Global Sensitivity Analysis in Industrial Application with LS-OPT

Björn Hohage*, Anja Förderer**, Gordon Geißler**, Heiner Müllerschön**

*Audi AG, Ingolstadt, Germany

**DYNAmore GmbH, Ingolstadt, Germany

Summary:

The presented paper is related to a global sensitivity analysis in the field of numerical simulations. The objects of investigation are different car side crash scenarios and the influence of model variable variation on the occupant safety.

The work is motivated by the enhancement of the understanding for complex system behaviour of occupant safety components. Therefore, the methodology of the investigation should be conditioned in such a way, that it can be used by engineers in the standard development process for car projects. For the systematisation of such studies it is essential to know how many experiments are necessary to approximate the global behaviour in a suitable manner. Therefore, a convergence study is examined. Other addressed questions are related to the comparability of different sensitivity measures and the influence of the distribution of experiments in the design space. To provide a comprehensive view on the results, different possibilities to compress the large amount of information are developed, tested and compared. Furthermore, the variable sampling in the design space and the influence of the type of the global response approximation model are addressed.

For the evaluation of the sensitivities between variables and system responses, different sensitivity analysis methods are applied. LS-OPT, the main software component for this analysis, provides different methods for the evaluation of dependencies between system input and output quantities. Namely, the linear ANOVA and the more general formulation of Sobol's indices are evaluated within this study to classify the individual variable influence. A number of design variables are selected for variation. In the vehicle model, the thickness of constructional elements from the side structure are introduced. Also some values of friction are considered as variables. As a value of model uncertainty, the positioning of the barrier is changed in the range of allowed deviations of the experimental execution. The selected set of variables is simultaneously used for the investigation of four side crash load cases.

Keywords:

Sensitivity Analysis, LS-OPT, ANOVA, Sobol, DOE, Robustness, Process and Data Management, Occupant Safety, Side Impact, CAx-LoCo

1 Introduction

The intension of this project is to understand the behaviour of complex systems and to educe a standard development process. With this standard process the engineer should be able to analyse the influence of model variables on occupant safety values with minimal effort. In this case some concrete questions have to be researched. For example how many simulations are necessary to approximate the global behaviour in a suitable manner. The influence of the sampling point distribution over the design space is investigated. Another question is if the results of various sensitivity analyses methods differ and what assets and drawbacks these methods have. In this context also the influence of a non-linear in contrast to a linear investigation is compared.

1.1 Terms and definitions

For the present investigation, variables are quantities that represent uncertainties that are given in the simulation model. This could be the thickness of a material or a coefficient of friction, for example. The influence of these uncertainties to various responses are analysed. Responses are various injury criteria that are consulted for rating the occupant safety. The rib intrusion is one of the responses, for example. Analytical functions that describe the coherence between variables and responses are not available. Therefore, meta models are build, that give an approximation for this coherences in the design space. The design space is a n-dimensional variable space, where n is the number of variables. In the design space a DOE (Design Of Experiments), that describes the sampling point selection, is set. In LS-OPT various methods for the point selection are available: Koshal, D-Optimal, Space Filling, Monte Carlo, etc. A sensitivity and a robustness investigation are rendered. In a sensitivity analysis the points are distributed preferable equally and deterministic over the design space. In contrast, the distribution is random in a robustness investigation and probabilistic aspects are taken into account. The variable ranges in a sensitivity analysis can be relativly large, whereas usually small ranges are considered in a robustness investigation.

1.2 Relevant load cases and variables

There are two types of side impacts: The vehicle-side-impact, i.e. the crash of another vehicle in the side, that is simulated with a moveable deformable barriere and the pole-side-impact, i.e. the crash towards a tree or column, where the vehicle is moved towards a fixed non deformable pole. Four different load cases are investigated. Two barrier (IIHS, US-SINCAP) and two pole load cases (EuroNCAP, FMVSS 214). The four load cases are illustrated in Figure 1.
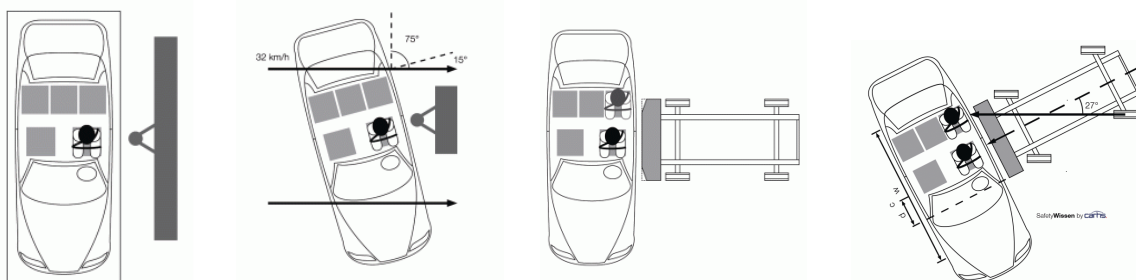

Figure 1: Load cases [9] (f.l.t.r.: EuroNCAP, FMVSS 214, IIHS, US-SINCAP)

In total, the influence of variables on the responses is investigated. The variables are dimensional quantities, coefficients of friction the moving direction and position of the barrier and mass distributions. The dimensional variables are thicknesses of iron sheets and carpetings of the side structure.

