10th World Congress on Structural and Multidisciplinary Optimization
Orlando Florida, 19 May - 24 May 2013, USA

Comparison of different global sensitivity analysis methods for aerospace vehicle optimal design

Loïc Brevault[1,2], Mathieu Balesdent[2], Nicolas Bérend[2], Rodolphe Le Riche[3]

[1] CNES, Launcher Directorate, Paris, France
[2] Onera, The French Aerospace Lab, Palaiseau, France
[3] CNRS LIMOS and Ecole des Mines de St Etienne, St-Etienne, France

1. Abstract

Space vehicle design is a process that induces numerous complex analyses and optimizations. Moreover, the design of aerospace vehicles involves a significant number of input variables. The combination of a large search space with multidisciplinary analyses leads to a computational burden that hinders design methods. A prescreening of the most important variables is required before simulation based optimization can be applied. Sensitivity analysis (SA) allows to study the impact of the variability in the inputs of the model on the outputs. A review of various global sensitivity methods (Sobol, ANalysis Of VAriance, Morris, Standard Regression coefficients) is given in this paper. It is substantiated by a comparison benchmark between the main SA methods on typical disciplines involved in launch vehicle design. The analysis outlines the numerical applicability of the methods and their advantages and drawbacks depending on the characteristics of the space vehicle design subproblem.

2. Keywords: Sensitivity analysis, Aerospace vehicle design, Sobol, ANalysis Of VAriance, Morris.

3. Introduction

For the design of complex systems such as aerospace vehicles, computer based models are essential to approximate real physical processes. Aerospace vehicle design and optimization involve intricate processes such as multidisciplinary and reliability analyses and multi-objective non-linear optimizations. However, due to the complexity of such systems, specific methods relying on Multidisciplinary Design Optimization (MDO) and Reliability Based Design Optimization (RBDO) strategies have been developed to facilitate the handling of the simulation and design problem [1,2]. The design of aerospace vehicles involves a significant number of input variables that result in a computational burden for the specific employed strategies (MDO, RBDO) [3]. To reduce this burden, a prescreening of the most important variables of the problem is required before the specific strategies can be applied.

The prescreening of the most important variables relies on Sensitivity Analysis (SA). SA is the study of how the variation in the model output can be apportioned, qualitatively or quantitatively to variations in the model inputs [4]. Two types of SA can be distinguished: local (around a point) and global (on the entire input factor domain of variation). In RBDO context, the use of global sensitivity analysis methods allows to characterize the input variables on the entire input space and to filter out the uncertainty factors with negligible effects on the output, decreasing the computational burden and the complexity of the model while allowing to allocate resources to the modeling of the most influential uncertainties. The screening of the most important effects helps the decision maker to settle on which factors will be considered. Saltelli [4] lists a set of reasons to use SA. It allows to determine

- if the computer model represents the system or physical processes under study,
- the factors that contribute the most to the output variability,
- insignificant model input factors,
- areas in the domain of variation of input factors for which the model variation is maximal,
- if and which factors interact with each other.

Different categories of SA methods exist: the variance decomposition methods (Sobol [5], ANalysis Of VAriance [6]), differential analysis (Morris [7]) and linear relationship measures (Correlation Coefficients, Partial Correlation Coefficients, Standardized Regression coefficients [8]).
In this paper, we propose to compare the SA methods, to analyze their features, domain of applicability, and to apply the different methods to disciplines involved in the design of a launch vehicle. Other reviews of SA methods exist [8,9] that involve simple analytical test problems. The objective of the paper is to identify the most suitable SA methods for analyzing the disciplinary design subproblems of

a launch vehicle. The remainder of the article is organized as follows: in Section 4, a review of various global sensitivity methods is provided. It highlights the assumptions of each method and compares their features. In Section 5, we introduce the four disciplines typically involved in the design of a launch vehicle (aerodynamics, propulsion, geometry and mass budget, trajectory). In Section 6, a benchmark of the SA methods introduced in Section 5 is performed on the different disciplines. This benchmark outlines the accuracy of the methods depending on characteristics of the disciplinary models, in parallel to a comparison of the advantages and drawbacks of each method. In Section 7, we synthesize important insights that emerge from the comparative study. We stress the trade-off between the accuracy and computational cost of the SA methods. Finally, in Section 8, we discuss the issues induced by the SA performed on a coupled multidisciplinary system and the difficulty to compute the global sensitivity indices of the entire system based on the disciplines SA.

4. Sensitivity analysis methods

This section is an overview of the different SA methods. The computer based model that has to be analyzed is represented as a black box function f, with k input factors $\mathbf{X} = [X_1, X_2, ..., X_k]$ and an output vector $\mathbf{Y} = [Y_1, Y_2, ..., Y_m]$ related by:

$$\mathbf{Y} = f(\mathbf{X}) = f(X_1, X_2, ..., X_k) \tag{1}$$

The term "input factor" stands for any quantity that can be modified in the model prior to execution (variable, parameter, initial condition, final condition, *etc.*). Here, we consider a scalar output Y but all the derivations can be generalized for an output vector.

In a first part, we highlight the quantitative methods. We start by the most general one: Sobol method. Then we detail SA methods that can be applied under certain assumptions to the model: ANalysis Of VAriance by Design of Experiment (DoE), Standardized Regression Coefficients and Partial Correlation Coefficients.

