2009

Sensitivity Analysis Framework for Large and Complex Simulation Models

Ghaith Rabadi
Old Dominion University

Shannon Bowling
Old Dominion University

Charles Keating
Old Dominion University

Resit Unal
Old Dominion University

Original Publication Citation
Rabadi, G., Bowling, S., Keating, C. & Unal, R. (2009) Sensitivity analysis framework for large and complex simulation models. *Proceedings of the 2009 Summer Computer Simulation Conference* (pp. 291-298). Simulation Councils, Inc.

Sensitivity Analysis Framework for Large and Complex Simulation Models

Ghaith Rabadi (grabadi@odu.edu), Shannon Bowling (sbowling@odu.edu), Charles Keating (ckeating@odu.edu), and Resit Unal (runal@odu.edu)

Engineering Management and Systems Engineering Department
Old Dominion University
241 Kaufman Hall, Norfolk, Virginia, U.S.A

Keywords: Simulation, Sensitivity Analysis Framework, Design of Experiments

Abstract

In this paper, a framework for conducting Sensitivity Analysis (SA) on large and complex simulation models is introduced. The framework consists of components that are designed to make the SA a systematic process that is easy to manage and follow by simulation analysts and practitioners. Unlike local SA (one-variable-at-a-time SA), the method presented here is variance-based and it is rooted in the field of Design of Experiments (DoE) where Input Variables are varied and Output Variables are measured. Based on the DoE results, a risk scoring system is developed to identify the sensitivity of the Input Variables, and as a result classify them into High, Medium, and Low risk variables. As such, decision makers can be aware of the most sensitive high-risk input variables in a simulation model to ensure they understand the value of data reliability in their model inputs.

1. INTRODUCTION

Sensitivity Analysis (SA) is the analysis of variability in input variables and their impact on the outputs of a certain system. Specifically, analysts and decision makers are typically interested in understanding how much output variation is produced by varying the inputs (or parameters) of a system. In this paper we focus on SA used with large and complex computer simulation models.

SA can be performed for various reasons including enhancing the validity of a model, deciding on the importance of certain inputs, and minimizing the risk rooted in models' inputs. By understanding the sensitivity of the outputs to the uncertainty in the inputs, one can pay special attention to input variables that are more sensitive than others. When highly sensitive input variables are identified, we can then closely analyze the data sources for these variables to, as much as possible, ensure that such sources are well founded and reliable. Ultimately, one can have groups of variables categorized as High Risk, Medium Risk, and Low Risk for decision makers and analysts to be aware of before making decisions.

When dealing with large simulation models with tens or hundreds of input variables and parameters, it becomes very difficult to understand how they affect the output metrics especially with the existence of input variables interactions. It will also be difficult to identify the most sensitive input variables. Therefore, a SA process needs to be designed to make it possible for analysts and decision makers to know the consequences of uncertainty in the model's inputs. In this paper we introduce a SA process-based framework to analyze input variables in large simulation models in order to understand the risk associated with such variables, and consequently, the analysts can pay close attention to variables with high risk. The proposed framework is generic enough to be applied to various types of simulation models.

The literature is rich with SA applications in various fields; however, the vast majority is in the form local sensitivity analysis (i.e., one-variable-at-a-time SA). This type of SA may result in misleading conclusions especially when it comes to large simulation models because of the likelihood of having interactions among different model inputs. Nevertheless, there does exist research in the literature that presents sophisticated SA methods for large simulation models (e.g., [Saltelli 1993], [Saltelli et al. 1999], [Chen et al. 2005], [Kleijnen 2005] and [Campolongo et al. 2007]). In this paper, we present a step by step framework for simulation analysts and practitioners to follow when conducting SA for large simulation models. The method used is variance-based, which means that variability in the inputs is induced to measure how much variance they cause in the outputs. The contribution of this paper lies in the process followed and its management, rather than introducing a brand new SA methodology. The framework is derived from the authors' experience in conducting real-life SA on large simulation models.