1.3 Evaluation of occupant safety values, injury criteria

Various injury criteria are evaluated. The injury criteria give a correlation between load measured at the dummy model and injury severity of an human beeing. Injury criteria can be accelerations (head, chest), forces (abdomen, pelvis), displacements (chest) and derived quantities (HIC, VC). The injury criteria of the ES2 dummy model, used in load cases EuroNCAP, FMVSS 214 and US-SINCAP, are illustrated in Figure 2, for example. A standard evaluation script is used, that automatically writes the injury criteria values in a response-file. For the US-SINCAP load case a star rating exists. This rating scheme gives an correlation between the values of the injury criteria and injury risks for various body regions (head, chest, abdomen, pubic). With this four injury risks a total injury risk can be computed by the following function ([10]):

$$P_{total} = 1 - (1 - P_{head})(1 - P_{chest})(1 - P_{abdomen})(1 - P_{pubic}) \tag{1}$$

The total injury risk is related to the star rating for the dummy over a relative injury risk value. The relation between injury criteria values and star rating is computed in LS-OPT. Therefore the star rating is defined as a dependent quantity of the injury criteria values. An example for the correlation between star rating and two injury criteria is shown in Figure 2.
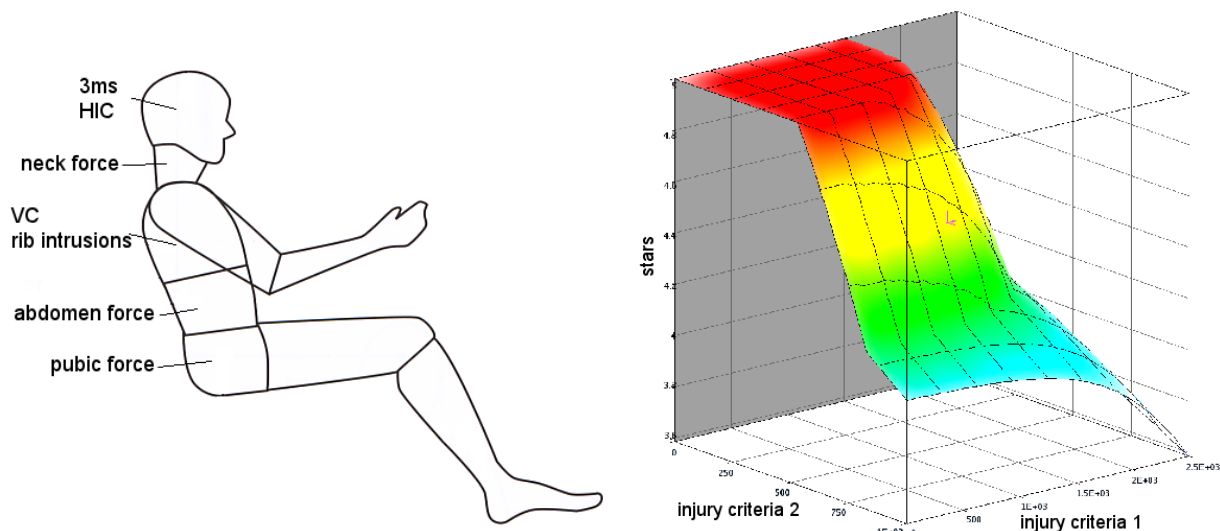

Figure 2: Injury criteria: example and correlation between injury criteria and star rating

