Comparison of Uncertainty and Sensitivity Analyses Methods Under Different Noise Levels

David Esh[a*] and Christopher Grossman[a]
[a] US Nuclear Regulatory Commission, Washington, DC, USA

Abstract: Uncertainty and sensitivity analyses are an integral part of probabilistic assessment methods used to evaluate the safety of a variety of different systems. In many cases the systems are complex, information is sparse, and resources are limited. Models are used to represent and analyze the systems. To incorporate uncertainty, the developed models are commonly probabilistic. Uncertainty and sensitivity analyses are used to focus iterative model development activities, facilitate regulatory review of the model, and enhance interpretation of the model results. A large variety of uncertainty and sensitivity analyses techniques have been developed as modeling has advanced and become more prevalent. This paper compares the practical performance of six different uncertainty and sensitivity analyses techniques over ten different test functions under different noise levels. In addition, insights from two real-world examples are developed.

Keywords: Uncertainty, Sensitivity, Probabilistic, Model, Radioactive.

1. INTRODUCTION

Decisions regarding the capability of radioactive waste disposal facilities to contain and isolate radioactive waste from the environment are typically informed by the results of performance assessments. Performance assessments of radioactive waste disposal facilities use modeling to estimate potential radiological risk in the distant future, because the facilities cannot be monitored indefinitely. Therefore, modeling is used and uncertainties must be accounted for to support reliable decision making. The performance assessments commonly have a large number of uncertain inputs, and complex, non-linear, and sometimes non-monotonic responses of the outputs. The performance assessment can be thought of as a complex, non-linear, multi-dimensional numerical experiment. A performance assessment can involve many models to represent a variety of different features, events, and processes. Commonly the performance and degradation of engineered barriers, geochemistry and release of radionuclides from a wasteform, hydrogeologic transport through environmental media, and exposure to humans in the environment are represented by models within a performance assessment. However there are no defined limits on the number of models, the types of models, or how the models are integrated and interact. These models can interact in complex and non-intuitive ways. Decision-makers using the results of performance assessments are often interested in discerning which features, events, or processes are expected to significantly affect waste isolation and the level of uncertainty in the understanding of those features, events, and processes in order to improve confidence in the decision.

Sensitivity and uncertainty analysis is important to be able to develop understanding of the behavior of the numerical models, such as in performance assessments, focus research on the real-world system to reduce uncertainties, as well as to efficiently focus review effort on the most significant aspects. Different researchers use different terminology to refer to the process of estimating the important input parameters in a computational model incorporating uncertainty. This paper examines the sensitivity of outputs to uncertain inputs for a variety of models.

A number of issues arise when sensitivity analyses are applied to performance assessments. First, the number of uncertain variables may be very large while the truly important variables from the standpoint of driving the variance in the output may be quite small. For many performance

[*] david.esh@nrc.gov

assessments of radioactive waste disposal facilities uncertain parameters often number in the hundreds or greater whereas only on the order of tens of them are typically believed to be “important” to waste isolation. The large numbers of inputs often result in significant noise when discerning the sensitivity of the complex performance assessments to input parameters. ‘Noise’ is defined here as anything that can confound the identification of the truly important uncertain variables.

There are various types of noise that can impact sensitivity analysis techniques. The first type of noise considered in this evaluation is unimportant input distributions. As mentioned previously, a performance assessment may have hundreds or in a few cases thousands of uncertain input distributions to represent variability and uncertainty in physical processes and phenomena. This is primarily driven by the radiological source term that has a large number of different isotopes (e.g. 50 to 100). Processes that represent the potential release and transport of these isotopes through the environment are uncertain, and the parameters used to represent the processes are different for different elements. Generally only a handful of the parameters will be driving the variance in the results at a particular time or for a particular scenario. A good sensitivity analyses technique should be able to avoid identifying unimportant inputs (confounding influences) as important while identifying the important inputs. The second type of noise (e.g. Gaussian noise) represents irreducible uncertainty in the output, such as a result of measurement uncertainty. Many different types of features and processes can contribute to this second type of noise. These sources of noise could lead to failure to identify an important input as well as improperly classifying an important input as non-important.

A second issue that arises when sensitivity analyses are applied to performance assessments results because performance assessments can be quite complex and the computational burden to execute them can be significant. A single probabilistic realization can take anywhere from seconds to days. At the higher end of the range of execution times, the input and output data produced would generally be considered to be sparse or very sparse. Finally, the output response is generally not normally distributed, may span many orders of magnitude, may contain many zero (null) results, and is dynamic with time. The sensitivity analyses techniques need to be able to handle these challenges. This paper compares the performance of six different uncertainty and sensitivity analyses techniques over ten different test functions under different noise levels. In addition to the ten analytical test functions, most of the methods are also evaluated against some real-world datasets from more complex computational models.

2. METHOD

Analytical test functions described in the literature were implemented in a probabilistic computational model using the GoldSim® dynamic simulation platform. Table 1 provides the test functions used in the analyses. The test functions were selected to provide different amounts of non-linearity and non-monotonicity in functions of different dimension. The dimensionality of the functions range from 2 to 10 as implemented in this analysis.

