CONF-850206--1

METHODS AND COMPUTER CODES FOR PROBABILISTIC SENSITIVITY AND UNCERTAINTY ANALYSIS

Jussi K. Vaurio
Fast Reactor Safety Technology Management Center
Argonne National Laboratory
Argonne, IL 60439

CONF-850206--1

DE84 015818

ABSTRACT

This paper describes the methods and applications experience with two computer codes that are now available from the National Energy Software Center at Argonne National Laboratory. The purpose of the SCREEN code is to identify a group of most important input variables of a code that has many (tens, hundreds) input variables with uncertainties, and do this without relying on judgment or exhaustive sensitivity studies. The purpose of the PROSA-2 code is to propagate uncertainties and calculate the distributions of interesting output variable(s) of a safety analysis code using response surface techniques, based on the same runs used for screening. Several applications are discussed, but the codes are generic, not tailored to any specific safety application code. They are compatible in terms of input/output requirements but also independent of each other, e.g., PROSA-2 can be used without first using SCREEN if a set of important input variables has first been selected by other methods. Also, although SCREEN can select cases to be run (by random sampling), a user can select cases by other methods if he so prefers, and still use the rest of SCREEN for identifying important input variables.

INTRODUCTION

This paper addresses the following two problem areas of sensitivity and uncertainty analysis:

A. How to determine a group of most important (influential) input parameters of a large computer code that has many input variables and is too expensive to run exhaustively through all parameter variations.

B. How to obtain probabilistic characteristics of the output variables of a large code.

Solutions to these problems can be used to focus both experimental and physical modeling work to important phenomena. It is sometimes possible to solve problem A by developing an adjoint code. By combining the forward and backward solutions one can calculate necessary sensitivity coefficients. Problem B can sometimes be solved by straightforward random sampling of input variables and running the code through sufficient number of cases to obtain distributions and/or moments of the output variables. However, this paper is focused on very large codes--cases when both of those methods as such are too expensive. The methods and application experience with two techniques (computer codes) are described, SCREEN[1] addressing problem A with statistical methods, and PROSA-2 addressing problem B by first solving a response surface, an analytical function to approximate an output variable as a function of the important input variables, and then random (Monte-Carlo) sampling of the response surface to obtain the probability distribution of the output variable.[2] Critical issues and comparisons between alternative methods are discussed, and the application experience with physical problems is summarized.

STATISTICAL SCREENING

It is often useful to determine the important or "effective" input variables of a complex computer code containing a great many input variables. It is advisable to do this as objectively as possible, without recourse to engineering judgement or physical intuition. Since computer runs with the code are time consuming and expensive, it is desirable to estimate the relative magnitude of the sensitivity coefficients by performing a numerical experiment using a limited number of runs followed by a statistical analysis of the results. Such a procedure is sometimes referred to as "statistical screening" of input variables. The problem is characteristically such that the number of input parameters, m (= number of unknown sensitivities in a linear case) is larger than the number of runs, n, that can be afforded in the study. The situation in a general form is that n output values y_i are available, i.e., can be calculated with a computer code being evaluated, each with different values x_{ij} of the m ($n < m$) input variables x_j:

$$y_i = f(x_{i1}, x_{i2}, \ldots, x_{im}), \quad i = 1, 2, \ldots, n . \qquad (1)$$

The task is to select values of x_{ij} and a procedure that will point out *important* input variables x_j without actually knowing the functional form of f. Important input variables are those that contribute most to the total variation of an output variable. Thus, the output variable is *sensitive* to these input variables, and significant *uncertainties/variations* are associated with these input variables so that the combination of the two characteristics contribute to the output uncertainty/variation.

There are two fundamentally different classes of statistical methods available, one based on overdetermined, the other on underdetermined systems of equations. The basic features of both are described and compared. There are two kinds of errors that can occur in the screening process: (1) false

(2) failure to identify an important parameter as such. Statistical tests are pointed out to assess the likelihood of such errors.

Overdetermined Methods

Overdetermined methods are defined here as methods that do not try to solve all sensitivities but to identify a group of most influential parameters by solving a series of overdetermined systems of equations. Strategies for selecting parameter groups and sequences based on best fits (smallest residual errors) are documented in the statistical literature.[3] These have been used and further developed for the SCREEN methodology.