2 Simulation process

The whole simulation process is rendered by three components: CAx-LoCo (CAx Load case Composer), the compute cluster and LS-OPT. The simulation model of a vehicle is devided into various components. All components are managed by CAx-LoCo in a component pool. First, the variables have to be defined in CAx-LoCo and the values in the components have to be changed into variables. Ranges for the upper and lower bound for each variable and each load case are defined in LS-OPT. Also a point selection scheme (experimental design method) can be selected in LS-OPT, for example Space Filling or D-Optimal. The simulation points are distibuted in the design space with respect to this point selection scheme. This distribution is also called sampling. For every simulation LS-OPT gives values of the variables to CAx-LoCo. This program changes the variables to the defined values, collects all components, copies them to the cluster, makes the assembly and starts the calculation. After the successful completion of the simulation the occupant safety values are automatically written in a response-file using the standard evaluation script and imported into LS-OPT. At the end the responses can be evaluated with LS-OPT. It establishes the relation between variables and simulation results. The whole simulation process with data flow is shown in Figure 3.
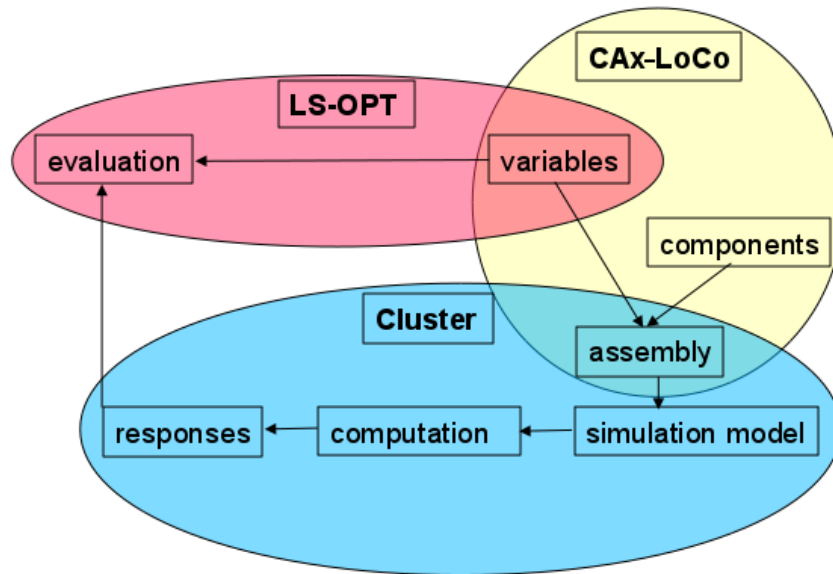
Figure 3: Simulation process

3 Sensitivity Analysis

3.1 Introduction

In a sensitivity analysis the influence of model variables with respect to specific responses is evaluated to filter variables, that have minor or no influence on the responses. This filtering happens through variation of the variables and comparison of the response values. In a sensitivity analysis usually rather large variable ranges are examined in order to reproduce design changes. The set of points is uniformly distributed on the design space (for example Space Filling or D-Optimal). LS-OPT provides several technologies for sensitivity analysis: ANOVA, Sobol's indices, correlation analysis, etc. In this project two sensitivity analysis methods, named ANOVA and Sobol's, are investigated and compared.

3.2 ANOVA - multi variate linear regression analysis

Based on a multi variate linear regression model, regression coefficients b_j and variances Δb_j for each response and variable are computed. The multi variate linear regression model for the response y_i is:

$$y_i = b_0 + \sum_{j=1}^{n} \frac{b_j}{\Delta x_j} x_j = b_0 + \frac{b_1}{\Delta x_1} x_1 + \frac{b_2}{\Delta x_2} x_2 + ... + \frac{b_{n-1}}{\Delta x_{n-1}} x_{n-1} + \frac{b_n}{\Delta x_n} x_n, \tag{2}$$

where i is the number of responses and n is the number of variables.

Δb_j gives an indication of the scatter (uncertainty) with respect to the linear regression. On details of computation of Δb_j it is refered to [2]. Using the regression coefficients a ranking can be disposed, that shows the influence of the variables on the responses. There are positive correlations, that means a larger value of the variable leads to a larger value of the response, and negative correlations, that means a larger value of the variable leads to a smaller value of the response. A positive correlation as well as the relation between theory and visualisation in LS-OPT is shown in Figure 4.

Usually ANOVA is used in conjunction with a D-Optimal DOE-sampling. An example for the visualisation of the regression coefficients and confidence intervals in LS-OPT is given in Figure 5 (see [3]).
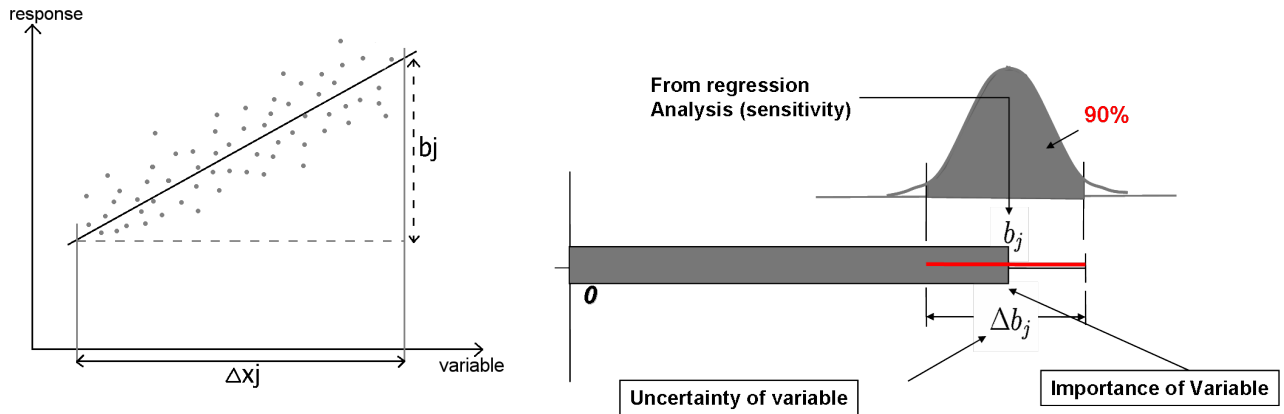
Figure 4: left: linear regression for the variable x_j with a positive correlation, right: display of sensitivities - regression coefficient b_j and confidence interval Δb_j according to the visualisation in LS-OPT (see Figure 5) [2]
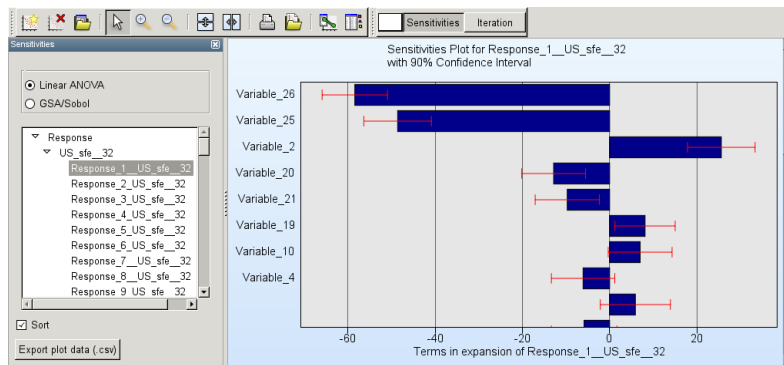

