Reliability Engineering and System Safety **21** (1988) 1–26

Comparative Study on Uncertainty and Sensitivity Analysis and Application to LOCA Model

Tae Woon Kim,* Soon Heung Chang
and Byung Ho Lee

Department of Nuclear Engineering, Korea Advanced Institute of Science and Technology, PO Box 150, Cheongryang, Seoul, Republic of Korea

(Received 2 August 1987)

ABSTRACT

Nowadays, statistical approaches are used in the design and operation stages of large engineering systems. In this context, this paper summarizes the characteristics of various conventional methods of uncertainty and sensitivity analysis, mainly in the point of view of uncertainty propagation, sensitivity analysis, and reduced model construction capabilities. Then, a new technique based on the Fourier amplitude sensitivity test and stepwise regression technique is suggested and applied to the thermal margin analysis for peak clad temperature at loss of coolant accident, which is one of the most important parameters in the design and operation of nuclear reactors.

1 INTRODUCTION

One of the most influential concepts on current thermal hydraulic design of a nuclear power plant is conservatism. Currently, the uncertainties of the mechanistic models and the design and operational parameters are conservatively treated. Thereby, the resulting operating margin of the nuclear power plant thus obtained is greatly reduced compared to that which really exists.

In 1975, Rasmussen first introduced a probabilistic approach to risk assessment of nuclear power plant in the report 'Reactor Safety Study'.[1]

* Address for correspondence: Reactor Safety Department, Korea Advanced Energy Research Institute, PO Box 7, Daeduk, Choongnam, Republic of Korea.

Reliability Engineering and System Safety 0951-8320/88/$03·50

This approach treats the uncertain quantities as realistically as possible so that the evaluated output value revealed is a realistic one. Since that time, the probabilistic or statistical approach to the risk, safety, and operating margin evaluation for large engineering systems is recognized to be a useful tool.

This approach is largely made up of three steps. In step 1, the values of input parameters needed in the system function evaluation are collected and evaluated. In step 2, the system functions are evaluated by an appropriate mechanistic model with the input set obtained in step 1. In step 3, the characteristics of the system are assessed with the output values. The above procedure may be interpreted as an uncertainty propagation analysis in a narrow sense.

The uncertainty is largely divided into three types;[2] modeling uncertainty, parametric uncertainty, and completeness. The modeling uncertainty is that which results from uncertainties of the mechanistic model, and it is analyzed by the change and development of the mechanistic model by experiment or numerical simulation. The parametric uncertainty is the one propagated from the uncertainties of the system model input parameters. The completeness is the extent to which the analyzer can recognize the nature of the system. Nobody knows the true value; thus, the confidence level is defined in a statistical manner depending on the size of the sample. Most researches in engineering belong to the mechanistic model development. In this paper, however, the main concern is limited to the parametric uncertainty. Therefore, the main purpose of this paper is to link the mechanistic approach and the statistical approach. There exist, at present, many methods to treat and analyze the parametric uncertainty. The parametric uncertainty analysis is usually performed with the following three aspects.

(1) To assess the propagated uncertainties from the uncertainties of the input parameters (uncertainty propagation analysis).
(2) To seek for the most dominant factors which contribute to the output uncertainty (sensitivity analysis).
(3) To construct the reduced system model composed of the most dominant input parameters (model reduction).

Most of the conventional methods, however, have different characteristics and are used only for local purposes. The conventional methods used up to now are reviewed here, and improved methods, which have the capability to consider the above aspects simultaneously, are developed with the following points of view for the application to the nuclear power plant design and operation stage.

In the design stage of a nuclear power plant, many large computer programs are needed. In practice, it is impossible to evaluate statistically the

outputs of these time-consuming computer programs. Therefore, it is desirable to extract information with the smallest number of runs possible.

In deciding the operating margin of a nuclear power plant, the most important thermalhydraulic factors are DNBR (departure from nucleate boiling ratio)[3,4] and PCT (peak clad temperature) at LOCA (loss of coolant accident).[5,6] To deal with the PCT problem statistically, various methods of uncertainty and sensitivity analyses are studied comparatively and a new technique based on the Fourier amplitude sensitivity test and stepwise regression technique is suggested.

2 METHODS OF UNCERTAINTY AND SENSITIVITY ANALYSES

2.1 Three features of the uncertainty analysis

Various conventional methods of uncertainty analysis are reviewed in this section. The main features of each method are described with the point of view of the three capabilities mentioned in section 1, i.e. the capabilities of uncertainty propagation analysis, sensitivity analysis, and reduced model construction.

The uncertainty propagation analysis includes the estimation of distribution function of the output variable and its representative values, for instance, mean, median, variance, standard deviation, and higher order moments. The Monte Carlo simulation technique is a representative of this type of analysis.

The measure of sensitivity is divided into two categories. One of these is the differential sensitivity measure,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x_i}\right|_{x_i = x_i^0}$$

The linear sensitivity theory and adjoint sensitivity theory are the representative ones of this category. The other is the global sensitivity measure,

$$\int \cdots \int \left|\frac{\partial y}{\partial x_i}\right| \mathrm{d}x_1 \cdots \mathrm{d}x_k$$

of which the Fourier amplitude sensitivity test (FAST) is representative. The adjoint sensitivity theory[7-9] is a powerful method when the number of input parameters is large, because this method gives the sensitivity of output variable for all the parameters with only single run. A major defect of this method is that a separate computer program has to be made for each

system model to be analyzed. This method is not considered here because the purpose of this paper lies in the development of statistical methods.

The methods which have the capability of the reduced model construction are the response surface method and the stepwise regression technique. These methods also provide the sensitivity measure as a by-product.

Latin hypercube sampling, the experimental design technique, and the Fourier amplitude sensitivity test are used to generate systematic sampling points. The least square method and analysis of variance are used as a basic tool for the response surface method and the stepwise regression technique.

2.2 Crude Monte Carlo method[10-12]

The crude Monte Carlo (CMC) method is the most primitive, but the most powerful, method in uncertainty propagation analysis. First, it generates random samples whose size is greater than a thousand according to the distribution shape of each input parameter. Each random combination of random samples $(x_1^i, x_2^i, \ldots, x_k^i)$, $i = 1, \ldots, n$ is passed through the system function to generate the corresponding output value y_i. The resultant y_is represent the uncertainty distribution of the output variable y. The mean, variance, and higher order moments of y are calculated from these y_is. The distribution function of y is obtained directly from these y_is by making histogram of tens or hundreds of equal intervals. Figure 1 illustrates this process. Sometimes the distribution function of y is matched to an appropriate empirical distribution shape by a moment matching technique using some lower order moments calculated above. The major deficiency of