In a second part, we detail a qualitative SA method relying on Morris method [7]. It is qualitative because the method provides a ranking in order of importance but no quantitative measures reflect how important a given input factor is compared to another in terms of influence on the output variability.

4.1 Quantitative methods

4.1.1 Decomposition of the variance

The variance decomposition methods for SA consist in a decomposition of the variance of the output into a sum of contributions due to the input factors and their interactions. These methods are called ANalysis Of VAriance (ANOVA) [6]. Two types of ANOVA can be distinguished: the functional ANOVA based on Sobol approach and the ANOVA by Design of Experiment [6]. The functional ANOVA method does not require any hypothesis on the form of the model except that the input factors are independent and $E(f^2(\mathbf{X})) < \infty$ (with E the expectation). Sobol [5] has demonstrated the following unique functional decomposition for the function f:

$$f(\mathbf{X}) = f_0 + \sum_{j=1}^{k} f_j(X_j) + \sum_{i<j}^{k} f_{ij}(X_i, X_j) + ... + f_{1...k}(X_1, ..., X_k) \tag{2}$$

where $f_0 = E(f(\mathbf{X})) = \int_\Omega f(\mathbf{X})p(\mathbf{X})d\mathbf{X}$. Ω is the k-dimensional cube $[0,1]^k$ where input factors are defined and assumed to have a uniform distribution U(0,1) represented by the probability density distribution p. Moreover, $f_j(X_j) = E(f(\mathbf{X})|X_j) - f_0$, $f_{ij}(X_i, X_j) = E(f(\mathbf{X})|X_i, X_j) - E(f(\mathbf{X})|X_i) - E(f(\mathbf{X})|X_j) + f_0$ and $f_{1...k}(X_1, ..., X_k)$ is defined as the difference between $f(x)$ and the sum of all the increasing dimension functions such that Eq.(3) is verified.

Furthermore, each function of the decomposition verifies [5]:

$$\int_\Omega f_{j_1,...,j_s}(x_{j1}, ..., x_{js})dx_{jl} = 0 \;; \forall l = 1, ..., s \;; \forall\{j_1, ..., j_s\} \subseteq \{1, ..., k\} \tag{3}$$

The orthogonality of the Sobol decomposition functions can be proved from Eq. (2) [5].

Sobol indices. Based on the functional decomposition Eq. (2), Sobol introduces the Sobol indices to quantify the partition of the output variance [5]. With the decomposition of the studied function f into the sum of functions of increasing dimensions (as exposed in the previous paragraph) and by using the decomposition of the variance [5], we have:

$$V(Y) = \sum_{j=1}^{k} V_j(Y) + \sum_{j<i}^{k} V_{ji}(Y) + ... + V_{123...k}(Y) \tag{4}$$

with V the variance, and $V_j(Y)$ as defined in Eq.(8). The variability of the output Y due to all input factors but X_j is analyzed by fixing the input variable X_j at a value x_j:

$$V(Y|X_j = x_j) = E(Y^2|X_j = x_j) - E(Y|X_j = x_j)^2 \tag{5}$$

The problem of this quantity is that it is dependent on the choice of x_j for X_j, therefore, the expected value for all the possibles x_j is taken into consideration:

$$E[V(Y|X_j = x_j)] = \int_{\Omega_{X_j}} V(Y|X_j = x_j)p_{X_j}(x_j)dx_j \tag{6}$$

Given the total variance:

$$V(Y) = V[E(Y|X_j)] + E[V(Y|X_j)], \tag{7}$$

the value $V[E(Y|X_j)]$ can be used for SA. It increases as the importance with respect to the variance of Y increases. To have a normalized quantity, the first order Sobol index S_j for the input factor X_j and second order Sobol index S_{ij} for the interaction between X_i and X_j, are defined by [5,10]:

$$S_j = \frac{V[E(Y|X_j)]}{V(Y)} \triangleq \frac{V_j}{V(Y)} \,;\, S_{ij} = \frac{V[E(Y|X_i, X_j)] - V_i - V_j}{V(Y)} \triangleq \frac{V_{ij}}{V(Y)} \tag{8}$$

The first order Sobol index quantifies the part of variance of Y due to X_j, refereed as main effect. The second order Sobol indices allow to measure the importance of the interaction between two input variables X_i and X_j. The same principle can be used to derive the Sobol indices of order 3, 4, *etc.* The total Sobol ST_i indices are the sum of all the Sobol indices relative to X_j:

$$ST_j = \sum_{j \sharp i} S_i \tag{9}$$

where $j\sharp i$ stands for all the $S_{i1,...,ik}$ terms that include the index *j*. For instance, ST_1 includes $S_1, S_{12}, ..., S_{1k}$, $S_{123}, ..., S_{123...k}$. Total Sobol indices measure the part of output variance explained by all the effects in which the index j plays a part (the first order and all the higher orders) [10]. For black box functions, Sobol indices cannot be defined analytically and have to be numerically estimated. Several methods can be employed. Monte Carlo (MC) method is traditionally used to estimate Sobol indices. Other sampling schemes can be performed such as Jansen's [11] or Fourier Amplitude Sensitivity Test (FAST) [12]. However, Sobol calculations are computationally expensive and require a large number of calls to the studied function. The Sobol method is applicable to every cases for which variances are finite (linear or nonlinear, monotonic or non monotonic functions). Another approach for SA based on the decomposition of the variance can be used. Instead of considering the functional ANOVA (Eq. (2)), the conditional variances are calculated with a Design of Experiment.

