THE TOOL SUSA 4 FOR PROBABILISTIC UNCERTAINTY AND SENSITIVITY ANALYSES

Martina Kloos
Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) gGmbH
Boltzmannstr. 14, 85748 Garching, Germany
e-mail: Martina.Kloos@grs.de

Keywords: SUSA, Uncertainty Analysis, Sensitivity Analysis, Monte Carlo Method.

Abstract. SUSA 4 (Software for Uncertainty and Sensitivity Analyses) is a powerful tool for uncertainty and sensitivity analyses of computational results and an important part of the GRS code system for safety analyses of nuclear power plants (NPPs). It provides support to quantify input uncertainties in terms of probability distributions, correlations and other appropriate dependence structures. For Monte Carlo simulation, the simple random and the Latin Hypercube sampling procedure are available. To prepare and launch computer code runs, a selection of code interfaces are implemented. Rather comfortable are the interfaces to selected codes in the field of nuclear safety analyses. But the application of SUSA 4 is not restricted to those codes. Many options exist for quantifying the uncertainty of a computational result. Well-known and increasingly applied in the frame of deterministic nuclear safety analyses are the tolerance limits of Wilks. Options for performing a sensitivity analysis are implemented as well. Various sensitivity indices may help to identify those input uncertainties which mostly contribute to the uncertainty of a computational result. SUSA 4 combines well established methods from probability calculus and statistics with a comfortable graphical user interface (GUI). The GUI guides the user through the main analysis steps and essentially contributes to comprehensibility and error prevention. The paper gives an overview on the main features and capacity of SUSA 4.
1 INTRODUCTION

In deterministic safety analyses for NPPs, it has become common practice to apply best estimate computer codes for calculating postulated accidents in a realistic and not in a conservative way. If a best-estimate code is used in combination with realistic input data, it is generally required that important epistemic (lack of knowledge) uncertainties which may affect the computational result shall be considered and that their influence on the result shall be quantified [1].

There are several sources of epistemic uncertainty which may influence a computational result. Important sources are the model formulations implemented in a computer code. They are mostly based on a scattering of measurements or may be rather simplified so that the level of predictive accuracy of a model – even a validated model – may not be precisely known. Further uncertainty sources are numerical solution algorithms which commonly include approximations and simplifications somehow affecting the result, or the initial and boundary conditions of the application case which are often not exactly known.

Also in probabilistic safety analyses for NPPs, uncertainties shall be considered and their influence on the result (e.g. core damage frequency, large early release frequency) shall be quantified. Relevant uncertainty sources in this context are the occurrence frequencies of initiating events, the reliability parameters (failure probabilities and rates) of systems and components, human error probabilities or the branching probabilities of a PSA level 2 event tree.

Intuitively the most appropriate method to account for the uncertainty sources of a computational result and to get a quantification of their influence is the Monte Carlo (MC) Method. It considers a range of values instead just one value for each input parameter of the computer code subjected to uncertainty. Each value selected for an uncertain input parameter is combined with a value selected for each other uncertain parameter and supplied as input to corresponding computer code runs. Based on the sample of values finally provided for a computational result, a quantification of the uncertainty of the result is obtained by applying statistical methods. Further information is given, for instance, in [2, 3, 4, 5].

To identify the main uncertainty sources of a computational result, an additional (global) sensitivity analysis is useful [6, 7]. It can show where to improve the state of knowledge in order to reduce the (epistemic) uncertainty of the computational result most effectively.

To facilitate the performance of uncertainty and sensitivity analyses based on the MC method, the tool SUSA was developed. First version of SUSA was available in the early 1990s [8]. Since then, SUSA has been constantly improved mainly according to the needs in the field of reactor safety analyses [9, 10, 11]. The currently available version is SUSA 4 [12].

SUSA 4 combines well established methods from probability calculus and statistics with a comfortable graphical user interface (GUI). The concept of SUSA 4 enables the user to fully concentrate on the analysis input including the identification of those input parameters of the applied computer code which represent the main uncertainty sources of the computational result and the formulation of the corresponding uncertainties. After this is done, SUSA provides support to quantify the uncertainties probabilistically and to perform the next steps of an uncertainty and sensitivity analysis.

