CDAP-TR-79-050
for U.S. Nuclear Regulatory Commission

CODE DEVELOPMENT AND ANALYSIS PROGRAM

DEVELOPMENTAL VERIFICATION OF THE FRAP-T5 UNCERTAINTY ANALYSIS OPTION

S. O. PECK   C. L. ATWOOD

January, 1979

EG&G Idaho, Inc.
IDAHO NATIONAL ENGINEERING LABORATORY
DEPARTMENT OF ENERGY
IDAHO OPERATIONS OFFICE UNDER CONTRACT EY-76-C-07-1570

NOTICE

EG&G Idaho, Inc.

FORM EG&G-398
Rev. 12-78)

INTERIM REPORT

Accession No. ________

Report No. CDAP-TR-79-050

Contract Program or Project Title: Fuel Behavior Model Development

Subject of this Document: Developmental Verification of FRAP Uncertainty Analysis Option

Type of Document: Status Report

Author(s): S. O. Peck, C. L. Atwood

Date of Document: January 1979

Responsible NRC Individual and NRC Office or Division: G. P. Marino, Reactor Safety Research

This document was prepared primarily for preliminary or internal use. It has not received full review and approval. Since there may be substantive changes, this document should not be considered final.

EG&G Idaho, Inc.
Idaho Falls, Idaho 83401

Prepared for the
U.S. Nuclear Regulatory Commission
and the U.S. Department of Energy
Idaho Operations Office
Under contract No. EY-76-C-07-1570
NRC FIN No.

INTERIM REPORT

E&B

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

CDAP-TR-79-050
January 1979

DEVELOPMENTAL VERIFICATION
OF THE FRAP UNCERTAINTY
ANALYSIS OPTION

S. O. Peck
C. L. Atwood

REVIEWED BY: W. F. Domenico

W. F. Domenico
Section Leader

Michael P. Bohn

M. P. Bohn
FARAD Branch Manager

APPROVED BY: Paul North

Paul North, Manager
Code Development and
Analysis Program

EG&G Idaho, Inc.

SUMMARY

An automated uncertainty analysis option has been added to the FRAP-T5 code which allows the user to obtain estimates of the uncertainty in computed code outputs as functions of known input uncertainties. Developmental verification, the subject of this report, is an ongoing process whereby the uncertainty subcode is checked for programming correctness and for validity of analysis results. Three developmental verification studies were undertaken. (1) A benchmark calculation using a simple equation replacing FRAP-T5 was performed to check the correctness of the subcode programming. (2) The adequacy of the fit of the response equations to the data was evaluated using a statistical criteria to guard against over and under fitting and a visual examination of residuals. An uncertainty analysis of a LOCA was used as an example. (3) The ability of the response equations to predict the true response surface was determined by generating new data and comparing the predictions of the original equations with the new data. Results indicated that the majority of the responses were fit acceptably well by just a linear response equation. The fit for a few responses could be improved by continuing the analysis to include higher order terms. Recommendations for future versions of the uncertainty analysis option are to include methods for assessing the validity of the response equations as an automatic user feature.

CONTENTS

TABLES

LIST OF FIGURES

1. INTRODUCTION

Part of the United States Nuclear Regulatory Commission's Water Reactor Safety Research Program is the development of analytical tools to accurately predict the response of nuclear reactor systems during off-normal or hypothetical accident operation. The development of the FRAP-T[1] (Fuel Rod Analysis Program-Transient) code is an important part of this program. FRAP-T is a best-estimate code designed to calculate fuel rod response (e.g., cladding temperature, cladding strain) during transient conditions in the reactor.

Recently an option has been added to the FRAP-T5 code which allows the user to automatically generate a complete uncertainty analysis. The uncertainty analysis option provides, at each problem time step, uncertainty limits on user selected fuel behavior responses as a function of uncertainties in code inputs and code models. The analysis also yields relative contributions to the total uncertainty of each of the inputs used at each time step. Knowing the total uncertainty of the FRAP-T calculated fuel behavior and knowing from where these uncertainties originate within the code is useful, both in code application as well as allowing rational direction of code development and refinement.

The Automated Uncertainty Analysis Option[2] is based on response surface methodology and second order error propogation techniques. The option has specifically been developed to be user oriented and will provide a complete uncertainty analysis in a single computer run. At each time step the option generates an approximation of the true response surface[*] for each fuel behavior response selected for analysis. Second order error propagation techniques are then used to determine uncertainties and relative contributions to the uncertainties based on the approximated response surface.

Developmental verification, the subject of this report, is an ongoing verification process whereby the uncertainty subcode is checked

---

* The concept of Response Surface is explained in Section 2.

for programming accuracy and for validity and applicability of analysis results. Specifically, the objectives of the FRAP-T5 uncertainty analysis developmental verification are:

1. To ensure the correctness of the subcode programming.

2. To assess the validity of the approximate response surface relative to the data used to generate it.

3. To assess how well the approximate response surface predicts the true response surface.

This report presents the developmental verification analyses comparisons and discusses the results of the analyses. The report includes an overview of the uncertainty analysis method (Section 2), the developmental verification studies (Section 3), a sample problem (Section 4), and conclusions and recommendations (Section 5).

2. OVERVIEW OF THE UNCERTAINTY ANALYSIS OPTION

The FRAP-T5 uncertainty analysis option has been developed so that a user may easily obtain estimates of the uncertainty in calculated code outputs. The option has specifically been designed so that any user may perform an uncertainty analysis on any FRAP problem in an understandable, systematic manner. To aid the user, features such as default uncertainty values for approximately 50 input variables have been built into the code. The option further provides for a sequential development of output complexity by allowing the user to restart and continue an analysis from intermediate points. One goal of the option is to provide to all users a straightforward technique based on sound methodology for estimating code uncertainties. A complete description of the option is available in Reference 2.