2. SENSITIVITY ANALYSIS FRAMEWORK FOR LARGE SIMULATION MODELS

The proposed framework takes the form of a process that can be followed to simplify the SA especially when there are many input and output variables in the model. This framework was developed for a simulation model that included several hundred input and output variables and was generalized to be used with similar large models. The framework is depicted in Figure 1 and includes the following processes and components:

(1) Input Variables Analysis,

(2) Output Variables Analysis,

(3) Experimental Design,

(4) Model(s) Execution, and

(5) Statistical Analysis

(6) Risk Scoring

2.1 Input Variables Analysis

One of the most challenging aspects of conducting SA is having large number of input and output variables. To conduct a manageable and meaningful SA, the number of input variables needs to be reduced. Reducing the number of variables may not be necessary for models that have a small number of variables even if these models were large. Typically, however, large models have large number of user input variables in addition to intermediate variables within the model that may have significant impact on the model's performance. Such intermediate variables need also to be considered in the pool of input variables to be analyzed.

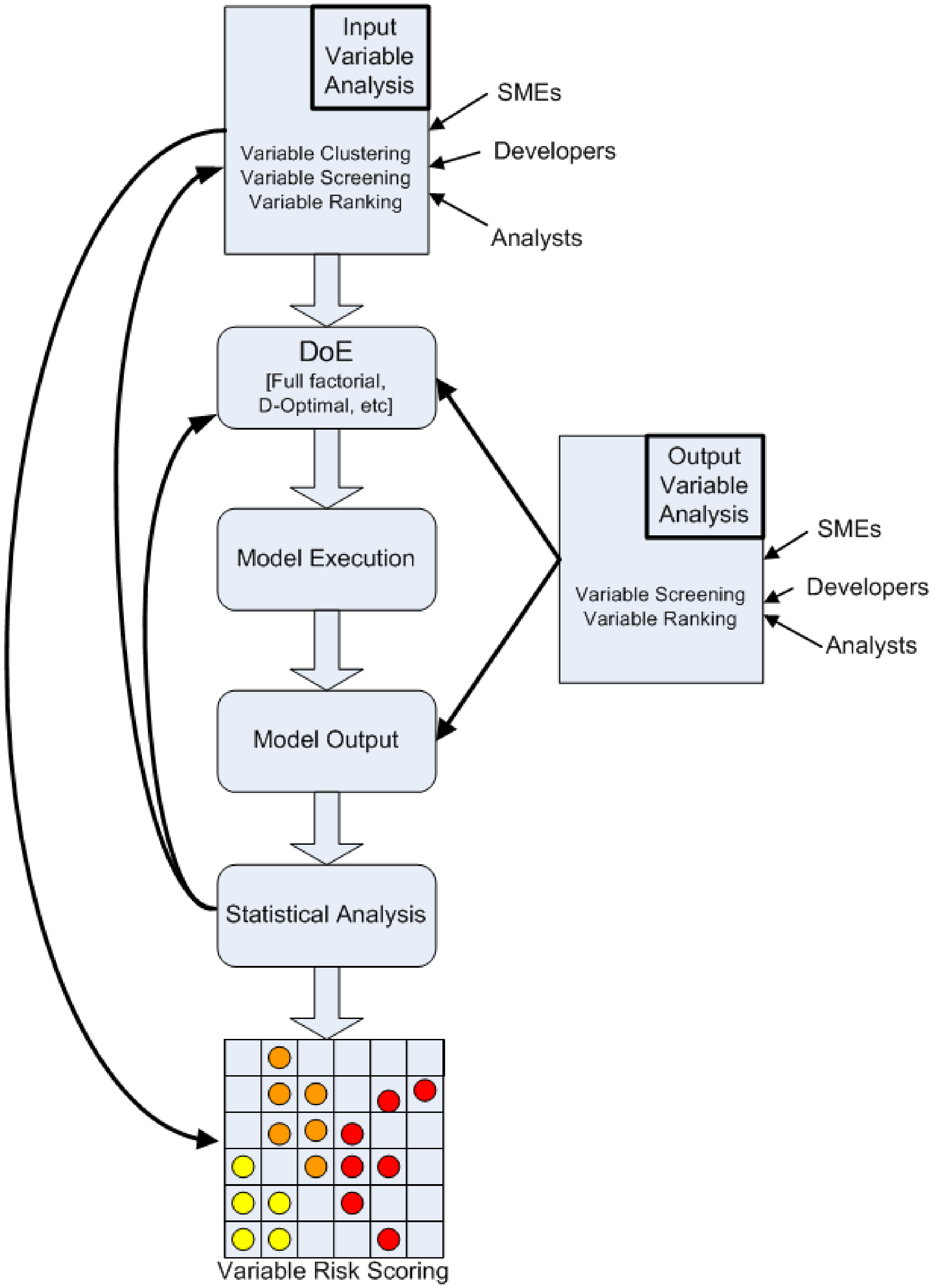

Figure 1. Sensitivity Analysis Framework

2.1.1. Variable Screening and Clustering

In our SA framework, input variables undergo a *screening* then a *clustering* process to make their size manageable. In the screening process, variables that are likely to remain constant or have little potential for variability are fixed to their most likely values and are removed from the pool. The remaining input variables are then clustered by categorizing them according to logical and programmatic criteria as defined by Subject Matter Experts (SMEs), developers, and analysts. The number of clusters must be limited to a relatively small number (10 or less). Within each cluster, the individual variables to vary will have to undergo a screening process by SMEs and analysts to decide on the minimum number of variables to be varied. For example, one criterion to limit the number of variables could be whether a variable can vary in reality by a wide range or not. Another could be the level of confidence in the data source for a variable. That is, if the data source is robust enough, then such a variable can be precluded.

Input variables are then grouped into clusters and within each cluster three levels for each changing variables are defined: *low*, *baseline* and *high*. The baseline variables in the input clusters take the values that the simulation model has for these variables as was determined by the SMEs and data analysts. The low and high values of these variables are respectively 30% below and above the baseline values to represent about ± one standard deviation. In some cases, the low and high levels may be selected based on expert recommendation. For example, in some cases a variable may become negative with 30% reduction in its value. If such a variable needs to remain nonnegative, then the lower level must be set manually. Another example of setting the low and high levels manually is if a 30% increase or decrease is unrealistic for a given variable. Therefore it is necessary to manually go through all variables to make sure that the one standard deviation rule is meaningful to apply. Note that if the number of input variables is small, the clustering step can be skipped and the top ranking variables can be screened directly.