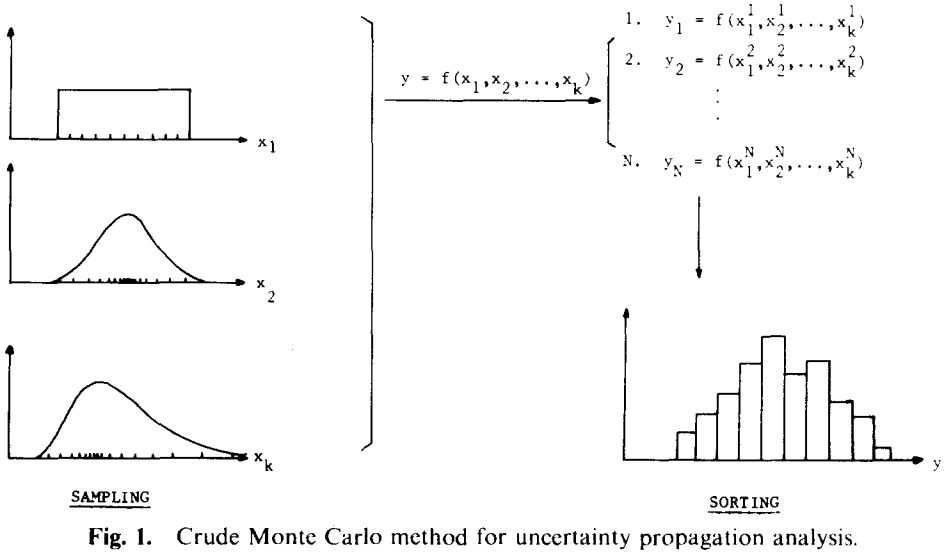


Fig. 1. Crude Monte Carlo method for uncertainty propagation analysis.

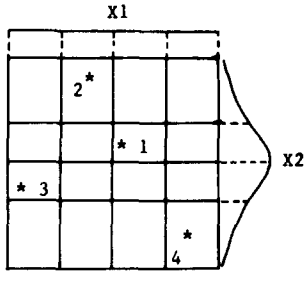
the CMC method is that the computing cost is too high due to the large sample size n. Thus, it is not used directly, in practice. To avoid this problem, the following procedure is usually used. First, the complex system function is simplified to an equivalent low order polynomial by the response surface technique and then the CMC is applied to this polynomial to construct the output distribution.

2.3 Latin hypercube sampling[13-15]

Latin hypercube sampling (LHS), which was originally proposed by McKay *et al.*,[13] is a kind of variance reduction technique to generate a sample of size n from k random variables, $X1$, $X2, \ldots, Xk$. The range of each variable is divided into n non-overlapping intervals of equal probability. One value from each interval is selected at random. The n values thus obtained for $X1$ are paired at random with the n values obtained for $X2$. These n pairs are combined in a random manner with the n values for $X3$ to form n triples. The process is continued until a set of k-tuples is formed. This resultant set of k-tuples is the Latin hypercube sample. Figure 2 illustrates this process in a two dimensional case with $n = 4$. This technique is good because it samples without a large sampling size n. Sample size n is sufficient if $n \geq 2k$.

2.4 Experimental design technique (EDT)[16-19]

There are several methods in designing experiments to establish a set of sampling points in the space of the Xs, at which Y will be observed. The most common of these are two and three level factorial design, two and three level fractional factorial design, and central composite design, which are discussed in this section.



X1 = uniform distribution

X2 = normal distribution

Fig. 2. An illustration of Latin hypercube sampling points in a two dimensional case with $n = 4$.

Two level factorial design utilizes two levels of X (coded value of ± 1). When the number of factors (here, number of input parameters) is k, the required number of code runs is 2^k, for the full (complete) factorial design and 2^{k-p} for fractional design, where 2^p is the fraction needed to reduce the required numbers. In the fractional factorial design, $k-p$ parameters are combined completely but remaining p parameters are composed of a product of already defined $k-p$ parameters levels. Obviously, the higher the degree of fraction is, the less is the degree of resolution in analyzing the lower order effects (main or cross term effects compared to the higher order interaction effects).

Three level design is the same as two level design except that it needs three levels of $X(-1, 0, +1)$. But in these designs the number of runs is increasing more rapidly than that of two level designs as the number of input parameters increases.

In central composite design, however, three distinct portions are included: (1) two level factorial points, (2) two axial points for each parameter, (3) one center point. Figure 3 illustrates three design points for two factors, where a is an arbitrary coded value.

Among the above mentioned type of experimental designs, the two level factorial design and the central composite design are most commonly used for response surface work. The reason is that the former can be used to fit a first order polynomial with a minimum number of code runs, while the latter is suitable for fitting the second order polynomial. The major advantage of the central composite design is utilizing the information of two level design to consider the non-linearity with a few additional data points in case the constructed linear model is not satisfactory.

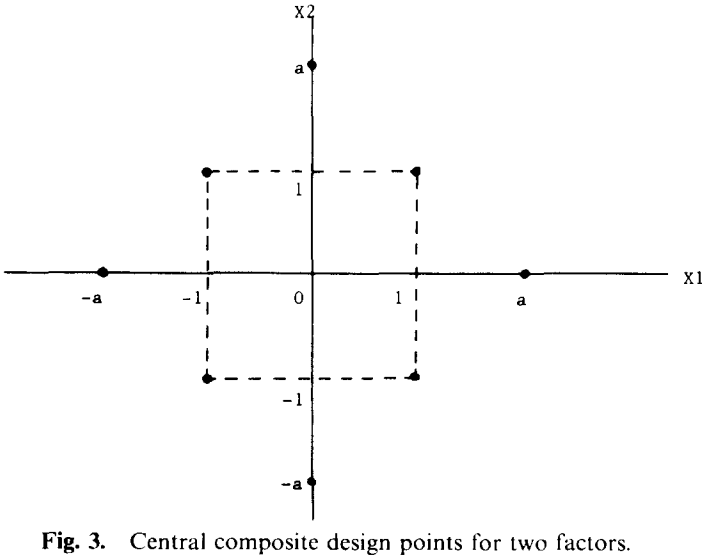


Fig. 3. Central composite design points for two factors.

TABLE 1
Comparison of the Number of Unknowns and Code Runs between Various Experimental Design Techniques

No. of factors k	*1st order polynomial*			*2nd order polynomial*			
	No. of unknowns	*No. of runs*		*No. of unknowns* $2k + {}_kC_2 + 1$	3^k	*No. of runs*	
		2^k	2^{k-p}			3^{k-p}	*Central composite design* $2^k + 2k + 1$
3	4	8	$2^{3-1} = 4$	10	27		15
4	5	16	$2^{4-1} = 8$	15	81	$3^{4-1} = 27$	25
5	6	32	$2^{5-1} = 16$	21	243	$3^{5-1} = 81$	43
			$2^{5-2} = 8$			$3^{5-2} = 27$	
7	8	128	$2^{7-1} = 64$	36	2 187	$3^{7-1} = 729$	143
			$2^{7-2} = 32$			$3^{7-2} = 243$	
			$2^{7-3} = 16$			$3^{7-3} = 81$	
			$2^{7-4} = 8$				

Table 1 shows the comparison of the number of unknowns and code runs required for these experimental design methods. It is obvious that the central composite design needs a relatively small number of runs compared to the three level factorial design.

2.5 Response surface method (RSM)[16,17,20]

This method is usually used to generate an approximate relationship between the input parameters and the output variable considering the system as a blackbox.