To represent the various types of noise that can impact sensitivity analysis techniques, seven additional uncertain input parameters that have no impact on the output of each test function were included in the determination of the sensitivity of the output to the input. The seven additional uncertain parameters were varied uniformly from zero to one. Further, to all of these test functions, a noise term, to represent irreducible uncertainty, was added to the output of the function prior to performing the sensitivity analysis. The noise term was developed using a discrete probability distribution that could be adjusted from 0 (no noise) to 5 (high noise) and a unit normal distribution with the standard deviation adjusted as follows. The range of each function without the noise was first determined, and then the standard deviation of the unit normal distribution was adjusted so that the noise would be of the same magnitude as the product of the range of the function without noise and the value of the discrete probability distribution mentioned previously (i.e. 0 to 5). Figure 1 shows the simple function output as a function of the two inputs (x_4 and x_5) for a noise level of 0.1 and a noise level of 1.

Table 1: Analytical Test Functions

Test Function Name	Description	Input Description[a]	Ref #
Simple	$2x_1 + 5x_2$	x_1 = U[-1,1] x_2 = U[-1,1]	NA
Moon	$x_1 + x_2 + 3x_1x_3$	x_1,x_2,x_3 = U[0,1]	[1]
Webster	$x_1{}^{\wedge 2} + x_2{}^{\wedge 3}$	x_1 = U[1,10] x_2 = N[μ=2,σ=1]	[2]
Eldred	$\frac{x_1}{x_2}$	x_1 = LogN[μ=1,σ=0.5] x_2 = LogN[μ=1,σ=0.5] r = 0.3 (x_1 and x_2)	[3]
Park	$\frac{x_1}{2}\left[\sqrt{1+(x_2+x_3{}^{\wedge 2})\frac{x_4}{x_1{}^{\wedge 2}}}-1\right]+exp[1+sin(x_3)]$	x_1,x_2,x_3,x_4 = U[0,1]	[4]
Currin	$\left[1-exp\left(\frac{-1}{2x_2}\right)\right]\frac{2300x_1{}^{\wedge 3}+1900x_1{}^{\wedge 2}+2092x_1+60}{100x_1{}^{\wedge 3}+500x_1{}^{\wedge 2}+4x_1+20}$	x_1,x_2 = U[0,1]	[5]
Ishigami	$sin(x_1) + a\, sin^2(x_2) + b\, x_3{}^{\wedge 4}\, sin(x_1)$	x_1,x_2,x_3 = U[-π,π]	[6]
G-Function	$\prod_{i=1}^{d}\frac{\lvert 4x_i-2\rvert+a_i}{1+a_i}$ where $a_i = \frac{i-2}{2}$ for all i = 1,…..,d	x_i = U[0,1]; d=3, 5, or 10	[7]

[a] U = uniform, N = normal, LogN = lognormal, r = correlation

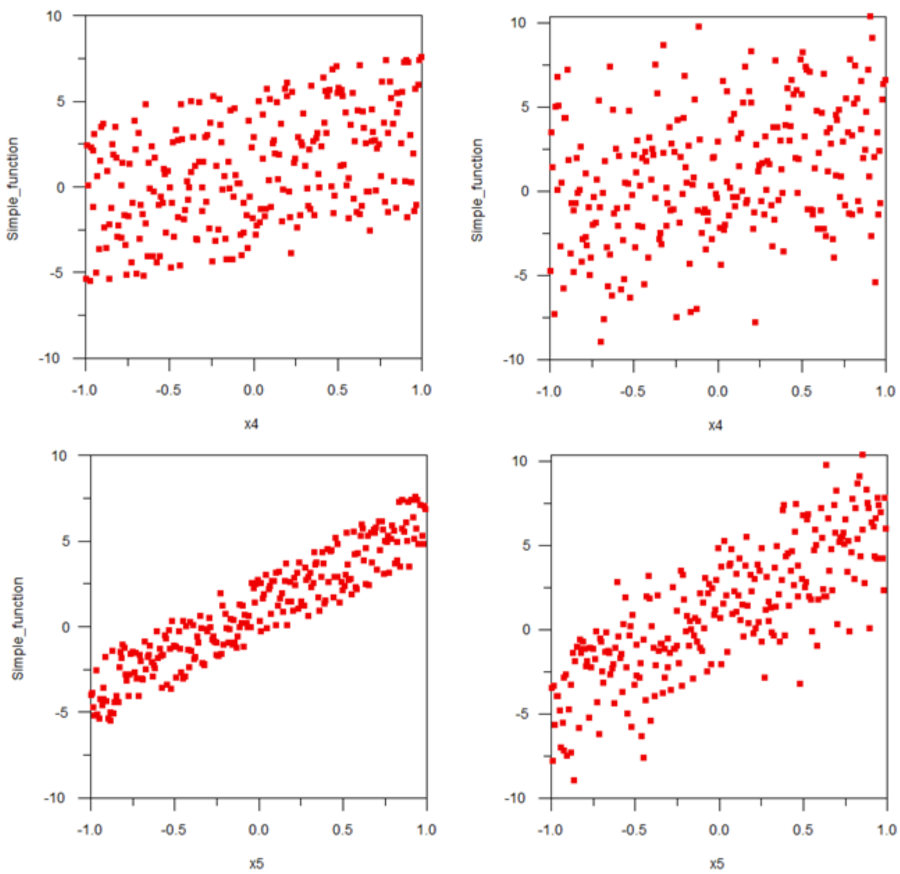


Figure 1: Comparison of the Simple Function Output for Two Different Noise Levels (a = 0.1, b = 1)