Figure 5: ANOVA: visualisation of sensitivities for various variables with respect to specific responses in LS-OPT

3.3 Sensitivity analysis according to Sobol

The Sobol's indices are a generally applicable non-linear sensitivity measure. The determination is computational expensive, because many sampling points (simulations) are required. To overcome this problem the Sobol's indices are computed based on meta models. In LS-OPT meta models can be for example Radial Basis Functions or Feedforward Neural Networks. The meta models give an approximation of the response function addicted to the variables.

Based on the evaluation of the meta models, the Sobol's indice S_i of variable v_i is computed as (see [1], [2], [4]):

$$S_i = \frac{variance\ caused\ by\ v_i}{total\ variance\ of\ response} \tag{3}$$

The sum of all Sobol's indices of one response is 1:

$$\sum_{i=1}^{n} S_i = 1, \tag{4}$$

where n is the number of variables.

In LS-OPT, the Sobol's indices are illustrated as horizontal bars, as shown in Figure 6 (see [3]). Behind the bars, in brackets, are two numbers listed. The first is the Sobol's indice of the variable and the second is the cumulative Sobol's indice. So the cumulative Sobol's indice of all variables is 1 (100 %). The Sobol's indices for more than one response is computed by building the mean of the Sobol's indices for one variable. These Sobol's indices are computed with the button "Multi" and the selection of the responses for which the Sobol's indices should be computed. The function "Transpose" shows the influence of each variable to each response in bars. The influence of one variable to all responses in the sum is 1, too.
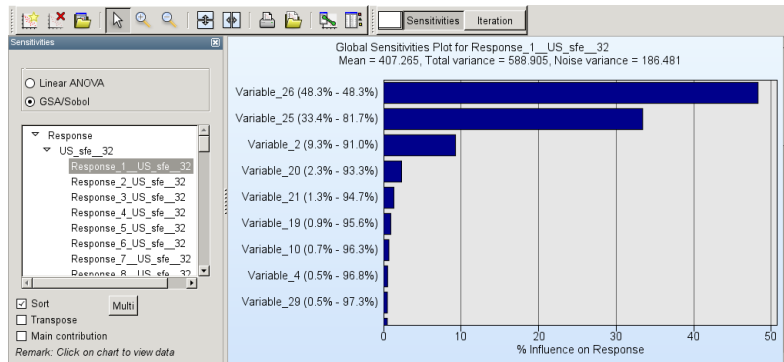

Figure 6: Sobol's indices: visualisation in LS-OPT

3.4 Evaluation of sensitivity analysis applied to the side impact loadcases

Intention of the evaluation is to find those variables, that have the greatest influence on the responses. In the following, the results of the two applied methods (ANOVA and Sobol's) of detecting these variables are quoted and compared.