This paper gives an overview on SUSA 4. Section 2 includes a description of the GUI and other software features of SUSA 4. Subject of Section 3 are the options available to quantify and document input uncertainties. The methods implemented to generate a sample of parameter values are outlined in Section 4. An overview on how SUSA 4 provides support to perform computer code runs is given in Section 5. Subject of Section 6 are the options implemented to quantify the uncertainty of a computational result. Section 7 deals with the options available for sensitivity analysis. The graphical techniques implemented in SUSA 4 to
support the uncertainty and sensitivity analysis are addressed in Section 8. The conclusions can be found in Section 9.

2 GUI AND OTHER SOFTWARE FEATURES

SUSA 4 has a comfortable graphical user interface (GUI) written in Visual Basic .NET (vb.net, Microsoft®). The GUI is menu-driven and guides the user through the analysis. The menu bar is displayed on top of the main window of SUSA 4 and consists of nine menus (Fig. 1). The first menu to select is the Project menu where a new project has to be defined or an existing project has to be selected. Selection of the other menus depends on the analysis status already achieved in the course of a project. Selection of a menu usually starts on the left and goes step-by-step further to the right.

![Figure 1: Main SUSA 4 window with focus on menu ‘Input Uncertainties’](image)

The menu ‘Input Uncertainties’ is to be selected in order to enter the uncertain input parameters of a project and to quantify the corresponding uncertainties probabilistically. In the menu ‘Sample Generation’, different sets of parameter values can be sampled. The menu ‘Computer Code Runs’ is designated to start a computer code run for each set of parameter values and to prepare the corresponding computational results in the format required in the subsequent analysis steps of SUSA 4. Menu ‘Uncertainty Analysis’ includes options for quantifying the uncertainty of a computational result. Sensitivity indices which help to identify the main uncertainty sources of a computational result are offered in the menu ‘Sensitivity Analysis’. The menus ‘Scatter Plot’ and ‘Cobweb’ provide additional graphical output supporting the uncertainty and sensitivity analysis. The Help menu currently includes the user’s guide document of SUSA 4, a context related online help function will be implemented in the next version.

Each menu of SUSA 4 consists of several items or submenus for specific tasks to be performed within the analysis step represented by the menu. For instance, the menu ‘Input Uncertainties’ includes the five items Documentation, Distribution, Dependency, Proportions and External Data (Fig. 1). Selection of the Documentation item is necessary in order to specify and document the uncertain parameters of a project (Fig. 2). The Distribution item should
be chosen to quantify the uncertainties of the documented parameters in terms of probability distributions (Fig. 3). Selection of the Dependency item is important, if dependencies between uncertain parameters are to be quantified. Quantifications can be made in terms of association measures, conditional distributions, inequalities and other functions appropriate for dependency modelling. The item ‘Proportions’ offers the possibility to categorize selected uncertain parameters as proportions of the same whole which have to sum up to 1. Selection of the last item ‘External Data’ allows for assigning values provided by external sources to uncertain parameters.

Figure 2: Dialog ‘Documentation’

The calculations of SUSA 4 are performed by Fortran program modules. For graphics generation, SUSA 4 applies the Java plotting tool AptPlot which is a free plotting tool designed for creating production quality plots of numerical data. AptPlot contains GUI support for the manipulation and analysis of data sets (cf. Fig. 4).

SUSA 4 is available for 32 bit and 64 bit computer systems.

3 QUANTIFICATION OF INPUT UNCERTAINTIES

SUSA 4 expects the uncertainty of an input parameter of a computer code to be primarily quantified as a probability distribution (section 3.1). If the value of a parameter is – to some extent or completely – dependent on the value of another parameter, an appropriate association measure, function, conditional distribution or inequality can be specified (section 3.2). Further options available for specifying input uncertainties are the characterization of uncertain parameters as proportions of the same whole which have to sum up to 1, or the assignment of values from external sources to uncertain parameters.

There is no limitation of the number of input uncertainties which can be considered within SUSA 4.

3.1 Probability Distributions

The probability distributions listed below are currently implemented in SUSA 4.