ANOVA by Design Of Experiment approach. ANOVA by DoE differs from the functional ANOVA in that it can only be used for qualitative or discrete input factors. A discretization of the continuous input factors is necessary. Eq.(4) is discretized according to a chosen DoE into several levels for each input factors. By choosing appropriate levels and DoE, the conditional variances can be approximated.

Each observation from the DoE can be modeled as [6,13]:

$$y_{X_{1h_1},X_{2h_2},...,X_{kh_k}} = \eta + \eta_{X_{1h_1}} + \eta_{X_{2h_2}} + ... + \eta_{X_{kh_k}} + \eta_{X_{1h_1}X_{2h_2}} + ... + \eta_{X_{1h_1}X_{2h_2}...X_{kh_k}} \tag{10}$$

with η the average effect, $\eta_{X_{1h_1}}$ the effect due to the level h_1 of the factor X_1, and $\eta_{X_{1h_1}X_{2h_2}}$ the effect due to the interactions between the level h_1 of the factor X_1 and the level h_2 of the factor X_2. Therefore, if the DoE requires N calls to the function:

$$\eta = \frac{1}{N}\sum_{j=1}^{k}\sum_{all\ levels\ h_j} y_{X_{1h_1},X_{2h_2},...,X_{kh_k}} = E\left(y_{X_{1h_1},X_{2h_2},...,X_{kh_k}}\right) \tag{11}$$

and for instance:

$$\eta_{X_{1h_1}} = E[y_{X_{1h_1},X_{2h_2},...,X_{kh_k}}|X_1 = h_1] - \eta \tag{12}$$

with the expectation representing the mean of all the observations in the DoE when $X_1 = h_1$.

ANOVA by Design of Experiment [6] allows to write the sum of squares decomposition:

$$SS(\tilde{\mathbf{Y}}) = SS_{X_1} + SS_{X_2} + ... + SS_{X_k} + SS_{X_1X_2} + ... + SS_{X_1X_2...X_k} \tag{13}$$

where $\tilde{\mathbf{Y}}$ is the vector of all the observations given by the DoE, SS_{X_j} the quadratic sum characterizing the main effect of the input factor X_j, $SS_{X_jX_r}$ the quadratic sum characterizing the interaction effect between the input factors X_j and X_r. Eq.(13) is a discrete version of Eq.(4).

$SS(\tilde{\mathbf{Y}})$ measures the total variability in the model output:

$$\begin{aligned} SS(\tilde{\mathbf{Y}}) &= \sum_{j=1}^{k} \sum_{all\ levels\ h_j} \left(y_{X_{1h_1}, X_{2h_2}, \ldots, X_{kh_k}} - \eta \right)^2 \\ &= NE\left(y^2_{X_{1h_1}, X_{2h_2}, \ldots, X_{kh_k}} \right) - 2N\eta E\left(y_{X_{1h_1}, X_{2h_2}, \ldots, X_{kh_k}} \right) + NE\left(y_{X_{1h_1}, X_{2h_2}, \ldots, X_{kh_k}} \right)^2 \\ &= NV(\tilde{\mathbf{Y}}) \end{aligned}$$

Moreover, each quadratic sum SS_{X_j} corresponds to the mean on all the levels taken by the input factor X_j of the conditional variance of the output $\tilde{\mathbf{Y}}$ for $X_j = h_j$:

$$SS_{X_j} = n_j \sum_{levels\ of\ X_j} \left(E[y_{X_{1h_1}, X_{2h_2}, \ldots, X_{jh_j}, \ldots, X_{kh_k}} | X_j = h_j] - \eta \right)^2 = NV(E(\tilde{\mathbf{Y}}|X_j)) \tag{14}$$

with n_j the level discretization number of the input factor X_j. Based on this decomposition of variance, the sensitivity index SI for the input factor X_j is defined by [6]:

$$SI_{X_j} = \frac{SS_{X_j}}{SS(\tilde{\mathbf{Y}})} \tag{15}$$

DoEs as fractional factorial, Latin square or full factorial allow to decrease the SA computational cost compared to Sobol approach [13]. However, the choices of the DoE and the different levels are crucial to approximate accurately the conditional variances.

Other SA methods exist under the assumptions of a linear model. Next, we detail two linear relationship measures: Standardized Regression Coefficients and Partial Correlation Coefficients.

4.1.2 Standardized Regression Coefficients (SRC)

In the case f can be approximated as a linear function, sensitivity measures for the model can be computed through linear regression. A linear model for the dependency of the outputs with respect to the k input factors N samples $(i = 1, ..., N)$ is considered:

$$Y_i = a_0 + \sum_{j=1}^{k} a_j X_{ij} + \epsilon_i \tag{16}$$

with a_i the regression coefficients and ϵ_i the errors due to the approximation (assumed $\epsilon_i \sim N(0, \sigma_i)$). Least squares fit between predicted and observed output data is typically used for the determination of the linear model. The coefficient of determination ($R^2 = \frac{\sum_{i=1}^{N}(\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{N}(Y_i - \overline{Y})^2}$, with N the sample size) allows to determine how well the linear model fits the data. The regression coefficients a_j measure the linear relationship between the input factors and the output. To avoid a unit dependence of the regression coefficients, it is possible to normalize the coefficients [14]:

$$\frac{\widehat{Y}_i - \overline{Y}}{\hat{t}} = \sum_{j=1}^{k} \frac{a_j \widehat{t}_j}{\hat{t}} \frac{X_{ij} - \bar{X}_j}{\widehat{t}_j} \tag{17}$$

with $\overline{Y}$ the mean of the output and $\bar{X}_j$ the mean of the input factor X_j over the N-sample. Furthermore:

$$\widehat{t} = \left(\sum_{i=1}^{N} \frac{(Y_i - \overline{Y})^2}{N-1} \right)^{1/2} ; \widehat{t}_j = \left(\sum_{i=1}^{N} \frac{(X_{ij} - \bar{X}_j)^2}{N-1} \right)^{1/2} ; SRC_j = \frac{a_j \widehat{t}_j}{\widehat{t}} \tag{18}$$

