Exploring Sensitivity Analysis Techniques for the Assessment of an Environmental Transport Model

D.M. Hamby
School of Public Health, University of Michigan
Ann Arbor, MI 48109-2029, USA

S. Tarantola
Institute for Systems, Informatics and Safety
Joint Research Center of the European Commission
TP. 361 - 21020 Ispra (VA) ITALY

ABSTRACT:

Mathematical models, designed to simulate complex physical processes, are often used in scientific and engineering studies. For example, modeling the movement and consequence of radioactive pollutants is extremely important in the nuclear industry for environmental protection and facility control. One of the steps in model development is the determination of the parameters most influential on model results. A sensitivity analysis of these parameters is not only critical to model corroboration, but also guides future research. We have examined various methods of sensitivity analysis using an environmental transport and dosimetry model for radiation dose assessment of atmospheric tritium. Sensitivity rankings of input parameters were determined for various one-at-a-time, local, and global techniques applied to the model. Software packages including Crystal Ball, SAS, and PREP/SPOP were used to perform operations ranging from the selection of random samples, to estimating regression coefficients, to computing FAST sensitivity indices. Conditionalized parameter sensitivities also were investigated to determine the influence of parameter variability on model output resulting in radiation dose estimates at the high-end of the distribution. These conditionalized sensitivities were determined using partitioned-data techniques including the Smirnov and Cramer-von Mises tests.

1 INTRODUCTION

Engineering and scientific phenomena are often studied with the aid of mathematical models designed to simulate complex physical processes. In the nuclear industry, modeling the movement and consequence of radioactive pollutants is extremely important for environmental protection and facility control. One of the steps in model development is the determination of the parameters most influential on model results. A sensitivity analysis of these parameters is not only critical to model validation, but also serves to guide future research toward increasing model accuracy.

This paper is a comparative study of a few of the sensitivity analysis methods deemed appropriate for an atmospheric tritium dose model (Hamby 1993). The work is intended to demonstrate the calculational rigor and compare parameter sensitivity rankings resulting from the various techniques.

2 SENSITIVITY ANALYSIS: THE METHODS EMPLOYED

2.1 *Partial Rank Correlation*

Strong correlations between input parameters may influence input/output correlations. Partial correlation coefficients (PCC) are calculated to account for correlations among other input variables (IAEA 1989). Given random variables X_1 and X_2 as input and the output variable Y, a partial correlation coefficient is a measure of the correlation between X_1 and Y, for example, while eliminating indirect correlations due to relationships that may exist between X_1 and X_2 or X_2 and Y. The PCC is defined as

$$r_{X_1 Y|X_2} = \frac{r_{X_1 Y} - r_{X_1 X_2} r_{X_2 Y}}{\sqrt{(1-r_{X_1 X_2}^2)(1-r_{X_2 Y}^2)}}, \quad (1)$$

where the notation $r_{X_1 Y|X_2}$ represents the partial correlation coefficient for X_1 and Y while accounting for the effects of X_2. The parameters of the generic model considered herein are assumed to be independent and no correlations have been assigned. Therefore, (1) reduces to,

$$r_{X_1 Y|X_2} = \frac{r_{X_1 Y}}{\sqrt{(1-r_{X_2 Y}^2)}} \quad (2)$$

where, again, X_1 and X_2 represent any two input variables and Y represents the output variable. The square of the partial correlation coefficient is useful in deter-

rankings based on the relative values of the partial correlation coefficients will not change from the rankings determined based on the simple correlation coefficients. Therefore, with no correlations existing between input parameters, there is no need for calculating partials to determine sensitivity rankings.

The rank transformation is applied to partial correlation as a test of monotonicity between input and output variables while accounting for relationships between input parameters. The partial rank correlation (PRCC) is widely utilized for sensitivity studies. The PRCC is reportedly more powerful at indicating the sensitivity of a parameter that is strongly monotonic yet highly nonlinear (Downing et al. 1985). The PRCC is used in this study as one method of determining model sensitivity.