The six sensitivity analysis techniques evaluated comprised regression-based and other methods. Regression-based methods are most useful for monotonic relationships (i.e., the dependent variable *y* only increases or decreases with increasing value of the independent variable *x*). When this is not the case, regression based methods are at a disadvantage because a non-monotonic relationship can be highly correlated but have a low correlation coefficient. The six uncertainty and sensitivity techniques evaluated were the importance measure, correlation coefficient, standardized regression coefficient (SRC), partial correlation coefficient, a genetic algorithms based measure (GA-m), and a method based on the Extended Fourier Amplitude Sensitivity Test (extended FAST) [8,9,10,11,12]. Outside of the GA-m measure, all of these methods are well-documented in the literature and therefore are not described in further detail here.

The genetic algorithm based measure was developed during the analysis of a performance assessment describing the degradation of a cement wasteform [12]. It was then evaluated on the data from performance assessment calculations associated with the proposed high-level waste repository at Yucca Mountain, Nevada [13]. In each of those cases the technique was found to perform well at identifying important inputs.

The GA-m approach used neural network software developed by Neuralware [14]. Neuralworks Predict® is an add-in to Microsoft Excel that can be used to build neural networks. The approach used was to export sampled stochastic input variables along with the pertinent output variables from the test functions or real-world datasets to Excel and then to build a neural network using Neuralworks Predict. The variable selection algorithms were used to select the most important input variables needed to develop a neural network to predict the output. The neural network itself was not used, just the variable selection algorithm. The variable select algorithm uses a genetic algorithm to search for synergistic sets of input variables that are good predictors of the output. The software also can perform a pre-selection of variables using a cascaded genetic algorithm approach. This method gives more consistent variable sets by eliminating variables that are consistently rejected during different invocations of the genetic algorithm. The sensitivity analysis measure (GA-m) was the frequency that a variable was retained by the GA over many iterations. The GA-m technique may work well on performance assessment-type data because of the relatively large (compared to other techniques) amount of variable transformations that are used by the technique.

The analyses in this paper for the first four techniques (importance measure, correlation coefficient, standardized regression coefficient (SRC), partial correlation coefficient) were performed on the raw data and rank transformed data. Rank transformation, a dimensionless transform, replaces the value of a variable by its rank (i.e., the position in a list that has been sorted from largest to smallest values). Analyses with ranks tend to show a greater sensitivity than results with untransformed variables. For performance assessments, if the distribution of results is skewed toward the low end, which is usually the case, rank transformation of the dependent variable can over-weight the lower results.

Variable transformation can be an important step in the analysis process. The correlation between input and output variables in statistical methods can often be enhanced by transforming the variables. Transformations are used to (i) eliminate dimensionality of the variables, (ii) reduce the role of points at the tails of the distributions, and (iii) properly scale the resulting sensitivities to the variability of the input variables. However, transformation of the dependent variable (usually peak dose in performance assessments) can skew the results in an undesirable way because the risk, being based on the average dose, is weighted heavily by the largest doses. Therefore, the sensitivities should reflect what matters most to the risk and give weights proportional to the doses. Although transformations of the dependent variable often will improve the goodness of fit in the sensitivity analysis, in general they should only be applied to the independent variables.

The ten test functions varied in complexity in terms of non-linearity and the dimensionality of the function. The coefficient of determination was calculated and used to determine the fraction of the variance in the output that can be explained by a linear relationship for each test function. Results were generated for all of the test functions, and were further categorized into the five functions that were more linear (i.e. a high coefficient of determination) and were more non-linear (i.e. a low coefficient of determination). In order to evaluate the performance of the techniques, the numerical value of the uncertainty measure was ranked and the statistics across all functions were compiled. For example, if a function had two uncertain inputs (A and B) and seven uncertain dummy variables (C through I), then a rank of 1 and 2 for the uncertainty measure of inputs A and B would mean the technique successfully identified the two important inputs and avoided spurious correlation with a noise input. The GA-m technique was different from the other techniques in terms of its output because inputs are screened to determine those that are suitable for building the neural network. The results are a confusion matrix whereas the other techniques only produce a ranking of the numerical values. For metrics that can have both positive and negative values as the output (e.g. correlation coefficient) the absolute value was used to determine the ranks for scoring purposes.