If the input factors are independent, SRC is a measure of the importance of the input factor on the variability of the output. Another interpretation comes from the decomposition of the variance for a linear function (by independence of input factors) as [8, 10]:

$$V(Y) = \sum_{i=1}^{k} a_i^2 V(X_i) \tag{19}$$

with $a_i^2 V(X_i)$, the part of variance due to the input variable X_i. Thus, it is possible to quantify the sensitivity of Y with respect to X_i, compared to the part of the variance due to X_i on the total variance.

$$SRC_j = a_j \sqrt{\frac{V(X_j)}{V(Y)}} \tag{20}$$

It quantifies the part of the variance of the output due to the variance of the variable X_i.

4.1.3 Correlation coefficients (CCs) and Partial Correlation Coefficients (PCCs)

Another SA measure in the case of linear model is given by Pearson's product moment correlation coefficients [14]. They measure the extent to which two variables can be assumed to have a linear dependency. In our case we are interested in measuring the dependency between the input factors X_j and the output Y. Considering N observations of Y for different X_j , CC is defined by [9]:

$$CC_j = \rho_{X_jY} = \frac{cov(X_j, Y)}{\sqrt{V(X_j)V(Y)}} = \frac{\sum_{i=1}^{N}(X_{ji} - \overline{X}_j)(Y_i - \overline{Y}))}{\left(\sum_{i=1}^{N}(X_{ji} - \overline{X}_j)^2\right)^{1/2}\left(\sum_{i=1}^{N}(Y_i - \overline{Y}))^2\right)^{1/2}} \tag{21}$$

with $\mathrm{cov}(X_j, Y)$ the covariance between X_j and Y. However, CCs do not take into account the possible effects that other variables might have. Partial Correlation Coefficient (PCCs) can be calculated to determine the strength of the linear relationship between the two factors when all linear effects from the other input factors are removed [8]. If we note $S_j = \{X_1, X_2, ..., X_{j-1}, X_{j+1}, ..., X_k\}$, then the PCC between X_j and Y with S_j fixed is given by:

$$PCC_{j|S_j} = \rho_{X_jY|S_j} = \frac{cov(X_j, Y|S_j)}{\sqrt{V(X_j|S_j)V(Y|S_j)}} \tag{22}$$

When the input factors are uncorrelated, we have SRC=CC. Note however that SRC and PCC measure different quantities. SRCs measure the effect on the output of the input factors in terms of a percentage of the outputstandard deviation. PCCs measure how linear is the relationship between one input factor and the output while removing the effect of other input factors.

The quantitative methods for SA quantify the importance of the variability of input factors and their interactions on the variability of the output. However, these methods tend to be computationally intensive. This is particularly true of Sobol's method, while ANOVA by DoE may be an exception if a small number of levels is chosen.

In the case where only few calls to the function are possible, screening methods can be employed.

4.2 Screening methods

Screening methods complement quantitative methods since the required number of model evaluations is low compared to other SA techniques [14]. For the design of aerospace vehicles which involves computationally expensive models and a large number of input factors, screening methods can identify the factors that have the strongest effects on the output variability.

4.2.1 One At a Time

The One factor At a Time (OAT) method is based on the variation of only one factor while the others are kept fixed at a baseline value. If we consider a baseline for the model (the nominal values of the input factors), we perform the OAT analysis by varying one of the inputs in an interval (for instance $\pm 10\%$) while the other input factors are fixed to the baseline value. The range of the output is analyzed for the set of the k-OAT computations. Another OAT technique consists in computing the partial derivatives of the model's function with respect to the input factors. This method is called *local sensitivity analysis* [7] as it depends on the choice of the point where the partial derivatives are calculated. The number of model evaluations is of the order of k [14]. While these OAT methods require few model evaluations, they are local and do not provide information on the entire range of variation of the input factors.

4.2.2 Morris method

Morris method [7] is based on a repetition of a set of randomized OAT design experiments. Morris method overcomes the limitation of the local sensitivity analysis by performing partial derivative calculations in different locations of the input factor range of variation chosen according to a random OAT experiment. The method is global because the input factors can vary over their entire domain of definition. Morris method consists in R repetitions of random DoEs with sequential OAT variation of the inputs [8]. The first point and the next direction for one experiment are chosen randomly. Morris method can establish if the effect of the input factor X_j on the output Y is important or negligible, linear or non linear, with or without interactions with other input factors X_{-j}. Morris distinguishes three ways an input factor

X_j can be important [7] depending on the nature of $\frac{f(X_j+\delta_j,X_{-j})-f(X_j,X_{-j})}{\delta_j}$, where δ is a variation in the input factor:

- if this term is non null, then X_j has an influence on the output,
- if this term is non null and does not vary as X_j varies, therefore, X_j has a linear influence on the output and has no interactions with other input factors,
- if this term varies as X_j varies, then X_j affects non linearly the output with or without interactions.