3.4.1 Convergence Investigation performed for ANOVA

Objective of the convergence investigation is to evaluate how many simulation points are required for a specific number of variables. For the first sensitivity analysis, 200 simulations for every load case in 4 iterations are examined. In the first analysis 50, in the second 100, in the third 150 and in the fourth 200 simulations are considered. In every iteration the regression coefficients are computed. These four regression coefficients for each response are compared. In Figure 7 the regression coefficients of one response (Response 1) for each iteration (50, 100, 150 and 200 simulation points) is exemplary diagrammed. On the abscissa the simulation points and on the ordinate the regression coefficients are plotted. After 100 simulation points the changes of the regression coefficients are marginal. With raising number of simulation points the variance Δb_j also decrease (this fact is not shown). The result of this investigation is that for this simulation model 100 simulation points are sufficient. This convergence investigantion is just made for the results of the ANOVA method. In the following investigations 100 simulation points are considered for ANOVA.
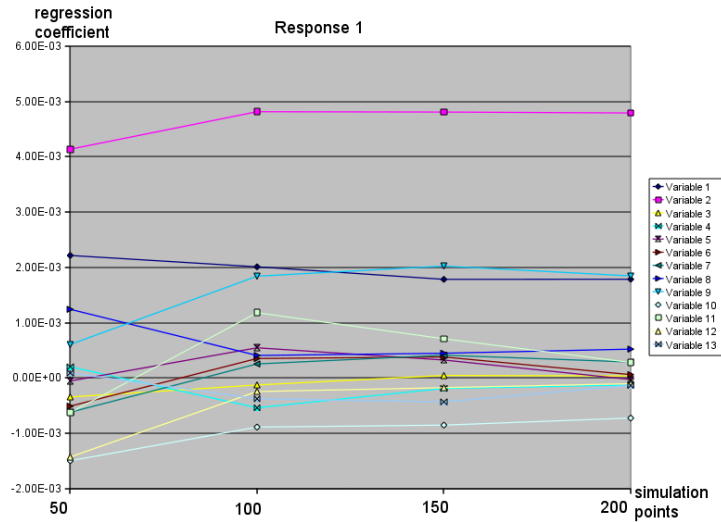
Figure 7: Convergence Investigation

3.4.2 ANOVA

The problem by evaluating the regression coefficients of the ANOVA method is that the responses, as well as the regression coefficients, have different units with values of different order of magnitude: forces in N, displacements in m, accelerations in m/s^2, etc. The regression coefficients give the influence that a variable has on each response. But for creating a ranking it is necessary to know the influence that one variable has on more than one response. To compare and summarise the results for one variable and to get the influence it has on more than one response, the regression coefficients are normalised. Through summarisation and averaging the normalised regression coefficients of various injury criteria one gets a ranking for a combination of several injury criteria.

In Figure 8 this problem is schematically illustrated for two variables and two responses. In this example the response 1 is in N and response 2 is in mm. In this case it is not obvious which variable has the greater influence on both responses. To get a combined averaged dimensionless regression coefficient for both responses, the regression coefficients are normalised. Through normalising the regression coefficients a weighting is rendered.
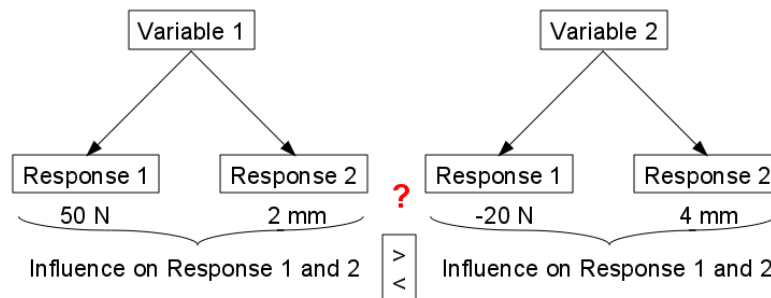

Figure 8: Getting the influence of one variable on various responses

In this way, rankings for each load case, for the two barrier and the two pole load cases, and for all four load cases can be generated. In this rankings it is observed, that the variables have different influences on the respective load cases. So the selection of the relevant variables is made in steps: First the relevant variables of the total ranking (all load cases) , then the variables of the barrier and pole load cases and at the end the relevant variables in each load case, are chosen. In Figure 9 the selection scheme is represented. A variable is defined as relevant, if it has an averaged influence greater or equal 4 percent on the injury criteria. Motivation for picking the relevant variables step by step is that there

could be a variable that has just in one load case a great influence. This one would not be selected by just looking on the ranking for all load cases. Otherwise there could be a variable that has in every load case a medium influence. This one would not be selected by just looking on the rankings for each load case.
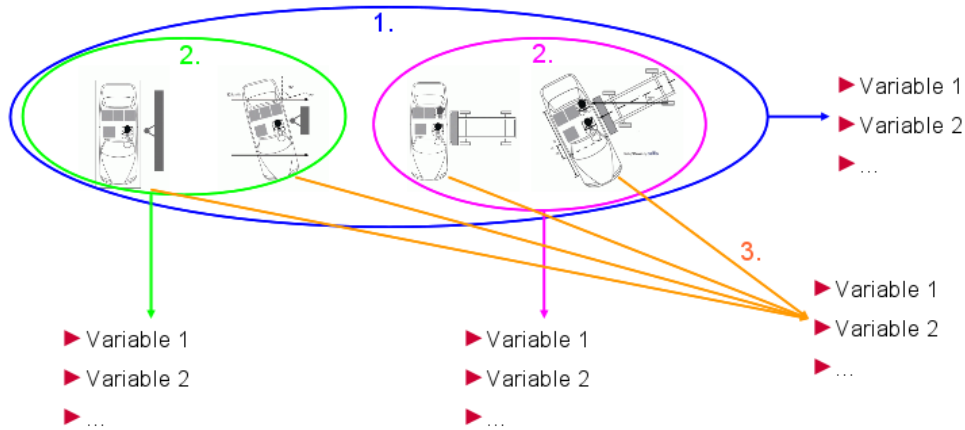