Case i is defined by the values of the input variables x_j, that yield the output y_i, i.e., by row i of the matrix $\underline{X}$ with elements x_{ij} ($i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$). Because it is important to obtain information from the whole range of values, a natural way to select cases is by randomly sampling input values from their distributions, from the variation interval. Eight distributions are available in the SCREEN code, including uniform, normal, exponential, truncated normal, beta and log-normal distributions. If a distribution is not well-known, or if one wants to make sure that cases are included from a far tail, one could use a wide uniform distribution. However, the rest of the SCREEN procedures are independent of how the cases are selected. If a user prefers to select his cases by systematic or stratified sampling, he can still use SCREEN for ranking input parameters.

A warning should be given against selecting input values at few discrete levels (rather than from a continuum of values). Such practice is common in many classical response surface techniques based on factorial designs. They tend to lead to singular matrices. Using two levels for each input variable, for example, allows linear sensitivities to be estimated, but they can be confounded with higher order terms, with little or no additional information available as to the degree of aliasing.

The screening is based on statistical methods available for selecting "best" regression equations. It starts with a stagewise correlation analysis. The correlation coefficient between an input variable and an output variable is proportional to both the sensitivity coefficient and the uncertainty of the input variable, which product correctly reflects the importance. The impact of the input parameter most correlated with the output variable is subtracted from the data, and a new correlation analysis is performed for the residuals. The impact of the input parameter most correlated with the residual data is then subtracted, and a new correlation analysis is performed for the new residuals, and so on. This correlation procedure identifies parameters with strong linear or monotonic effect on the output. Statistical F-tests are available to assess the probability that an unimportant input variable has a specified level of correlation when calculated from the sample.[1] Students' t-tests can be defined to assess the probability that a significant input parameter (with certain true correlation) appears as an insignficant one.[4]

Screening continues with a stepwise regression analysis employing multivariate second-degree functions to identify parameters that may be misranked with the correlation procedure. The regression is performed first with quadratic functions of two parameters, going through all possible parameter pairs and identifying pairs that give a good fit to the data. The results obtained are used to select a seed group of parameters for the next phase. The regression analysis is then performed successively with overlapping groups of parameters, retaining parameters that have been found to be important and replacing the rest, going repeatedly through all input variables. At each step the residual variance of the regression model of the parameter group is used as a criterion in selecting parameters for the next step. A measure of "spuriousness" has been developed that is the product of the residual variance and the total sensitivity of the model. It is also used as a selection criterion because the standard method of using residual errors alone sometimes fails to identify an important input variable.

For a general nonlinear regression analysis the significance of adding an input variable to the model is tested by comparing the relative improvement in the residual sum of squares to an $F(M-K, n-M)$-variable, when K is the number of coefficients in the seed model and M the number of coefficents in the tested (larger) model.

The SCREEN code also includes techniques to identify input variables that contribute to threshold effects (discontinuities) of an putput variable. A threshold means that the total sample of n cases splits into subsets that have different mean values and other statistics for each of the input variables. Student's t-tests and extreme value tests can be used to detect significant differences between the subsets for input variables that determine which side of the threshold a particular case belongs to.

Underdetermined Methods

Underdetermined methods are defined here as methods that try to solve all sensitivities as a special solution to a fundamentally underdetermined system of equations. The additional criterion used in Refs. 5 and 6 to select the special solution from an infinity of possible solutions was to minimize the sum of the squares of the estimated sensitivities, partial derivatives (i.e., the length of the sensitivity vector). Although a version of this technique has been included in the current SCREEN code, several limitations for this technique have been identified recently:

(1) The method tends to yield sensitivities that are too small by a factor n/m in the average. Unbiased estimates are obtained by multiplying by a factor $\alpha = m/n$. However:[7]

(2) The unbiased estimate does not satisfy the original equations between input and output values;

(3) The length of the unbiased estimation vector is by a factor of $\sqrt{\alpha}$ larger in the average than the true sensitivity vector (when one exists);

(4) The error variance of the estimation vector is $(\alpha-1)$ times the square length of the true sensitivity vector;

(5) There is no residual error criterion or statistical test to measure the quality of the solution obtained.