2.2 *Regression methods*

Regression methods are often used to replace a highly complex model with a simplified response surface (Helton et al. 1991). The response surface is simply a regression equation that approximates model output using only the most sensitive model input parameters. Stepwise regression procedures are utilized to ensure that the final regression model provides for the best fit of raw data. The stepwise regression may involve higher ordered equations, quadratic terms, and parameters as functions of other parameters. The generalized form of a simple regression equation is,

$$\hat{Y} = b_0 + \sum_k b_k Z_k \qquad (3)$$

where each Z_k is a predictor variable and a function of (X_1, ..., X_n) and each b_k is a regression coefficient. The use of the regression technique allows the sensitivity ranking to be determined based on the relative magnitude of the regression coefficient. This value is indicative of the amount of influence the parameter has on the whole model. Because of units and the relative magnitude of parameters, a standardization process is sometimes warranted.

Standardization takes place in the form of a transformation by ranks or by the ratio of the parameter's standard deviation to its mean. The effect of the standardization is to remove the influence of units and place all parameters on an equal level. The calculation of a rank regression coefficient is a fairly simple procedure, however, the standardized regression coefficient (SRC) is slightly more rigorous and is achieved by,

$$\frac{(\hat{Y} - \bar{Y})}{s} = \sum_k \frac{b_k s_k}{s} \frac{(Z_k - \bar{Z}_k)}{s_k} \qquad (4)$$

where each Z_k is a function of (X_1, ..., X_n), s is the standard deviation of the output, and s_k is the standard deviation of the inputs. If each Z_k is a function of only one parameter in X, then the value of $b_k s_k / s$ is the standardized regression coefficient for parameter X_k, where k = 1 to n.

2.3 *Partitioning techniques*

These statistical tests involve some form of dividing or segmenting input parameters into two or more empirical distributions based on an associated partitioning of the output distribution (Crick et al. 1987). The tests are utilized to compare the characteristics of the input distributions created by the partitioning.

Standard parametric tests are not reasonable on input data sets generated by random sampling methods because of our limited knowledge of the input variables and their associated distributions. Nonparametric statistical tests, therefore, are used where the data are considered to be distribution-free. The convention stated earlier, that Y is a function of X, is no longer appropriate; a new notation is used and specified for each test. The two partitioning techniques considered here, the Smirnov test and the Cramer-von Mises test, operate on ranks of the raw data.

The Smirnov test operates on the two empirical distributions $S_1(X)$ and $S_2(X)$, generated as a result of partitioning the input parameter values. The degree of similarity between distributions, measured by the test statistic, is used to indicate the degree of sensitivity between the input and output values.

The Smirnov test statistic can be measured directly as the greatest vertical distance between two distribution functions plotted on the same graph or the test statistic can be calculated using,

$$T_1 = \sup \left| S_1(X) - S_2(X) \right| \qquad (5)$$

where "sup" is the abbreviation for supremum and the equation represents the greatest absolute difference between $S_1(X)$ and $S_2(X)$ (Conover 1980).

The Cramer-von Mises test is very similar to the Smirnov test in that its purpose is to determine whether two empirical distributions are statistically identical. The computation of the test statistics is slightly more complicated, yet there is little difference in the test's power compared to the Smirnov statistic (Conover 1980). The Cramer-von Mises statistic, T_2, is the sum of all squared vertical distances between the two empirical distributions,

$$T_2 = \frac{m\,n}{(m+n)^2} \left[S_1(X) - S_2(X) \right]^2 \qquad (6)$$

where the values of n and m are the number of samples utilized to estimate the distributions. It is expected that the parameter rankings based on the Smirnov and Cramer-von Mises tests will be very similar since the two tests show little difference in their statistical power. In either case, a large statistic is indicative of a larger difference in the two empirical distributions generated by the division of input data based on some output criteria. This large difference indicates a greater correlation between the independent and dependent variables.