Figure 9: Selection of relevant variables

3.4.3 Comparison of ANOVA and Sobol based on a linear meta model

The sensitivity of the variables on one response is obviously the same for both methods since Sobol's method on linear meta model is similar to ANOVA with a linear regression analysis. This is shown in Figure 10. Per contra, the rankings for more than one response are different for the evaluation by ANOVA and Sobol. The differences are a result of the normalisation of the regression coefficients with the occupant limit values.
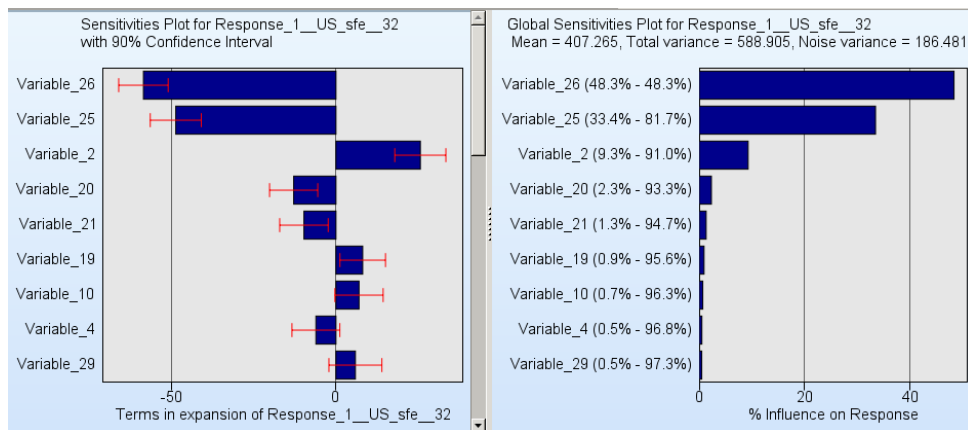

Figure 10: Comparison of the sensitivities of ANOVA and Sobol for one response

3.4.4 Sobol's method

The Sobol's indices are normalised quantities, so it is possible to summarise them for more than one response directly in LS-OPT (shown in Figure 6).
First the Sobol's indices are evaluated based on a linear meta model, because it is not possible to build a non-linear meta model for such a great variable space. Therefore, 100 simulation points per load case are considered. The relevant variables are selected by using the same scheme as shown in Figure 9. The variables that are identified as relevant from both methods (ANOVA and Sobol's) are used to make a second investigation. In this, the Sobol's indices are computed based on a non-linear meta model (Radial Basis Functions). The meta model is build based on 100 simulation points for each load case. The results of the linear and non-linear investigation are compared. The result is that the non-linear fraction has a marginal influence. One reason could be, that 100 simulation points are insufficient to build a non-linear meta model for the variable space. Furthermore, the variables for the non-linear analysis are picked by using linear sensitivity measures. Thus, these variables have a great linear and just a marginal non-linear coherence to the responses. This could be another reason for the rather similar results of the linear and non-linear analysis.

3.5 Summary of Sensitivitiy Analysis with ANOVA and Sobol

In the following, the assets and drawbacks of both methods are listed.

- ANOVA
    - + positive or negative correlation between variable and response available
    - + values b_i give the absolute value of the change
    - + normalisation: weighting of the responses
    - - regression coefficients of various responses are not compareable: normalisation
    - - just linear analysis possible
- Sobol
    - + Sobol's indices are normalised (dimensionless): directly comparable
    - + linear and non-linear analysis possible
    - - positive or negative correlation between variable and response is not available
    - - no absolute value of the change
    - - no weighting of the responses

4 Robustness investigation

4.1 Introduction

In a robustness analysis one investigates how sensitive the responses react to small changes of the variables. If small changes of the variables does not effect relevant changes in the responses, the system is called robust. The sampling of the variables is performed by the consideration of probabilistic distribution functions. Most popular is the normal (Gaussian) distribution. The distribution of the points has a random component with respect to the given density function ([2], [8]).
In the present project normal distributions are specified over the design ranges of the variables. A normal distribution is defined through the mean value μ and the standard deviation σ.

