CONF-800331--1

AUTOMATED UNCERTAINTY ANALYSIS METHODS IN THE FRAP COMPUTER CODES

S. O. Peck

EG&G Idaho, Inc.
Idaho Falls, Idaho U.S.A.

I. INTRODUCTION

A user oriented, automated uncertainty analysis capability has been incorporated in the Fuel Rod Analysis Program (FRAP) computer codes. The FRAP codes have been developed for the analysis of Light Water Reactor fuel rod behavior during steady state (FRAPCON)[1] and transient (FRAP-T)[2] conditions as part of the United States Nuclear Regulatory Commission's Water Reactor Safety Research Program. The objective of uncertainty analysis of these codes is to obtain estimates of the uncertainty in computed outputs of the codes as a function of known uncertainties in input variables. This objective has been accomplished through development of an option that allows a user to perform an uncertainty analysis on any FRAP problem for any choice of probabilistic inputs and outputs desired. This capability will facilitate the following traditional analyses:

(1) Sensitivity Studies: Most phenomenological models are developed independently before being incorporated into a complex computer code. Their actual contributions to the final code output are not necessarily well known. An uncertainty analysis can be used as a sensitivity study in that the influence or importance of each model may be ranked by its effect on the output.

(2) Experimental Data Needs: Determination of the relative contribution of the uncertainty in various models and input variables to the total uncertainty will provide guidelines for future experimental work. For example, if it is found that the uncertainty in fuel thermal conductivity contributes 80% of the uncertainty in cladding temperature, but the uncertainty in fuel Poisson's ratio contributes less than 1%,

then this would suggest that future experimental programs be aimed at refining the uncertainty in thermal conductivity. Thus, an uncertainty analysis will provide the means for the selection of cost effective experiments.

(3) Code Assessment: Determination of the agreement between code predictions and experimental data requires comparing the uncertainty in the code output with the uncertainty in the experimental measurement. In this way only can systematic differences be detected and evaluated. An uncertainty analysis thus provides one half of the necessary information required for code assessment.

This paper presents the methods used to generate an uncertainty analysis of a large computer code, discusses the assumptions that are made, and shows techniques for testing them. An uncertainty analysis of FRAP-T calculated fuel rod behavior during a hypothetical loss-of-coolant transient is presented as an example and carried through the discussion to illustrate the various concepts.

II. EXAMPLE PROBLEM

A FRAP-T analysis of a pressurized water reactor (PWR) hypothetical loss-of-coolant accident (LOCA) was chosen as the example problem. The nominal case was a PWR fuel rod subjected to thermal hydraulic conditions resulting from a 200% cold leg break at 100% power. Beginning-of-life conditions were assumed. The thermal-hydraulic boundary conditions through blowdown were supplied by RELAP4.[3] The thermal-hydraulic conditions from blowdown through reflood were calculated by FRAP-T. The nominal calculations predicted that the fuel rod did not fail during the course of the two hundred and fifty second transient.

The emphasis of the uncertainty analysis was on variables thought to significantly affect the fuel rod behavior during reflood. The variables chosen were core flooding rate, coolant channel flow blockage, FLECHT heat transfer coefficient, moisture carryout fraction, and the ANS decay heat curve. Two additional variables (gap heat transfer coefficient and fuel thermal conductivity) were included because of their known importance to rod behavior during blowdown. The variables and their respective uncertainty estimates are shown in Table 1. Cladding surface temperature was chosen as the dependent variable not only for its importance to licensing but also because it serves as a good example of the uncertainty analysis process.

III. RESPONSE SURFACE METHOD

The uncertainty analysis option is based upon the response surface method.[4] A response is any calculated output of the code. If it were possible to evaluate the response over a range of input variable values a surface could be constructed describing the relationship between the inputs and the response, hence the term "response surface". For very simple codes such a surface could potentially be analytically constructed. However, in the case of the FRAP codes the range of problem types and input values is very large and the response surface can only be evaluated at discrete points. The complete response surface cannot be determined analytically. The response surface method of uncertainty analysis is therefore based on a systematic sampling of the true surface to generate a set of data. These data are used to fit polynomial approximations to the true surface. The polynomials, or response equations, are used to study the effect of propagating errors through the inputs to determine their effect on the output, or response. It is extremely important to note the assumption that the polynomials reasonably approximate the behavior of the true surface. Techniques for testing this assumption will be discussed in Section IV.

The polynomial equation form chosen most frequently to approximate the response surface is a truncated Taylor's series expansion. The expansion is carried out about the nominal code calculation and is usually truncated at second order terms. The generated data are then used to estimate the coefficients of this polynomial through a least squares approach.

The systematic choosing of a pattern for perturbing the independent input variables is known as experimental design. An experimental design may be envisioned as a matrix where the rows correspond to individual analyses to be performed and the columns to the values of the inputs for each analysis. One particular class of experimental designs, known as two level fractional factorial designs, is well suited to generating data to fit truncated Taylor's series expansions. Two level refers to the fact that each independent variable will be input at two different values or levels. Factorial implies that all possible combinations of the independent variables may be estimated (excluding powers of an input such as squares), and fractional is used to indicate that frequently only a portion of the total design is necessary. For example, consider the problem posed in Section II of a seven variable FRAP-T uncertainty analysis of a PWR LOCA. All possible combinations of possible equation coefficients, from a constant to a seven factor crossproduct would equal two raised to the seventh power, or one hundred twenty eight coefficients. However, a Taylor's series truncated at second order will contain only terms up to two factor crossproducts. Furthermore, for the purpose of this example only a linear expansion shall be considered, that is, a constant and seven linear coefficients. Therefore, a one-sixteenth fraction of the design totalling eight analyses will be chosen as the experimental design. This design is shown in Table 2 where the values of the inputs are shown as normal deviates.

The choice of which one-sixteenth fraction to choose is not arbitrary. In Table 2 the columns of the experimental design are orthogonal. This will ensure that the coefficients will be estimated independently of one another. The design has purposefully been chosen for this property, however, the coefficients may be biased or confounded by the remaining one hundred twenty unestimated coefficients. Essentially, the problem is one of too many unknowns and too few equations. In order to solve for a specific set of unknowns the balance must be assumed to be zero. None of the coefficients of interest depend on each other, however. Quite frequently the assumption that higher order interactions are zero is justified, but the user must be aware that the assumption has been made.

Once the design and variables have been chosen, the actual experiment can be performed. The FRAP-T code automatically executes the required number of cases, each time perturbing the input variables according to the experimental design. The resulting data is illustrated in Figure 1, where the chosen response, cladding surface temperature, is plotted for all eight analyses versus time. This is the raw data that will be used to fit response surface equations for fixed time points throughout the problem history.

IV. RESPONSE EQUATION FITTING AND VALIDATION

Response surface equations are obtained by fitting coefficients of a Taylor's series expansion by a least squares technique. The coefficients are determined by minimizing the sum of the squared differences between the predicted and observed data values. This is a commonly accepted and well documented practice for fitting equations to data. However, an examination of some of the usual assumptions made for least squares reveals certain differences which must be considered when the method is applied to computer code response data. In particular, it is often assumed that (1) the model being fit is the true model, (2) the independent variables are known exactly, and (3) the dependent variable observations contain an element of uncorrelated random error with zero mean and constant variance. In this case, (1) the model being fit is at best an approximation over a specific region of interest, (2) the independent variables are input to the code as exact values, but it is the purpose of the work to propagate errors in the inputs through the response equations and (3) the output of the code can be observed without any random error whatsoever. This leads to the conclusion that the residuals of a response surface equation fit to computer generated data are due to lack of fit only, where a residual is defined to be the difference between the response surface equation and the code calculations at each data point. These residuals may be examined to determine the adequacy of fit of the response surface equation.