When the output variable y is a complex function of a number of input parameters, $x_i (i = 1, 2, \ldots, k)$, say

$$y = f(x_1, x_2, \ldots, x_k) \tag{1}$$

the input–output relationship can be approximated in the form of polynomial as

$$y = b_0 + \sum_i^k b_i x_i \qquad \text{(1st order regression)} \tag{2}$$

or

$$y = b_0 + \sum_i^k b_i x_i \sum_{i<j}\sum^k b_{ij} x_i x_j \quad \text{(2nd order regression)} \tag{3}$$

where

$$x_i = \frac{z_i - z_i^0}{\Delta z_i} = \text{coded value or level} \tag{4}$$

z_i = real value

z_i^0 = nominal value

Δz_i = unit deviation

The term response surface refers to the geometrical interpretation of a function of several independent variables. If the functional relationship is not highly non-linear in the interested region of independent variables, then the first order regression equation can be used. Otherwise, the second or higher order regression equation will fit the relationship reasonably well. Sometimes, even if the nonlinearity is obvious in the wide range of independent variables, the first order model can be used in the narrow interested region of independent variables.

It is the purpose of RSM to generate a simplified polynomial equation from a sampled data set and to extract information for the unknown system from this response surface equation. The coefficients are obtained by the least square method and the characteristic features of the system are extracted by using an analysis of variance (ANOVA). ANOVA is described in the following section. To generate a sample without computer simulation, EDT, described in the previous section, is usually used.

2.6 Multiple least square and analysis of variance[16,21,22]

Equations (2) and (3) can be represented by matrix form as

$$\mathbf{y} = X\mathbf{b} + \mathbf{e} \tag{5}$$

where

$\mathbf{y} = n \times 1$ observed data vector

$X = n \times m$ design matrix

$\mathbf{b} = m \times 1$ coefficient vector

$\mathbf{e} = n \times 1$ error vector, i.e. the difference between observed data and estimated response

n = number of observations (runs)

$$m = \begin{cases} 1 + k & \text{for 1st order regression} \\ 1 + 2k + {}_kC_2 & \text{for 2nd order regression} \end{cases}$$

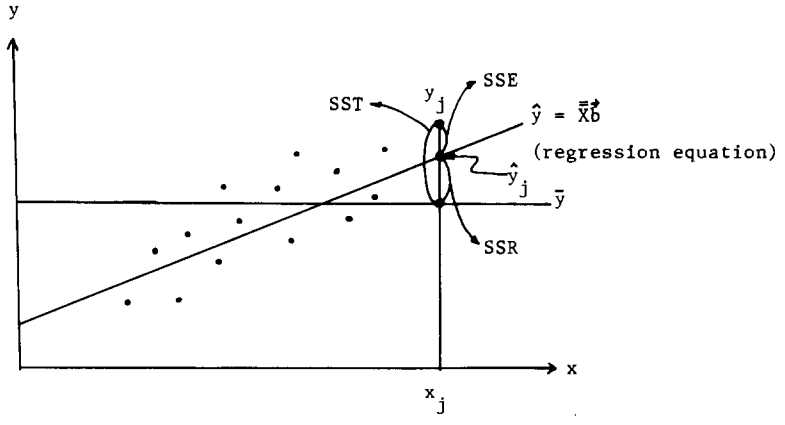


Fig. 4. An illustration of analysis of variance for a single factor case.

The best estimated regression coefficients b can be found by the method of least squares. If the sum of the squares of the error

$$L = \mathbf{e}^{\mathrm{T}}\mathbf{e} = (\mathbf{y} - X\mathbf{b})^{\mathrm{T}}(\mathbf{y} - X\mathbf{b}) \tag{6}$$

is minimized, i.e.

$$\frac{\partial L}{\partial b} = 0$$

then the regression coefficient is represented by

$$\mathbf{b} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\mathbf{y} \tag{7}$$

The analysis of variance (ANOVA) is used to distinguish between the variation due to regression and the variation due to residual errors from the total variation. Figure 4 illustrates this for the single variable case.

Total variation is usually called total sum of squares (SST) and is defined by

$$\mathrm{SST} = \sum(y_j - \bar{y})^2 = \sum y_j^2 - n(\bar{y})^2 = \mathbf{y}^{\mathrm{T}}\mathbf{y} - n(\bar{y})^2 \tag{8}$$

Variation due to regression equation is called sum of square due to regression (SSR) and is defined by

$$\mathrm{SSR} = \sum(\hat{y}_j - \bar{y})^2 = \sum \hat{y}_j^2 - n(\bar{y})^2 = \hat{\mathbf{y}}^{\mathrm{T}}\hat{\mathbf{y}} - n(\bar{y})^2 \tag{9a}$$

Since $\mathbf{y} = X\mathbf{b}$

$$\mathrm{SSR} = \mathbf{b}^{\mathrm{T}}X^{\mathrm{T}}\mathbf{y} - n(\bar{y})^2 \tag{9b}$$

Finally, the sum of squares due to residual errors (SSE) can be represented by

$$\mathrm{SSE} = \mathrm{SST} - \mathrm{SSR} \tag{10}$$

TABLE 2
ANOVA Table for Multiple Regression

	Source					
	SS	*DF*	*MS*	*F*	*Critical-F*	R^2
Regress.	SSR	k	MSR	(MSR/MSE)	$F(k, n-k-1; \alpha)$	(SSR/SST)
Error	SSE	$n-k-1$	MSE			
Total	SST	$n-1$				

where SS = Sum of squares
DF = Degrees of freedom
MS = Mean of squares = (SS/DF)
n = Number of observations (runs)
k = Number of regression factors
R^2 = Coefficient of determination

The degrees of freedom for total variation, the regression equation, and residual errors are $n-1$, k, and $n-k-1$, respectively. From this information, an ANOVA table for multiple regression can be made as shown in Table 2.

There are three measures which describe the accuracy of regression equation compared to the observed data. These are MSE (mean of squares due to residual error), F (=MSR/MSE), and R^2 (=SSR/SST, the so called coefficient of determination). The smaller the MSE is, the better is regression, while the reverse situation is applied for F and R^2. It is said that the regression is significant at a level of $100(1-\alpha)\%$ if

$$F > F(k, n-k-1; \alpha) \tag{11}$$

When the significance of each component (main effect term, quadratic term, cross-term) of regression is considered, a partial F-test can be applied, which is defined by

$$F_i = \frac{b_i^2}{(X^T X)_{ii}^{-1}(\text{MSE})} \tag{12}$$

It is also said that the component X_i is significant at a level of $100(1-\alpha)\%$ if

$$F_i > F(1, n-k-1; \alpha) \tag{13}$$

2.7 Stepwise regression technique (SRT)[16,21,22]

SRT is used to select the most important input parameters, thereby to build a regression equation, composed of a not undue number of input

parameters, while the constructed regression equation reveals the input–output relationship. This procedure selects or removes the most important, or the least important, input parameter sequentially. At each step to decide the adequacy of the constructed regression model composed of the selected input parameters, ANOVA is used.

In selecting input parameters at each step, the partial F value is used. The input parameter which has the largest partial F value is selected. For the selected input parameter, the F test is performed. If $F_i > F(1, n-p-1; \alpha)$, where p is the number of input parameters included in the regression model, the input parameter is included in the model and then goes on to the next step. Otherwise, the procedure is stopped here. The meaning is that the input parameter inclusion is significant at a level of $100(1-\alpha)\%$.

