006 | 8 | 0.0000 | 0.0003 | $\sim 0$ | 0.0075 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Fifth-order regression model. $R^2_{adj} = 0.938$

Table 3.15 Comparison of sensitivity analysis methods. Problem 5 (see Eq. (3.40)).

As reported in Table 3.15, the method of Morris presents the same type of problem encountered with the SRCs. However, a certain qualitative distinction between the factors' importance may still be identified. This is mainly due to the asymmetry of the absolute value of Eq. (3.40) in the variability interval determined by the variable ranges. The results are obtained with $R = 10,000$, thus a sample size of $90,000$. In Figure 3.24 the trends for increasing sample size are shown. It is very evident that for increasing value of $R$, the measures identified by the method of Morris converge towards zero.

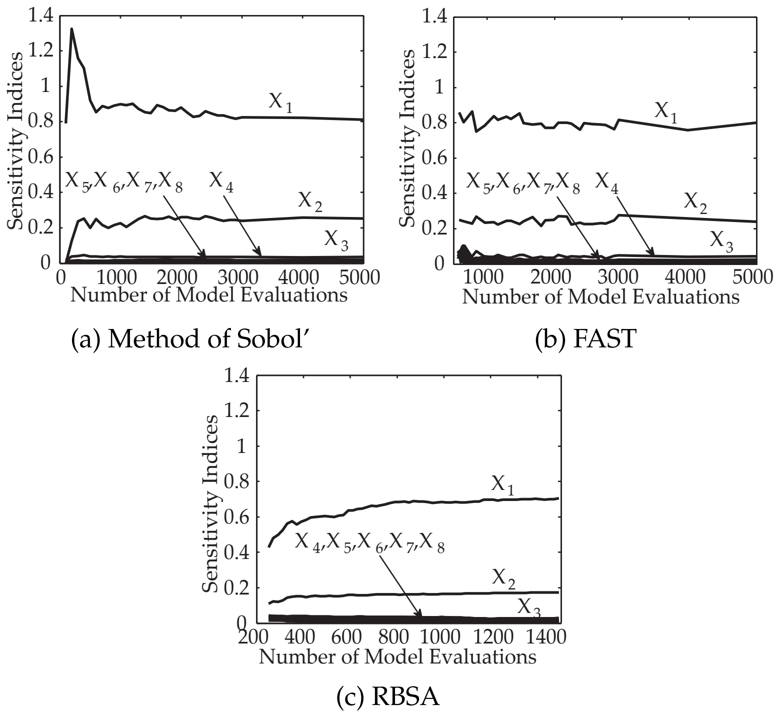



Figure 3.23 Total-order sensitivity indices as a function of the sample size. Problem 5 (see Eq. (3.40)).

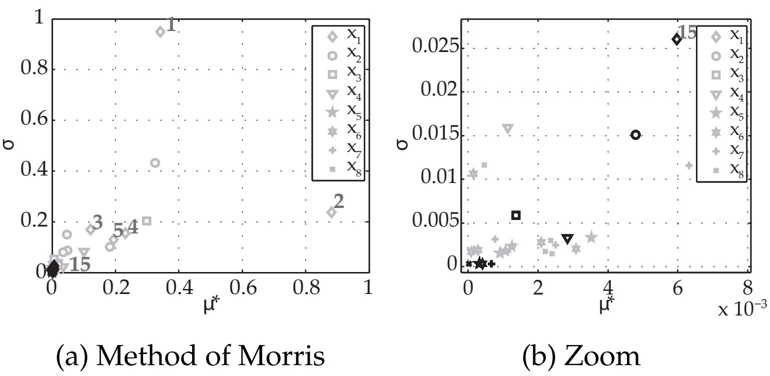



Figure 3.24 Sensitivity Analysis using the method of Morris. Problem 5 (see Eq. (3.40)).
$1 \Rightarrow R = 10$, $2 \Rightarrow R = 110$, $3 \Rightarrow R = 210$, $4 \Rightarrow R = 310$, $5 \Rightarrow R = 410$, $15 \Rightarrow R = 10,000$

The last problem is non-monotonic with 3 uniformly distributed variables (Problem 6, Helton and Davis (2002)):

$$f(\mathbf{x}) = \sin x_1 + A \sin^2 x_2 + B x_3^4 \sin x_1 \qquad \mathbf{x} = [x_1, x_2, x_3]\,, \tag{3.41}$$