The uncertainty analysis option is based on the response surface method. Any of the output variables of a computer code may be termed a response. There is some functional relationship between a response and the input variables. In the space of the input variables, this relationship defines a surface, and hence the term "response surface". When the code is rather sample, this surface may be determined analytically over the entire range of the input values. More often, as in the case of the FRAP code, the surface may be known only through the code, and the range of inputs and problem types is very large. Thus, the complete true response surface cannot be determined analytically. The response surface method of uncertainty analysis is based on a systematic sampling of the true surface which is then approximated by a polynomial equation in the independent (input) variables. In effect, the true surface is approximated by a smooth surface[3,4].

The polynomial equation approximating the true surface is derived as follows. Let $Y(x_i)$ denote the code response as a function of x_i = x_1, x_2, . . ., x_k inputs. The Taylor's series expansion about any point μ_i is then given by:

$$Y(x_i) = Y(\mu_i) + \sum_{i=1}^{k} \frac{\partial Y(u_i)}{\partial x_i} (x_i - \mu_i) + 1/2 \sum_{i=1}^{k} \frac{\partial^2 Y(u_i)}{\partial_i^2} (x_i - \mu_i)^2$$

$$+ \sum_{\substack{i,j \\ i<j}}^{k} \frac{\partial^2 Y(\mu_i)}{\partial x_i \partial x_j} (x_i - \mu_i) (x_j - \mu_j) + \text{higher order terms.} \qquad (1)$$

Truncating the Taylor's series at second order terms, the desired polynomial equation is obtained by identifying the coefficients of the polynomial with the partial derivatives of the series expansion. The coefficients are estimated from sample values of the true response surface obtained by perturbing the nominal inputs. For a second order polynomial to reasonably approximate the true response surface, the region of the surface being sampled must be small enough so that large irregularities are not present. Experience has shown that a range of plus and minus one standard deviation ($\pm 1\sigma$) in the input variable uncertainties will usually satisfy this requirement for the FRAP code.

The polynomial approximation to the true response surface may be used to examine the behavior of the true surface in the region of the sample space without the burden of excessive cost. In particular, the polynomial can be used to study the propagation of errors through the code and their effect on the uncertainty in computed outputs. Thus, an estimate of response uncertainty and the relative contributions of input variables to this uncertainty may be obtained using the response surface method.

Once the user has selected a base case problem and made a choice of output responses and input variables, the following procedures will be followed by the code to obtain the desired final results, the estimates of response uncertainties.

(1) An experimental design will be chosen. This is simply a pattern for perturbing the independent variables of the problem. The pattern is obtained in matrix format where the columns correspond to inputs and the rows correspond to the individual analyses that must be computed. The problem is run as many times as the design dictates, each time varying the input variable perturbations according to the pattern.

(2) The response surface equations are then generated using the information derived from step one. Basically, a multiple regression routine is used with certain simplifications arising from orthogonal properties of the experimental design.

(3) The response surface equations are used to generate uncertainty distributions for the response parameters. Second order error propagation analysis is used to estimate the means and variances of the responses.

(4) Finally, estimates of the fractional contributions to the response variances are made to indicate the relative importance of individual input variables.

3. DEVELOPMENTAL VERIFICATION STUDIES

The principle objective of this developmental verification of the uncertainty analysis option is to demonstrate the conditions which lead to useful and valid uncertainty estimates. In particular, the user must be able to identify these conditions and assess the validity of his results based on standard criteria. The following developmental verification studies and the particular illustrative examples were chosen to demonstrate the wide variety of possible uncertainty analysis results and the methods used to judge their adequacy. Some of these methods appear well suited for future inclusion in an optimized version of the option, and this point will be addressed in Section 5, Conclusions and Recommendations. At present they will be considered only as tools of developmental verification.

Three developmental verification studies were performed. First, a benchmark calculation was performed using the uncertainty subcode to ensure the accuracy of the subcode programming. Second, the validity of the response surface equations and the assumed models were addressed by independently fitting a least squares model to the data using a statistical criteria to guard against under and overfitting. Lack of fit (important terms not included) was addressed as well. Finally, the range of applicability of the response equations was investigated by obtaining a variety of off-design data points (points other than those used to fit the equations) and comparing the predictions of the original response equations with the new data.

3.1 Benchmark Calculation

The objective of the benchmark calculation is to demonstrate the ability of the uncertainty analysis subcode to perform accuratly. To this end a simple equation was substituted for FRAP-T5 in the uncertainty option and an uncertainty analysis performed on it. The option generated results were then compared to hand calculations. The results demonstrate not only the accuracy of the coding but specific features of the response surface method of which the user should be aware.

The simple equation chosen was

$$Y(x_1,x_2) = 1 + x_1 + x_1^2 + x_1x_2 + x_2^4$$

This equation was chosen because it contained constant, linear, cross product, quadratic and a higher order term plus more than one variable. Uncertainty distributions (assumed normal) for each variable were assigned with a standard deviation of one for each ($\sigma_1 = 1$, $\sigma_2 = 1$) and respective mean values of 1 and 10 ($\mu_1 = 1, \mu_2 = 10$). Values of the response $Y(x_1,x_2)$ were obtained according to an experimental design in nine runs that allowed the estimates of all terms up through quadratic without any confounding between them.