In spite of these limitations, the screening efficiency of the underdetermined method is about the same as that of simple correlation coefficients between input and output values.

Oblow and Perey[8] have recently modified the underdetermined method by assuming joint-normal prior distributions for the sensitivities to be estimated, and calculating the posterior distributions after n observations (output values). This approach yields the original Durston-Krieger result[6] as a special case; it does not in general correct the problems identified above (bias and a large error vector), unless fortuitious prior information is available. There is no method or rules available on how the prior should be selected to correct the problems.

There is one kind of problem that is difficult to solve with any statistical methods. That is, when all (or many) variables have equal sensitivity/importance. In such a case models of one or few input variables dont "explain" the output variable, which is reflected by the fact that the residual error is not significantly smaller than the total variance of original data. This feature and statistical F-tests available for the overdetermined method reveal (i.e., "warn" the user) that the tried models may not be significant. Other methods should then be seeked to solve the problem.

RESPONSE SURFACE TECHNIQUE

The role of the PROSA-2-code[2] is first to fit a response surface, an analytical function of the important input variables (identified by SCREEN or other means) to the observed (calculated) output values, and then perform random sampling of the input values and calculate a large sample with the response surface to obtain a histogram for the output variables. In summary, the features of the code are:

(1) Several ways of selecting the knot-points, i.e., cases on which the fitting is done. This includes random selection, i.e., same cases that may have been used for screening.

(2) Global and regional quadratic polynomial response surfaces, and global third degree surfaces. A weighting scheme is available to synthesize a smooth global surface from regionwise surfaces.

(3) Residual errors of the fitted surfaces are available to assess the quality and accuracy of the fit.

(4) Functional transformations can be input-specified for both input and output variables to "linearize" the surface in terms of transformed variables. Improvement of trials can be monitored from the residual errors.

(5) Several distributions and correlated sampling of input variables are available for the Monte-Carlo simulation phase.

(6) Analytical functions can be sampled from a subroutine, without external calculations.

(7) Conditional distributions of output (consequence) variables can be calculated. Either fixed or random conditions can be specified for one or more variables, and the distributions of other output variables are then calculated (sampled) under such conditions.

Inverse polynomial surfaces have been found useful for certain applications. Even though they are not internally part of possible functions in PROSA, they can be found by defining a new output variable (= product of input variables divided by the actual consequence value at that point) and using its values instead of the original ones as input to PROSA.

APPLICATIONS

Code Validations

The first applications of the screening methodology were mathematical problems for which the order of importance among 500 input parameters was known in advance. These sample problems basically validated the code and methods used.[9] Correct ranking was approached with increasing number of cases n to the level 40 to 80, compared to the number of variables ($m = 500$). Stagewise correlation and successive regression methods were found to be more reliable and stable than simpler correlation and underdetermined methods. Test runs with the original Durston-Krieger sample problem also revealed severe limitations when selecting input values from two discrete levels only. Consequently, the method was modified for the SCREEN-code to use input values selected from a continuum of values.

Next SCREEN was applied to VENUS-II, a code that calculates energetic consequences that might result from a hypothetical core disruptive accident in a fast breeder reactor.[10] Thirteen material properties and initial conditions were varied for screening purposes. Five different output variables were studied, each measuring accident energetics in some integral manner (work potential, mass of vaporized and molten fuel, etc.). Results based on 40 cases were consistent with one-at-a-time sensitivity studies. PROSA-2 was used to find response surfaces as functions of the six most important parameters identified by SCREEN, and the surfaces were used in Monte-Carlo simulation to obtain histograms. A number of validation studies with PROSA-2 have been described in Ref. 11.

Design-basis Loss of Coolant Accident (LOCA)

The probability of fuel pin failures and the expected amount of radioactive release in case of a PWR LOCA have been studied with PROSA-2.[12] The sensitivities and distributions of the number of failed pins and the amount of release were obtained using key features of the code, in particular, the capability to calculate at the same time the probability of failures and the distribution of release in case of failures (= conditional distribution). This kind of information cannot be obtained with analytical stress-strength overlapping methods. Other features of this application included:

- Parametrization of core-wide ("global") variables through equations like

$$F = F_0 + x F_1, \text{ and} \tag{2}$$

$$G = yG_u + (1 - y) G_\ell , \tag{3}$$

where x and y are variable parameters with distributions used in PROSA, F_0 is a nomial value and F_1 a standard deviation of a physical quantity (e.g., fission gas release fraction), and G_u and G_ℓ are the upper and lower limits of a quantity or function (e.g., axial or radial power distribution), respectively.