In traditional least squares estimation a well proven result is that the expected values of the estimated parameters are the parameters themselves. In other words, the parameters are unbiased. Under the present conditions, however, it is easily shown[5] that the parameters or coefficients are in fact biased by the exclusion of significant terms from the response equations (or, conversely, the inclusion of unimportant terms). By postulating a true model in addition to the fitted model, postulated residuals may be determined for any input values. These may be compared to observed residuals to determine if the postulated model has merit. The only difficulty with this scheme is that the residuals must be independent of the data used to generate the response equation. Practically, this presents a problem since the analyst frequently cannot afford to generate such an additional data set. One independent residual is always available, however, the nominal case. If a linear model is fitted and a second order equation postulated, the difference between the response equation and the nominal will be equal to the sum of missing quadratic coefficients. Thus, possibly important terms that might bias the response equation may be examined through this independent residual.

Practically, the nominal is the only truly independent residual available to the user. Fortunately, a technique has been devised for choosing terms to include in the response equation that sequentially uses each data point available for independent residual analysis. Known as Prediction Error Sum of Squares[6] (PRESS), the method removes one data point at a time from the data set and fits an equation based on the remaining data. The residual associated with the excluded data point is then squared and the process repeated for all other data points. The equation that minimizes PRESS is selected. This method has the advantage of using a form of independent residuals without the need of generating additional data.

Residuals of a PRESS selected response equation are shown, for example, in Figure 2, where cladding surface temperature residuals at time step forty (seventy eight seconds) are plotted versus the response equation prediction. Interpretation of the plot must be undertaken with care. In this instance the residuals are very small, of like magnitude, and alternating sign. The residual at the nominal (not shown) is also small (about eight degrees) indicating that quadratic terms are probably not important. Thus the linear approximation appears reasonable and the very small residuals are due to the inclusion of unimportant terms in the equation. In fact, the last term added had a coefficient an order of magnitude smaller than the next smallest coefficient. In summary, this example response equation appears to reasonably approximate the true response by a linear fit, has not omitted higher order terms, and has included one unimportant term in the equation. Estimated uncertainties obtained from this equation should be reasonably free of bias and so give useful results.

V. RESPONSE UNCERTAINTY ESTIMATION

Once a response surface equation has been determined to adequately approximate the true response, the equation may be used to infer information about the uncertainty of the response. The method for doing this is known as second order error propagation. Second order refers to the order of the Taylor's series expansion truncation. The method simply finds the expected values of the first four moments of the response as a function of the response equation coefficients and input distribution moments. Since the equation is up to second order, the first eight central moments of the inputs are required. The method is exact and the only approximation is that inherent in the response equations. Figure 3 illustrates the estimated cladding surface temperature mean and standard deviation during the course of the sample problem. Figure 3 actually illustrates the net result of one hundred twenty six response equations that approximated the true response at two second intervals.

Using the four moments of response, the probability density function for that response may be estimated through a technique known as moment matching. The moments are compared to the Pearson family of distributions and a suitable distribution is selected.[7] The density function and cumulative distribution function for the response may then be approximated. Figure 4 shows the probability density function of cladding surface temperature at time step forty of the sample problem. The distribution is a normal or Gaussian distribution. This is not surprising since the inputs were all assumed normally distributed and a linear response equation fit. The method is, however, entirely free of assumptions about the form of the input distributions and the user may input arbitrary distributional forms. In fact, so long as the standard deviation is not changed, different assumptions about the form of the input distributions may be evaluated after the basic experiment has been performed. This brings out a curious but important fact. Inferences about uncertainty in the code outputs are not actually made until the last stages of an analysis. Up until that point the perturbations specified by the user can be entirely arbitrary as, for example, a sensitivity study. Only at the end when inferences are made about the output uncertainty do the input perturbations assume meaning in terms of the input uncertainties.

The final task in achieving the purpose of the uncertainty analysis is to determine the sensitivity of the response uncertainty to the various input uncertainties. This is accomplished by calculating the fractional contributions to the variance (the square of the standard deviation) of the response from each input. Figure 5 illustrates the fractional contributions of each of the seven input variables to the estimated cladding surface temperature uncertainty. Note that not only does the estimated uncertainty vary during the problem history, as shown in Figure 3, but the fractional contributions to that uncertainty also vary significantly during the problem. In this case, the gap heat transfer and fuel thermal conductivity uncertainties are important during the blowdown phase of the transient while the flooding rate clearly dominates the reflood portion. Reductions in the calculated uncertainty of cladding surface temperature will be effected by a reduction in the uncertainty in flooding rate, whether it be by improved modeling or refined experimental data. Conversely, moisture carryout fraction and coolant channel flow blockage uncertainties do not appear to affect the uncertainty in cladding surface temperature for this problem and further work on these models and uncertainties is not justified. Thus, the sensitivity of the various models and the direction of experimental data needs have been defined through the example uncertainty analysis.

VI. SUMMARY

A user oriented, automated uncertainty analysis capability has been incorporated in the FRAP computer codes. The option uses the response surface method to generate equations that approximate the code behavior. The equations are first subjected to a validation procedure to determine how suitable the approximations are, then used to propagate errors in the inputs so that estimates of the response uncertainty can be made. Probability density functions for the responses are determined and finally fractional contributions to the variance from each input uncertainty calculated. All of the above functions are combined in a flexible code package that can significantly reduce the time normally required to conduct an uncertainty analysis, while providing a consistent, well-documented methodology. The option will facilitate traditional analyses such as sensitivity studies, analysis of experimental data needs, and code assessment in such a way that the mechanics of conducting these analyses will be routine and emphasis can be placed more appropriately on the interpretation and application of results.

VII. REFERENCES

1. FRAPCON MOD1/VER4: EG&G Idaho, Code Configuration No. H00730IB.

2. L. J. Siefken, et al, FRAP-T5: A Computer Code for the Transient Analysis of Oxide Fuel Rods, EG&G Idaho, Inc., NUREG/CR-0840, TREE-1281, June 1979.

3. K. R. Katsma, et al, RELAP4/MOD5: A Computer Program for Transient Thermal-Hydraulic Analysis of Nuclear Reactors and Related Systems, Users Manual, ANCR-NUREG-1335, September 1976.

4. N. D. Cox, "Comparison of Two Uncertainty Analysis Methods", Nuclear Science and Engineering, 64 (September 1977) pp 258-265.

5. R. H. Myers, Response Surface Methodology, Allyn and Bacon, Inc., pub., Boston, 1971.

6. D. M. Allen, The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables, University of Kentucky Statistics Department Technical Report No. 23 (1971).

7. C. F. Miller, N. D. Cox, A. J. Nelson, C. L. Atwood, User's Guide to PDFPLOT, A Code for Plotting Certain Statistical Distribution Functions, EG&G Idaho, Inc., TREE-1290, September 1978.

TABLE 1

ASSUMED LOCA UNCERTAINTY FACTORS*

1.	Flooding Rate	10%
2.	Flow Blockage Percentage	5%
3.	Flecht Heat Transfer	10%
4.	Carryout Fraction	10%
5.	Gap Heat Transfer	25%
6.	Fuel Thermal Conductivity	0.4 (W/m·K)
7.	ANS Decay Heat Curve	6.7%