In order to best compare results of the two calculations, the equation for $Y(x_1,x_2)$ should be transformed to standard normal form since the uncertainty analysis results are given in this format. This is accomplished by the change of variable

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \qquad (2)$$

where z_i is now a variable with zero mean and unit standard deviation. Alternatively,

$$x_i = \sigma_i z_i + \mu_i \qquad (3)$$

substituting into equation (1) gives the following expression for $Y(z_1,z_2)$.

$$Y(z_1,z_2) = (1 + \mu_1 + \mu_1^2 + \mu_2^4) + (\sigma_1 + \mu_2\sigma_1 + 2\mu_1\sigma_1)z_1$$
$$+ (\mu_1\sigma_2 + 4\sigma_2\mu_2^3)z_2 + \sigma_1\sigma_2 z_1 z_2 + \sigma_1^2 z_1^2 + 6\sigma_2^2\mu_2^2 z_2^2 \qquad (4)$$
$$+ 4\sigma_2^3\mu_2 z_2 + \sigma_2^4 z_2^4$$

Putting in the values for μ_1, μ_2, σ_1 and σ_2 yields

$$Y(z_1,z_2) = 10013 + 13z_1 + 4001z_2 + z_1z_2 + z_1^2 + 600z_2^2 + 40z_2^3 + z_2^4 \qquad (5)$$

This is still the true function that was to be approximated by the response surface equation. The response equation as determined by hand calculation using the uncertainty analysis method and designated here by $\hat{Y}(z_1,z_2)$, was determined to be

$$\hat{Y}(z_1\ z_2) = 10013 + 13z_1 + 4041z_2 + z_1z_2 + z_1^2 + 601z_2^2 \qquad (6)$$

This equation was predicted identically by the uncertainty subcode, thus verifying that portion of the programming.

Notice that the linear and quadratic terms for z_1 were estimated exactly as expected while the higher order terms of z_2 are approximated by slightly different coefficients for the linear and quadratic terms from the exact values. That is, the coefficients in the equation for $\hat{Y}(z_1,z_2)$ are estimates of the true coefficients as expressed in equation (5). When the approximation is exact, so are the estimates. But when the approximation is not exact the estimates are biased by the exclusion of higher order terms.

Perhaps the most important results of an uncertainty analysis are the estimates of the mean and variance of the response. The following discussion shows how the mean and variance are estimated by a second order approximation, and how they differ from the true values. The following results were also predicted identically by the subcode.

Given a response Y(x) where the independent variable x has a known uncertainty distribution (as specified by its probability density function), the uncertainty distribution of the output Y(x) may be determined. The mean of this distribution, called the Expected Value of the response, is computed from

$$\mu_y = E(Y) = \int_{-\infty}^{\infty} Y(x)\ f(x)\ dx \qquad (7)$$

where f(x) is the probability density function of the independent variable x. The variance, defined as the expected value of the second central moment of the distribution, is given as

$$\sigma_y^{\,2} = E(Y-\mu_y)^2 \qquad (8)$$

Estimates of the mean and variance are obtained by similarly determining the mean and variance, respectively, of the approximating response function $\hat{Y}(z_i)$,

$$\hat{\mu}_y = E(\hat{Y}) \qquad (9)$$

$$\hat{\sigma}_y^{\,2} = E(\hat{Y}-\hat{\mu}_y)^2 \qquad (10)$$

These are the values returned by the uncertainty analysis subcode. For example, refering to equation (6), the estimated mean would be

$$E(\hat{Y}) = E(10013) + E(13z_1) + E(4041z_2) + E(z_1z_2) + E(z_1^2) + E(601z_2^2) \quad (11)$$

The expected values of the independent variable z_i, which was assumed to have a normal probability density function, are easily shown to be

$$\begin{aligned} E(z) &= 0 & E(z^5) &= 0 \\ E(z^2) &= 1 & E(z^6) &= 15 \\ E(z^3) &= 0 & E(z^7) &= 0 \\ E(z^4) &= 3 & E(z^8) &= 105 \end{aligned} \quad (12)$$

So, it can easily by seen that

$$E(\hat{Y}) = 10013 + 0 + 0 + 0 + 1 + 601 = 10615 \quad (13)$$

when terms to second order are included. When only terms to first order are included, the estimate is

$$E(\hat{Y}) = 10013 \quad (14)$$

Now, referring to the true response function, equation (5), the exact value of the mean can be seen to be

$$E(Y) = 10013 + 0 + 0 + 0 + 1 + 600 + 40 + 3 = 10617 \quad (15)$$

Thus, the second order estimate is better than the first order estimate since the model approximation is more accurate. The first order estimate is said to be biased by the exclusion of important higher order terms.

One further important comment should be made regarding the mean. The nominal value of the response that is obtained when $Y(x_i)$ is evaluated at the mean values of the x_i, is not the same as the mean, μ_y. In the example the nominal value of $Y(z_i)$ is 10013 yet the second order estimate of the mean is 10615. This indicates the presence of positive curvature in the sample space about the nominal.

The estimated value of the variance of the response as determined from equations (6) and (10) is given as

$$\begin{aligned}\hat{\sigma}_y^2 = E(\hat{y}-\hat{\mu}_y)^2 &= (13)^2E(z_1^2) + (4041)^2E(z_2^2) \\ &+ E(z_1^2)E(z_2^2) + E(z_1^4) + (601)^2E(z_2^4) \\ &- E(z_1^2) - (601)^2(z_2^2)\end{aligned} \tag{16}$$

dropping terms that contribute only zero. Substituting in the values from equation (12) gives the following estimate

$$\hat{\sigma}_y^2 = 1.70 \times 10^8 \tag{17}$$

The true value, σ_y^2, is determined as

$$\begin{aligned}\sigma_y^2 = E(Y-\mu_y)^2 &= (13)^2E(z_1^2) + (4001)^2E(z_2^2) + E(z_1^2)E(z_2^2) \\ &+ E(z_1^4) + (600)^2E(z_2^4) - E(z_1^2) - (600)^2E(z_2^2) + E(z_2^8) \\ &- E(z_2^4)E(z_2^4) + (40)^2E(z_2^6) + (2)(4001)(40)E(z_2^4) \\ &+ (2)(600)(1)E(z_2^6) - 2(600)(1)E(z_2^2)E(z_2^4)\end{aligned} \tag{18}$$

again dropping terms that only contribute zero. Substituting in from equation (12) gives the true value of the variance as

$$\sigma_y^2 = 1.77 \times 10^8 \tag{19}$$

Thus, the second order estimate is not seriously in error. In fact, the estimate of the variance obtained by including only first order terms is

$$\hat{\sigma}_y^2 = 1.63 \times 10^8 \tag{20}$$

which differs from the exact value by less than 8 percent.