It is not possible with Morris method to distinguish between non linearity and the interactions with other input factors. Morris method is applicable in dimension n. The mean of the absolute value of the different partial derivatives is a measure of the sensitivity. The variance is a measure of both the interactions and the non linear effects. The main advantage of Morris method is the low computational cost, requiring only about one model evaluation for each elementary effect per replication. However, its drawback is that it is not possible to distinguish non linearity from interactions, which might be essential for the designer [14]. Furthermore, the sensitivity measures cannot be quantitatively interpreted beyond their definition.

5. Disciplinary analyzes for the design of a launch system

In this section we present the four typical disciplines involved in the design of a launch vehicle: the propulsion, the geometry and mass budget, the aerodynamics and the trajectory. These four disciplines will be used for a benchmark and a numerical comparative analysis of the SA methods presented in the previous section. We base our analysis on the baseline of the Vega European launcher [17].

5.1 Propulsion. The propulsion discipline provides essential characteristics for a launch vehicle such as: the thrust produced by the engines, the specific impulse, the combustion dynamics, *etc.* For our SA, we use the first stage of Vega launch vehicle (P80-FW)[17]. The P80-FW is a solid propellant motor. We focus our analysis on the maximal thrust produced by the motor by considering five input variables: the initial propellant mass, the stage diameter, the maximal admissible chamber pressure, the nozzle expansion ratio and the relative length of the two types of grain for the propellant: tube and star (Figure 1 and Table 1).

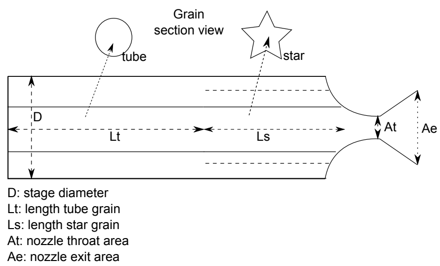

Figure 1: Propulsion discipline

The propulsion calculations are based on the NASA computer program CEA (Chemical Equilibrium with Applications)[16] which calculates chemical equilibrium compositions and properties of complex mixtures and on flow dynamics calculations of the expansion of gases through the nozzle of the motor. A uniform distribution (defined in Table 1) is assumed for the input variables.

Table 1: Sensitivity analysis input variable definition for propulsion and mass budget

Input variables	Min value	Baseline	Max value
Nozzle expansion ratio $\epsilon = \frac{At}{Ae}$	14.4	16	17.6
Initial propellant mass (kg)	79200	88000	96800
Stage diameter (m)	2.70	3.005	3.31
Maximum admissible chamber pressure (bar)	85.5	95	104.5
Relative length grain tube $R = \frac{Lt}{Lt+Ls}$	65%	70%	75%

5.2 Geometry and Mass budget. The geometry and mass budget discipline consists in computing the overall geometry and dry mass of the rocket. For our SA, we are interested in the P80-FW dry mass. In the case of a solid propellant motor, the dry mass of the motor can be mainly decomposed into three parts: the mass of the motor case (the structure around the propellant), the mass of the insulation and the mass of the nozzle. The input considered variables and their distributions are the same as the ones of the propulsion discipline (Table 1).

5.3 Aerodynamics. The aerodynamics discipline consists in computing the aerodynamics coefficients such as the drag and lift coefficients to characterize the aerodynamics of the rocket. The calculations of the drag and lift coefficients are based on the US AirForce computer program MissileDATCOM [15] relying on an experimental data base to determine the aerodynamics forces and coefficients for complex rocket geometry. It takes as input the geometry of the launcher (Figure 2 and Table 2), the angle of attack, the Mach and the altitude. A uniform distribution (defined in Table 2) is assumed for the input variables.

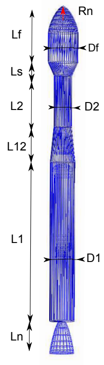

Figure 2: Aerodynamics discipline

Table 2: Sensitivity analysis input variable definition for aerodynamics

Input variables	Min value	Baseline	Max value
Length stage 1 L1 (m)	6.93	7.7	8.47
Diameter stage 1 D1 (m)	2.70	3.005	3.31
Length nozzle Ln (m)	0.78	0.87	0.95
Length interstage 1-2 L12 (m)	3.15	3.5	3.85
Length stage 2 L2 (m)	11.26	12.51	13.76
Diameter stage 2 D2 (m)	1.71	1.9	2.09
Length fairing Lf (m)	7.09	7.88	8.67
Diameter maximum Df (m)	2.34	2.6	2.86
Length nose Rn (m)	2.84	3.16	3.48
Length fairing adapter Ls(m)	1.38	1.53	1.69
Angle of attack (deg)	0	2	4
Mach	0.4	4	5
Altitude (m)	10000	25000	30000

5.4 Trajectory. The trajectory discipline consists in computing the trajectory of the rocket. For the SA, we simulate only the first stage P80-FW trajectory of the launcher Vega until the separation. The trajectory is simulated with the Andromede (Perseus, Cnes) computer program. We analyze the maximal axial load, the altitude and speed of separation with the first stage. A thrust profile is parametrized by two parameters: thrusts 1 and 2. A uniform distribution (defined in Table 3) is assumed for the input variables.