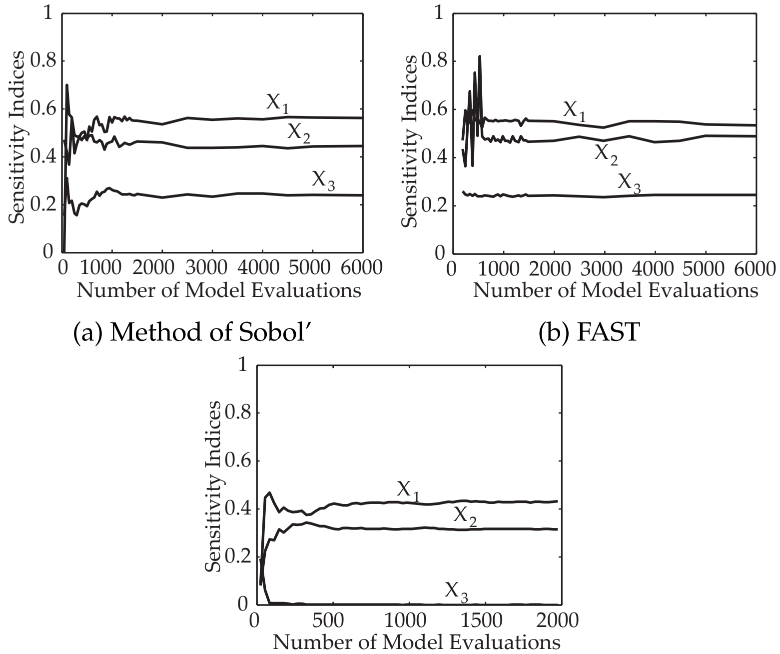
with $x_i : U(-\pi, \pi)$ for $i = 1, 2, 3$ and $A = 7, B = 0.1$.

Also in this case, the SRCs and the method of Morris are not able to detect the correct contribution of the factors to the variability of the performance, see Table 3.16. The method of Sobol′ and FAST confirm the fact that the results they provide are not sensitive to the nature of the underlying model. Indeed in Table 3.16 and in Figure 3.25 it is shown that they provide a stable estimate of the sensitivity indices for $3,000$ and $5,000$ model evaluations respectively. The RBSA, using a seventh-order model provided a coefficient of determination of $0.75$. In

| | Sobol′ | FAST | Morris method | | | SRC[a] | RBSA[b] |
|---|---|---|---|---|---|---|---|
| | $N = 3,000$ | $N = 5,000$ | $N = 40,000$ | | | $N = 100$ | $N = 1,000$ |
| Variable name | $S_{Ti}$ | $S_{Ti}$ | Rank | $\mu^*$ | $\sigma$ | Value | $S_{Ti}$ |
| $x_1$ | 0.556 | 0.536 | 1 | 7.99 | 0.0988 | $\sim 0$ | 0.417 |
| $x_2$ | 0.445 | 0.487 | 3 | 0.0055 | 0.0284 | $\sim 0$ | 0.330 |
| $x_3$ | 0.237 | 0.242 | 2 | 0.1157 | 0.0781 | $\sim 0$ | 0.0054 |

[a] Standardized Regression Coefficients. Data adapted from Helton and Davis (2002)
[b] Seventh-order regression model. $R^2_{adj} = 0.75$

Table 3.16 Comparison of sensitivity analysis methods. Problem 6 (see Eq. (3.41)).




(c) RBSA

Figure 3.25 Total-order sensitivity indices as a function of the sample size. Problem 6 (see Eq. (3.41)).

this case this result is not as good as the previous examples, leading to a misleading value for the absolute sensitivity index of the variable $x_3$. However, at least a correct ranking of the importance of the factors can be identified.

Concluding, the Regression-Based Sensitivity Analysis method has shown good performance with different types of models. The sensitivity indices for linear and non-linear monotonic models can be precisely computed with a very reduced number of model evaluations, when compared to other methods. In the case of non-monotonic problems, the polynomial representation shows its limitations. RBSA provides less accurate quantitative results in these cases but still it provides insight in the ranking the factors according to their importance, also when other qualitative methods fail. A polynomial function does not cope well with terms like $\sin x$, $\cos x$, $e^x$, and $\frac{1}{x}$, for instance. Therefore it is hard to obtain a value for the coefficient of determination that is close *enough* to one. These non-polynomial terms could be included in the representation of the model of Eq. (3.15), but then the sensitivity indices would indicate the effect of the terms $\sin x$, $\cos x$, $e^x$, and $\frac{1}{x}$ rather than the effect of the factor $x$, which is what