3.0 RESULTS

3.1. Test Function Analyses

In addition to the magnitude of the noise added to the output of the functions, the analysis also considered the number of probabilistic realizations. As previously discussed, performance assessments can be complex and computationally expensive. In some cases generating more than a few hundred realizations may be impractical. Figure 2 provides the results of the sensitivity analyses techniques for 250 realizations with no noise added to the output function. The metric plotted is the average classification rate over the ten test functions. The extended FAST and importance factor consistently rated highest, followed by the GA-m. However, the results for the total sensitivity index for the extended FAST method (STi) may not be directly comparable to the other methods. To perform the analyses to produce Figure 2, all of the other methods were given 250 realizations of results. For the extended FAST method shown in the chart, 65 outputs were generated for each input and 4 replicates were performed. For a function with 10 inputs and 7 "dummy" inputs, a total of 4420 calculations of the function would be performed. Calculations were performed with 10,000 realizations for the other methods and the classification percentages increased 10-15%, closing the gap on the STi of the extended FAST method though it was still superior.

Figure 3 provides the results of the sensitivity analyses techniques as a function of the number of realizations with no noise function added to the output of the functions. The more common (and simpler) measures show limited sensitivity to the number of realizations over the range investigated. However, the importance factor showed a strong decrease in performance as the number of realizations decreased. Even at the lowest end of the range shown in Figure 3 the ratio of the number of realizations to inputs (~ 25) is much higher than many performance assessment applications. In some cases the ratio of the number of realizations (calculations of model output) to uncertain inputs can be less than unity.

The results were further broken down by separating the results into those associated with the five simpler test functions and those associated with the more complex test functions (the last five in Table 1). The distinction between simpler and complex was made on the basis of the resultant coefficient of determination and the dimension. The mean coefficient of determination of the simpler functions was 0.87 compared to 0.19 for the more complex test functions.

The mean dimension of the simpler functions was 2.6 compared to 4.6 for the more complex functions. Table 2 provides the difference in overall classification percentages for the sensitivity analyses methods between the simpler and more complex test functions as a function of noise level (250 realizations). In general, the performance on the complex test functions is lower, as would be expected. For all the methods at moderate to low noise levels the performance is comparable. However at high noise levels the importance method and GA-m show deteriorating performance for the simple test functions. For the complex test functions, the performance of the importance measure and GA-m is better than the simpler methods, with the importance measure really standing out. The simpler methods struggle with the complex functions.

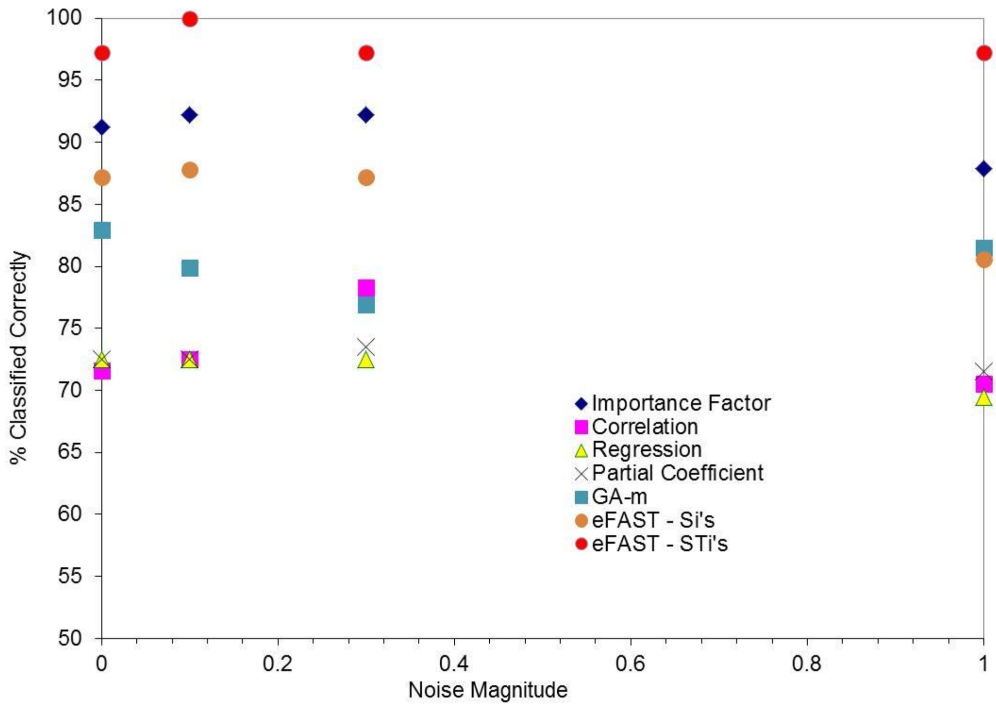


Figure 2: Impact of the Magnitude of Noise on Average Classification Rate (Number of Realizations = 250)[†]