2.8 Fourier amplitude sensitivity test (FAST)[23-28]

Consider a system that is described by an ordinary differential equation containing k input parameters, $x_1, x_2, \ldots, x_k$,

$$\frac{dy}{dt} = f(y; x_1, \ldots, x_k) \tag{14}$$

where

y = interested system output at time t

$\mathbf{x}$ = k-dimensional parameter vector

The basic problem is to determine the sensitivity of y to simultaneous variations in all the parameters $\mathbf{x}$. This is done by considering that $\mathbf{x}$ has a distribution of values resulting from either imprecision or uncertainty in their definition. The ensemble mean value of y is then given by

$$\langle y \rangle = \int \cdots \int y(t; x_1, \ldots, x_k) p(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \tag{15}$$

where

$y(t; x_1, \ldots, x_k)$ = the solution of eqn (14)

$p(x_1, \ldots, x_k)$ = joint probability density function of $\mathbf{x}$

The central idea of the FAST method is to convert the k-dimensional integral of eqn (15) into an equivalent one-dimensional form by using the transformation

$$x_i = G_i[\sin(w_i s)],\ i = 1, \ldots, k \tag{16}$$

where

G_i = a set of known functions

w_i = a set of frequencies

s = a scalar variable

By means of this transformation, the variations of the k parameters are transformed into variations of the single scalar variable s. By variation of s over the range $-\infty \leq s \leq \infty$, eqn (16) traces out a space-filling curve in the m-dimensional parameter space. For a suitable choice of G_i, which transforms the probability density function $p(\mathbf{x})$ into s-space, Weyl demonstrated that

$$\bar{y} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y(t; x_1(s), \ldots, x_k(s))\, \mathrm{d}s \tag{17}$$

is identically equal to $\langle y \rangle$ in eqn (15). Equation (17) is the fundamental expression in the FAST method for computing the mean value, variance, and other properties of the output y.

By using an appropriate integer frequency set $\{w_i\}$, the parameters, x_i are periodic in s on the finite interval $(-\pi, \pi)$, in which case eqn (17) becomes

$$\bar{y} = \frac{1}{2\pi} \int_{-\pi}^{\pi} y(t; x_1(s), \ldots, x_k(s))\, \mathrm{d}s \tag{18}$$

The variance of y is then

$$\sigma_y^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} y^2(t; x_1(s), \ldots, x_k(s))\, \mathrm{d}s - (\bar{y})^2 = \overline{y^2} - (\bar{y})^2 \tag{19}$$

The evaluation of σ_y^2 can be carried out by using the s-space Fourier coefficients of y. From Parseval's theorem

$$\overline{y^2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} y^2\, \mathrm{d}s = \sum_{j=-\infty}^{\infty} (A_j^2 + B_j^2) \tag{20}$$

where the Fourier coefficients A_j and B_j are defined by

$$A_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} y \cos(js)\, \mathrm{d}s \tag{21}$$

$$B_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} y \sin(js)\, \mathrm{d}s \tag{22}$$

Thus, from eqns (21) and (22)

$$(\bar{y})^2 = A_0^2 + B_0^2 = A_0^2 \tag{23}$$

Then

$$\sigma_y^2 = 2 \sum_{j=1}^{\infty} (A_j^2 + B_j^2) \tag{24}$$

The variances due to w_i and its harmonics are expressed by

$$\sigma_{wi}^2 = 2\sum_{k=1}^{\infty}(A_{kwi}^2 + B_{kwi}^2) \tag{25}$$

The normalized sensitivity measure, partial variance, S_{wi}, is defined by the ratio of the variance due to the frequency w_i to the total variance as follows

$$S_{wi} = \frac{\sigma_{wi}^2}{\sigma_y^2} \tag{26}$$

Restricting the frequency set to odd integers reduces the range of s to $[-\pi/2, \pi/2]$. In this case

$$\begin{aligned} y(\pi - s) &= y(s) \\ y(\pi + s) &= y(s) \\ y(\pi/2 + s) &= y(\pi/2 - s) \\ y(-\pi/2 + s) &= y(-\pi/2 - s) \end{aligned} \tag{27}$$

and the Fourier coefficients can be expressed as

$$A_j = \begin{cases} 0 & j \text{ odd} \\ \dfrac{1}{\pi}\displaystyle\int_0^{\pi/2} [y(s) - y(-s)]\cos(js)\,\mathrm{d}s & j \text{ even} \end{cases} \tag{28}$$

$$B_j = \begin{cases} 0 & j \text{ even} \\ \dfrac{1}{\pi}\displaystyle\int_0^{\pi/2} [y(s) - y(-s)]\sin(js)\,\mathrm{d}s & j \text{ odd} \end{cases} \tag{29}$$

This can be further reduced by a simple numerical quadrature as

$$A_j = \begin{cases} 0 & j \text{ odd} \\ \dfrac{1}{2n+1}\left\{y_0 + \displaystyle\sum_{k=1}^{n}(y_m + y_{-m})\cos\left(\frac{jm\pi}{2n+1}\right)\right\} & j \text{ even} \end{cases} \tag{30}$$

$$B_j = \begin{cases} 0 & j \text{ even} \\ \dfrac{1}{2n+1}\left\{\displaystyle\sum_{k=1}^{n}(y_m - y_{-m})\sin\left(\frac{jm\pi}{2n+1}\right)\right\} & j \text{ odd} \end{cases} \tag{31}$$

where $y_m(-n \le m \le n;\ n = \omega_{\max})$ is the y which is calculated at an input set

$$x_i^m = G_i\left\{\sin\left(\omega_i \frac{m\pi}{2n+1}\right)\right\}, \qquad i = 1, 2, \ldots, k$$

TABLE 3
Integer Frequency Sets and Number of Sampling Points

Number of variables	*Integer frequency set* $\{\omega_i\}$	*Minimum sampling points*
5	11, 21, 27, 35, 39	79
6	1, 21, 31, 37, 45, 49	99
7	17, 39, 59, 69, 75, 83, 87	175
8	23, 55, 77, 97, 107, 113, 121, 125	251
9	19, 59, 91, 113, 133, 143, 149, 157, 161	323
10	25, 63, 103, 135, 157, 177, 187, 193, 201, 205	411
11	41, 67, 105, 145, 177, 199, 219, 229, 235, 243, 247	495
12	31, 87, 113, 151, 191, 223, 245, 265, 275, 281, 289, 293	587
13	23, 85, 141, 167, 205, 245, 277, 299, 319, 329, 335, 343, 347	695
14	87, 133, 195, 251, 277, 315, 355, 387, 409, 429, 439, 445, 453, 457	915
15	67, 143, 189, 251, 307, 333, 371, 411, 443, 465, 485, 495, 501, 509, 513	1 027
16	73, 169, 245, 291, 353, 409, 435, 473, 513, 545, 567, 587, 597, 603, 611, 615	1 231
17	85, 145, 241, 317, 363, 425, 481, 507, 545, 585, 617, 639, 659, 669, 675, 683, 687	1 375
18	143, 229, 289, 385, 461, 507, 569, 625, 651, 689, 729, 761, 783, 803, 813, 819, 827, 831	1 663
19	149, 275, 361, 421, 517, 593, 639, 701, 757, 783, 821, 861, 893, 915, 935, 945, 951, 959, 963	1 927

It is sufficient if the total sampling points are determined by $N_s = 2\omega_{\max} + 1$ by the Nyquist criterion in the digital signal processing theory.