- Application of response surface techniques on core-wide variables,

- Direct Monte Carlo sampling method for variability between nominally-like fuel pins ("local" statistics), and

- Synthesis of the resulting disributions by convolution.

Functional transformations available in PROSA-2 were also used for input and output variables in this application to improve the accuracy of response surfaces (by making them smoother, more linear, in terms of the transformed variables). Results indicated that the probability of exceeding the peak cladding temperature 2200°F in case of a design basis LOCA for a particular PWR design is less than 0.01. The probability of releasing more than 10% of the available I-131 inventory is about 1%.

Fuel Behavior in Normal Operation

The thermal characteristics and fission gas release in a Halden BWR fuel pin as functions of irradiation level (burnup) have been studied with PROSA.[13] Uncertainties in input variables associated with the fission gas release model and fuel thermal conductivity were found to be most influential for fuel temperature, stored energy, gap conductance, fission gas release and rod internal pressure. Less important were cladding manufacturing tolerances and fuel densification and shifting. Uncertainties are growing in all output consequences as the burnup increases. Global and regionwise response surfaces were used in this study, and significant differences were observed mostly in the high burnup region (>10 MWD/kg).

Time to Core Uncovery

The systematic way of selecting knot-points, and the simple analytical (non-matrix) equations for the coefficients of a quadratic response surface, as implemented in PROSA, can be used easily in manual calculations, outside of the code for which they were originally developed by the author. They have been found useful for obtaining a response surface for the time-to-core-uncovery in case of a small LOCA. It has been used in calculating distributions of the core uncovery times in TMI-type accidents at various plants,[14] and the core uncovery probability under the influence of human errors.[15] The initial screening of input variables in these studies (elimination of 10 out of 16 variables) was based on one-at-a-time variation sensitivities rather than regression analysis.

Design Conservatism and Safety

General design criteria, regulatory guides and current licensing practice, 10CFR Part 50 (Appendices A, K) generally require applicants to analyze postulated accidents using conservative assumptions and values for physical, geometric, and modeling parameters. Statistical combination of uncertainties is allowed because combining all conservative assumptions at the same time would lead to unnecesarily large margins and penalize plant operational and thermal economy. Reactor vendors have started to develop techniques to realistically calculate distributions for critical quantities such as the maximum cladding temperature under LOCA conditions, or the minimum Departure from Nucleate Boiling Ratio (DNBR). Because of a large number of uncertain input variables, these efforts so far have used engineering judgment (or one-at-a-time sensitivity studies) to eliminate some variables from probabilistic considerations, to determine "conservative" values for others, and then use response surfaces for the remaining relatively few parameters in probabilistic calculations. There is a need to use a more unified, systematic approach to this problem, with less reliance on subjectivity and judgment. Review of earlier results also indicates that problems arise from using few (two or three) discrete levels of values for input variables as customary in classical fractional factorial designs. They tend to lead to singular matrices (that would not arise if values are sampled from a continuum of values) and confounding of lower and higher order terms of the response surfaces. The codes described here, SCREEN and PROSA-2, should be able to contribute to a systematic treatment of significant portions of this problem, "search of reasonable margins." Efforts in this direction have been initiated.[16]

Special Problems

Theoretical controversies that have surrounded statistical screening methods earlier have been largely solved.[7] However, certain practical aspects have limited their usefulness, especially for codes that analyze very complex physical phenomena. Most difficult is a case that has several threshold effects within the domain of a problem. A fast reactor accident analysis is a typical such problem: variations in input parameters can cause a case to fall on either side of a multitude of discontinuities, thresholds: prompt criticality or not, energetic fuel coolant interactions in a channel or not (in multiple different channels) due to variations in coolant liquid/vapor boundary location, fuel-cladding gap closed or not, deviation from nucleate boiling, etc. Different response surfaces (or regression models) would be valid on different sides of each threshold, and a large sample might be needed to have enough cases in each region for sensitivity screening. In such a case one tries to select an integral, robust output quantity (consequence), but there is no guarantee that any of the known sensitivity methods would be successful. Even if adjoint methods would yield sensitivities, they are valid only around a specific point. They do not tell how far from that point the sensitivities are valid in different directions, or which input variables most likely bring a case across a threshold.