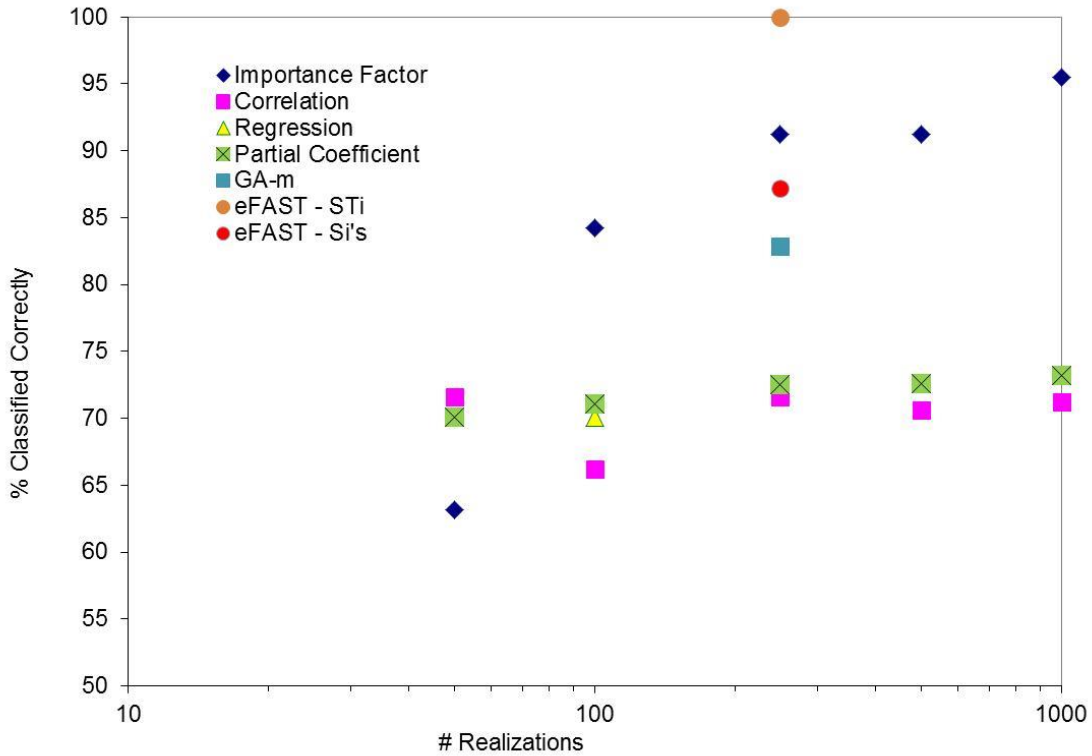


Figure 3: Impact of the Number of Realizations on Average Classification Rate (Noise level = 0)

[†] As discussed in the text, the results for the extended FAST method may not be directly comparable. Not all methods were executed for each comparison provided in this paper.

Table 2: Performance for Simple vs. Complex Functions (% Classified Correctly)

	Noise	Importance	Correlation	Regression	Partial Coefficient	GA-m
Simple	0	92	100	92	92	92
	0.1	92	100	92	92	92
	0.3	92	100	100	100	92
	1	85	92	92	100	92
	3	62	85	85	85	54
	5	54	69	85	77	46
	Noise	**Importance**	**Correlation**	**Regression**	**Partial Coefficient**	**GA-m**
Complex	0	83	48	52	52	65
	0.1	87	57	52	52	70
	0.3	87	57	52	52	65
	1	83	43	43	43	65
	3	78	52	48	48	57
	5	65	43	48	48	48

For performance assessment, it is very important that the sensitivity analysis method can determine the important inputs when there are a very large number of non-contributing inputs and a low ratio of results to uncertain inputs. To evaluate the sensitivity of the results to the number of non-contributing inputs, the number of "dummy" inputs was increased from 7 to 20 for each test function. For 250 realizations and a noise level of 1.0, the performance of the importance measure, correlation coefficient, standardized regression coefficient, and partial correlation coefficients did not change significantly for the simple test functions (compared to the results in Table 2). However, for the complex test functions the importance measure held up well (65% vs. 83%), but the correlation coefficient, standardized regression coefficient, and partial correlation coefficients performance decreased substantially. In fact the classification percentages are no better than would be expected by chance (22%). Resources did not allow the authors to complete similar analyses with the GA-m and extended FAST methods though they may be evaluated for sensitivity to increasing the number of inputs in future research.