This benchmark calculation has shown how the uncertainty analysis subcode approximates the true response behavior and estimates distributional parameters for the response. Hopefully this has shed light on how the method works as well as proving the correctness of the programming.

3.2 Analysis of Response Equation Fit to Data

The present version of the uncertainty analysis subcode constructs response equations based on the full assumed model. That is, every term in the approximate response equation which can be estimated from the chosen design is included in the equation. This leads to the possibility of overfitting the data. Overfitting may affect the ability of the equation to predict off design responses in the region of the sample space. At the other extreme, the response equation may underfit the data due to the exclusion of significant terms in the assumed model. In this case the response equations may tend to underestimate the mean and variance of the response. Therefore, a program called ANYOLS, devised by the Reliability and Statistics Branch of EG&G Idaho, Inc. was used to examine the possible over or underfit of the response equations to the data and thereby assess their adequacy in modeling the true response behavior.

3.2.1 Sample Problem. An uncertainty analysis on an example problem was conducted to generate response surface equations. A complete description of the sample problem is presented in Section 4. Briefly, the problem consisted on a nominal PWR fuel rod in a loss of coolant accident (LOCA) through blowdown at beginning of life conditions. The fuel rod power was artifically raised to increase the likelihood of rod failure but otherwise all parameters were left at their nominal values. The uncertainty analysis considered ten factors in sixteen FRAP-T5 executions. Five crossproduct terms were thus estimable and the factors were ordered so that those crossproducts thought most likely to be influential were estimated. Ten responses were chosen including cladding surface temperature, fuel centerline temperature, gap heat transfer coefficient and cladding strains. The output was taken at half second intervals throughout the history and so 10 responses x 59 data sets = 590 response equations were generated.

3.2.2 ANYOLS, PRESS, and Residual Plots. ANYOLS[5] is a computer code that allows the user to fit one or more ordinary least squares regression models according to predetermined statistical criteria. In this case the criterion chosen was the prediction error sum of squares (PRESS) which guards against both lack of fit and overfitting regression models. The criterion will become large if too few or too many variable are included in the model. ANYOLS with PRESS was used to select a suitable regression model from the whole model originally assumed for the 590 sets of FRAP-T5 generated data. In practically all cases the PRESS criterion sequentially chose variables to add to a given model in the order of absolute magnitude of the response coefficients. This was to be expected since the variables have been transformed to standard normal dimensionless form. However, the final model chosen by PRESS frequently contained fewer than the full model. As a general observation those responses that were strong functions of only a few variables contained the fewest terms in the PRESS equation and, vice versa, those response equations that were weak functions of many variables were fit by PRESS with practically the full model.

One of the best methods for judging the adequacy of a response equation is to look at a residual plot. A residual plot is a graph of the differences between the observed data and the response equation prediction for the same data (residuals) plotted versus the response equation predictions. An equation well fit to the data would not show up and down trends in the residuals which would otherwise indicate biasing due to an inadequate assumed model. Furthermore, the residuals should be fairly well dispersed across the graph indicating a reasonably constant variance. Examples of residual plots taken from the sample problem are discussed in the following subsections.

3.2.3 Example 1: Good Fit. Cladding surface temperature at 5.0 seconds in the transient (response 1 at time increment 11) illustrates a response that is well fit by the PRESS generated response equation. PRESS chose a constant and eight variables (two were crossproducts) to model the response out of fifteen possible. The residuals, Figure 1, are reasonably uniform and well distributed. The residuals are fairly