Parametric probability distributions:
– Uniform Distribution
- Log. Uniform Distribution
- Triangular Distribution
- Log. Triangular Distribution
- Beta Distribution
- Normal Distribution
- Log. Normal Distribution
- Weibull Distribution
- Gamma Distribution
- Exponential Distribution
- $X^2$ Distribution
- F Distribution
- Extreme Value Distribution Type I (Gumbel)
- Extreme Value Distribution Type II (Frechet)

Non-parametric probability distributions:
- Discrete Distribution
- Histogram
- Log. Histogram
- Polygonal Line

Figure 3: Dialog ‘Distribution’
If the distribution to be quantified is parametric, SUSA offers a selection of distributions which might be appropriate dependent on what is known on the range and other characteristics of the distribution such as quantiles, expectation and std. deviation, or median and k95-factor. If the parameters of a probability distribution are not known, SUSA calculates them from what is known (Fig. 3). This is done either analytically or by applying a random search algorithm. Distributions which normally have an infinite minimum and/or maximum can be truncated at the values indicated as minimum and maximum (Fig. 4). The distribution assigned to an uncertain parameter can be plotted separately or compared with other distributions in a single diagram.

Figure 4: Truncated Normal distribution shown in the AptPlot GUI window

Figure 5: Table of specified uncertain parameters and probability distributions
All uncertain parameters and corresponding probability distributions entered within SUSA 4 are displayed in a clear table (Fig. 5) which can be printed or easily copied into a paper or report.

3.2 Dependencies

Information on (epistemic) dependencies between uncertain parameters can be provided in terms of association measures, full dependencies, conditional distributions, functions or inequalities (Fig. 6). Association measures can refer either to the population or to the sample of potential values. In the latter case, the sample to be generated by SUSA 4 (section 4.1) will strictly reproduce the degree of specified association.

Following population measures of association are implemented in SUSA 4:
- Pearson’s ordinary correlation coefficient
- Blomqvist’s medial correlation coefficient
- Kendall’s rank (tau) correlation coefficient
- Spearman’s rank correlation coefficient

As sample (empirical) measure of association, Spearman’s sample rank correlation coefficient is implemented.

Besides the population and sample related association measures, following other options are available to quantify dependencies between uncertain parameters:
- Full (positive or negative) dependence, if the potential value of a parameter is assumed to satisfy an unknown monotone relationship to the potential value of another parameter.
- Conditional distribution, if the distribution to be quantified for an uncertain parameter depends on the potential value assumed for another parameter.
- Function, if the potential value of a parameter is a function of the values of other parameters.
- Inequality, if the dependence relationship between two parameters X and Y is due to the inequality $X > a \cdot Y$ with $a$ being a multiplication factor.

If dependencies are not provided, SUSA 4 assumes independencies between the uncertain parameters.

[Figure 6: Dialog ‘Dependence’]

Dependence relationships such as association measures, full dependencies or inequalities can be represented in a scatter plot (Fig. 7). Similar to the specified probability distributions
Martina Kloos
(Fig. 5), the specified dependencies are provided in a clear table which can be printed or easily copied into a paper or report.

![Figure 7: Scatter plot](image)

**4 SAMPLE GENERATION**

To sample the values of uncertain parameters, two procedures are currently implemented in SUSA 4, the simple random sampling (SRS) and the Latin Hypercube sampling (LHS) procedure [6, 13, 14].

The SRS procedure provides values randomly sampled from a multivariate probability distribution defined by the distributions and dependencies specified as input. The implemented sampling algorithm initially considers standard normal distributions instead of the specified distributions. All types of correlations entered as input are handled by applying the ordinary Pearson correlation. Transformation of the actual specifications into corresponding ordinary correlations of standard normals is done either analytically or by applying the bisection method. The values obtained for the standard normals are finally transformed into the values of the specified distributions.

The LHS procedure provides values covering the range of each uncertain parameter evenly in probability. To this purpose, the range of each parameter is divided into equiprobable intervals and, then, one value is selected from each interval. This value is either the median or a randomly selected value of the interval. The values selected for each parameter are finally permuted in order to comply with specified correlations between parameters. More information on the sampling algorithm can be found in [15, 16].

For the sample finally generated, SUSA 4 can provide the sample-related correlation coefficients between the uncertain parameters. This information is useful to detect unintentional (spurious) correlations due to numerical effects which may affect the results of the uncertainty and sensitivity analysis.
The random values to be generated by SUSA 4 can be controlled by the initial seed value of the (pseudo) random number generator. This value is required as input to the sample generation program.

There are no limitations of the number of uncertain parameters which can be considered and the number of values which can be generated for each parameter. Limitations are set only by available computer memory.