* NORMAL DISTRIBUTIONS, ONE STANDARD DEVIATION ASSUMED

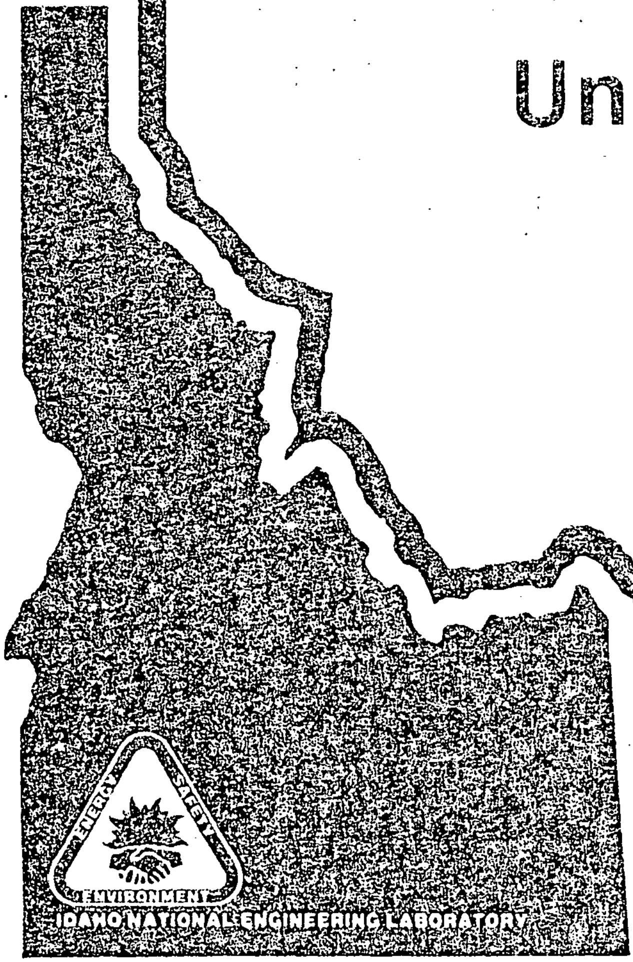
TABLE 2

DESIGN MATRIX FOR A 2^{7-4} FRACTIONAL FACTORIAL

	A	B	C	C	E	F	G
1	1	1	1	1	1	1	1
2	-1	1	1	-1	-1	1	-1
3	1	-1	1	-1	1	-1	-1
4	-1	-1	1	1	-1	-1	1
5	1	1	-1	1	-1	-1	-1
6	-1	1	-1	-1	1	-1	1
7	1	-1	-1	-1	-1	1	1
8	-1	-1	-1	1	1	1	-1

Automated Uncertainty Analysis for the FRAP Codes

Presented by
S.O. Peck



EG&G Idaho, Inc.

Outline

- Uncertainty analysis objective
- Example problem - PWR LOCA
- Response surface method
- Response equation validation
- Uncertainty estimation
- Features of an automated option
- Summary

INEL-S-24 532

Uncertainty Analysis

- Estimate the uncertainty in code outputs as a function of known uncertainties in code inputs

- Determine fractional contributions to the total uncertainty of individual inputs

INEL-S-18 763

Purpose

- Sensitivity studies
- Experimental data needs
- Code assessment

INEL-S-24 540

Example Problem

- Zion PWR LOCA
- 200% cold leg break
- Calculations by FRAP-T5 through reflood
- Uncertainty emphasis on reflood variables

INEL-S-24 533

Assumed LOCA Uncertainty Factors*

1.	Flooding rate	10%
2.	Flow blockage percentage	5%
3.	FLECHT heat transfer	10%
4.	Carryout fraction	10%
5.	Gap heat transfer	25%
6.	Fuel thermal conductivity	0.4 (W/m-K)
7.	ANS decay heat curve	6.7%

*Normal distributions, assumed one standard deviation

INEL-S-24 539

Response Surface Methodology

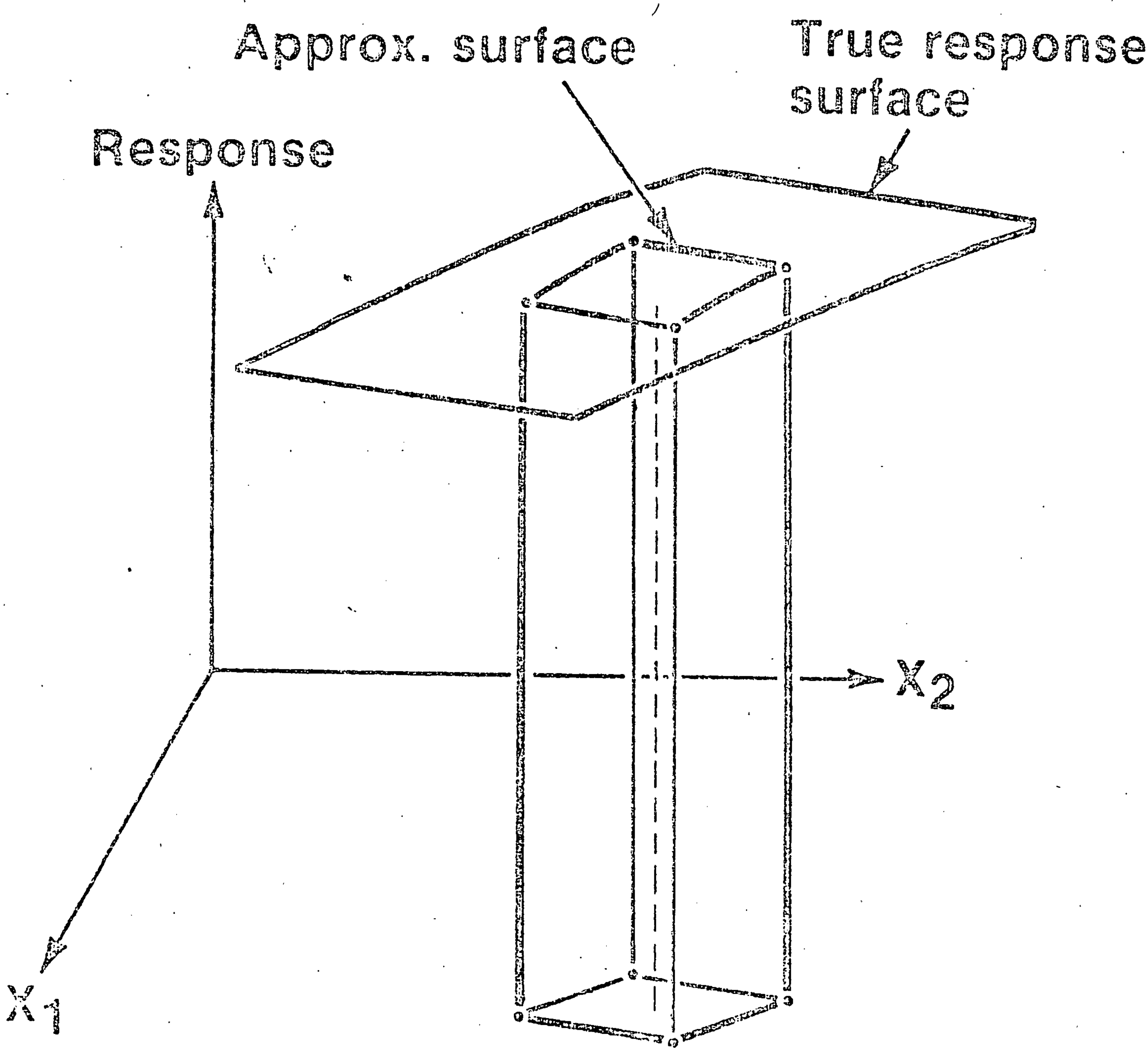
- "Responses" are the FRAP outputs
- Independent variables are code inputs
- The error analysis procedure determines a relationship between each response and the independent variables.

INEL-S-10 674

Response Surface Methodology (cont'd)

- RSM fits a simple polynominal to a portion of the response "surface" about a nominal point via a Taylor series
- Propagation of error and response uncertainty is then inferred from the approximate surface