For this example, the output data have been partitioned at the 90th percentile to show how the rank order may change when analyzing conditionalized out-

put. Input values resulting in a given output then will be used to generate the Smirnov and Cramer-von Mises test statistics. Partitioning the data at some other point, i.e. the mean, may result in a different ranking.

2.4 *Variance-based sensitivity analysis techniques*

In ANOVA-like sensitivity analysis techniques (see (Archer et al., 1997) for a review) the total variance V of the model output is apportioned to all the input factors X_i, i =1,2,…,k as

$$V = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + \ldots + V_{12..k} \qquad (7)$$

where

$$V_i = V(E(Y|X_i = x_i^*)),$$

$$V_{ij} = V(E(Y|X_i = x_i^*, X_j = x_j^*)) - V(E(Y|X_i = x_i^*)) + -V(E(Y|X_j = x_j^*))$$

and so on. In the above formulas, Y denotes the output variable, X_i denotes the i[th] input factor, $E(Y|X_i = x_i^*)$ denotes the expectation of Y conditional on X_i having a fixed value x_i^*, and $V_i = V(E(Y|X_i = x_i^*))$ stands for variance over all the possible values of X_i. Conditional variances V_i are estimated for the purpose of sensitivity analysis where the sensitivity index S_i for the factor i is given by V_i/V. The reason for that is intuitive: if the inner mean varies considerably with the selection of a particular value x_i^* for X_i, while all the effects of the X_j's, $j \neq i$ are being averaged, then surely factor X_i is an influential one. Estimation procedures for S_i are classical FAST (Fourier Amplitude Sensitivity Test) (Cukier et al. 1973), the method of Sobol' (Sobol' 1990), and others (Iman and Hora 1990; McKay 1996).

Higher order sensitivity indices, responsible for interaction effects among factors, are usually not estimated as in a model with k factors, the total number of terms in the decomposition (7) is as high as 2^k-1. However interactions may have a strong impact on the output uncertainty especially when k is large and factors are varied on a wider scale, as happens in numerical modelling.

The $V(E(Y|X_i = x_i^*))$ estimate, although powerful, does not satisfy the requirements for a quantitative global sensitivity analysis method: it falls short of accounting for interactions. A global method should be efficient, i.e. capable to cope with the curse of dimensionality mentioned above.

An example of global method is the extended FAST (Saltelli et al. 1999), which is capable of estimating the total sensitivity index, S_{Ti}, defined as the sum of all effects (first and higher order) involving factor X_i (see Saltelli 1999 in these Proceedings). The computation of the total sensitivity indices S_{Ti} makes the analysis affordable. Further, the extended FAST method allows the simultaneous computation of the first and total effect indices.

2.5 *A special technique: the extended FAST*

In the extended FAST each uncertain input factor X_i is associated to a frequency ω_i, and a set of standardised parametric equations

$$X_i = G_i(\sin \omega_i s) = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin \omega_i s) \qquad (8)$$

allows each factor to be explored globally across its range of variation, as the parameter s is varied over $(-\pi;\pi)$. The parametric equations define a curve that systematically explores the unit hypercube $\Omega = (\mathbf{X}|0 \le X_i \le 1; i = 1,\ldots,k)$, from which standard samples of non-correlated input factors that are uniformly distributed in the range [0,1] can be generated. Using the space-filling parametric curve given in (8) summary statistics on the output can be computed, by integrating either along the curve itself instead over Ω.