5 COMPUTER CODE RUNS

SUSA 4 can automatically start a computer code run for each set of parameter values sampled inside or outside of SUSA 4. If the computer code is run on the basis of an input file, SUSA 4 is able to automatically generate the input files related to the different sets of parameter values and to start the corresponding runs. The computational result to be analysed can be selected via a key-file or the address of the result in the corresponding output file. SUSA 4 automatically transfers the values of all selected results from all runs to a file adequately formatted in order to be accessible by the subsequent analysis steps.

A special feature of SUSA 4 is that it can automatically prepare a Fortran code template including appropriate instructions for considering the input uncertainties specified in SUSA 4. This template can be extended by Fortran instructions for any desired model. For compiling the completed Fortran code, SUSA 4 loads the free GNU Fortran Compiler (Gfortran) which is automatically installed with SUSA 4.

In principal, any computer code can be coupled with SUSA 4. Comfortable interfaces are implemented so far for codes frequently used in GRS for safety analyses of NPPs.

6 UNCERTAINTY ANALYSIS

SUSA 4 provides various options to quantify the uncertainty of a computational result.

![Figure 8: Set of curves expressing the uncertainty of a time-dependent result](image)

The uncertainty analysis can be performed for scalar and time/index-dependent computational results. A scalar result is just a single value per computer code run, whereas a
time/index-dependent result is a series of values over time, space, etc. Figure 8 exemplarily shows the curves of a time-dependent result from 1000 computer code runs.

For quantifying the uncertainty of a computational result, SUSA 4 can calculate estimators of key statistics such as mean, standard deviation and coefficient of variation. If the uncertainty is to be quantified in terms of quantiles, SUSA 4 can calculate empirical quantiles and the tolerance limits of Wilks [17] (Fig. 9, Fig. 10). These ($\beta \cdot 100 \% ; \gamma \cdot 100 \%$) tolerance limits are increasingly applied in the frame of a deterministic nuclear safety analysis using $\beta = 0.95$ and $\gamma = 0.95$ [11]. The two- or one-sided limits of a bounded or half-bounded ($\beta \cdot 100 \% ; \gamma \cdot 100 \%$) tolerance interval are estimates of the left and/or right endpoint of an interval covering a proportion of at least $\beta \cdot 100 \%$ of the potential values of the computational result at a confidence level of at least $\gamma \cdot 100 \%$. The information on the confidence level is quite useful, if an estimate (e.g., of the interval covering a $\beta \cdot 100 \%$-proportion) can be derived from only a small sample of values as it is the case with long running computer codes which do not allow for performing many runs. The calculation of tolerance limits is independent of the number of uncertain parameters. It just requires a minimum number of runs to be performed. For instance, at least 59 runs must be performed to get a one-sided ($95 \% ; 95 \%$) tolerance limit [18].

If the uncertainty of a computational result is to be quantified in terms of a probability distribution, SUSA 4 helps to find an appropriate distribution. It can perform Lilliefors- and Kolmogorov-Smirnoff tests which inform on the goodness-of-fit of a selected parametric distribution to the (empirical) distribution of a code result [19].

SUSA 4 can even build a regression model which is linear in the uncertain parameters and which may serve as a so-called response surface. Such a response surface may be of interest, if many runs of a computer code are to be performed and just a single run is found to be very time consuming. SUSA 4 can perform Monte Carlo Simulation based on the simplified model and compare the (empirical) distribution of the results with the distribution of the results of the original model. If the simplified model provides acceptable results, it can be used instead of the original model in order to provide a large data basis for quantifying the uncertainty of the computational result.

![Figure 9: Empirical distribution function and two-sided (95 %; 95 %) tolerance limits of a scalar result](image-url)
SENSITIVITY ANALYSIS

An extra sensitivity analysis or, more precisely, an uncertainty importance analysis can be supplemented to an uncertainty analysis to identify those uncertain input parameters which mostly contribute to the uncertainty of the computational result. It can show where to improve the state to knowledge in order to reduce the uncertainty of the computational result most effectively.

The sensitivity analysis with SUSA 4 can be performed with the same sample data as already generated for the uncertainty analysis. Like the uncertainty analysis, it can be performed for scalar as well as time/index-dependent computational results.