The rapidly-decaying properties of the Fourier amplitudes ensure that the summation to $k = 2$ is sufficient in the calculation of σ_{wi}^2 in eqn (25), that is

$$\sigma_{wi}^2 = 2(A_{1wi}^2 + B_{2wi}^2) \tag{32}$$

The interference problem, which is caused by multiples of harmonics between different integer frequencies, can be avoided up to a certain degree (order) if the integer frequency sets are selected as follows

$$i \cdot \omega_l + j \cdot \omega_m \neq k \cdot \omega_n, \qquad i, j, k = 0, 1, 2, 3, 4, \ldots \tag{33}$$

The integer frequency sets which avoid the interference problem up to the 4th order are shown in Table 3.

3 LOCA UNCERTAINTY ANALYSIS

3.1 A new method using FAST and SRT

A new method which uses the Fourier amplitude sensitivity test and the stepwise regression technique was developed. This method can estimate

total variance of the output variable, partial variances contributed by each input parameter, and finally, can construct a linear or second order regression polynomial.

The Monte Carlo simulation method is the most powerful to estimate the output distribution itself. However, this method needs more than a thousand sampling points, i.e. system function calculations. If the computing time for one calculation is very large as in the case of the design stage of nuclear power plant, it is impractical to use the Monte Carlo method. The method proposed in this section can estimate the sensitivity of each input parameter with much reduced sampling points.

Figure 5 shows the procedure proposed for uncertainty and sensitivity analysis using FAST and SRT. The procedure is largely composed of three parts: the FAST sampling and variance calculation part, the code run part, and the response surface construction by SRT part.

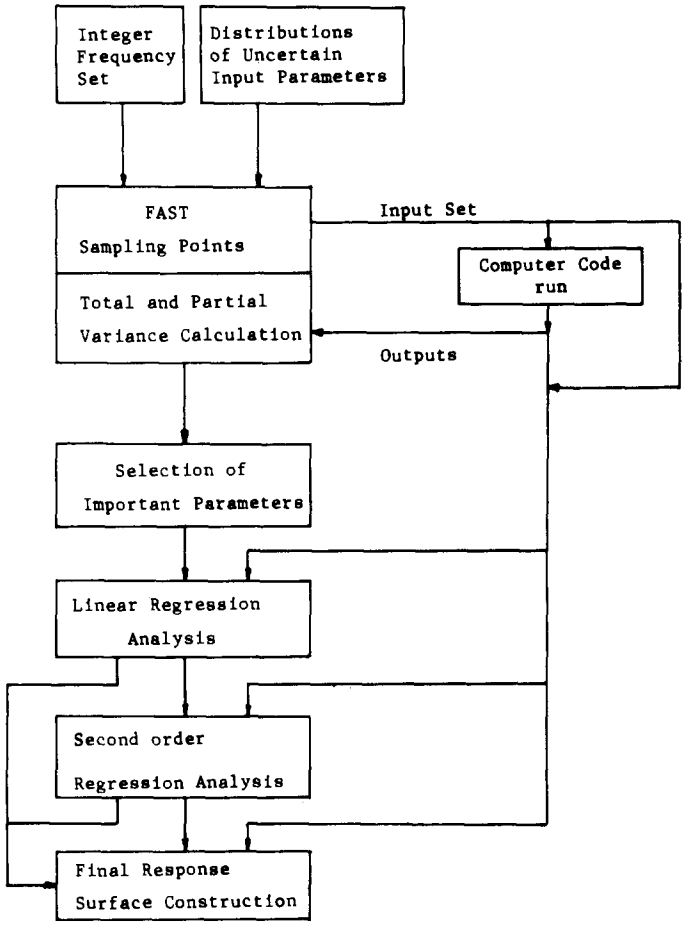


Fig. 5. Proposed procedure of uncertainty and sensitivity analysis using the Fourier amplitude sensitivity test and stepwise regression technique.

First, the FAST generates sampling points using an integer frequency set and the distribution types of input parameters as described in section 2.8. Next, computer code runs are performed according to the input set generated in the above. Then, the FAST calculates total variance and partial variances using the input set and the outputs obtained by the computer code.

Next, the important parameters which contribute to the output variance more than to a certain percentage are selected according to the partial variances calculated in the FAST part. Using these important parameters, the linear regression analysis is performed to construct a first order response surface. If the result reveals that the first order response surface is sufficient, the procedure stops here. Otherwise, the procedure is continued to construct a second order regression model composed of factors which have partial F-values greater than critical F-values. Using the factors selected in the linear and quadratic regression analysis, the final second order regression model is constructed. The stepwise regression technique used here is modified from the original one described in section 2.7. When the model has roughly more than 10 input parameters this modified procedure is more useful because it preselects the important input parameters by FAST and the linear regression part.

The method proposed is applied to the uncertainty analysis of peak cladding temperature during hypothetical loss of coolant accident (LOCA). The operating condition of light water reactors is limited by the peak cladding temperature not exceeding 2200°F. Although the LOCA model described in the subsequent section is a much simplified one, it is used only for demonstration purposes of the proposed method, and to gain some insights through the application.

3.2 Analytic LOCA model[5]

During full-power operation of a nuclear power plant, the temperature of the fuel pellet is considerably greater than that of the cladding material, which drives the heat flux from the fuel to the coolant. If coolant flow is lost, as in an accident situation, the cladding is rapidly heated even though the power generation may drop to a low level. The simple model presented here attempts to approximately handle this process.

Assume that the fuel region is at uniform temperature T_f, the clad is at uniform temperature T_c, and the coolant is at T_w, as shown in Fig. 6. The energy equation for the fuel rods is

$$\rho_f C_f V_f \frac{dT_f}{dt} + \rho_c C_c V_c \frac{dT_c}{dt} = V_f q''' + \frac{T_w - T_c}{R_c + R_s} \tag{34}$$

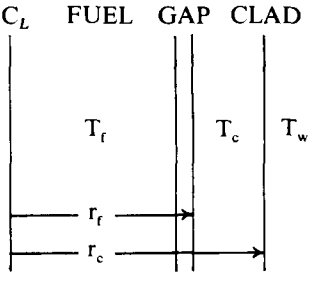


Fig. 6. Schematic of fuel rod.

where

ρ = Density
C = Heat capacity
V = Volume; $V_f = \pi r_f^2$, $V_c = \pi(r_c^2 - r_f^2)$
q''' = Heat generation rate per unit volume
R_c = One half of the heat transfer resistance across the clad
R_s = Convective resistance at rod surface

and the subscripts f, c, and w denote the fuel pellet, the clad, and the coolant water, respectively.

The first two terms refer to the change in stored energy in the fuel and clad, respectively. The $V_f q'''$ is the total heat generation and the last term is the heat flux at the clad surface.