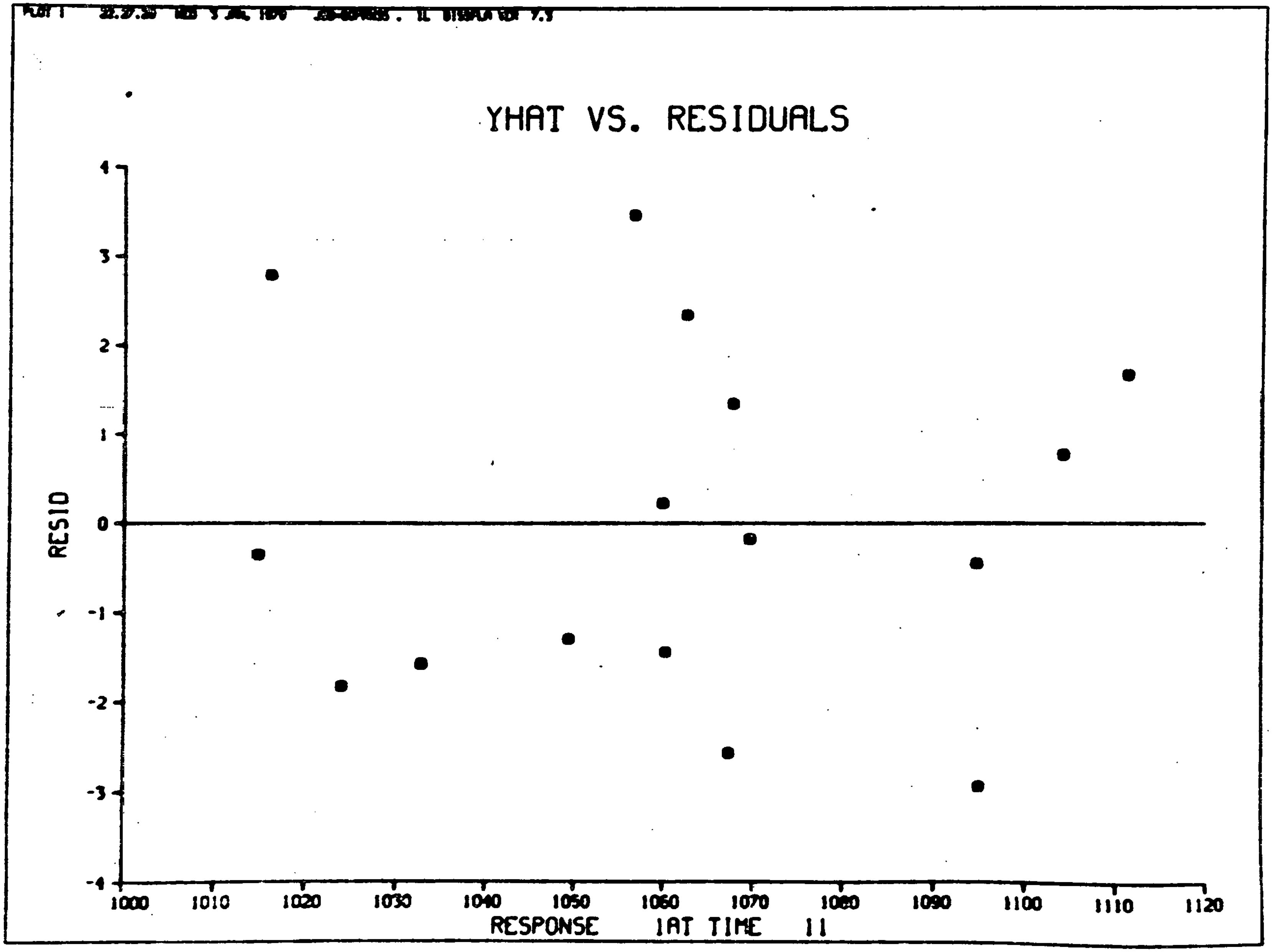

Fig. 1 Cladding surface temperature (K) residuals at 5 seconds (ANYOLS).

small but this is because they still represent only lack of fit to the data. The user should feel confident that this equation adequately models the true response in the region about the nominal.

3.2.4 Example 2: Overlift. Fuel centerline temperature at 15.0 seconds is a response that has been overfit to the data by PRESS. The residual plot, Figure 2 (response 3 at time increment 31), shows very small residuals of equal magnitude alternating in sign. This indicates the addition of terms to the response equation that just counter the effect of other terms without actually contributing a new effect. The PRESS equation included fifteen of a possible sixteen terms in the equation. Curiously, one term contributed 82% of the calculated variance and one might have expected PRESS to stop adding terms shortly after adding that one. It appears instead that most of the remaining terms were of roughly equal significance and so PRESS just continued to add them. A prudent user might be tempted to examine the equation that includes only those terms contributing a significant contribution to the variance as a possible fitting criteria.

3.2.5 Example 3: Underfit. Cladding permanent hoop strain at 5.0 seconds (response 7 at time increment 11) illustrates a response that definitely suffers from an inadequate model approximation. The PRESS equation chose only eight terms out of sixteen to fit the data, yet the residual plot, Figure 3, shows highly grouped data and a large difference in the magnitude of the residuals. The equation fit the data as well as could be expected with the terms available but obviously higher order terms would provide a better fit and equation. This is not an unexpected result since cladding permanent strain is known to be a very sensitive function of local rod conditions. The user would be advised to expect large biases to enter into any estimates about distributional parameters made by this equation.