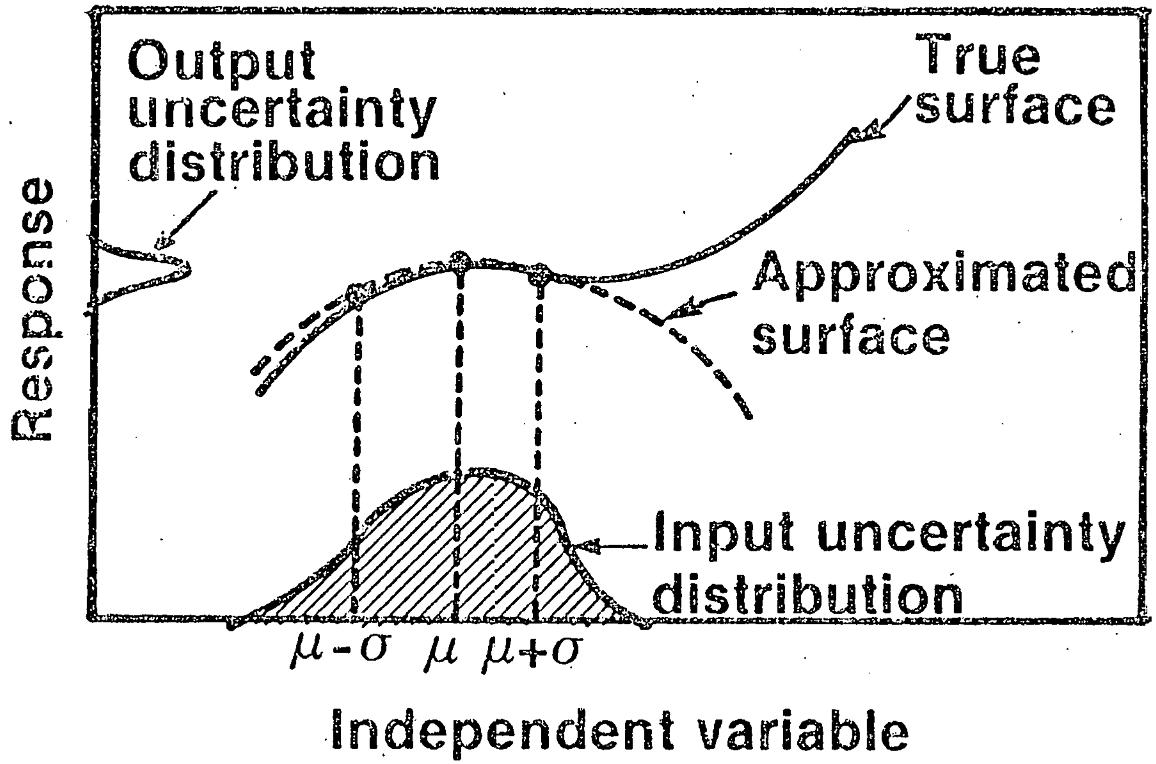
INEL-S-10 675



$$\text{Reponse} = C_0 + C_1X_1 + C_2X_2 + C_3X_1X_2 + C_4X_1^2 + C_5X_2^2$$

INEL-S-10 672

Response Uncertainty Determination



INEL-S-17 028

Experimental Design

- Provides a systematic pattern for perturbing the independent variables so that the maximum information is obtained with fewest FRAP runs

- Basic designs - Two level fractional factorial - for estimating linear and (some) two-factor coefficients

INEL-S-24 542

Experimental Design (cont'd)

- Plackett-Burman designs used for numbers of factors near 12 or 20
- Foldover design used to eliminate two-factor confounding of linear terms
- Quadratic terms may be added to above designs

INEL-S-24 543

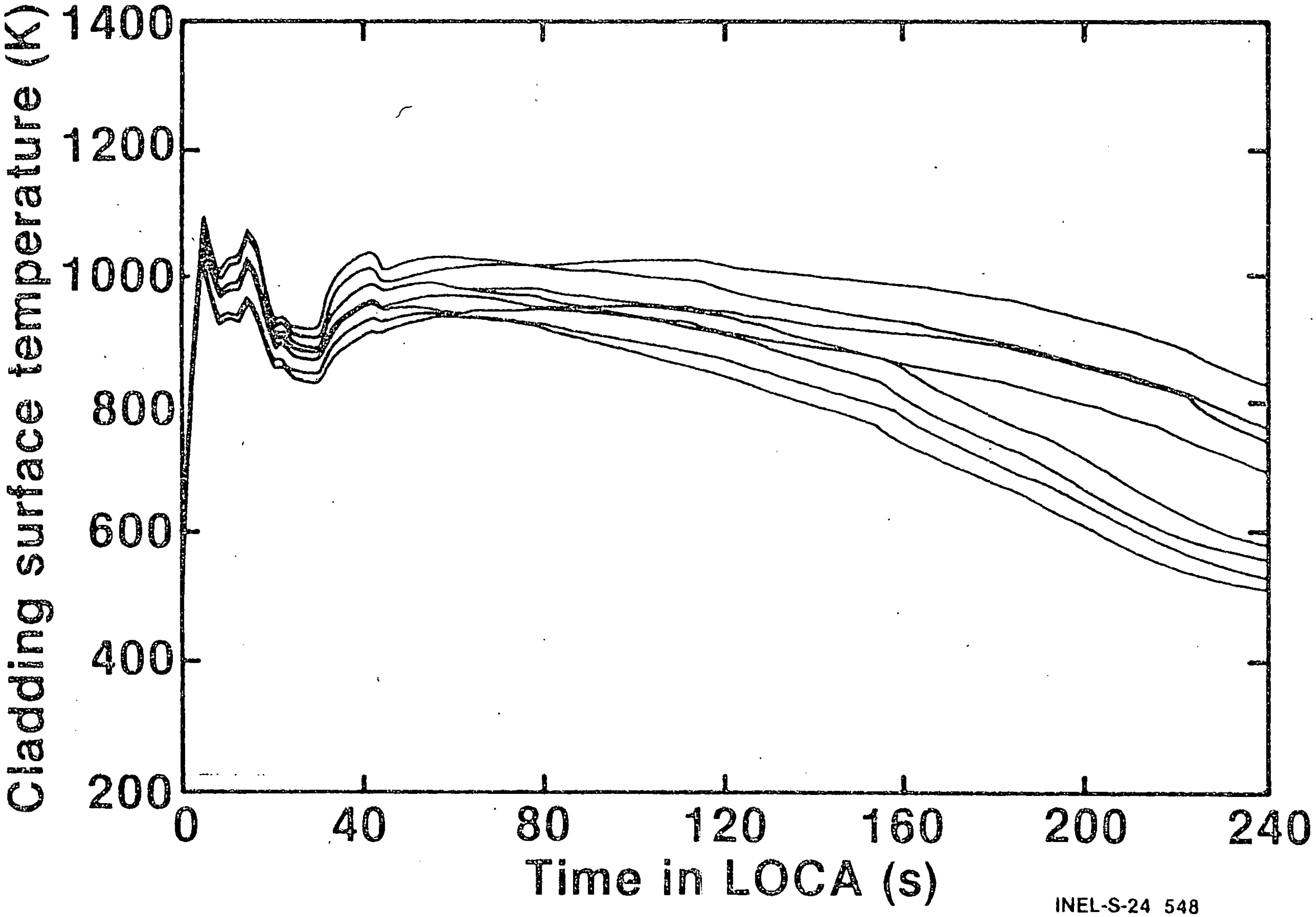
Design Matrix for a 2^{7-4} Fractional Factorial

Factor Levels

Runs	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	-1	1	1	-1	-1	1	-1
3	1	-1	1	-1	1	-1	-1
4	-1	-1	1	1	-1	-1	1
5	1	1	-1	1	-1	-1	-1
6	-1	1	-1	-1	1	-1	1
7	1	-1	-1	-1	-1	1	1
8	-1	-1	-1	1	1	1	-1

INEL-S-24 545

Cladding Surface Temperature



INEL-S-24 548

Definitions

- A response surface equation is a polynomial that approximates the code calculations over a given region

- A residual is the difference between the response surface equation and the code calculations at each data point