The second equation needed is the energy equation for the fuel region, given as

$$\rho_f C_f V_f \frac{dT_f}{dt} = V_f q''' + \frac{T_c - T_f}{R_f + R_g + R_c} \tag{35}$$

where

R_f = Resistance to heat transfer across the fuel region
R_g = Gap resistance

These heat transfer resistances are defined as

$$R_c = \frac{1}{2}\frac{\ln(r_c/r_f)}{2\pi K_c}$$

$$R_s = \frac{1}{2\pi r_c h_c}$$

$$R_g = \frac{1}{2\pi r_f h_g}$$

$$R_f = \frac{C_0}{2\pi K_f}$$

where

h_g = Gap conductance
h_c = Convective heat transfer coefficient
K = Thermal conductivity
$C_0 = 0{\cdot}295$, A constant chosen to give a reasonable value for R_f

Substituting eqn (34) into eqn (35) and solving for T_f, yields

$$T_f = -(R_f + R_g + R_c)\left[\frac{T_w - T_c}{R_c + R_s} - \rho_c C_c V_c \frac{dT_c}{dt}\right] + T_c \tag{36}$$

Differentiating and substituting eqn (36) into eqn (34) yields

$$\frac{d^2 T_c}{dt^2} + C_1 \frac{dT_c}{dt} + C_2 T_c = C_3 \tag{37}$$

where

$$C_1 = \frac{(\rho_f C_f V_f)\left[1 + \dfrac{R_f + R_g + R_c}{R_c + R_s}\right] + \rho_c C_c V_c}{(\rho_c C_c V_c)(\rho_f C_f V_f)(R_f + R_g + R_c)}$$

$$C_2 = \frac{1}{(R_c + R_s)(\rho_f C_f V_f)(R_f + R_g + R_c)(\rho_c C_c V_c)}$$

$$C_3 = \frac{V_f q''' + \dfrac{T_w}{R_c + R_s}}{(\rho_f C_f V_f)(R_f + R_g + R_c)(\rho_c C_c V_c)}$$

Equation (37) is simply a second order ordinary differential equation whose solution is

$$T_c(t) = A_1 \exp(m_1 \cdot t) + A_2 \exp(m_2 \cdot t) + \frac{C_3}{C_2} \tag{38}$$

where

$$m_1 = \frac{-C_1 + \sqrt{C_1^2 - 4C_2}}{2}$$

$$m_2 = \frac{-C_1 - \sqrt{C_1^2 - 4C_2}}{2}$$

The coefficients A_1 and A_2 may be evaluated by the initial conditions

$$T_c(0) = T_c^0 \tag{39}$$

$$T_f(0) = T_f^0 \tag{40}$$

Substituting eqn (39) into eqn (38), gives

$$T_c^0 = A_1 + A_2 + \frac{C_3}{C_2} \tag{41}$$

Substituting eqn (40) into eqn (37), we get

$$\left.\frac{dT_c}{dt}\right|_0 = \frac{T_f^0 - \dfrac{R_f + R_g + R_c}{R_c + R_s}(T_c^0 - T_w) - T_c^0}{(R_f + R_g + R_c)(\rho_c C_c V_c)} = m_1 A_1 + m_2 A_2 \tag{42}$$

Solving eqn (41) and eqn (42) for A_1 and A_2 yields

$$A_2 = \frac{(T_f - T_c) - \dfrac{R_f + R_g + R_c}{R_c + R_s}(T_c - T_w) + (R_f + R_g + R_c)(\rho_c C_c V_c) m_1 \left(\dfrac{C_3}{C_2} - T_c^0\right)}{(R_f + R_g + R_c)(\rho_c C_c V_c)(m_2 - m_1)}$$

$$A_1 = \frac{(T_f^0 - T_c^0) - \dfrac{R_f + R_g + R_c}{R_c + R_s}(T_c^0 - T_w) + (R_f + R_g + R_c)(\rho_c C_c V_c) m_2 \left(\dfrac{C_3}{C_2} - T_c^0\right)}{(R_f + R_g + R_c)(\rho_c V_c C_c)(m_1 - m_2)}$$

Now that every quantity in eqn (38) is known, the transient value of T_c may be found for the initial conditions of specified initial clad and fuel temperature. Implicitly assumed in the analysis is that the geometry, the material properties, and the heat transfer coefficients all remain constant throughout the transient.

A rapid nuclear reactor transient may be simulated with this model by inputting the initial fuel and clad temperatures with the representative ones of full power operation, and the heat generation rate and the heat transfer coefficients corresponding to a loss-of-flow condition. For instance, the transient of the instantaneous loss-of-flow at time equal to zero, followed by flow resumption after 4 s, may be handled by inputting a low heat transfer coefficient at the clad surface corresponding to the loss-of-flow condition. The power generation rate may be input at approximately 10% full power to simulate an instantaneous scram. Since the flow is resumed after 4 s, the maximum clad temperature, the desired output variable, is merely the value of the clad temperature, T_c, at time equals 4 s. This calculation corresponds to the blowdown peak clad temperature observed in LOCA analysis.

The proper values of the initial fuel and clad temperatures may be calculated by solving eqn (34) and eqn (35) for T_w and T_c at steady state condition, i.e. dT_w/dt and dT_f/dt equal to zero. The appropriate expressions are

$$T_c(0) = V_f q'''(0)(R_c + R_s) + T_w \tag{43}$$

$$T_f(0) = V_f q'''(0)(R_f + R_g + R_c) + T_c(0) \tag{44}$$

The full power value of the heat generation rate and a convective heat transfer coefficient corresponding to normal operation should be used.

3.3 Results and discussions

The model outlined above requires 14 inputs. The mean and the assumed one-side limit values of these parameters are represented in Table 4. It is assumed that all the input parameters have uniform uncertainty distributions and one-side uncertainty limits of 10%. The exceptions are 5% for two manufacturing parameters, fuel radius and clad thickness and 30% for the unmeasurable, thus highly uncertain, parameter, gap conductance.

Table 5 shows the means and variances of the cladding temperature up to 4 s estimated by the FAST and Monte Carlo methods. In the FAST method

$$G(X_i) = \bar{X}_i + \Delta X_i \frac{2}{\pi} \sin^{-1}\{\sin(w_i s)\} \tag{45}$$

where

$\bar{X}_i$ = mean value of X_i
ΔX_i = one-side limit value of X_i

is used for the X_i to result in uniform distribution. Here the Monte Carlo simulation is performed with the purpose of benchmarking with sampling size $n = 2000$. The resulting variances estimated by these two methods are in good agreement.

The time dependent mean, mean +/− one standard deviation, and minimum and maximum output values generated by FAST are illustrated in Fig. 7. The partial variances shown in Fig. 8 represent the time dependent contributions of each input parameter to the total variance of the output variable.