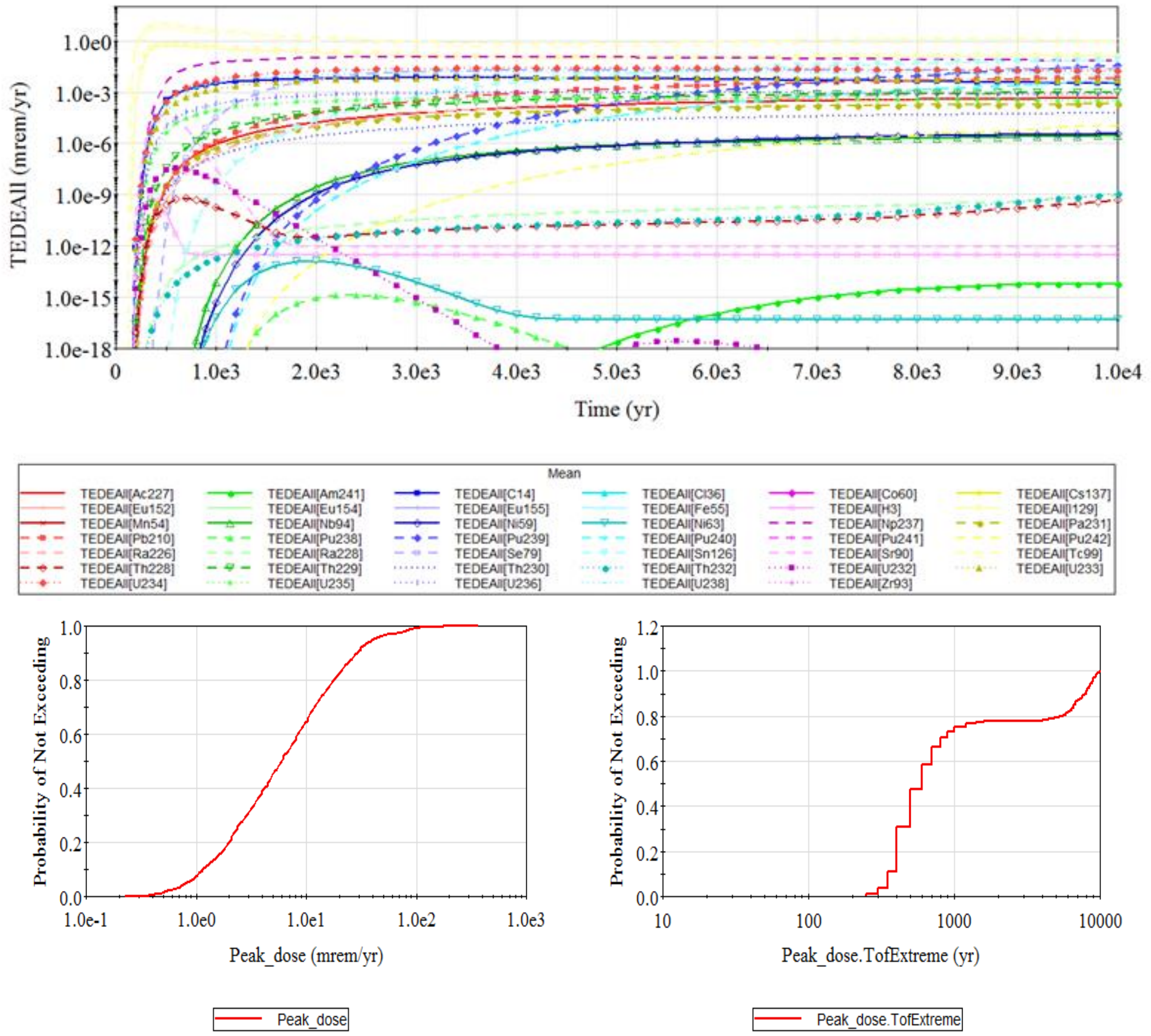
3.2. Performance Assessment Analyses

The final analysis that was performed was comparison of some of the methods on real world datasets from performance assessments. The performance assessments are very complex, can take an extremely long time to execute, and have a very large number of uncertain inputs. The output can have many zero results and dynamic behavior that change over the many thousands or hundreds of thousands of years of simulated performance. Many methods struggle with this type of problem due to the sparseness of data to analyze or inherent limitations of the methods with complex datasets. In comparison to a performance assessment, the test functions evaluated in the analysis are extremely computationally efficient. Producing 10,000 realizations of output required only a second of computer time.

The Nuclear Regulatory Commission (NRC) staff has a long history of performing sensitivity analysis on performance assessment and other complex models. Much of that work has recently been summarized in Mohanty et al. [13]. Many methods were applied to the same dataset. The dataset was produced by an assessment describing the performance of a proposed high-level nuclear waste repository, the Total-system Performance Assessment code (TPA) [15]. The TPA code has on the order of 1,000 parameters, of which 200-400 are typically sampled in a given calculation and 40-50 are cross-correlated to other sampled parameters. A large variety of sensitivity analyses techniques were applied to TPA code output. NRC staff experience has found that the best way to identify important parameters in these large complex models is to run a variety of techniques and rank the parameters based on how many of the techniques identify the parameters as important. The GA-m method was then run on the same dataset produced by the TPA code that was evaluated with the other methods. The selection algorithm identified six parameters, all with a high frequency. The parameters pertained to the degradation of the waste form, the quantity of radionuclides directly consumed in drinking water, or the dilution of contaminants in the saturated zone during transport and transport time. From a physical standpoint, the parameters selected were in strong agreement with the analysts' conceptual understanding of the performance of the disposal system. Using modified raw data, the GA-m method successfully identified 6 of the 10 parameters ranked as most significant by compiling the results across all methods. Screening out realizations that had zero doses (about 200 of the 512) identified 7 of the 10 parameters. Using both sets of results identified 8 of the 10 parameters, which was a very strong result for an individual method.

All of the methods except the extended FAST were applied to a second performance assessment model. This model describes the degradation of a cement wasteform and potential release of radioactivity into the environment for 41 species (isotopes) including decay chains [12]. Figure 4a shows the time histories of projected dose (mean values) for 1,000 realizations. Figure 4b is the distribution of peak doses, and Figure 4c is the distribution of the time of peak doses. The peak dose and the time of the peak dose were moderately correlated (-0.26). The sensitivity analysis techniques

a)



b) c)

Figure 4: a) Mean Time History of Projected Doses by Isotope; b) Cumulative Distribution Function of Peak Doses; c) Distribution of the Time of Peak Doses.

were applied to this model. The model had 374 uncertain input parameters, some of which were correlated. Latin Hypercube Sampling (LHS) was used to sample the parameters for probabilistic simulations.