INEL-S-18 764

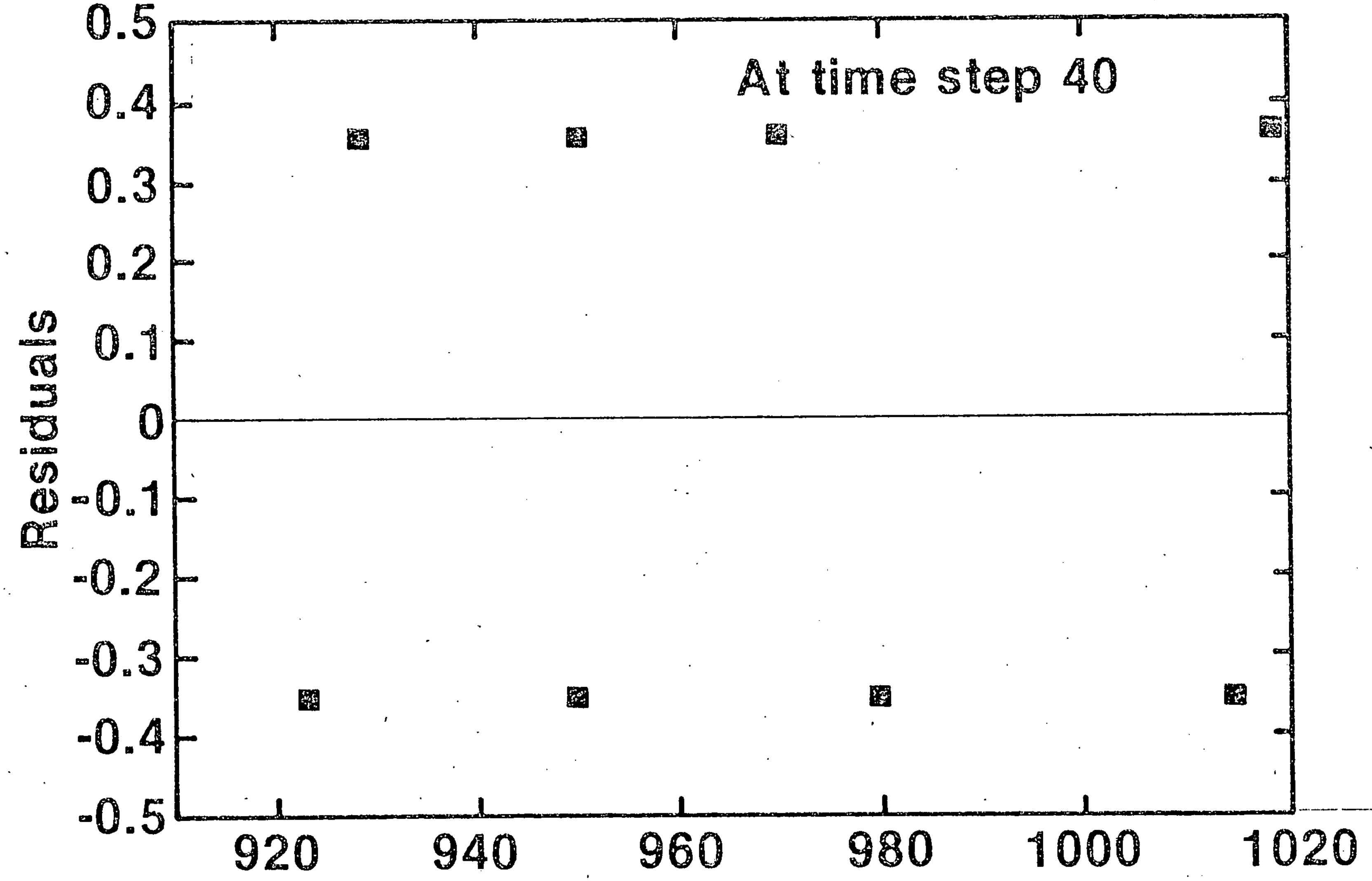
Response Equation Validation

- Determine whether the response surface equations adequately approximate the unknown functional form of the code response

- Poor approximations will bias estimates of uncertainty

INEL-S-18 762

Cladding Surface Temperature (K)



INEL-S-24 541

Least Squares Assumptions
General Regression

- The model fit is the true model
- The independent variables are known exactly
- The dependent variable observations contain random error

INEL-S-24 538

Least Squares Assumptions (cont'd.)

Regression on Computer Data

- The model is at best an approximation
- It is the purpose of the work to propagate errors in the inputs
- The output of the code can be observed without error

Conclusion

- These residuals are soley due to lack of fit

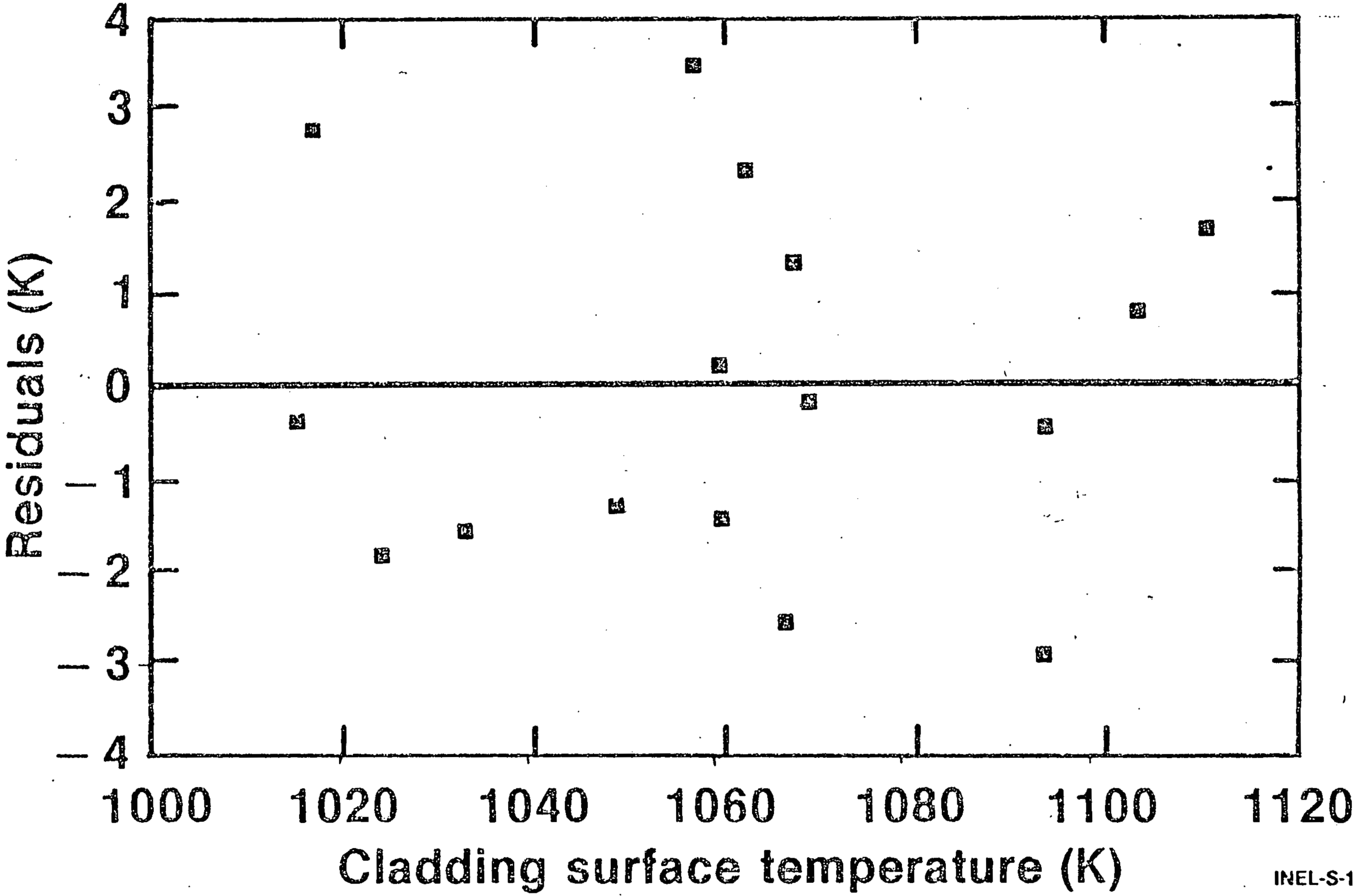
INEL-S-24 536

Analysis of Residuals

- Evenly dispersed and well distributed residuals indicate good fit
- Very small residuals of like magnitude and alternating sign indicate overfit
- Highly grouped residuals indicate underfit, that is, significant terms omitted
- Well dispersed residuals with one or two outliers indicate a threshold response