Table 3: Sensitivity analysis input variable definition for the trajectory

Input variables	Min value	Baseline	Max value
Payload mass (kg)	1350	1500	1650
Propellant mass stage 1 (kg)	79200	88000	96800
Combustion time (s)	95.4	106	116.6
Dry mass (kg)	43730	48589	53448
Thrust 1 (kN)	2475	2750	3025
Thrust 2 (kN)	2115	2350	2585
Reference area (m^2)	6.19	7.21	7.93
Gravity turn angle (deg)	22.5	25	27.5
Time of gravity turn (s)	8	9	10

6. Disciplinary SA results

To compare the different SA methods we focus on the following criteria:

- Domain of applicability (linear model, non linear, differentiable model),
- Information provided (qualitative, quantitative, main effects, interactions),
- Robustness with respect to the sampling size and computational cost.

To compare each SA method, the same number of calls to the function is implemented, except for ANOVA-DoE, in which the DoE is taken according to a 2-level orthogonal array based on a latin square [13], with the minimum and maximum of the uniform distribution taken for the levels. The Sobol indices are estimated by the EASI method [9] derived from the FAST method [12]. The following paragraphs are organized based on the input variable sets.

6.1 Aerodynamics. The aerodynamics discipline is a highly non linear and non monotonous model. The model coefficient of determination is $R^2 = 0.254$, the regression model used by SRC and PCC calculations is not accounting for most of the variability in the drag coefficient. SRC, ANOVA-DoE, Morris and Sobol do not provide the same ranking for the most influential parameters (Figures 3 & 3a). Only Sobol and Morris rank the five first most influential parameters the same way. SRC method ranks the fairing diameter and the Mach as the two most influential variables and ANOVA-DoE ranks the Mach and the length of the interstage 1-2 as the two most influential variables. However, due to the non linearity of the model, SRC and ANOVA-DoE do not provide accurate ranking and relative contributions of the input factors to the variability in the drag coefficient ((Figure 3 & 3a)). Morris and Sobol provide the same ranking. However, only Sobol offers the possibility to quantify the relative contributions of the input factors to the variability of the drag coefficient.

6.2 Propulsion and mass. The propulsion and mass disciplines rely on complex and computationally intensive models. However, the relationship between the input factors and the maximal thrust or the dry mass can be approximated by a linear regression (respectively $R^2 = 0.995$ and $R^2 = 0.9942$). The four SA methods provide the same ranking and the same relative contributions (within 1%) to the variability of the maximum thrust (Figures 4 & 4a) and the dry mass (Figures 7 & 7a). SRC, Morris and Sobol methods rely on 5000 calls to the function whereas ANOVA-DoE relies on 8 calls to the function.

6.3 Trajectory. The trajectory discipline can be relatively accurately approximated by a linear model considering the studied outputs ($R^2 = 0.916$) in which the regression model accounts for most of the variability in the maximum axial load. SRC, Sobol and Morris methods provide the same ranking and SRC and Sobol give the same relative contributions (within 1%) to the variability of the maximum axial load. However, ANOVA-DoE does not provide the same ranking and relative contributions. The approximation by a linear model based on the minimum and maximum levels does not accurately capture the variability of the maximum axial load resulting in wrong relative contributions (Figures 5, 5a, 6, 6a).

7. Synthesis

7.1 Domain of applicability. In the case of a relatively linear model, the four SA methods provide the same ranking and relative contributions to the variability of the output. However, in the case the model can only be approximated by a linear regression, the ANOVA-DoE by 2 levels is not robust compared to the SRC method and provides wrong relative contributions. In the case of a highly non linear model, only Sobol and Morris provide the same ranking for the most important input factors. ANOVA-DoE could provide more accurate results if the DoE was chosen by increasing the number of level according to the variations of the function. Morris method requires the function to be $f \in \mathcal{C}^1$ which is not always the case.

7.2 Information provided. Morris provides only qualitative information through the ranking of the input factors based on their contributions to the output variability. Recent works [6] studied the link between derivative-based global sensitivity measures (as Morris) and Sobol indices, but further efforts are still expected to examine the link between the two SA measures. Sobol, ANOVA-DoE and SRC provide ranking and relative contributions to the output variability. Eventually, Sobol and ANOVA-DoE allow to compute the relative contributions of the interaction between the input factors. Therefore, depending on the context, if we want to pre-screen the non influential input factors in a complex computation code (as in a MDO framework), ANOVA-DoE 2 levels for linear model or Morris for non linear model are sufficient. However, if we are interested to know the exact contributions of each input factor and their interactions, SRC can be used for linear model (but without knowing the relative contributions of the interactions) and Sobol for non linear model.

7.3 Robustness to sample sizing and computational cost. As disciplines involved in the design of a launch vehicle required intensive computation, it is often necessary to limit the number of calls to the functions. Therefore, the robustness to the sample size is studied on the trajectory discipline (Figure 8 & 8a), by decreasing the number of call to the function from 9000 to 500. Morris provides the same ranking and SRC method provides the same relative contributions (within 3%). However, Sobol method is not robust to the reduction of the sample size and provides differences in the ranking and in the relative contributions (as high as 15%). ANOVA-DoE 2-levels requires a number of calls to the function in order of the number of input factors which is interesting since complex disciplines as the propulsion are expensive to evaluate. Moreover, for complex disciplines with a large number of input variables (more than 10)[6], Sobol indices cannot be evaluated by the traditional Monte Carlo method due to the unacceptable computational cost and more advance methods as EASI [9] are required to limit the number of call to the function while providing accurate results.

