SAND 97-2800A

2/18/98

Quantitative Sensitivity and Importance Analysis of Parameters in Monte Carlo Models

QA: N/A

Sean A. McKenna and Bill W. Arnold
Sandia National Laboratories
P. O. Box 5800, MS-1324
Albuquerque, NM 87185-1324
(505) 848-0713
e-mail: samcken@nwer.sandia.gov

I. INTRODUCTION

In a performance assessment of a nuclear waste repository, the modeling of many subsystems and processes is accomplished using Monte-Carlo methods. Monte-Carlo techniques are necessary to assess the magnitude of uncertainty in system or subsystem performance due to uncertainty in parameters. In all cases, it is desirable to determine the relative importance of the various input parameters through a sensitivity analysis. While results may be sensitive to various parameters, those parameters only become important if they also control regulatory failure or acceptance of the model results. The sensitivity and importance analysis results can then be used to guide further site characterization activities and/or to determine if the given uncertainty in a parameter is consequential to the overall performance assessment.

II. SENSITIVITY ANALYSIS

In many cases, sensitivity analyses performed on the results of nuclear waste repository performance assessment calculations have focused on parameter sensitivity to all model results regardless of any regulatory driver (e.g., ¹). The difficulty with this approach is that while model results may be sensitive to a given parameter, if all results are still below (or above) the regulatory limit, then the final regulatory decision is insensitive to that parameter, and the importance of that parameter in future modeling or site characterization studies may be negligible.

A smaller volume of work (e.g., ²) has considered a regulatory constraint in determining the sensitivity and importance of parameters in Monte Carlo models. Generalized sensitivity analysis (GSA)³ is a technique that considers the sensitivity of model parameters to the final regulatory decision against which the model results are being judged. GSA has recently been used in a decision analysis of different approaches to aquifer remediation⁴ but has not been applied to sensitivity analysis of nuclear waste repository performance assessment models.

The essence of GSA is the placement of each vector of model input parameters into one of two sample sets: those that produced model results that failed the regulatory

requirement and those that produced model results that passed the regulatory requirement. This behavioral classification of the input vectors into two sample sets allows for the quantitative examination of the differences between the two sample sets as a function of any parameter contained within the input vectors. This process allows for determination of whether or not the differences between a "fail" sample set of input parameters and a "pass" sample set is significant and also for the relative ranking of all input parameters in terms of the sensitivity of the results to the given parameter. For those parameters where the distribution of values in the fail sample set is statistically indistinguishable from those in the pass sample set, the model results are insensitive to that parameter.

III. EXAMPLE

An example GSA of input parameters to an analytical solution for solute transport through a single fracture with matrix diffusion and sorption onto the matrix is presented. Given a continuous source, the expression for the relative concentration at the fracture outlet is ⁵:

$$\frac{C}{C_o} = \operatorname{erfc} \left[\left(\frac{D_e K_d \rho_B}{t - t_{adv}} \right)^{0.5} \frac{LW}{Q} \right]$$

The hypothetical example problem to be addressed is the time to a relative concentration (C/C_o) of 0.90 when there is uncertainty in the values of the distribution coefficient (K_d), the bulk density (ρ_B), the length and width of the fracture (L and W) and the flux through the fracture (Q). The advective travel time (t_{adv}) is calculated using Q , L , and the cross-sectional area of the fracture where the cross-sectional area is calculated as ($W * \text{aperture}$) and aperture is also considered an uncertain parameter. For this example, 10,000 vectors of the 6 uncertain parameters were drawn using the latin hypercube sampling program LHS⁶ and the program BETA*. The type of distribution and the

*. BETA: A visual BASIC program for calculation of beta distributions, L. Yarrington, 1995, Sandia.

parameters describing each distribution are given in Table 1. The solute of concern is Cesium (Cs) and the free water diffusion coefficient, D_e , of Cs is held constant at $20.7 \times 10^{-10} \text{ m}^2/\text{sec}$.

The regulatory constraint for this example problem is that the hypothetical site will "fail" if the time to a relative concentration of 0.90 at the outlet is less than 100 years. The equation above was solved for t when $C/Co = 0.90$ for each of the 10,000 parameter vectors and each of these resulting times was compared to the regulatory limit. The results show that 7768 vectors failed and 2232 passed the regulatory constraint.

GSA provides both qualitative and quantitative techniques to examine the sensitivity of results to a given parameter. The qualitative results are a graph of cumulative density functions (cdf's) for each parameter where the distribution of the parameter within the vectors that passed the regulatory constraint is compared to the distribution within the vectors that failed. If there is a large separation between the two cdf's, then the results, in terms of the decision point defined by the regulator, are sensitive to that parameter. Example cdf's for both a relatively sensitive (fracture width) and an insensitive (K_d) parameter are shown in Figure 1, A and B.