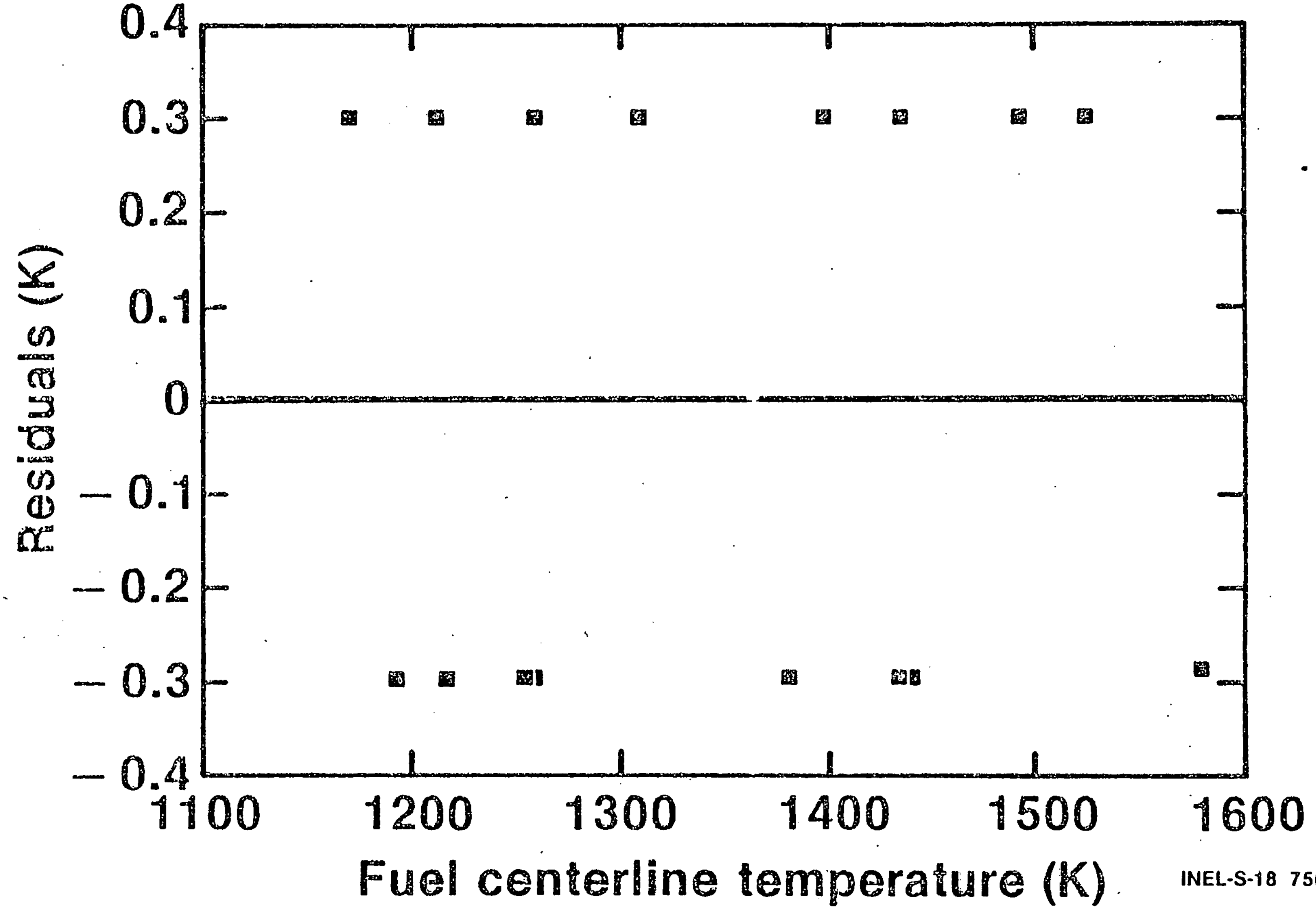
INEL-S-18 767

Cladding Surface Temperature Residuals



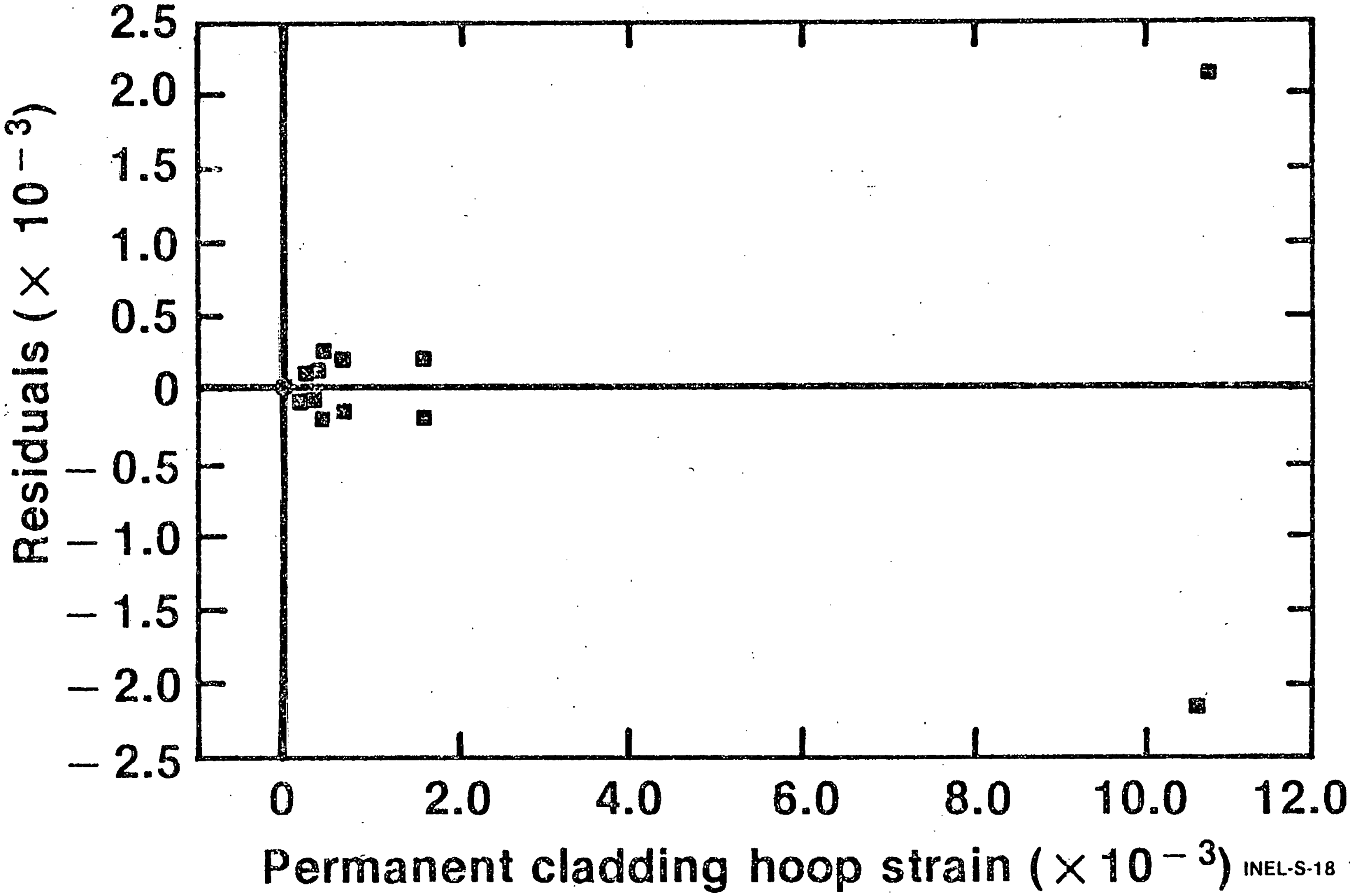
INEL-S-18 757

Fuel Centerline Temperature Residuals



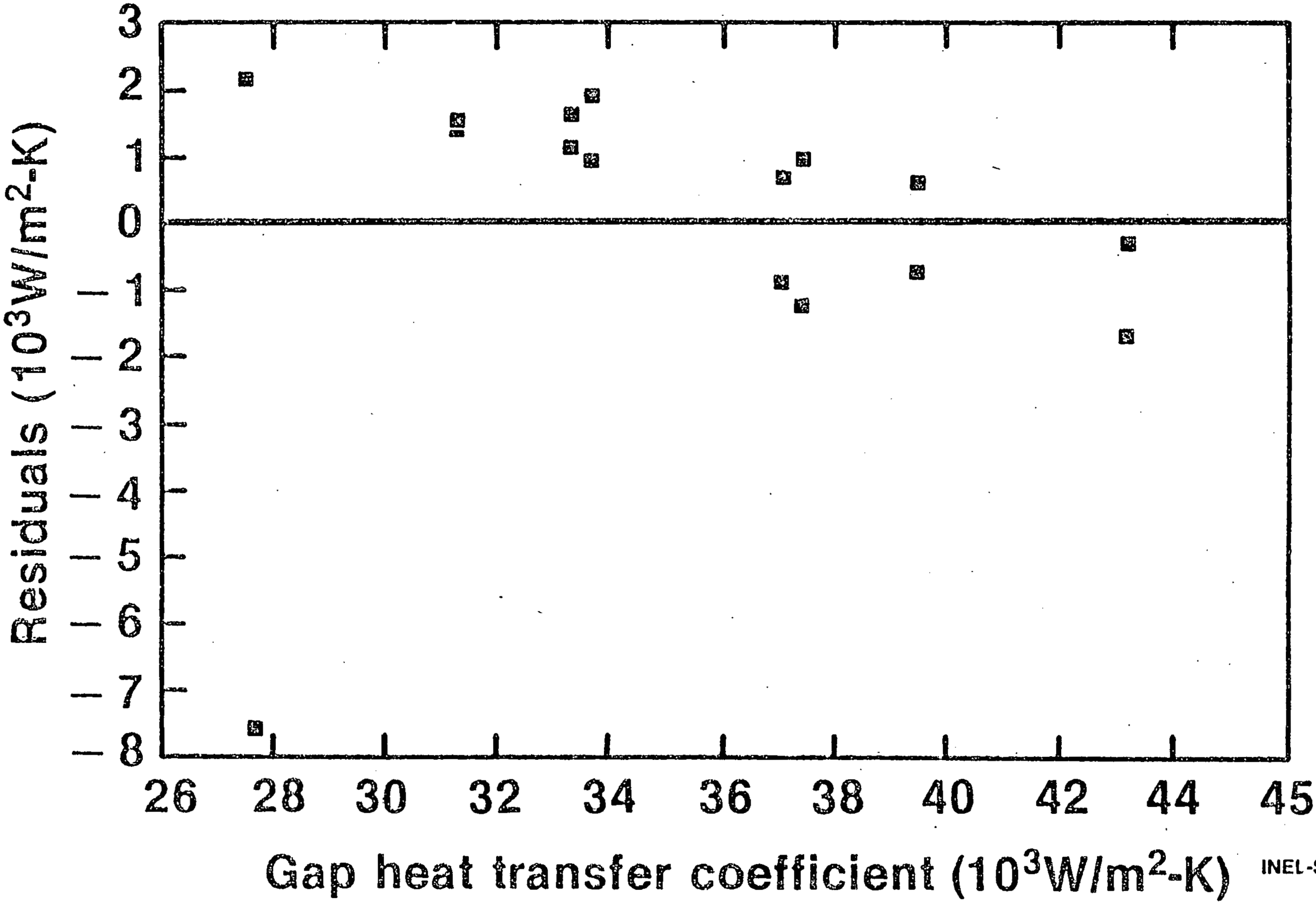
INEL-S-18 756

Cladding Hoop Strain Residuals