Table 4: Summary of the sensitivity analysis methods, advantages and drawbacks

Methods	Advantages	Drawbacks
Sobol	• No hypothesis on the form of the model • Converge to the exact relative contributions of the input factors and their interactions to the output variability • Quantification of interactions between the input factors • Methods like FAST allow to compute more efficiently the sensitivity indices than MC	• Non robust to decrease of the sample size • Computationally intensive, need of replications to compute the indices • Often intractable
ANOVA-DoE	• Few calls to the function, can quantify input factor interactions • Easy to implement	• Restricted to qualitative or discrete input factors • Choice of DoE and levels are critical
SRC/PCC	• Less computationally intensive than Sobol, robust to sample size • Easy to compute	• No information on the contribution of input interactions • Limited to linear model
Morris	• No assumptions on the linearity of the model • Robust to few calls to the function	• Not a relative contribution of the input factors measure • No distinction between non linearity and factor interactions • Calculation of derivatives and $f \in \mathcal{C}^1$

8. Sensitivity analysis in a multidisciplinary design framework

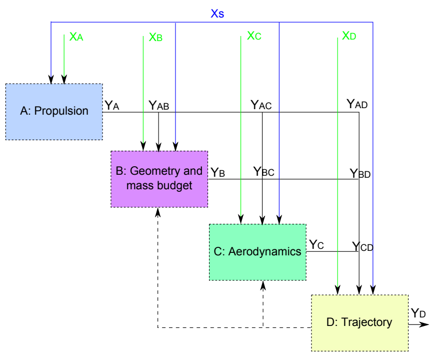

Figure 3: Multidisciplinary design of launch vehicle

In the previous sections, we detailed and applied different SA methods on several disciplines involved in the design of launch vehicles. However, the design of such aerospace vehicle is a coupled multidisciplinary problem and disciplinary SA are not sufficient to represent the global contributions of input factors on the global system performances [18]. Indeed, the contributions of shared variables $\mathbf{Xs}$ are not fully captured by disciplinary SA. For instance, considering Figure 3, the geometry and mass budget discipline SA has three category of inputs: the discipline inputs $\mathbf{X_B}$, the shared inputs $\mathbf{Xs}$ and the coupling inputs with propulsion discipline $\mathbf{Y_{AB}}$. However, the sensitivity indices representing relative contributions of the shared variables $\mathbf{Xs}$ on the geometry and mass budget outputs $\mathbf{Y_B}$ are not taking into account all the contributions due to $\mathbf{Xs}$ since the input coupling variables $\mathbf{Y_{AB}}$ are also affected by the shared variables. To overcome this problem, global sensitivity indices applicable to hierarchical systems have been propounded [18-19]. Assuming that no feedback loops (dotted arrows in Figure 3) exist in the multidisciplinary problem, the design can be viewed as a hierarchical system design with the trajectory discipline being the lower level depending on the upper levels (aerodynamics, propulsion and geometry and mass budget). The approach of computing the sensitivity indices for a multidisciplinary system based on disciplinary SA is interesting for three reasons: first, it would allow to avoid to perform a single SA with a large number of uncertain input variables on the entire system but rather to perform several disciplinary SAs with less uncertain input variables as the number of uncertain variables is a key aspect in SA. Secondly, it would avoid to perform computationally expensive Multi-Disciplinary Analyses (MDAs) by using disciplinary models. Eventually, it could provides disciplinary SA information allowing the decision maker to decide to refine disciplinary models with large influences on the system performances.

However, the multidisciplinary aspect of the problem rises several challenging issues to allow the computation of the sensitivity indices of the performance of the entire system based on disciplinary SAs:

- If separated discipline SA are performed, the uncertainty in the coupling variables Y_{ij} between discipline i and j are unknown and only estimated, affecting the accuracy of the sensitivity indices of the lower disciplines.
- When the disciplines share common input factors, the resulting coupling variables, which are inputs of the lower level, are not independent. The relative contributions of the input dependent coupling variables to the lower level output variability are affected by the modeling of the input dependent coupling variable distributions. An estimation of the joint probability distribution of the dependent coupling variables is necessary before performing the lower level discipline SA. The computation

of the joint probability distribution of the dependent coupling variables requires the simultaneous simulation of the lower discipline models [18] and therefore is numerically costly since it involves repeated MDAs of the entire system. For instance, in Figure 3, the coupling variables between trajectory and aerodynamics $\mathbf{Y_{CD}}$ and trajectory and propulsion $\mathbf{Y_{AD}}$ are dependent variables. To compute the relative contributions of the dependent coupling variables ($\mathbf{Y_{AD}}$,$\mathbf{Y_{BD}}$,$\mathbf{Y_{CD}}$) to the variability of the trajectory output $\mathbf{Y_D}$ an estimation of the joint probability of these dependent variables is necessary. The calculation of this joint probability requires the simultaneous simulation of the three upper level disciplines which is computationally intensive and limits the interest of performing separated discipline SA to compute the sensitivity indices of the performance of the entire system.