| Design Variables | | Code | Intervals Min | Max | Levels |
|---|---|---|---|---|---|
| Number of days (rep. ground track) | [-] | A | 1 | 3 | 3 |
| Number of orbits (rep. ground track)[a] | [-] | B | 1 | 3 | 3 |
| Instrument aperture diameter | [m] | C | 0.3 | 1 | – |
| Min. $\epsilon$ | [deg] | D | 5 | 50 | – |
| Max. slewing angle | [deg] | E | 0 | 50 | – |
| Min. maneuver time | [s] | F | 60 | 180 | – |
| Number of slew maneuvers | [-] | G | $10k$ | $30k$ | – |
| Transmitting output RF power | [W] | H | 5 | 30 | – |
| Antenna diameter | [m] | I | 0.1 | 1 | – |
| Type of solar array | [-] | J | 1 | 2 | 2 |
| Type of thrusters | [-] | K | 1 | 2 | 2 |
| Payload heritage | [-] | L | 1 | 2 | 2 |

Table 3.17 Settings of the design variables.[a] When $A = 1$, $B = 13, 14$ or $15$. When $A = 2$, $B = 28, 29$ or $30$. When $A = 3$, $B = 43, 44$ or $45$.

we are interested in. In the next section, we provide an example of the utilization of the RBSA method on the complete mathematical model of the *satellite system for Earth observation*.

### 3.2.6 Test case: satellite system for Earth-observation, sensitivity analysis

Earth-observation satellites can observe areas over a wide range quickly. It is expected that their observation data combined with information obtained by aircraft and helicopters will be useful for a regular disaster condition assessment. This would make rescue operations more effective, it would allow for extracting topographical information reflecting latest land-usage changes, and identifying disaster risks. In this section we use the mathematical model of a satellite to show the performance of RBSA with a relatively complex model, and to show the benefits of having important factors identified already in the preliminary phases of the design process. The mathematical model used for the analysis is discussed in Appendix B. The settings and the results are discussed by the author also in **?**.

The main purpose is to achieve a compromise between the design variables in such a way to obtain the best possible image resolution, at minimum cost. These are the two objectives: the satellite shall revisit the same area on the Earth surface within 24 hours, and shall be able to send the acquired data back to any equipped ground station (the reference ground station is considered with $1$ m aperture antenna diameter) with a link margin of at least $4$ dB. The selected launcher is of the class of the *Delta II 6920/25*, with a maximum payload for polar orbit of $2950$ kg. These are the constraints of the analysis: a highly inclined, circular orbit has been selected, with $i = 98°$.

In Table 3.17 the design variables taken into account in the analysis, their type and intervals or levels (in case of discrete variables) are summarized.

In Figure 3.26 the first-order sensitivity indices are visualized for the constraints (top three graphs) and for the objectives (lower two graphs). The results are obtained using a second-order model, see Eq. (3.15), and re-sampled for additional cubic terms of the factors. Two full-factorial designs (3-level and 2-level) have been used for the discrete factors *A* and *B*, and *J*, *K*, and *L*, respectively (Table 3.17). Regarding the continuous variables, instead, the Sobol' sequence required $60$ samples. Therefore the total number of model evaluations is 4320. The results shall be discussed in a similar manner as these in Figures 3.14 and 3.15.

The first conclusion is that the factors *E*, *F*, *G*, *J*, and *K* have a limited effect on the objectives and constraints, probably less than one would expect since some of them are related to the

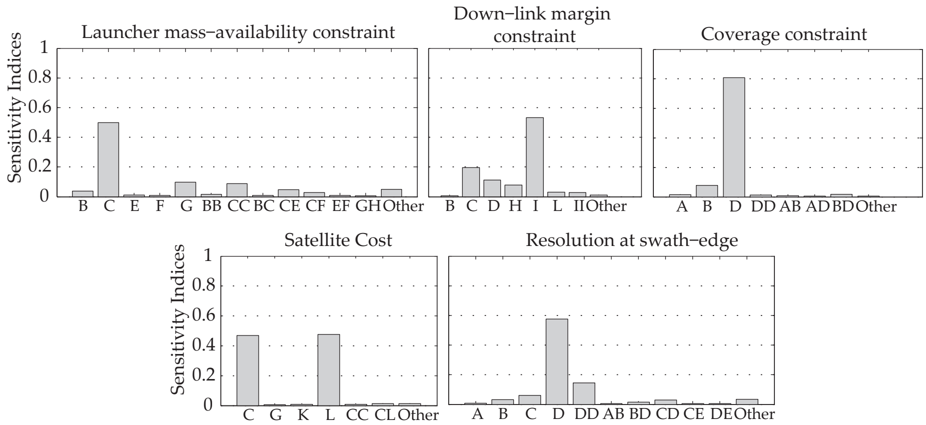



Figure 3.26 Bar plots indicating the first-order sensitivity indices computed with the RBSA method.