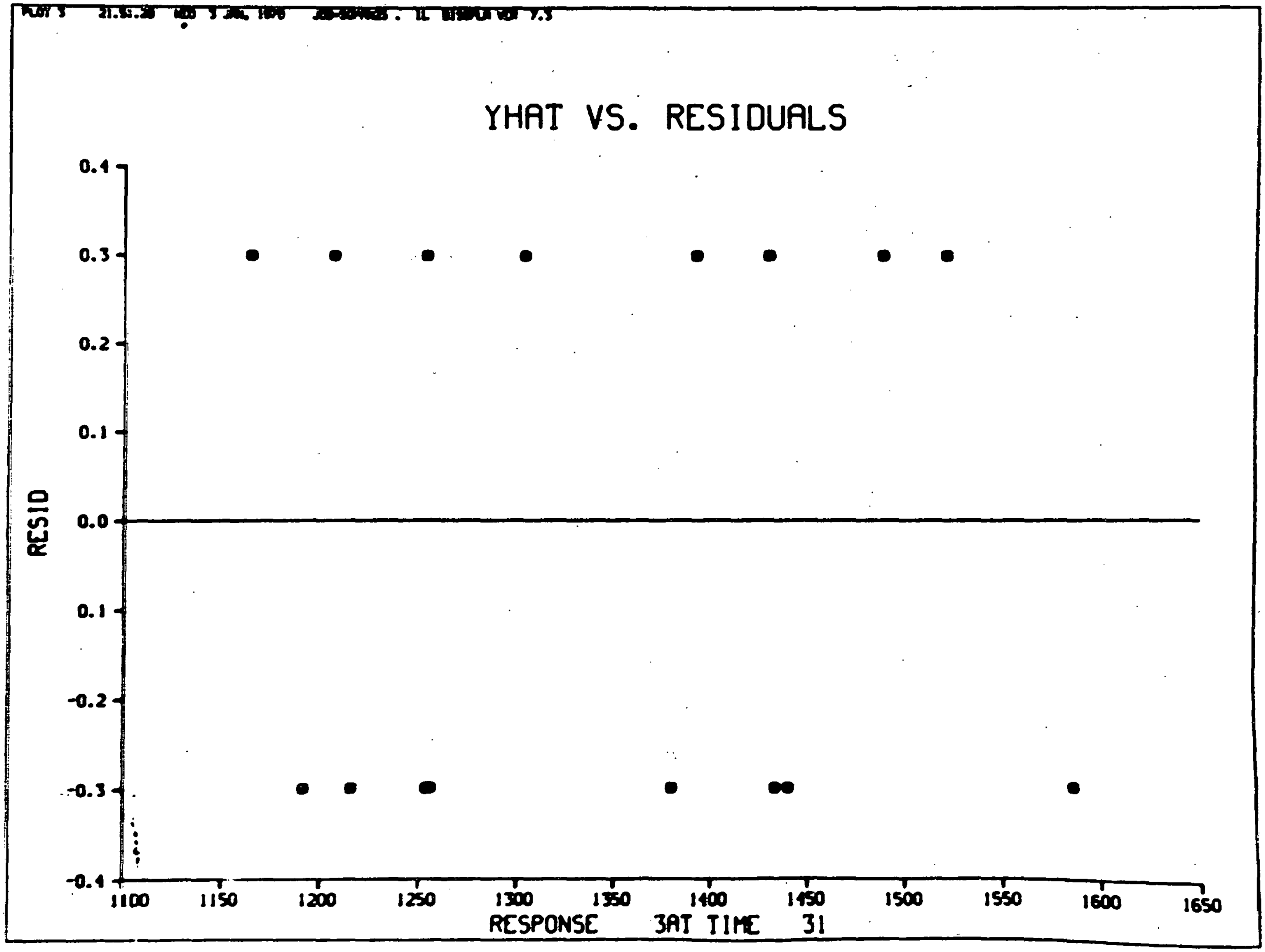

Fig. 2 Fuel centerline temperature (K) residuals at 15 seconds (ANYOLS).

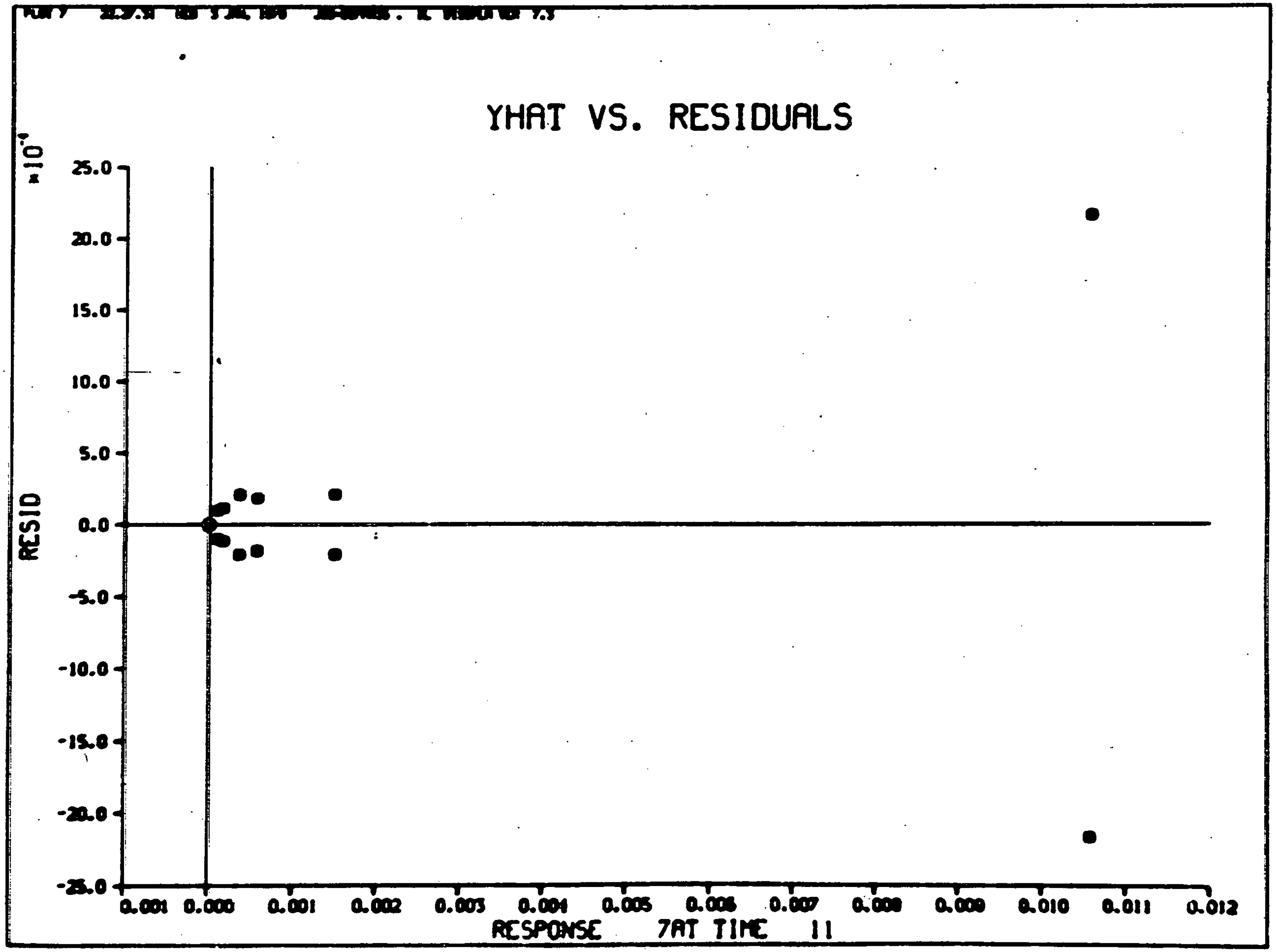

Fig. 3 Permanent cladding hoop strain residuals at 5 seconds (ANYOLS).

3.2.6 Example 4 = Outliers in the Data. Heat transfer across the gap at 5.0 seconds (response 6 at time increment 11) is an example of a data set that includes an outlier, or obviously different point from most of the data. The PRESS equation has fit the data in only four terms yet the residual plot, Figure 4, shows that the entire fit has been biased by one very low data point.

In a traditional analysis the experimenter might be tempted to throw the point out of the data set as possibly erroneous. That is not possible here and this case in fact represents a slice in time near the beginning of a physically real phenomena that exhibits a threshold behavior. That is, the gap heat transfer in a LOCA will markedly degrade when the gap opens up during the course of the transient. Just when that occurs appears to be a function of some of the relative values of the input parameters. The problem here is that while a first or second order response equation may reasonably approximate the true behavior on either side of the threshold, across the threshold the approximation is inadequate. Including higher order terms and the necessary runs to estimate them may not solve the problem since in some physically real cases these thresholds approach step functions. An alternative is to divide the data into two groups when this happens and analyze each individually. This is the practice recommended to the user when this occurs.