- If separated discipline SA are performed, it is essential to ensure that the couplings between the disciplines are satisfied to guarantee the feasibility of the system and the accuracy of the SA on the performances at the system level.
- One way to compute the global sensitivity coefficients is to aggregate the different disciplinary sensitivity indices, as proposed in [18] for the main effect indices. The method is only viable for hierarchical systems with no feedbacks loops. For a hierarchical system, two cases are distinguished, depending on if there is a linear or non linear relation between the lower level outputs and the upper level dependent coupling variables. Sobol [20] method assumes the independence in the input variables of a discipline. In the case of dependent input variables, Sobol method has to be performed on all the independent variables and on an artificial subset variable $\mathbf{U}_{\mathbf{Y}_i,\mathbf{Y}_j}$ which includes all the dependent coupling variables [18]. Then, the covariance of the dependent coupling variables needs to be determined to compute the global sensitivity main effect indices. A correction coefficient is used in [18] to compute the SA of the entire system performance in case of non linear dependency. The method is limited to the estimation of the main effect of the entire system and further efforts are necessary to extend this methodology for interaction effect terms.
- If feedback loops exist in the multidisciplinary design problem, the system cannot be considered hierarchic and other methods have to be developed to compute the global sensitivity indices of the entire system performances.

These difficult issues have to be solved to be able to compute the sensitivity indices of a multidisciplinary coupled system based on disciplinary SA. Due to the challenges imposed by the shared variables and the resulting couplings, it is not clear that such approach will be computationally more efficient than a single SA on the full multidisciplinary system.

9. Conclusions and further work

This paper provides a theoretical and numerical comparison of the main global SA method in industrial benchmark test cases. It highlights when to apply each SA methods depending on the complexity and characteristics of the model and on the numerical applicability (Table 4). In case we are only interested in ranking the most important variables, Morris method is robust for any differentiable function (linear or non linear). In case we are interested in the relative contributions of the input factors to the output variability and the function is linear, ANOVA-DoE is appropriate and requires few calls to the function. In case the function is non linear Sobol method provides accurate relative contributions but is computationally intensive. Further efforts and researches on DoE configurations could be interesting to improve the ANOVA-DoE method allowing to capture the variation of the function providing relative contribution of the main effects and the interaction effects while keeping low the number of calls to the function compared to Sobol method. Additional researches on SA in a multidisciplinary system with feedback loops are necessary to compute the global sensitivity indices of the entire system based on disciplinary SA.

Acknowledgment: The work presented in this paper is part of a CNES/ONERA PhD thesis.

9. References

[1] J. Sobieszczanski-Sobieski *et al.*, Multidisciplinary aerospace design optimization: survey of recent developments, *Structural and Multidisciplinary Optimization*, 1997, 14(1): 1-23.

[**2**] W. Kuo *et al.*, An annotated overview of system-reliability optimization, *Reliability, IEEE Transactions*, 2000, 49(2): 176-187.

[**3**] Yao *et al.*, Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles, *Progress in Aerospace Sciences*, 2011, 47(6): 450-479.

[**4**] A. Saltelli *et al.*, Sensitivity analysis as an ingredient of modeling, *Statistical Science*, 2000:377-395.

[**5**] I.M. Sobol, Sensitivity estimates for nonlinear mathematical models, *Mathematical Modelling Computational Experiments*, 1993: 407-414.

[**6**] M. Lamboni, Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models, *Reliability Engineering & System Safety*, 2011, 96(4): 450-459.

[**7**] M. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics*, 1991, 33(2): 161-174.

[**8**] B. Iooss, Revue sur l′analyse de sensibilité globale de modèles numériques, *Journal de la Société Française de Statistiques*, 2011, 152(1): 3-25.

[**9**] E. Plischke *et al.*, Sensitivity analysis benchmark based on the use of analytic and synthetic PA cases milestone (N: M2. 1. D. 11), European commission report, PAMINA project, 2009.

[**10**] J. Jacques, Pratique de l′analyse de sensibilité: comment évaluer l′impact des entrées aléatoires sur la sortie d′un modèle mathématique, 2011, Université de Lille.

[**11**] M. Jansen, Analysis of variance designs for model output, *Computer Physics Communications*, 1999, 117(1): 35-43.

[**12**] A. Saltelli *et al.*, A quantitative model-independent method for global sensitivity analysis of model output, *Technometrics*, 1999, 41(1): 39-56.

[**13**] H. Monod *et al.*, Uncertainty and sensitivity analysis for crop models, 2006: 55-100, Working with Dynamic Crop Models, Elsevier.

[**14**] P. Ekström *et al.*, Sensitivity analysis methods and a biosphere test case implemented in EIKOS, *Posiva Working Report*, 2006.

[**15**] W. Blake, Missile Datcom: User's Manual-1997 FORTRAN 90 Revision, 1998, DTIC Document.

[**16**] B. McBride *et al.*, Computer program for calculation of complex chemical equilibrium compositions and applications, II. *NASA reference publication*, 1996, 1311: 84-85.

[**17**] Arianespace, Vega′s User Manual, 2006.

[**18**] Y. Liu *et al.*, A hierarchical statistical sensitivity analysis method for multilevel systems with shared variables, *Journal of Mechanical Design*, 2010, 132, 031006.

[**19**] X. Yin *et al.*, A hierarchical statistical sensitivity analysis method for complex engineering systems design,*Journal of Mechanical Design*, 2008, 130, 071402.

[**20**] I.M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Mathematics and computers in simulation*, 2001, 55(1-3), 271-280.

A. Appendix: Sensitivity analysis results for the launch vehicle design disciplines