4.2 Evaluation and visualisation in LS-OPT

LS-OPT provides three possibilities to show the results of a robustness investigation, called "Histogram", "Summary" and "Bounds".
At the histogram illustration the range of the response is disposed in a defined number of intervals. The number of responses in each intervall is counted and plotted as columns. So there is one column in each interval. The values for μ and σ are displayed on a base of a normal distribution. An example of a Scatter Plot and Histogram is shown in Figure 11.
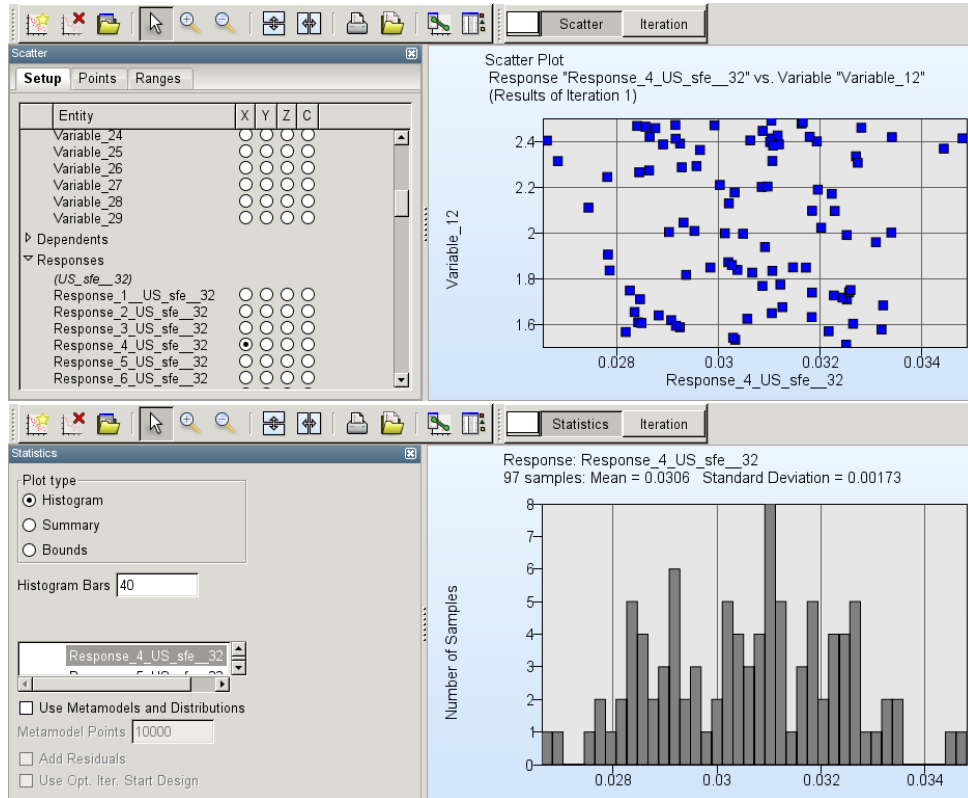

Figure 11: Coherence between Scatter Plot and Histogram

In the visualisation display "Summary", shown in Figure 12, the mean value and the standard deviation are displayed as grey bars. The 95%ile confidence interval is displayed as a thin red bar.
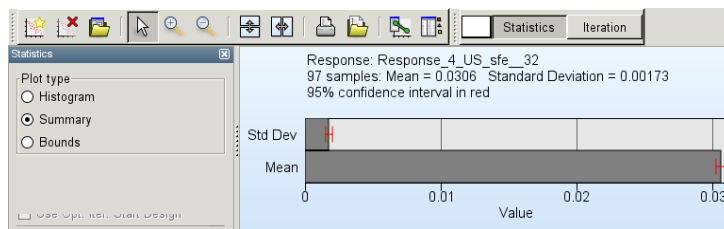

Figure 12: Robustness visualisation: Summary

In the visualisation display "Bounds", shown in 13, a lower and a upper bound have to be defined. Displayed are the fraction of values that is larger than the upper bound and smaller than the lower bound. The 95%ile confidence interval is displayed, too.
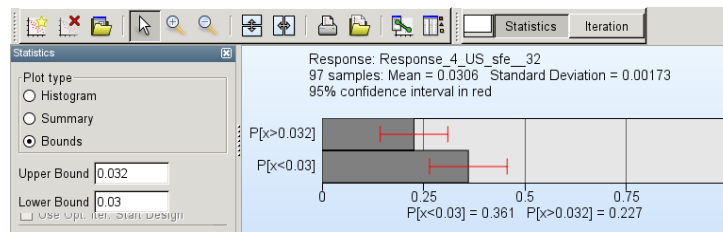

Figure 13: Robustness visualisation: Bounds

In the example shown in Figure 13, 36.1% of the responses are smaller than 0.03 and 22.7% are larger than 0.032. Thus, 41.2% of the responses range between 0.03 and 0.032.

5 Conclusions

For the sensitivity analysis performed in this study, the influence of various variables on several occupant safety responses are analysed. On this, four load cases - two pole (EuroNCAP and FMVSS 214) and two barrier load cases (US-SINCAP and IIHS) - are considered.

The variables are moving directions and positions of the barrier/pole , friction coefficients ,thicknesses of the body shell, thicknesses of the covering and mass distributions.

The sensitivitiy analysis is performed with the regression analysis ANOVA and the sensitivity analysis by Sobol. For comparing the regression coefficients a normalisation has to be applied, because they have different units with values of different order of magnitude. Three different methods of normalisation are analysed and compared.

At the beginning, a convergence investigation is examined, to get an idea about the number of simulation points, that are necessary to approximate the global behaviour in a suitable manner. Therefore, 4 analysis steps with 50, 100, 150 and 200 simulation points are examined. The regression coefficients and the variances (ANOVA) of this 4 investigations are compared. As a result it turned out that 100 simulations for each load case are sufficient. A rule for the correlation between number of variables and number of simulation points is developed. Hence, in all following simulations 100 points are considered.