GSA provides quantitative results by using a non-parametric statistical test of the difference in cumulative probability between two distributions. The test used here is the Kolmogorov-Smirnov (K-S) test where the maximum difference between two cdf's, d , is determined. This maximum can occur at any value of the selected parameter (e.g., the entire range of values in Figure 1 A or B). The K-S test also provides a determination of the probability, P , that the maximum distance between the passing and failing cdf's could have occurred if the two distributions were in fact obtained from a single population. The lower this probability, the more certainty there is that the cdf's are significantly different. In order to demonstrate the effect of the number of runs on parameter sensitivity and importance, results of the K-S test are given for each parameter in Table 2 for runs with both 100 and 10,000 vectors.

IV. SUMMARY

GSA is a quantitative technique for determining the sensitivity of Monte Carlo model results to the input parameters. Importance of parameters is determined in GSA by explicit consideration of the regulatory constraints. As exhibited by the results in Table 2, the calculated probability values (P 's) are a function of the number of Monte Carlo runs. Relative to regression techniques for sensitivity analysis, GSA is focused on behavior about a regulatory decision point and does not require the fitting of a regression model to the resulting data. For the example

problem demonstrated here, fracture transport results are sensitive to all six parameters for the given regulatory cut-off (100 years to 0.90 relative concentration) when 10,000 vectors are analyzed. We are currently engaged in applying GSA to results of sub-site scale, saturated zone, transport modeling at Yucca Mountain and suggest that GSA be considered for sensitivity analysis of the final TSPA-VA calculations.

ACKNOWLEDGEMENTS

This work was supported by the Yucca Mountain Site Characterization Office as part of the Civilian Radioactive Waste Management Program which is managed by the U.S. Dept. of Energy, Yucca Mountain Site Characterization Project. Sandia is a multiprogram laboratory operated by the Sandia Corp. a Lockheed Martin Company for the U.S. Dept. of Energy under contract DE-ACO4-94AL85000.

REFERENCES

- ¹Wilson, M.L., 1993, Sensitivity Analyses for Total-System Performance Assessment, in Proceedings of the Fourth Annual International High Level Radioactive Waste Management Conference, Las Vegas, Nevada, April 26-30, 1993, American Nuclear Society, La Grange Park, IL, Vol. 1 pp. 14-21.
- ²Reeves, M., V.A. Kelley and J.F. Pickens, 1987, Regional Double-Porosity Transport in the Culebra Dolomite: An Analysis of Parameter Sensitivity and Importance at the Waste Isolation Pilot Plant (WIPP) Site, SAND87-7105, Sandia National Laboratories, Albuquerque, New Mexico.
- ³Spear, R.C. and G.M. Hornberger, 1980, Eutrophication in Peel Inlet. II. Identification of Critical Uncertainties via Generalized Sensitivity Analysis, *Water Research*, Vol. 14 (1), pp. 43-49.
- ⁴James, B.R., J.-P. Gwo and L. Toran, 1996, Risk-Cost Decision Framework for Aquifer Remediation Design, *Journal of Water Resources Planning and Management*, Vol. 122 (6), pp. 414-420.
- ⁵Carslaw, H.S. and J.C. Jaeger, 1959, Conduction of Heat in Solids, Oxford Univ. Press, Oxford, 2nd Edition, 510 pp..
- ⁶Iman, R.L. and M.J. Shortencarrier, 1984, A FORTRAN 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models, SAND83-2365, Sandia National Laboratories, Albuquerque, New Mexico, 50 pp.

Table 1. Parameter distributions for Monte Carlo fracture transport model.

Parameter	Distribution	Minimum	Maximum	Alpha	Beta
Kd (m ³ /kg)	Uniform	100	200	NA	NA
Bulk Density (kg/m ³)	Beta	1800	2510	1.246	0.221
Length (m)	Beta	100	300	0.8	3.0
Width (m)	Log-Normal	0.01	10.0	NA	NA
Flux (m ³ /day)	Uniform	0.1	5.0	NA	NA
aperture (microns)	Beta	50	1000	0.8	3.0

Table 2. Maximum separation distances (d) between passing and failing cdf's and the probability (P) that those differences could be caused by sampling from the same population as calculated with 100 and 10,000 input vectors.

Parameter	Width	Flux	Aperture	Length	Bulk Density	Kd
d (n = 100)	0.54	0.59	0.19	0.19	0.19	0.09
P (n = 100)	5.8×10^{-05}	5.6×10^{-05}	0.54	0.54	0.54	0.99
d (n = 10,000)	0.539	0.485	0.137	0.136	0.136	0.061
P (n = 10,000)	0.00	0.00	1×10^{-28}	2×10^{-28}	3×10^{-28}	1×10^{-06}

Figure 1. Cumulative distributions of the fracture width (A) and the distribution coefficient (B) after classification into passing and failing sample sets. Note the greater sensitivity of the results to fracture width relative to distribution coefficient. Results shown are based on 10,000 input vectors.