INEL-S-18 759

Gap Heat Transfer Coefficient Residuals



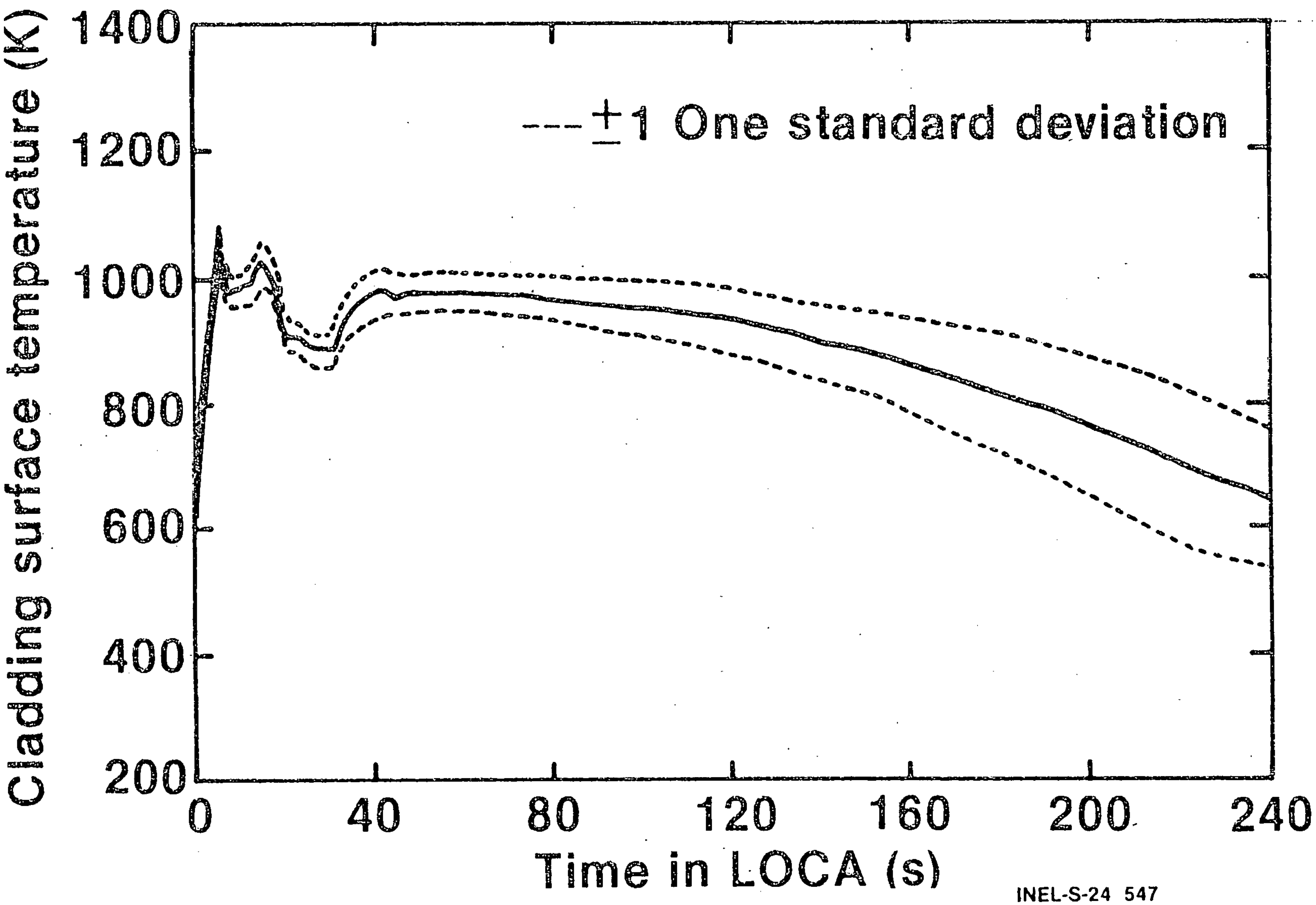
INEL-S-18 758

Response Uncertainty Estimation

- Second order error propagation
- Requires first eight moments of the input distributions
- Estimates first four moments of the output distribution