To evaluate the pair of indices (S_i,S_{Ti}) for the factor i, a curve is defined by choosing a 'high' value for the frequency ω_i, associated to X_i and a set of 'low' values for the other frequencies, $\{\omega_{(-i)}\}$, usually $\{\omega_{(-i)}\} = 1$, which are associated to the remaining factors $\mathbf{X}_{(-i)}$. A better coverage of the complementary space $\mathbf{X}_{(-i)}$ (the index -i stands for 'all but i') can be obtained using a different strategy (Saltelli et al. 1999), which consists in adopting different frequencies for the factors $\mathbf{X}_{(-i)}$, provided that they are similar and, in any case, much lower than ω_i. The output $Y = f(X_1(s), X_2(s),\ldots, X_k(s))$ is evaluated along the curve and is considered as a function of

s. The spectrum $\Lambda^2(\omega)$ of $f(s)$ at each frequency ω is computed as $\Lambda^2 = A^2 + B^2$, where

$$A(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\cos\omega s\, ds \text{ , and}$$

$$B(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\sin\omega s\, ds$$

are integrals numerically evaluated over s. The above integrals are computed using quadrature formulas that operate on a set of N points, which are selected along the curve and are equally spaced. N is related to the frequency ω_i through the Nyquist theorem $N \geq 2M\omega_i + 1$, where M, usually set to 4, is the number of higher harmonics considered in the computation of sensitivities. Hence, fixing ω_i, the minimum number of points along the curve is determined. With N points on a curve is possible to derive estimates for D_i, $D_{(-i)}$ and D:

$$\hat{D}_i = 2\sum_{p=1}^{M}\Lambda^2(p\omega_i)\text{ , }\hat{D}_{(-i)} = 2\sum_{j=1}^{\omega_i - 1}\Lambda^2(\omega_j)\text{ ,}$$

$$\hat{D} = 2\sum_{j=1}^{N/2}\Lambda^2(\omega_j)$$

Finally, $\hat{S}_i$ and $\hat{S}_{Ti}$ are obtained, respectively, by $\hat{D}_i / \hat{D}$ and $1 - \hat{D}_{(-i)} / \hat{D}$. To estimate the pair of sensitivity indices for another factor j, a permutation of the frequencies is necessary, because a 'high' frequency value must be assigned to the factor of interest. This computation requires a new sampling of the input space Ω. Hence, the total cost of the analysis for computing all the pairs of indices is $k \leftrightarrow N$.

3 *The simulation model*

The sensitivity of a tritium dose model to nineteen input parameters has been determined using five different methods generally accepted as appropriate for simple multiplicative and additive models. The sensitivity methods include partial rank correlation, standardized regression techniques, two data partitioning methods, and the extended FAST procedure. The model on which these techniques were executed is described below.

Total atmospheric radiation dose to a downwind receptor is the sum of the inhalation and ingestion pathway doses, given a known concentration of atmospheric tritium, C^a,

$$D_{total} = \frac{DFC^a Rf_w}{H}\left(\frac{1.5BH}{Rf_w} + 1000(U_p f_p) + 1000(U_l f_l) + (2.74F_b f_b Q_b e^{-\lambda t_b} U_b) + (2.74F_m f_m Q_m e^{-\lambda t_m} U_m)\right)$$

where the constants account for unit conversions. Parameter characteristics are given in Table 1. The five additive components in the set of parentheses represent the five primary exposure pathways of 1) inhalation, 2) produce consumption, 3) leafy vegetable consumption, 4) beef consumption, and 5) milk consumption, respectively.

Table 1. Parameter distribution characteristics for the atmospheric tritium dose model.