Following 4 groups of correlation related sensitivity measures are implemented in SUSA 4:
- Pearson's ordinary correlation
- Blomquist's medial correlation
- Kendall's rank correlation
- Spearman's rank correlation

Within each group, SUSA 4 can calculate the (sample) ordinary and partial correlation coefficients as well as the standardized regression coefficient. The (sample) coefficient of determination ($R^2$) is additionally provided to inform on the usefulness of the calculated coefficients as sensitivity indices. $R^2$ is the fraction of the variability of the computational result explained by the combined influence of the uncertain input parameters [6, 20]. In this context, combined influence means a multiple linear regression function of the parameters.

Besides correlation related sensitivity measures, SUSA 4 can provide estimates of the classical correlation ratio from original and rank transformed data [21]. The square of the correlation ratio is equivalent to the variance based first order sensitivity index [22, 23, 24]. For scalar results, SUSA 4 can additionally provide association measures from 2x2 contingency tables (Goodman and Kruskal’s $\gamma$-coefficient [25]) or regression coefficients derived from a stepwise (rank) regression.

All results of the sensitivity analysis can be graphically represented. Examples are given in Figure 11 and Figure 12.
SCATTER & COBWEB PLOTS

Besides the diagrams shown in the previous sections, SUSA 4 can produce scatter and cobweb plots to visualize dependencies and sensitivities. A scatter plot is a XY-plot of a selected pair of parameters and/or code results. Figure 7 shows an exemplary scatter plot of two uncertain parameters. In the same way, scatter plots between two computational results or between an uncertain parameter and a computational result can be represented (Fig. 13).
A cobweb plot allows for examining the local sensitivity of a computational result. For instance, it can show how high (or low) potential values of a computational result are linked to the values of the uncertain input parameters. A cobweb plot may also help to detect sensitivities with respect to interactions between uncertain input parameters.

An exemplary cobweb plot is shown in Figure 14. Each dotted vertical line in this figure represents an uncertain input parameter except the left-most one which represents the computational result. Each polygonal curve links each of the 10 largest values of a computational result (ranks 91-100) to the corresponding set of parameter values represented by their ranks.
9 CONCLUSIONS

The tool SUSA 4 essentially facilitates the performance of probabilistic uncertainty and sensitivity analyses based on MC method. The comfortable menu-driven GUI of SUSA 4 guides through the main analysis steps and contributes to comprehensibility and error prevention and, thus, to the quality assurance of an uncertainty and sensitivity analysis.

A wide range of probability distributions and association measures are available for quantifying input uncertainties probabilistically. Besides association measures, SUSA 4 even implements techniques to account for conditional distributions, inequalities and other functions for modelling dependencies between input uncertainties. The simple random and the Latin Hypercube sampling procedure are applicable for selecting sets of parameter values fulfilling the probability distributions and dependencies specified as input. SUSA 4 can automatically transfer these sets of values to the input decks of a computer code and start the corresponding runs.

For quantifying the uncertainty of a computational result, SUSA 4 can calculate estimators of key statistics such as mean, standard deviation, median or other quantiles. Wilks’ tolerance limits which are increasingly applied in the frame of deterministic nuclear safety analyses are implemented as well. If the uncertainty of the result is to be quantified in terms of a probability distribution, SUSA 4 helps to find an appropriate distribution by applying statistical goodness-of-fit tests. SUSA 4 can even build a regression model which may serve as a so-called response surface.

The sensitivity analysis with SUSA 4 can be performed with the same sample data as generated for the uncertainty analysis. Different groups of correlation related sensitivity indices are implemented. Within each group, SUSA 4 can calculate the ordinary and partial correlation coefficient as well as the standardized regression coefficient. SUSA 4 can also provide an estimate of the classical correlation ratio which is equivalent to the square root of the variance based first order sensitivity index (Sobol sensitivity index). Some other coefficients which are useful as sensitivity indices are implemented as well (e.g. Goodman-Kruskal association coefficient).

All results of the uncertainty and sensitivity analysis can be graphically represented. Additional scatter and cobweb plots help to understand existing relationships and sensitivities.

The calculations of SUSA 4 are performed by Fortran program modules. Graphics generation is done by the free Java plotting tool AptPlot. Besides graphics, SUSA 4 provides tables of the input specifications and data files of analysis results which can be easily inserted into reports and papers.

SUSA 4 does not have any limitations of the numbers of uncertain input parameters and computational results and is available for 32 bit and 64 bit systems. As an important part of the GRS code system for safety analyses of NPPs, it is constantly improved and validated. Of course, SUSA 4 can also be applied in other fields outside of nuclear safety analyses.

REFERENCES