The results of FAST are represented in Table 6, where seven input parameters (initial heat generation rate ($X1$), gap conductance ($X5$), coolant temperature ($X6$), fuel radius ($X7$), fuel conductivity ($X10$), clad density ($X12$), and clad heat capacity ($X13$)), whose partial variances are greater than 3%, are selected to be the major uncertainty sources. The linear regression analysis for these parameters results in an R^2 value of 0·967. The summation of the partial variances greater than 3% results in 97·32%. The difference between these two values is thought to be caused by some non-linear sources. To investigate the non-linearity of the model, the quadratic regression analysis is performed at step 2 for the seven parameters selected above. Resulting major sources of non-linearity are $X5^2$, $X10^2$, $X1X5$, $X1X7$, $X1X10$, $X1X12$, $X1X13$, $X5X12$, $X5X13$, and $X7X10$. These factors are those whose partial F-values are greater than the critical F-value of 3·84. At step 3,

TABLE 4
Input Parameters of LOCA Model

Input parameter	*Name*	*Nominal value*	*Uncertainty (%)*
$Z1$ $q'''(0)$	Initial heat generation rate (Btu/h ft^3)	$5{\cdot}56 \times 10^7$	10
$Z2$ $q'''/q'''(0)$	Transient power level	0·1	10
$Z3$ $h_c(0)$	Initial convective coefficient (Btu/h ft^2 °F)	10 000	10
$Z4$ h_c	Transient convective coefficient (Btu/h ft^2 °F)	10	10
$Z5$ h_g	Gap conductance (Btu/h ft^2 °F)	1 000	30
$Z6$ T_w	Coolant temperature (°F)	588	10
$Z7$ r_f	Fuel radius (ft)	0·152 9	5
$Z8$ ρ_f	Fuel density (lb/ft^3)	640	10
$Z9$ C_f	Fuel heat capacity (Btu/lb °F)	0·09	10
$Z10$ K_f	Fuel conductivity (Btu/h ft °F)	2	10
$Z11$ $r_c - r_f$	Clad thickness (ft)	0·002 29	5
$Z12$ ρ_c	Clad density (lb/ft^3)	500	10
$Z13$ C_c	Clad heat capacity (Btu/lb °F)	0·12	10
$Z14$ K_c	Clad conductivity (Btu/h ft °F)	13	10

TABLE 5
Mean and Variance Estimated by FAST and CMC Methods

Time (s)	*FAST*			*Monte Carlo*	
	Mean	*Var.*	*Uncer.*[a]	*Mean*	*Var.*
0	660	1 167	5·17	661	1 162
1	1 224	4 471	5·46	1 232	4 596
2	1 494	7 806	5·91	1 502	8 192
3	1 626	10 200	6·21	1 632	10 800
4	1 696	11 823	6·41	1 702	12 570

[a] Uncertainty (%) $= \dfrac{(\text{Var.})^{\frac{1}{2}}}{\text{Mean}} \times 100.$

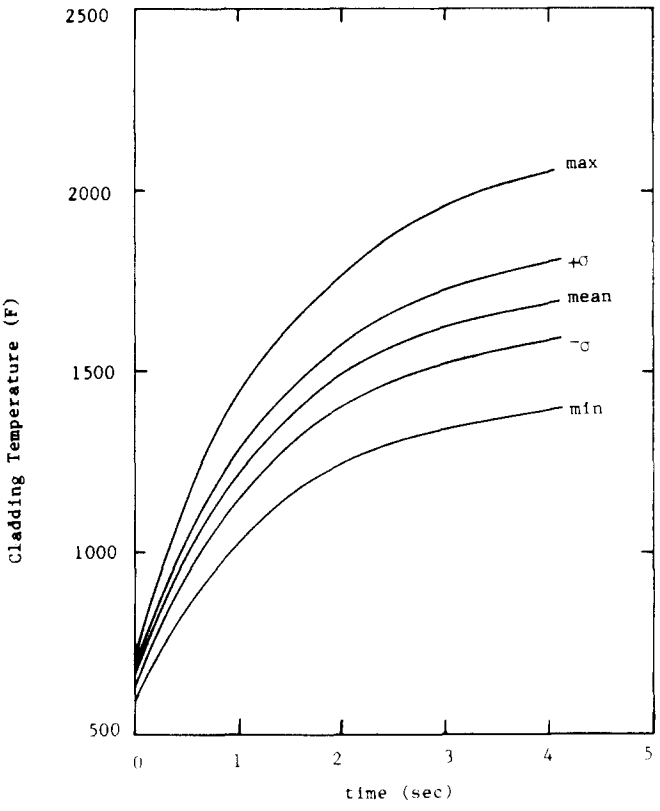


Fig. 7. Time dependent mean and uncertainty of clad temperature estimated by FAST

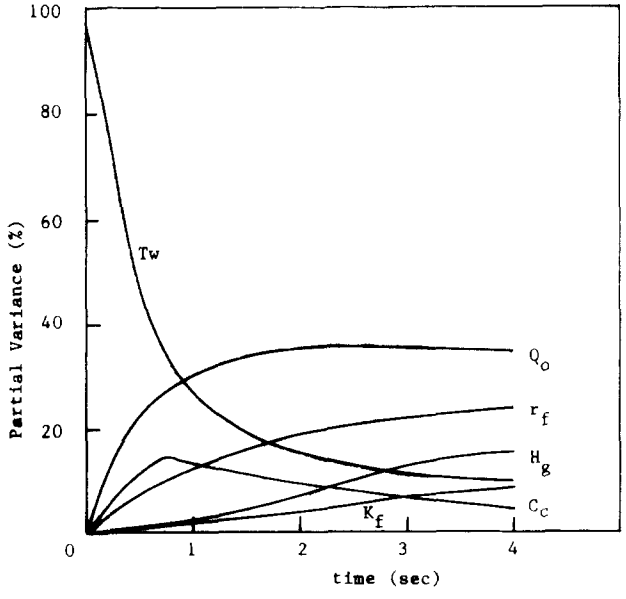


Fig. 8. Time dependent partial variances of clad temperature.

TABLE 6
Results of FAST at 4 seconds

Parameter X	*Partial variance* (%)	
$X1$	4010	33·91[a]
$X2$	9·5	0·08
$X3$	3·2	0·03
$X4$	4·5	0·04
$X5$	1856	15·70[a]
$X6$	1051	8·89[a]
$X7$	2808	23·75[a]
$X8$	93	0·78
$X9$	93	0·79
$X10$	951	8·05[a]
$X11$	105	0·89
$X12$	373	3·15[a]
$X13$	457	3·87[a]
$X14$	9·3	0·08
Sum	11823	100

[a] Represents the factors whose partial variances are larger than 3%.

the final second order regression analysis is performed, including the elements selected from the linear and the quadratic regression analysis. The result is shown in Table 7 and the final regression model is constructed as follows

$$\begin{aligned} Y = 1688 &+ 110{\cdot}8\,X1 - 74{\cdot}6\,\mathrm{X5} + 58{\cdot}7\,\mathrm{X6} \\ &+ 92{\cdot}45\,X7 - 53{\cdot}74\,X10 - 36{\cdot}69\,X12 - 37{\cdot}3\,X13 \\ &+ 20{\cdot}71\,X5 + 4{\cdot}391\,X10 - 7{\cdot}703\,X1X5 \\ &+ 9{\cdot}201\,X1X7 - 5{\cdot}562\,X1X10 - 3{\cdot}695\,X1X13 \\ &+ 4{\cdot}593\,X5X12 + 5{\cdot}0\,X5X13 - 4{\cdot}954\,X7X10 \end{aligned} \tag{46}$$

Note that $X1X12$ is not included in this regression model because the partial F-value of this element becomes less than the critical F-value of 3·84 in this final model.

The resulting R^2 value in this case is 0·972, which is nearly the same as the summed value of 97·32% in FAST as mentioned above. It can be said that this regression model covers 97·2% of the total variation of data points. Thus, it can be concluded that this regression model is well fitted to data points.