No.	Param.	Dist. Type*	Units	Mean[Ü]	Std. Dev.[ß]
	C^a	constant	Bq/g	100	-
	λ	constant	1/d	1.54×10^{-4}	-
1	H	N	g/m^3	11.3	0.53
2	DF	LN	Sv/Bq	2.2×10^{-11}	1.6
3	f_w	T	-	0.86	0.77; 0.95
4	R	N	-	0.54	0.1
5	f_b	U	d/kg	-	0.002; 0.02
6	f_m	U	d/L	-	0.002; 0.02
7	Q_b	N	kg/y	19	4.2
8	Q_m	N	kg/y	13	2.9
9	F_b	T	-	0.75	0; 1
10	F_m	T	-	0.56	0; 1
11	t_b	LN	d	6.4	1.4
12	t_m	LN	d	3.1	1.5
13	B	N	m^3/y	8500	1700
14	U_b	LN	kg/y	79	2.0
15	U_m	LN	L/y	77	2.1
16	U_p	LN	kg/y	150	2.0
17	U_l	LN	kg/y	47	1.8
18	f_p	T	-	0.75	0.5; 1
19	f_l	T	-	0.75	0.5; 1

*Distribution type: N=normal; LN=lognormal; T=triangular; U=uniform

ÜArithmetic mean for normal, geometric mean for lognormal, and mode for triangular distributions

ßStandard deviation for normal, geometric standard deviation for lognormal, and minimum and maximum values for triangular and uniform distributions

The radiation dose to downwind receptors from releases of tritium is calculated by this method at many nuclear facilities in the United States, and throughout the world. It is the same specific-activity method prescribed by the U.S. Nuclear Regulatory Commission (1977).

4 RESULTS

The sensitivity rankings for each test are presented in Table 2, with the top 5 parameter rankings given in

bold. The differences in ranks are clearly visible, however, the same parameters consistently appear as those which are more influential on model results.

Table 2. Sensitivity rankings of the nineteen parameters in the atmospheric tritium dose model. Total sensitivity indices by the extended FAST are given in the last column.

Param.	PRCC	SRC	Smirnov	Cramer	FAST	$\hat{S}_{Ti}$
H	13	17	18	16	8	.08
DF	**1**	**2**	**2**	**2**	**1**	**.78**
f_w	14	19	9	10	15	.06
R	**3**	**5**	**5**	**5**	8	.08
f_b	10	18	13	13	15	.06
f_m	7	13	11	8	7	.09
Q_b	8	7	12	12	19	.04
Q_m	17	11	17	18	8	.08
F_b	17	14	8	7	18	.05
F_m	17	14	16	17	**4**	**.13**
t_b	14	10	15	15	8	.08
t_m	14	12	10	11	13	.07
B	**3**	**3**	**4**	**4**	**3**	**.19**
U_b	**5**	**4**	7	9	13	.07
U_m	10	7	6	6	15	.06
U_p	**2**	**1**	**1**	**1**	**2**	**.30**
U_l	**5**	16	13	14	8	.08
f_p	10	6	**3**	**3**	6	.12
f_l	8	9	19	19	**4**	**.13**

In the last column of Table 2 the estimates of total sensitivities, $\hat{S}_{Ti}$, obtained by using the extended FAST, are given. These estimates are much more informative than simple qualitative parameters' ranking. The computational cost of this analysis can be quantified in N*k=1843 model executions, where k=19 and N has been fixed to 97. Extended FAST also allows estimating sensitivities when parameters are grouped according to different logical levels. For instance, in this study the 19 parameters could be partitioned into the five primary exposure pathways plus another group accounting for dose unit conversion (the parameter *DF* alone). The results would be a set of total sensitivity indices for the six groups.

5 DISCUSSION

The five sensitivity methods presented herein provide similar results in terms of rank-ordered parameters. No two techniques offer the same result, however, the methods are consistent in indicating that the most influential input parameters are DF, U_p, and B. Significant discrepancies in rank order are generally not present. However, the results from the conditionalized sensitivity analyses, Smirnov and Cramer-von Mises tests, show that the model becomes more sensitive to some parameters (e.g., f_p and F_b) when the output is conditionalized on the high estimates (90th percentile) of atmospheric tritium dose. This suggests that, for routine assessments, one set of parameters is most important, while for worst-case scenarios, a different group of inputs may dominate the uncertainty in the dose calculation.