3.2.7 Summary. One further observation regarding ANYOLS, PRESS and residual plots is in order. PRESS, or for that matter most other statistical criteria, was developed assuming that in general the number of data would largely exceed the terms of an assumed model. In particular, the data were assumed to contain a random element of error and replicate experiments might not produce identical results. In our case this is not true. The output of a computer code can be observed without error and replication only produces identical results. Furthermore, the number of data generated are ususally close to the number of terms being fitted. The mean squared deviation of the residuals is thus purely due to fitting and could potentially be made arbitrarily small by including terms up to the full model. The result is that the user should treat any response

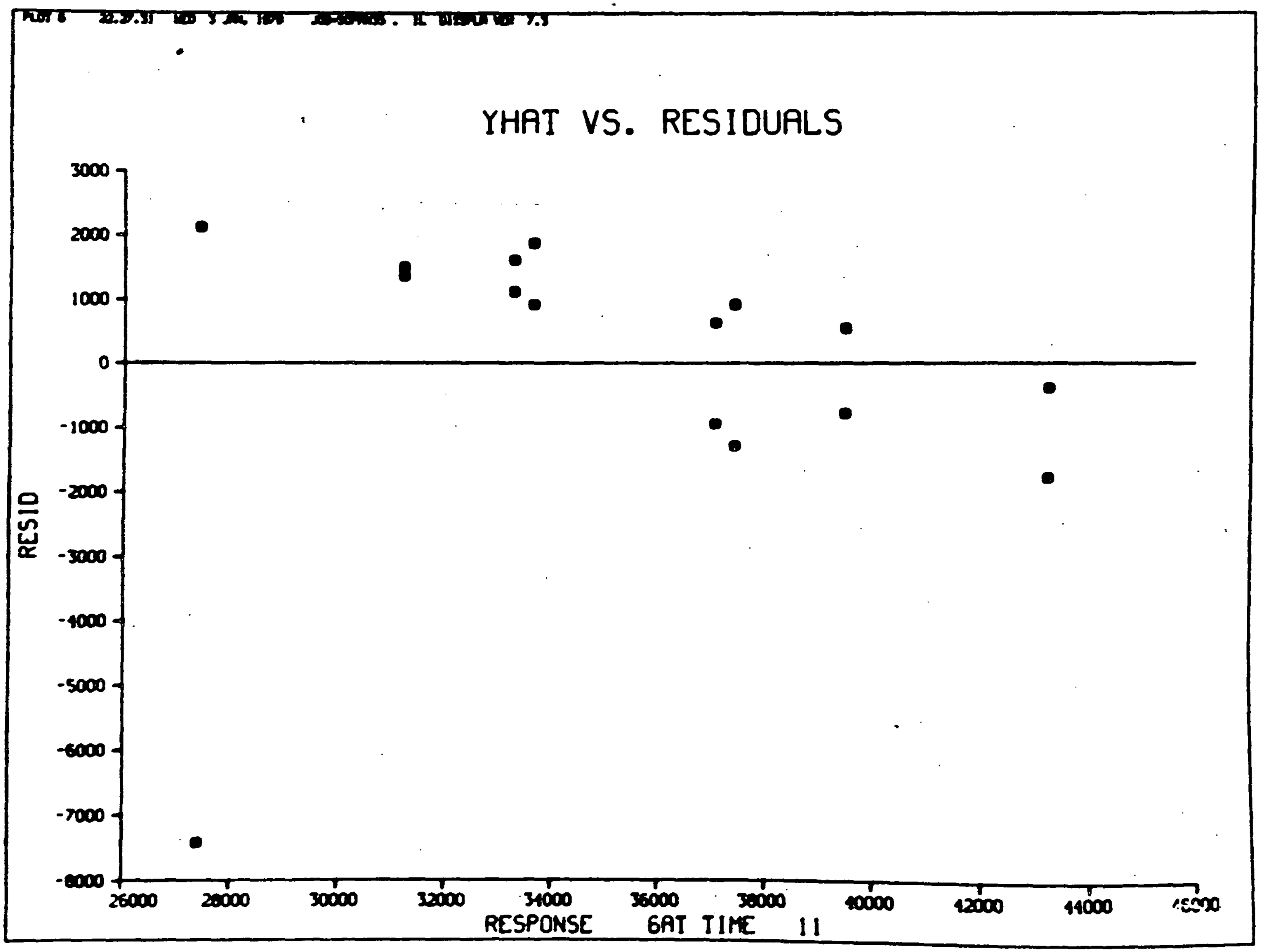

Fig. 4. Gap heat transfer coefficient (W/m^2-K) residuals at 5 seconds (ANYOLS).

equation with caution until the residuals have been examined and the equation checked for physically meaningful behavior over the region in which it was generated. Once this has been done, the validated results may be reported with confidence.

3.3 Analysis of Response Equation Interpolation and Extrapolation

The previous section has illustrated how the response equations fit the calculated data. The next logical question is how well the fitted equations predict the true responses at points other than the design points. That is, how well do the equations interpolate or extrapolate? In one dimension interpolation or extrapolation is well understood. In more than one dimension, however, the issue is not as clear, especially with the sparse experimental point distribution of a fractional factorial design. For example, points that are a multiple (for example 2) of the original fraction are clearly extrapolated while points that are a fraction (for example, 1/2) are interpolated. Not as clear are points that are taken from another fraction of the fractional factorial (multipliers of - 1/2, or -1, for example). Keeping in mind the previous analyses of equation fit, it seems worthwhile to determine the performance of the response equations at off-design points and so the range of their validity.