One of the problems analyzing performance assessments is that the "true" answer is not known. The models provide projections into the future but validation cannot be achieved in the strict sense. Rather, indirect methods must be used to provide confidence in the model results. To evaluate if the sensitivity analysis methods results are reasonable, the analyst has to use the information such as that shown in Figure 4 combined with interaction with subject matter experts (if the analyst is not familiar with the features, events, and processes represented in the model). Analysts can use their experience to determine if the results of the sensitivity analysis methods make sense. For instance, if the sensitivity analysis method identifies parameters associated with an isotope that contributes a trivial amount to peak dose, it is likely that sensitivity analysis results at and below that level are indeterminate at best.

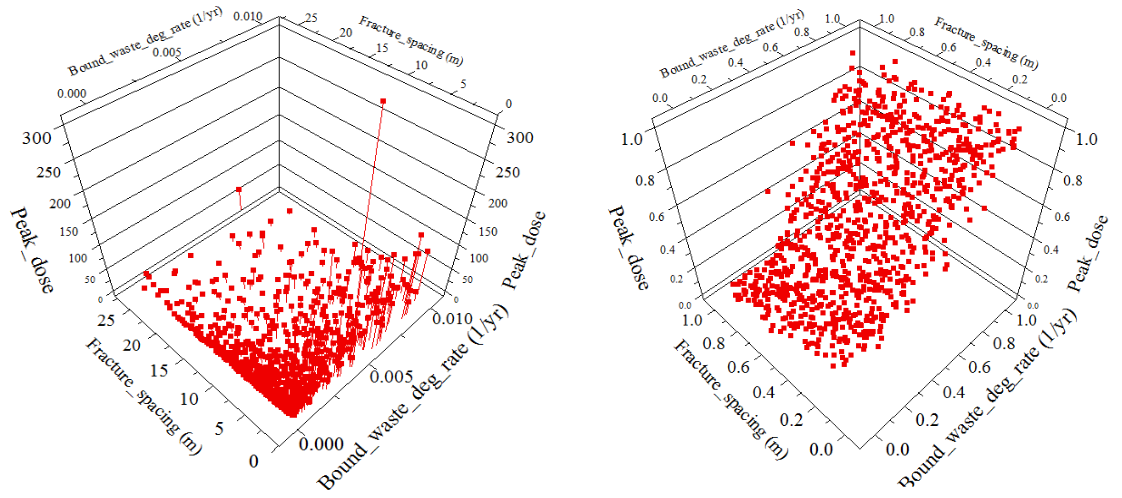
Table 3 provides the list of the top ten parameters identified by the sensitivity analysis techniques. The parameters highlighted in green are, based on experience, believed to be accurate determinations of important parameters. The parameters highlighted in blue are parameters that can confidently be determined as being insignificant. The parameters highlighted in yellow are plausible that can't be confirmed or eliminated. The standardized regression coefficient method provided the poorest results, whereas the GA-m was the best performer. There was a rather substantial improvement in the results for the importance measure when the ranks were used as compared to the raw data.

To test the reasonableness of the results, a multivariate plot of the top two parameters from the GA-m method was developed. For large, complex datasets visualization can be useful especially to confirm or refute the results of analysis methods. Figure 4a and 4b are the multivariate plots of the raw data and ranks of the total dose at 10,000 years as a function of the bound waste degradation rate parameter and the fracture spacing parameter. The plots confirm the importance of these parameters to the results.

By cautiously eliminating most of the inputs that contribute very little to the output variance, the ability of the sensitivity analysis methods to identify the important parameters can be significantly enhanced. This is in agreement with the results presented earlier that showed as the number of "dummy" inputs was increased the classification percentages decreased. In other words, and experience-based screening step of the initial results can greatly improve the identification of important parameters. As discussed in [12], the number of important parameters that were identified with confidence was doubled when the same technique was applied on a shortened variable list.

Table 3: Performance Assessment Model Top Ten Parameters Identified

Importance	**Importance - Ranks**	**Correlation**	**Regression**	**PCC**	**GA-m**
Frac Spacing	Frac Spacing	Tc	SZ pipe length	Tc	Bnd W deg rate
H Sol	GW flow	Frac Spacing	Kd UZ Np	Nm	Frac Spacing
Kd W Pu Ox	Water Intake	GW flow	Cap life	Kd SZ Cs	GW flow
Kd Soil Cl	Nm	Kd UZ Se	U Sol	Erosion rate	Infiltration rate
Kd W Am deg	Soil T lateral	Water Intake	Kd SZ I	SO4 conc	Water Intake
Pb Sol	Kd Sz Ac	Kd W U Ox	Sr Sol	Np Sol	Kd W Tc
Kd Soil Cs	Kd UZ Sn	C	Frac Diet Animal	Kd UZ Tc	Nm
Kd Garden C	Kd W Pb Ox	Th Sol	Tc	Kd SZ Pa	SO4 conc
Rn	Soil int farm	Cl Sol	Th Sol	Fe	Kd W Np
Kd UZ Cs	Kd W Mn deg	Kd SZ Mn	Porous length	Co Sol	Kd W Pu



a) b)