Extended FAST ranks f_l as the fourth most important parameter, whereas the other methods attribute a lower rank to f_l (see Table 2). This means that f_l contributes to the model output uncertainty through interactions with other parameters. In fact, the $\hat{S}_{Ti}$'s provided by FAST account for interaction effects among parameters, while the other sensitivity methods presented herein do not supply this information. An indicator of how far the model is from being additive is given by $1 - \sum_i S_i$, which is 0.21 in this study. This means that 21% of the output variance *V* is accounted for by interactions among the parameters.

Percentile cobweb plots (Cooke 1998) for the output of atmospheric tritium dose and the nineteen input parameters were generated and are given below for output groups of 0-5, 50-55, and 95-100th percentiles, respectively. These plots are intended to provide a visual indication of the important parameters for a given output range. Each line represents one iteration of the tritium dose calculation, with a data point at each parameter indicating its value (in percentile) for that calculation. A qualitative assessment of parameter sensitivity is conducted by noting areas of high and low density of intersecting lines.

Based on the results of Table 2, the model is most sensitive to parameters 2, 4, 13, and 16. The cobweb plots, however, seem to indicate that parameters 3, 11, and 12 are important for very low values calculated for dose (0-5th percentile), and parameters 14, 16, and 17 are important for very high calculated dose values (95-100th percentile). Intermediate values of dose, as shown by the 50-55th percentile plot, do not seem to suggest that any particular parameter is highly influential.