propellant utilization on-board, which is usually a mass driver, thus with an effect on the cost. They can eventually be fixed to a certain level/value and this will have a limited impact on the mission, regarding the objectives and constraints that we are analyzing. The other design variables, instead, present contrasting behaviors. The instrument aperture diameter (factor *C*), for instance, affects the mass of the satellite and the satellite cost (the larger the diameter the larger the mass and the cost, reasonably) but also the down-link margin. The *minimum elevation angle* for the observation (factor *D*) has an effect on *coverage* (the smaller *D* is, the better) and on the *resolution at the edge of the swath* (the larger *D* is, the better). However, factor *D* also has some influence on the down-link margin constraint (for this analysis we imposed that communication takes place with a ground station at the edge of the instrument swath width). The effect of factors *C* and *D* on the *down-link margin* constraint, rather than the more obvious impact of the antenna diameter (factor *I*) and the transmitter RF power output (factor *H*), can be explained as follows. After these results were obtained, a close investigation on the model lead us to the relationship between the instrument aperture diameter and the *angular resolution*, that is related to the *pixel angular resolution*, thus to the *number of pixels* and finally to the *real-time data rate*, which causes the influence on the link margin. The elevation angle, instead, is related to the atmospheric attenuation that increases as the path to the receiver increases (so as the minimum elevation angle decreases). The presence of non-linear terms, such as *BB*, *CC*, and *DD* means that the behavior of the *launcher-mass availability* constraint is not linear with the *instrument aperture diameter*. The mass of the satellite does not scale linearly with the aperture diameter of the instrument, therefore it is somehow expected that also a constraint related to the mass would not scale linearly. The non-linearities will be clear in a follow-up study in Section 3.3.3.

## 3.3 Graphical support to the engineering team

Sensitivity analysis marks out the road map for the engineering team to efficiently direct the design effort. The non-influential design factors can be fixed to a pre-determined level, because they will not affect the performance much, *de facto* reducing the dimensions of the design search-space. The influential design variables and the behavior of the system under the effects caused by their variation and their interactions shall be investigated in more detail. In this subsection we present some visualization techniques that allow to extract additional information from the simulations performed to do sensitivity analysis with RBSA. Indeed, the results from those simulations are used again to compute and present the response surfaces and the variable-trends linking the most influential design factors to the performance. For dis-

| | | Levels of Factor B | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | $Y_{i.}$ |
| Levels of Factor A | 1 | 0 | 1 | 2 | 3 |
| | 2 | 2 | 6 | 10 | 18 |
| | 3 | 4 | 11 | 18 | 33 |
| | $Y_{.j}$ | 6 | 18 | 30 | 54 |

Table 3.18 Matrix design, 2 factors at 3 levels. Performance of the model.

crete variables, instead, we introduce the concept of linear and interaction graphs and show their utilization in combination with contour plots and variable trends.

### 3.3.1 Response surfaces for continuous variables

The subject of Response Surface Methods (RSMs) includes the procedures of sampling the design space, doing a regression analysis, testing for model adequacy, optimizing the response, and then validating (Kuri and Cornell, 1996). At this stage of the design process the first three steps of the RSM are already in place, as previously discussed. The iterative approach of RBSA, besides giving quantitative information on the sensitivity indices, also provides the regression coefficients, computed with Eq. (3.18), related to the best-found sample-fitting regression model. Thus, at this stage of the methodology, a surrogate model that links the design variables to the performance is already available, see Eq. (3.19). Therefore, it is possible to visualize the trends of the objectives and the constraints as a function of the continuous design variables for each combination of discrete-variable levels. Response surfaces, and their bi-dimensional representation called contour plots, can effectively represent the shape of the subspace formed by two continuous variables. When only one continuous variable is of interest, single-variable trends are a valid alternative to contour plots.

Contour plots and single-variable trends could in principle also be computed for discrete variables, since the regression coefficients are available from the RBSA. However, the regression of a continuous function for intermediate discrete-variable levels would not be realistic. To visualize the average effect of the discrete variables on the objectives and the constraints, linear and interaction graphs can be computed instead with the method shown in the following subsection.

### 3.3.2 Linear and interaction graphs for discrete variables

The main purpose of the linear and interaction graphs is to show the engineering team the behavior of the system when the levels of the discrete variables are changing. Studying the effect of a factor on the response, and showing it to the engineering team, means studying the variation in the response caused by a change in the level of the factor itself. Let us consider the example presented in the previous section, when discussing ANOVA.