Figure 4: a) Multivariate Plot of the Top Two Parameters Identified with the GA-m Method for a) the Raw Data and b) the Ranks of the Data

4. CONCLUSIONS

Developing sensitivity analysis methods for application to performance assessments is a challenging problem. Computational limitations may limit the number of realizations (calculations) that can be performed to provide data to the methods. In addition, the number of uncertain inputs can be many hundred or more. The results of the analyses identified methods that had more computationally efficient performance over others (i.e. the extended FAST). Some methods were more sensitive to noise, and some showed greater sensitivity to the amount of data available. Some techniques were much more favorable to use when non-linearity was present. Overall conclusions were:

- Simpler methods are reliable for simpler functions.
- The total sensitivity index (STi) of the extended FAST method produced very strong results for the cases where it was evaluated. Future research may involve more extended evaluation of this method on performance assessment models if computational resources allow.
- Applying numerous methods and averaging or ranking the results appears to lead to the most reliable identification of important parameters.
- The GA-m method produced adequate results on the test functions with moderate to low noise levels, but much stronger results on performance assessment data. At this time it is not clear why and may be a subject of future research.
- The importance measure using ranks performed much better on performance assessment data compared to other methods and compared to using unranked data.
- A screening step to remove clearly unimportant uncertainties can greatly improve the results for many techniques, especially when the number of inputs is large relative to the number of probabilistic realizations. Using expert judgment as a screening step to remove unimportant variables may improve results as much or more than advances in different sensitivity analyses techniques.

References

[1] H. Moon, "*Design and Analysis of Computer Experiments for Screening Input Variables*", Doctoral dissertation, Ohio State University, Columbus, OH, (2010).
[2] M. D. Webster, M. A. Tatang, and G. J. McRae, "*Application of Probabilistic Collocation Method for an Uncertainty Analysis of a Simple Ocean Model*", Technical report, MIT joint program on the science and policy of global change, report series no. 4, MIT, (1996).

[3] M. S. Eldred, H. Agrawal, V. M. Perez, S. F. Wojtkiewicz, and J. E. Renaud, "*Investigation of Reliability Method Formulations in DAKOTA/UQ*", Structure and Infrastructure Engineering, 3(3), pp. 199-213, (2007).
[4] D. D. Cox, J. S. Park, C. E. Singer, "*A Statistical Method for Tuning a Computer Code to a Data Base*", Computational Statistics and Data Analysis, 37(1), pp. 77-92, (2001).
[5] C. Currin, T. Mitchell, M. Morris, and D. Ylvisaker, "*Bayesian Prediction of Deterministic Functions, with Applications to the Design and Analysis of Computer Experiments*", Journal of the American Statistical Association, 86(416), pp. 953-963, (1991).
[6] T. Ishigami and T. Homma, "*An importance quantification technique in uncertainty analysis for computer models*", Proceedings of the First International Symposium on Uncertainty Modeling and Analysis, IEEE, pp. 398-403, (1990).
[7] A. Marrel, B. Iooss, B. Laurent, and O. Roustant, "*Calculation of Sobol Indices for the Gaussian Process Metamodel*", Reliability Engineering and System Safety, 94(3), pp. 742-751, (2009).
[8] J. R. Benjamin and C. A. Cornell, "*Probability, Statistics, and Decision for Civil Engineers*", McGraw-Hill, 1970, New York.
[9] A. Saltelli and S. Tarantola, "*On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal*", J. Am. Stat. Ass., Vol. 97, No. 459, (2002).
[10] R. L. Iman, et al., "*A FORTRAN Program and User's Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients*", NUREG/CR-4122, SAND85-0044, US Nuclear Regulatory Commission, Sandia National Laboratory, (1985).
[11] A. Saltelli, S. Tarantola, and K. P.-S. Chan, "*A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output*", Technometrics, 41:1, pp. 39-56, (1999).
[12] D. W. Esh, A. C. Ridge, and M. Thaggard, "*Development of Risk Insights for Regulatory Review of a Near-Surface Disposal Facility for Radioactive Waste*", Proceedings of the WM'06 Conference, February 26-March 2, 2006, Tuscon, AZ, (2006).
[13] S. Mohanty, et al., "*History and Value of Uncertainty and Sensitivity Analyses at the Nuclear Regulatory Commission and Center for Nuclear Waste Regulatory Analyses*", ML112720123, US Nuclear Regulatory Commission, (2011).
[14] Neuralware, NeuralWorks Predict® Product Version 2.40, Carnegie, PA, (2001).
[15] B. W. Leslie, C. Grossman, and J. Durham, "*Total-System Performance Assessment (TPA) Version 5.1 Code Module Descriptions and User Guide*", ML080510329, The Center for Nuclear Waste Regulatory Analyses, San Antonio, TX, (2007).