2.2. Output Variables Analysis

This step is necessary when a model produces large number of output variables. In case of a small number of output variables, all of them can be taken into consideration; otherwise, the number of output variables must be reduced. The first step in identifying output variables is to probe the clients and stakeholders for a small set of metrics or output variables (Y) in which they have great interest (e.g., $Y \leq 10$ variables) along with the priorities associated with these metrics. If there is an absolute need to include larger number of output variables (for example the stakeholders do not have a clear way of selecting the top Y variables), some analysis will be necessary to identify the top Y output variables. The approach followed here is based on an initial set of simulation runs where output variables that could potentially vary the most are selected. The rationale here is that an output variable that does not change much when the inputs are changed by ±30% will most likely be insensitive and is not interesting to track from sensitivity analysis perspective. That of course does not necessarily mean that such a variable is not important; they are simply less sensitive to change in the input variations.

The subset of output variables selected is then subjected to correlation analysis to determine if some of these variables are highly correlated with each other. The output variables were further reduced by removing one of the highly correlated variables and keeping the other. The rationale here is if two variables are strongly correlated, then we can predict one from the other and there is no need to consider both.

2.3 Experimental Design

Design of Experiment (DoE) is a structured approach to varying input variables, observe output variables and extract the most information (and knowledge) possible depending on the design applied and the number of runs. Full Factorial design is the most comprehensive DoE approach as it runs all combinations and produces the most amount of information. It is, however, the most time consuming as it entails a large number of runs. If the number of input variables and their levels to vary is small, and also the model is not computationally demanding to run, then the number of runs will be manageable and thus a Full Factorial design could be appropriate. This is not the case, however, in most large models and therefore partial factorial design will be more reasonable. *D-Optimal* design is a good candidate design as it minimizes the variance of the model coefficient estimates and is suitable for screening situations where the objective is to find which factors are significant along with their parameter estimates [Montgomery 2000]. In addition, it reduces the number of experiment runs significantly.