Suppose now that these results are coming from a simulation with discrete variables only, or from a simulation with a mixed hypercube approach as shown in Figure 3.6. In the last case, the performance in Table 3.18 can be considered average performance values, each computed from the sample points coming from the Sobol' sampling method for each combination of discrete-variable levels.

The effect of factor A when it is at *low* level is computed as the average of the performance when factor B varies over its full range. In this case it can be computed as follows:

$$A_1 = \frac{1}{3}(0 + 1 + 2) = 1 \tag{3.42}$$

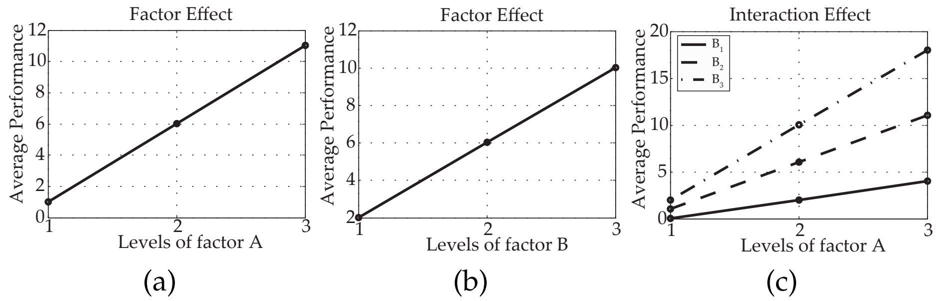



Figure 3.27 Linear graphs for visualizing the effect of the discrete variables on the performance. (a) Linear graph factor A. (b) Linear graph factor B. (c) Interaction graph representing the interaction between factor A and factor B.

In the same way we can compute the effect of factor A at level 2 and level 3:

$$A_2 = \frac{1}{3}(2 + 6 + 10) = 6 \tag{3.43}$$

$$A_3 = \frac{1}{3}(4 + 11 + 18) = 11 \tag{3.44}$$

These results can be presented in the form of a graph as shown in Figure 3.27(a). This will give the engineering team a clear picture of the system behavior under the effect of factor A, on average. Similarly, we can compute the effect of Factor B. The results are presented in Figure 3.27(b). The variation of Factor A has a (slightly) larger influence on the performance if compared to the variation of Factor B. The line is *steeper* in the case of Factor A. The sensitivity analysis computed in the previous section, with the same example proposed here, gave us already the same information. Further, we discovered that there is interaction between these two factors. The interaction between Factor A and Factor B can be presented to the engineering team as shown in Figure 3.27(c). In this case, the circles in Figure 3.27(c) represent the performance as presented in Table 3.18. The fact that the lines are diverging is a clear indication that the two variables are interacting. This means that the variation of Factor B (*e.g.,* from *level 1* to *level 3*) is enhanced when Factor A is at 3.

A general approach to compute linear and interaction graphs is presented hereafter. Consider the analysis of a system with $M$ discrete factors $[A, B, \cdots, M]$, each with a different number of levels $[a, b, \cdots, m]$, and $L$ continuous ones. Thus, there are $M + L = K$ design variables that form a $k$-dimensional design space. Referring to Figure 3.6, the matrix-design for the discrete variables would be an $a \times b \times \cdots \times m$ hypercube (considering a full-factorial). Concerning the continuous variables, let us assume that $l$ sample points are required for each combination of discrete design-variable levels. Once the design space has been sampled and the simulations executed, the responses of the system's model can be analyzed.

Let $Y_{...}$ represent the sum of all the responses obtained during the simulations, $Y_{...} = \sum y = \sum_{i=1}^{a} \sum_{j=1}^{b} \cdots \sum_{w=1}^{m} \sum_{s=1}^{l} y_{ij...ws}$. Let $Y_{i...}$ represent, the sum of all the responses with the factor $A$ at level $i$, $Y_{i...} = \sum_{j=1}^{b} \cdots \sum_{w=1}^{m} \sum_{s=1}^{l} y_{ij...ws}$.

Considering the values of $Y_{i...}$ normalized with the number of experiments, $n = b \times \cdots \times m \times l$, for which the variable $A$ is at level $i$, we compute the average value of the performance for $A$ at level $i$:

$$C_{A_i} = \frac{Y_{i...}}{n} \tag{3.45}$$

The values of $C_{A_i}$ plotted for all the $i$ levels of Factor A against the objective values provide the linear graphs.