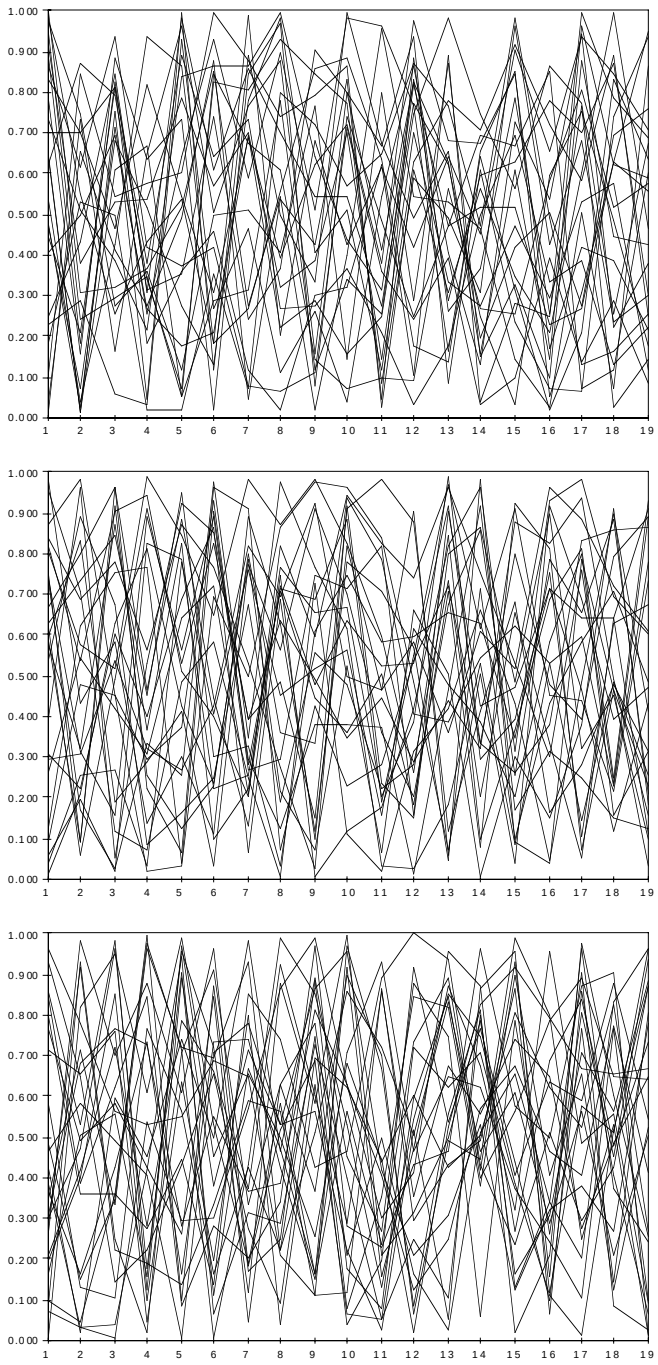

Figure 2. Cobweb plots for the 19 input parameters and output percentile ranges of 0-5, 50-55, and 95-100th percentile, respectively.

REFERENCES

Archer, G., A. Saltelli, and I. M. Sobol', 1997, Sensitivity measures, ANOVA like techniques and the use of bootstrap, *Journal of Statistical Computation and Simulation*, **58**, 99-120.

Conover W.J., 1980, Practical nonparametric statistics, 2nd ed., New York, John Wiley & Sons.

Cooke, R.M., 1998, Local probabilistic importance measures for comparing form and Monte Carlo calculations illustrated with dyke ring calculations, In: *Proceedings of the Second International Symposium on Sensitivity Analysis of Model Output*, Joint Research Center, European Commission, EUR 17758 EN, 71-74.

Crick, M.J., M.D. Hill, D. Charles, 1987, The role of sensitivity analysis in assessing uncertainty, In: *Proceedings of an NEA Workshop on Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems*, Paris, OECD, 1-258.

Cukier, R. I., C. M. Fortuin, K. E. Schuler, A. G. Petschek, J. H. Schaibly, 1973, Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory, *The Journal of Chemical Physics*, **59** (8), 3873-3878.

Downing, D.J., R.H. Gardner, F.O. Hoffman, 1985, An examination of response-surface methodologies for uncertainty analysis in assessment models, *Technometrics*, **27**, 151-163.

Gardner, R.H., R.V. O'Neill, J.B. Mankin, J.H. Carney, 1981, Application of error analysis based on a stream ecosystem model, *Ecological Modelling*, **12**, 173-190.

Hamby, D.M., 1993, A probabilistic estimation of atmospheric tritium dose. *Health Physics*, **65**, 33-40.

Helton, J.C., J.W. Garner, R.D. McCurley, D.K. Rudeen, 1991, Sensitivity analysis techniques and results for performance assessment at the Waste Isolation Pilot Plant, Albuquerque, NM, Sandia National Laboratory, SAND90-7103.

International Atomic Energy Agency, 1989, Evaluating the reliability of predictions made using environmental transfer models. Vienna, Safety Series No. 100, STI/PUB/835, 1-106.

Saltelli, A., S. Tarantola, K. Chan, 1999, A quantitative, model independent method for global sensitivity analysis of model output, *Technometrics*, **41** (1), 39-56.

Saltelli, A. 1999, The role of sensitivity analysis. Transparency, relevance, robustness and parsimony in modelling, *Proceedings of ESREL '99*

Sobol', I. M., 1990, Sensitivity estimates for nonlinear mathematical models, *Matematicheskoe Modelirovanie*, **2**, 112-118, translated as: I. M. Sobol', 1993, Sensitivity analysis for nonlinear mathematical models, *Mathematical Modelling & Computational Experiment*, **1**, 407-414.

Iman, R. L., and S. C. Hora, 1990, A robust measure of uncertainty importance for use in fault tree system analysis, *Risk Analysis*, **10** (3), 401-406.

McKay, M. D., 1996, Variance-Based Methods for Assessing Uncertainty Importance in NUREG-1150 Analyses, *LA-UR-96-2695*, 1-27.

U.S. Nuclear Regulatory Commission, 1977, Calculation of annual doses to man from routine releases of reactor effluents for the purpose of evaluating compliance with 10 CFR Part 50, Appendix I, Washington, DC: Regulatory Guide 1.109.