If we have enough confidence to assume that there is no interaction among input variables, one-at-a-time Sensitivity Analysis can be conducted to simplify the process. To have no interaction in large simulations is typically unusual.

2.4 Model Execution

Once a DoE has been determined, the model will then be run accordingly, and the output values of interest (based on Component in 2.2 above) will be collected, and formatted for analysis. Also, depending on the complexity of the model, the amount of data, and the number of runs, it will most likely be necessary to automate the execution and output collection process. This may have tremendous impact on the effectiveness of the approach as well as the accuracy of the results especially that manual execution and data collection are susceptible to human errors. Another advantage of an early investment in automation is that if it turns out there is a need to add or remove any of the input/output variables, or for whatever reason reruns are necessary, it will be much easier and less time consuming to rerun the experiments. It is very common for analysts to make changes to the analysis, or for the client to modify the requirements, and so, automation is highly recommended for large models with large number of runs.

In most simulations, a warm-up period of time is necessary to bring the model up to a steady state phase. In the preliminary runs of the model, analysts must observe when output variables reach a steady-state phase and eliminate the output data prior to that warm up cut-off point.

2.5 Statistical Analysis

The objective of the DoE is to measure the statistical variability in the output when the selected input variables are varied. If the change in the output when varying the inputs is statistically significant, then the input variables and/or their interactions that account for most of the variability will be considered sensitive, and therefore, will be selected for Risk Scoring (Component 2.6). ANOVA statistical tests can be used to measure the significance depending on the application, data, and objective of the experiments. Also different measures, such as the Range and Variance, can be used to measure the variability in the outputs.

The first step in the analysis is to study the fitness of the regression model for each output variable, usually by inspecting the R-Squared values obtained. High R-Squared values mean that the regression model explains most of the variability in output variables indicating that input variables were selected meaningfully as they influence the output variables. The second step is to study and identify the significant input factors (or clusters) and/or their statistically significant 2^{nd} order interactions. This can be determined based on the *p*-values of the parameter estimates for each output variable under consideration for each time period of the analysis. At 95% confidence level, a *p*-value of 0.05 or less for the main effects of the input clusters or their 2^{nd} order interactions indicates statistically significant impact on the output variability.

The significance of the main effects is considered only when they had no significant 2^{nd} order interactions. An interaction between two input factors means that the value of the output variable is influenced by the setting of both factors simultaneously, and making conclusions about the main effects independently could be meaningless and possibly misleading. Interactions higher than the 2^{nd} order are typically not considered as in many cases they are considered as part of the error and they are difficult to interpret.

Depending on the results of the statistical analysis, the variable clustering, screening and ranking may have to be modified or a different DoE may have to be selected and the components in 2.3, 2.4, and 2.5 must be repeated. Hence, the feedback arrows in Figure 1 feeding back into the Input Variable Analysis and DoE processes were deemed necessary.

2.6 Risk Scoring

In this component of the framework, the idea is to classify different input variable clusters according to the magnitude of sensitivity in the context of the amount of risk they may have on the simulation model's output. As mentioned earlier, the impact on output variables can be measured by utilizing a measure of variability such as the Range or Standard Deviation in the output variables. To categorize the input variables into *High*, *Medium*, or *Low* Risk variables, the following criteria can be used:

- Based on the results of the statistical analysis, the statistical significance of the input clusters and/or their interactions will be identified and can be ranked according to their level of significance, which can be determined by the *Least Square Mean* and *p-values*.