INEL-S-24 534

Cladding Surface Temperature



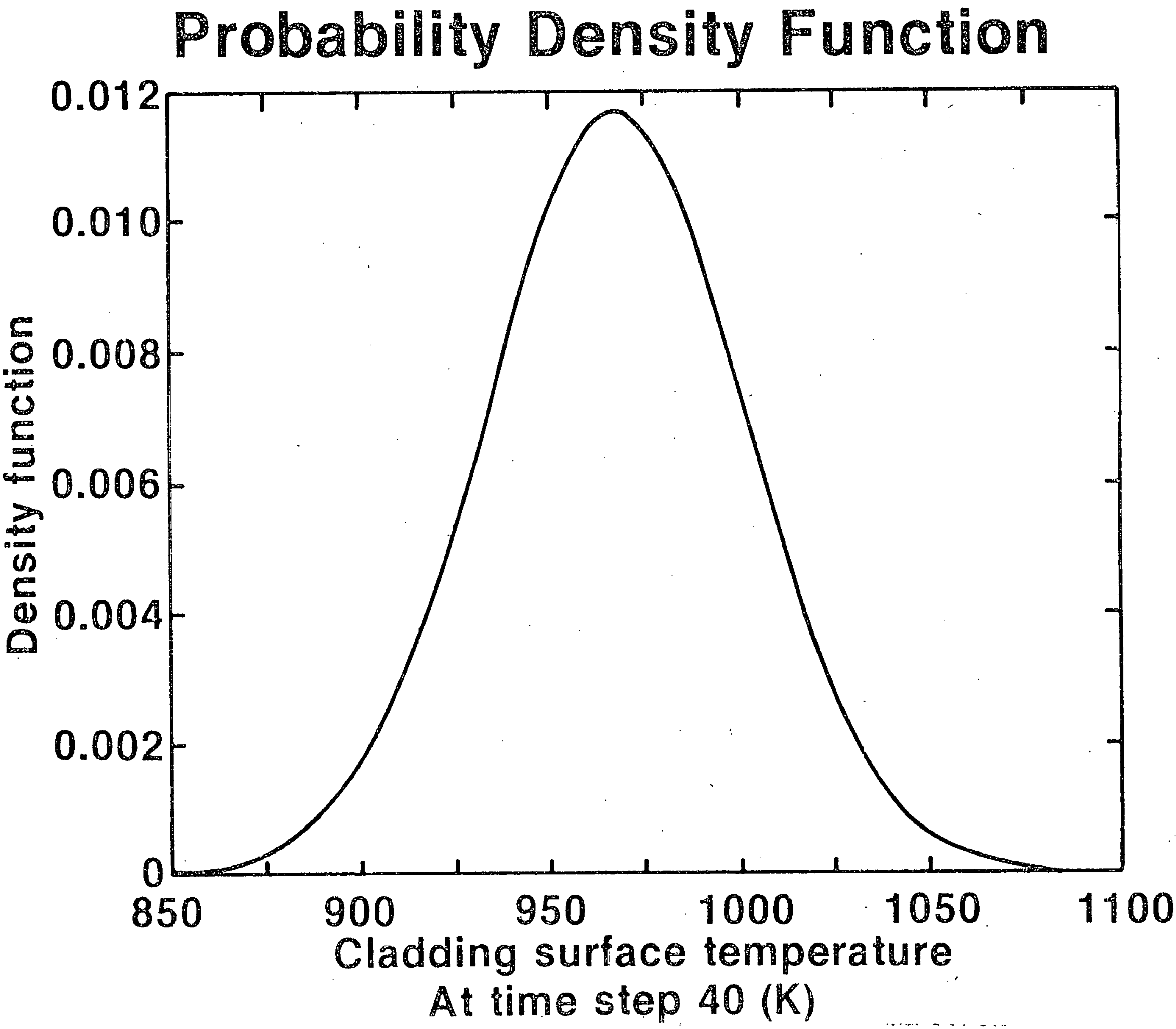
INEL-S-24 547

Response Uncertainty Estimation (cont'd.)

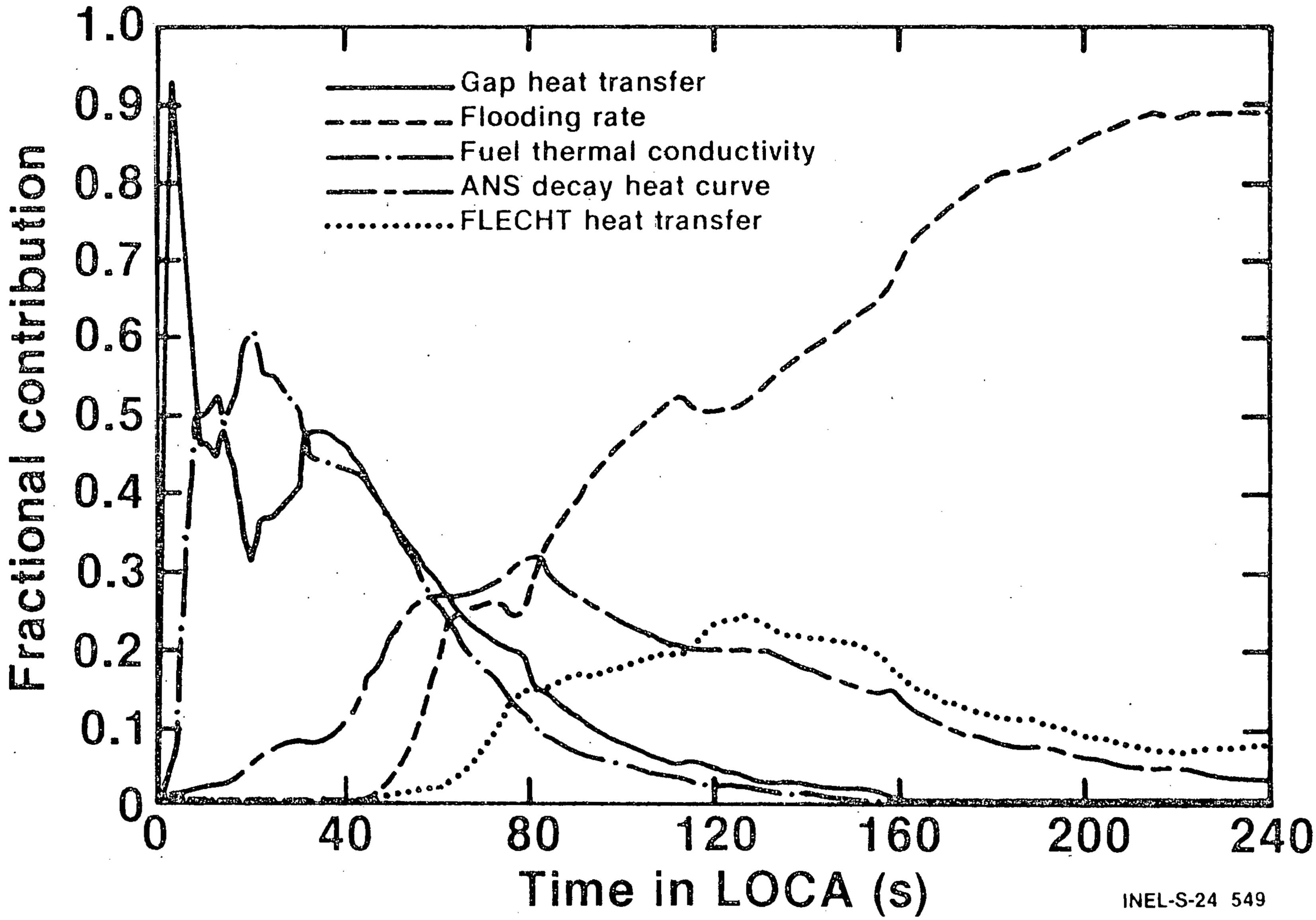
- Estimate the probability density function of the response at a point in time

- Moment matching technique

INEL-S-24 535


Probability Density Function
Density function
0.012
0.010
0.008
0.006
0.004
0.002
0
850
900
950
1000
1050
1100
Cladding surface temperature
At time step 40 (K)

Cladding Surface Temperature



Automated Error Analysis

User specifies inputs to vary and responses to analyze

User specifies degree of polynomial

Code automatically

- Determines experimental design and confounding pattern
- Calls FRAP
- Fits response polynomial for all responses
- Estimates means and variances
- Computes fractional contribution to variance

Summary

- A user oriented, automated uncertainty analysis option is available in the FRAP codes
- It is based on response surface methodology
- The response equations must undergo validation
- The option can be used for:

    1) Sensitivity studies
    2) Experimental data needs
    3) Code assessment