CONCLUSION

In summary, statistical methods for parameter importance screening and uncertainty propagation have been reviewed and illustrated by several applications. Two computer codes are now available, fully tested and documented, from the National Energy Software Center:

SCREEN: to identify a group of most important input variables of a code that has many (tens, hundreds) input variables with uncertainties.

PROSA-2: to propagate uncertainties and calculate the distributions of interesting output variable(s) of a safety analysis code (using response surface techniques).

These codes are completely generic, not tailored to any specific safety application code. They are also independent of each other, e.g., PROSA-2 can be used without first using SCREEN, if a set of important input variables has first been selected by other methods. Also, although SCREEN can select cases to be run (by random sampling), a user can select cases by other methods if he so prefers, and still use the rest of SCREEN for identifying important input variables.

REFERENCES

1. J. K. Vaurio, "Statistical Identification of Effective Input Variables," ANL-82-57, Argonne National Laboratory (1982).
2. J. K. Vaurio, "PROSA-2: A Probabilistic Response-Surface Analysis and Simulation Code," ANL-81-33, Argonne National Laboratory (1981).
3. N. R. Draper and H. Smith, Applied Regression Analysis, John Wiley, (1966).
4. Z. W. Birnbaum, Introduction to Probability and Mathematical Statistics, pp. 226-233, Harper & Brothers, (1962).
5. C. Durston, T. Krieger, and D. C. Albright, "Sensitivity Analysis Methods Application to Code Validation," BNL-NUREG-23430, Brookhaven National Laboratory (1977); see also T. J. Krieger, C. Durston, and D. C. Albright, Trans. Am. Nucl. Soc., 23, 515 (1978).
6. T. J. Krieger, Nucl. Sci. Engng., 81, 281 (1982).
7. J. K. Vaurio, "On Statistical Identification of Important Input Variables to Large Codes," to be published in Nucl. Sci. Engng., (1984).
8. E. M. Oblow and F. G. Perey, "Screening Sensitivity Theory," to be published in Nucl. Sci. Engng., (1984).
9. J. K. Vaurio, Trans. Am. Nucl. Soc., 32, 296 (1979); 35, 263 (1980).
10. J. K. Vaurio, J. M. Kyser, and L. Mync, "Statistical Sensitivity Analysis of a Safety Evaluation Code," Proc. Intl. Fast Reactor Safety Technology, Seattle, WA, August 19-23, 1979.
11. J. K. Vaurio, "Response Surface Techniques Developed for Probabilistic Analysis of Accident Consequences," Proc. ANS Topical Mtg. Probabilistic Analysis of Nuclear Reactor Safety, May 8-10, 1978, Los Angeles, CA; Vol. 2.
12. R. Sairanen, L. Mattila, and J. K. Vaurio, "Probabilistic Analysis of Fuel Damage and Radioactivity Release During an LWR LOCA;" Proc. ANS Mtg., Reactor Safety Aspects of Fuel Behavior, August 2-6, 1981, Sun Valley, ID.
13. J. Rossi and R. Sairanen, "Uncertainty Analysis of Fuel Rod Steady-state Thermal Behavior by the Probabilistic Response Surface Method," presented at the IAEA Specialists' Meeting on Water Reactor Fuel Element Performance Computer Modeling, Bowness-on-Windermere, U.K., April 9-13, 1984.
14. T. L. Chu and G. E. Apostolakis, "Assessment of the Uncertainties Associated with the Core Uncovery Time in TMI-type Accidents," Reliability Engineering, 8, 23-56 (1984).
15. G. Apostolakis and T. L. Chu, "Time-Dependent Accident Sequences Including Human Actions," Nucl. Tech., 64, 115-126, (1984).
16. C. Lee and L. W. Deitrich, private communication, Argonne National Laboratory, (1984).