TABLE 7
Final Result of Stepwise Regression Analysis

Step 3: Final regression ANOVA table

	SS	*DF*	*MS*	*F*	R^2
Regr.	0·1071E + 08	17·00	0·6299E + 06	1846	0·972
Error	0·3060E + 06	897·0	341·2		
Total	0·1101E + 08	914·0			

	Partial *F*-value	Regression coeff.
X(0) *X*(0):	$F = 0{\cdot}2102E + 07$	$B =$ 1688
X(1) *X*(0):	$F = 0{\cdot}1097E + 05$	$B =$ 110·8
X(5) *X*(0):	$F =$ 4967	$B =$ −74·60
X(6) *X*(0):	$F =$ 3074	$B =$ 58·70
X(7) *X*(0):	$F =$ 7640	$B =$ 92·45
X(10) *X*(0):	$F =$ 2577	$B =$ −53·74
X(12) *X*(0):	$F =$ 1201	$B =$ −36·69
X(13) *X*(0):	$F =$ 1244	$B =$ −37·30
X(5) *X*(5):	$F =$ 90·39	$B =$ 20·71
X(10) *X*(10):	$F =$ 4·583	$B =$ 4·391
X(1) *X*(5):	$F =$ 17·65	$B =$ −7·703
X(1) *X*(7):	$F =$ 25·22	$B =$ 9·201
X(1) *X*(10):	$F =$ 9·208	$B =$ −5·562
X(1) *X*(13):	$F =$ 4·073	$B =$ −3·695
X(5) *X*(12):	$F =$ 5·710	$B =$ 4·593
X(5) *X*(13):	$F =$ 7·449	$B =$ 5·000
X(7) *X*(10):	$F =$ 7·312	$B =$ −4·954

Critical *F*-value = $F(1; 897) = 3{\cdot}840$.

Some points obtained and recommended through this study are as follows

(1) FAST gives the partial variance for each input parameter, which can be used as a sensitivity ranking between input parameters, with moderate sampling points compared to the crude Monte Carlo method.
(2) Modified SRT is a good tool to construct the later-used first or second order response surface model consisting of relatively important parameters.
(3) The combined uncertainty analysis method using FAST and SRT can be used for uncertainty/sensitivity analysis of the large computer codes at moderate cost.
(4) The method proposed in this study would be a useful tool to perform

a structural analysis for the newly developed, highly uncertain models.

(5) For the sake of generality, the FAST method, which can sample other than a uniform distribution type, is required.

REFERENCES

1. Anon., *Reactor Safety Study*, WASH-1400, US Nuclear Regulatory Commission, 1975.
2. Anon., in Uncertainty and Sensitivity Analysis, *Probabilistic Risk Assessment Procedure Guide*, NUREG/CR-2300, US Nuclear Regulatory Commission, Jan. 1983, Ch. 12.
3. H. Chelemer, L. H. Boman and D. R. Sharp, Improved Thermal Design Procedure, WCAP-8567, Westinghouse, Pittsburgh, PA, USA, July 1975.
4. E. Oelkers, Statistical Core Design, NPGD-TM-416, Babcock and Wilcox, Oct. 1977.
5. M. Mazumdar, J. A. Marshall, P. A. Awate, S. C. Chay and D. K. McLain, Review of the Methodology for Statistical Evaluation of Reactor Safety Analysis, EPRI-309, Electric Power Research Institute, Palo Alto, CA, USA, Sept. 1975.
6. M. Mazumdar, J. A. Marshall and S. C. Chay, Methodology Development for Statistical Evaluation of Reactor Safety Analysis, EPRI NP-194, Electric Power Research Institute, Palo Alto, CA, USA, 1976.
7. E. M. Oblow, Sensitivity Theory for Reactor Thermal-Hydraulic Problems, *Nucl. Sci. Eng.*, **68** (1978), pp. 322–37.
8. D. G. Cacuci *et al.*, Sensitivity Theory for Nonlinear Systems: I. Nonlinear Functional Analysis Approach, II. Extensions to Additional Classes of Responses, *J. Math. Phys.*, **22**(12) (1981), p. 2794(I), p. 2803(II).
9. S. J. Wilderman and G. S. Was, Application of Adjoint Sensitivity Analysis to Nuclear Reactor Fuel Rod Performances, *Nucl. Eng. Des.*, **80** (1984), pp. 27–38.
10. Anon., SAMPLE—Monte Carlo Fault Tree Analysis Program, in Reactor Safety Study, WASH-1400, App. II, US Nuclear Regulatory Commission, 1975.
11. SPASM User's Manual, EPRI, RP-2507-1, Electric Power Research Institute, Palo Alto, CA, USA, 1983.
12. T. W. Kim and S. H. Chang, MOCUP—Monte Carlo Uncertainty Propagation program, KAIST, Dept. Nucl. Eng., Seoul, Korea, Oct. 1985.
13. M. D. Mckay, W. J. Conover and R. J. Beckman, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics.*, **21** (1979), 239.
14. R. L. Iman, J. C. Helton and J. E. Campbell, An Approach to Sensitivity Analysis of Computer Models, Parts I and II, *J. Quality Tech.*, **13** (1981), p. 174.
15. R. L. Iman and W. J. Conover, Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment, *Comm. Stat.*, **A9** (1980), p. 1749.
16. S. H. Park, *Regression Analysis*, Daeyoungsa, Seoul, Korea, 1984.
17. S. H. Park, *Modern Experimental Design*, Daeyoungsa, Seoul, Korea, 1984.
18. D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, 1984.

19. J. P. C. Kleijnen, *Statistical Techniques in Simulation, Part 2*, Marcel Dekker, New York, 1975.
20. R. H. Meyers, *Response Surface Methodology*, Allyn and Bacon, Boston, MA, USA, 1976.
21. N. R. Draper and H. Smith, *Applied Regression Analysis*, John Wiley & Sons. New York, 1981.
22. G. A. F. Seber, *Linear Regression Analysis*, John Wiley & Sons, New York, 1977.
23. R. I. Cukier, C. M. Fortuin and K. E. Shuler, Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients; I. Theory, *J. Chem. Phys.*, **59** (1973) p. 3873.
24. J. H. Schaibly and K. E. Shuler, Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients; II. Applications, *J. Chem. Phys.*, **59** (1973) p. 3879.
25. R. I. Cukier, J. H. Schaibly and K. E. Shuler, Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients; III. Analysis of the Approximations, *J. Chem. Phys.*, **63** (1975), p. 1140.
26. R. I. Cukier, H. B. Revine and K. E. Shuler, Nonlinear Sensitivity Analysis of Multiparameter Model Systems, *J. Comp. Phys.*, **26** (1978), p. 1.
27. M. Koda, G. J. McRae and J. H. Seinfeld, Automatic Sensitivity Analysis of Kinetic Mechanisms, *Int. J. Chem. Kinetics*, **XI** (1979), p. 427.
28. G. J. McRae, J. W. Tilden and J. H. Seinfeld, Global Sensitivity Analysis A Computational Implementation of the Fourier Amplitude Sensitivity Test (FAST), *Comp. & Chem. Eng.*, **6** (1982), p. 15.