The results are evaluated by linear and non-linear sensitivity measures. First, a linear analysis with all variables is examined. The results are evaluated with the regression analysis ANOVA and the Sobol's method based on a linear meta model, which leads obviously to similar results. In a second step the resulting relevant variables are selected for a non-linear sensitivity analysis according to Sobol, based on a non-linear meta model (Radial Basis Functions). Therefore, 100 simulation points are considered, too. The results of the linear and non-linear analyses are rather similar. One reason could be, that 100 simulation points are insufficient to build a non-linear meta model for the considered variables. Furthermore, the variables for the non-linear analysis are picked by using linear sensitivity measures. Thus, these variables have a great linear and just a marginal non-linear coherence to the responses. This could be another reason for the rather similar results of the linear and non-linear analysis.

The point selection and the normalisation methods show just marginal influence. The choice of the variable ranges has major influence.

A robustness investigation is rendered. In this, small changes of the variables that represent uncertainties are investigated. The distribution of the points over the design ranges of the variables has a random component with respect to the given density function. LS-OPT provides three possibilities to show the results. Examples for the illustration of this three alternatives - "Histogram", "Summary" and "Bounds" - are given in this paper.

5.1 Outlook

The robustness analysis could be considered more exactly. For example, the coherence between uncertainties and star rating could be evaluated. That would give the information how reliable the design is with respect to an expected star ranking. In the example given in Figure 14 80 percent of the simulations lead to 5 stars and 20 percent lead to 4 stars. That means, if all uncertainties are considered the probability to get 5 stars at the star rating is 80 percent.
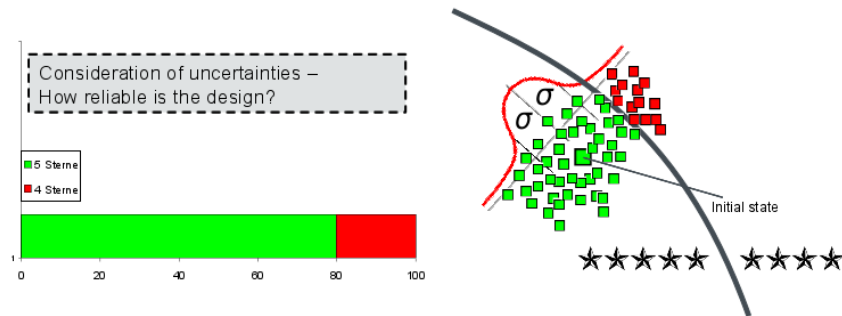

Figure 14: Example for a possible reliability of the design

A References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J.Cariboni, D. Gatelli, M. Saisana, S. Tarantola, *Global Sensitivity Analysis - The Primer*, John Wiley & Sons, Ltd, Chichester, 2008

[2] N. Stander, W. Roux, T. Goel, T. Eggleston, K. Craig, *LS-OPT User's Manual - A design optimization and probabilistic analysis tool for the engineering analyst*, Februar 2010, Version 4.1, Livermore Software Technology Corporation

[3] K. Witowski, H. Müllerschön, *Capabilities of Result Visualization in LS-OPT V4.1 - Demonstrated by Means of Industrial Problems*, 11th International LS-DYNA Users Conference, Detroit, 2010

[4] U. Reuter, M. Liebscher, *Global sensitivity analysis in view of nonlinear structural behaviour*, 7. LS-DYNA Anwenderforum, Bamberg, 2008

[5] U. Reuter, M. Liebscher, H. Müllerschön, *Global Sensitivity Analysis in Structural Optimization*, 7th European LS-DYNA Conference, Salzburg, 2009

[6] H. Müllerschön, M. Liebscher, W. Roux, N. Stander, U. Reuter, *Statistics and Non-Linear Sensitivity Analysis with LS-OPT and D-SPEX*, 10th International LS-DYNA Users Conference, 2008

[7] T. Goel, C. Maurath, *Variable Screening Using Global Sensitivity Analysis*, 11th International LS-DYNA Users Conference, Detroit, 2010

[8] W. Roux, *Robust Parameter Design In LS-OPT*, 6th European LS-DYNA Users Conference, 2008

[9] carhs GmbH, http://www.carhs.de/de/training/safetywissen

[10] carhs GmbH, Wissen für die Fahrzeugentwicklung von morgen, *Safety Companion 2010*

[11] Cukier, R.I., *Nonlinear sensitivity analysis of multiparameter model systems*, Journal of Computational Physics, 26, 1978

[12] Xu, C., G.Z. Gertner, *A general first-order global sensitivity analysis method*, Reliability Engineering & System Safety, 93, 2008

[13] Montaño, J.J., A. Palmer, *Numerical sensitivity analysis applied to feedforward neural networks*, Neural Computing & Applications, 12, 2003