- Frequency and significance of input interactions among input variable clusters
- Number of output variables significantly impacted by input variable clusters
- The actual output variable values produced by different input clusters
- Other measures that may emerge during method design and implementation

Further reduction of the previous factors can be performed depending on the results. For example, if the number of significant factors and/or their interactions is too large, the most significant ones will be selected by performing Pareto Analysis, which considers only the top 20% of the significant factors the affect 80% of the results. Similarly, if the frequency of interactions among the input factors is high, this can be reduced by selecting the most frequent input factors to interact with other inputs. The same principle can be applied to any other criteria used in ranking and scoring Risk.

A point that is worthwhile mentioning here based on the authors' experience conducting studies in risk and SA is to take into consideration normalizing the results. For example, the nature of the output measures may require that we normalize the results to be able to conduct valid comparisons. While the use of actual range values is meaningful when working with one output variable, it may not be that consistent when working across multiple output variables. For example if x_1 and x_2 are both significant terms with respect to output variables y_1 and y_2 respectively, and the Least Squared Means ranges for them are 1,000,000 and 1000 respectively, the significance of the latter value may be undermined because it is much lower than the former, while in realty x_2 may have more impact on the system than x_1. Therefore normalizing the outputs is necessary to be able to compare input significance across different output variables.

3. FRAMEOWRK LIMITATIONS

The SA framework presented in this paper is considered a Global SA approach as its focus is on studying the impact of variations over the entire range of model inputs as opposed to local SA in which one-variable-at-a-time analysis is mostly employed [Chen et al. 2005]. Our approach does capture higher level interaction effects in additions to the main effect; however since it utilizes Analysis of Variance (ANOVA) methods, the underlying assumption is that the relationships among variables are linear. An additional limitation to the method is the need for potentially intensive computation depending on the complexity of the model, the number of inputs/outputs, and number of runs that may be needed to reach sound conclusions.

4. CONCLUSIONS AND FUTURE WORK

In this paper a Sensitivity Analysis (SA) framework for large simulation models was presented. The framework was motivated by large system dynamics simulation models that were studied in various projects; however, the framework was generalized to fit other large simulation models. The framework was based on the authors' experiences with various real-life simulation models that they conducted. While there are significant amount of literature on SA methods available, the focus of this paper was more on the management aspects of the SA through a process-based framework when applied to large models with hundreds of input variables and tens of output variables. The SA framework consists of several components including clustering variables, designing experiments, running the model, analyzing the results and finally ranking the risk of the sensitive input variables. Identifying input variables with high, medium, or low risks, gives decision makers the knowledge they need about various inputs in terms of the risk associated with their variability. As a result, data sources of high risk variables can be made more reliable if possible or at least when conclusions are made, they are made cautiously given the amount of risk in the variables.

Among the most frustrating aspects of conducting SA is the large amount of manual work that analysts need to perform. Therefore, it will be interesting to develop SA tools based on this framework that can import the input variables from the users and intelligently run through the different stages and present the analysts with the results.

It may also be worthwhile to make quantitative or semi-formal connections between SA, risk, and uncertainty to help others integrate the proposed approach and its associated framework to risk assessment models that are widely used in many application domains.

References

Campolongo, F., Cariboni, J., and Saltelli, A. (2007) “An Effective Screening Design For Sensitivity Analysis Of Large Models”, *Environmental Modelling & Software,* Vol. 22, P. 1509-1518

Chen, W., Jun R., and Sudjianto, A. (2005) **“**Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design under Uncertainty”, Journal of Mechanical Design, Transactions of the ASME, Vol. 127, P.875 – 886

Kleijnen, J., (2005) “An Overview of the Design and Analysis of Simulation Experiments for Sensitivity Analysis”, *European Journal of Operational Research*, 164, P. 287–300

Montgomery, D.C. (2000), *Design and Analysis of Experiments*, 5th ed. NY: John Wiley & Sons,

Saltelli, A., Andres, T.H., and Homma, T. (1993) “Sensitivity Analysis of Model Output: An Investigation of New techniques”, *Computational Statistics & Data Analysis*, Vol 15., P. 211 – 238

Saltelli, A., Tarantola, S., Chan, K. P.-S. (1999) “A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output”, *Technometrics*, Vol. 41,1, P. 39-56.