3.3.1 Generation of Off-Design Points. To generate off-design data, the sample problem discussed in Section 4 was run an additional 16 times. The original design was divided into four groups of four runs each and the new runs were made at (1) one-half the original design, (2) minus one-half the design, (3) minus one times the design, and (4) twice the design. The runs for each group were chosen so that each factor always had plus and minus values within a group. This way a distinction could be made within categories as well as between them. The object in choosing these particular off-design points was to reasonably cover the sample space from interpolation to extrapolation. The ability of the response equations to predict new data is illustrated in the following examples continued from the previous section.

3.3.2 Examples of Off-Design Residuals. Figure 5 shows cladding surface temperature (5 seconds) off-design residuals. The square symbols corresponds to half design points, circles to minus half design points, triangles to minus design points, and pluses to twice design points. In Subsection 3.2.3 this response equation was judged a good fit. The off-design residuals show that even at twice the design the maximum residual is only about 2% of the predicted mean. The judgment therefore continues to hold and the response equation appears useful as a predictor within the sample space.

Fuel centerline temperature off-design residuals (15 seconds) are shown in Figure 6. This equation overfit the available data and cautious use of extrapolation was recommended. In fact, the first three kinds of off-design points consistently overpredict the response (residual = response - equation) while the twice design points seriously underpredict the response. This indicates the presence of positive upward curvature in the response not previously detected by PRESS. This curvature can be detected by comparing the nominal run with the constant term in the response equation. The nominal is an off-design point included in every analysis and the constant is the response equation evaluated at nominal conditions. The response equation overpredicts the nominal (it is important to remember that the nominal was not used in this case to fit the response equation) in this example and, in fact, for most of the LOCA history for this response. Within the sample space the magnitude of the bias appears to be reasonable (~2%) but any extrapolation is very risky due to the possibility of curvature in the response.

Figure 7 shows permanent cladding hoop strain (5 seconds) off-design residuals. This equation already lacked needed higher order terms and the off-design residuals fared no better. The residuals are the same magnitude as the predictions, both over and underpredicted, and show no discernable trends toward which kind of off-design points fit better. Thus, the previous conclusion that large biases could be expected still holds.

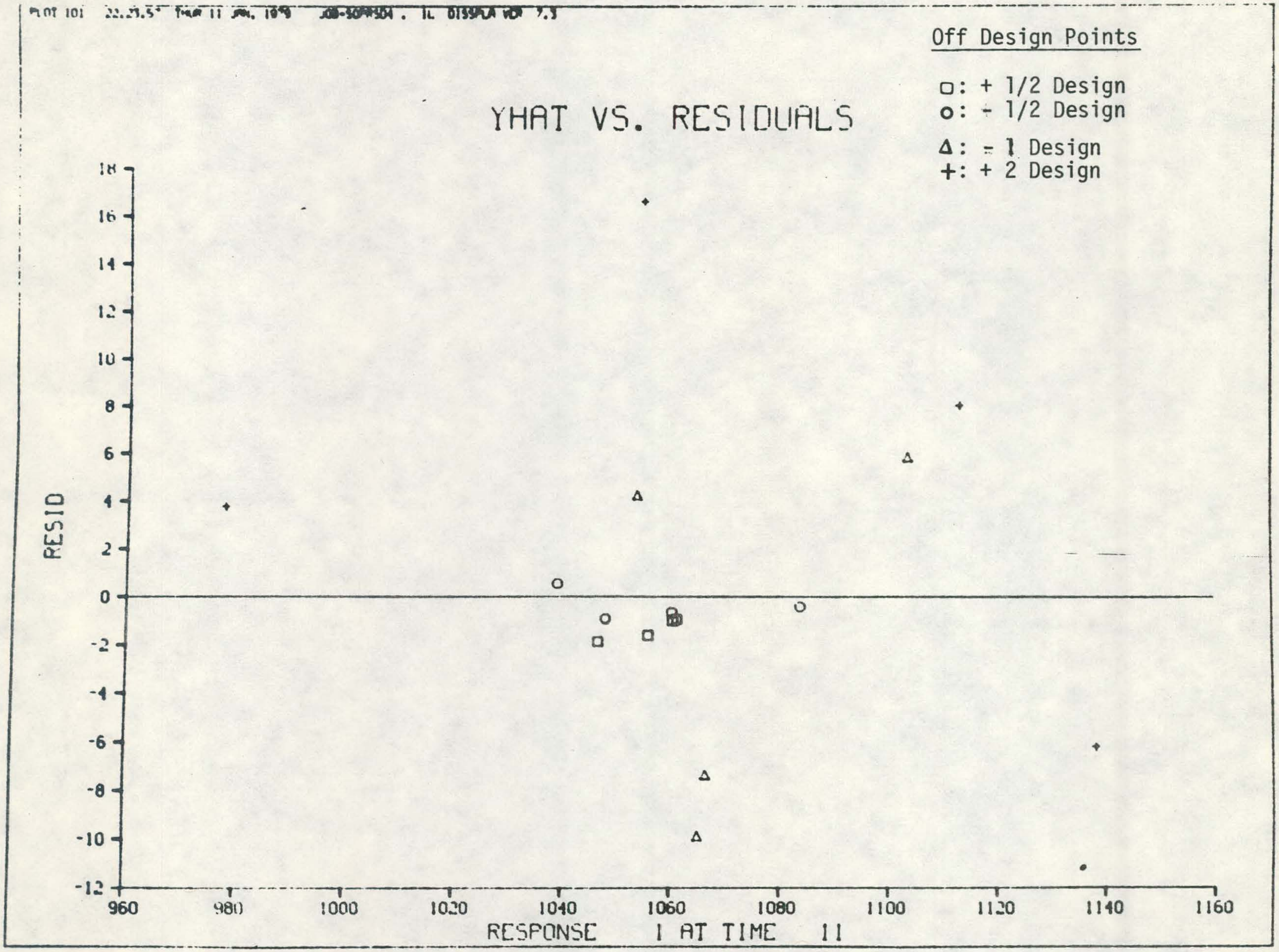

Fig. 5 Cladding surface temperature (K) off design residuals at 5 seconds.

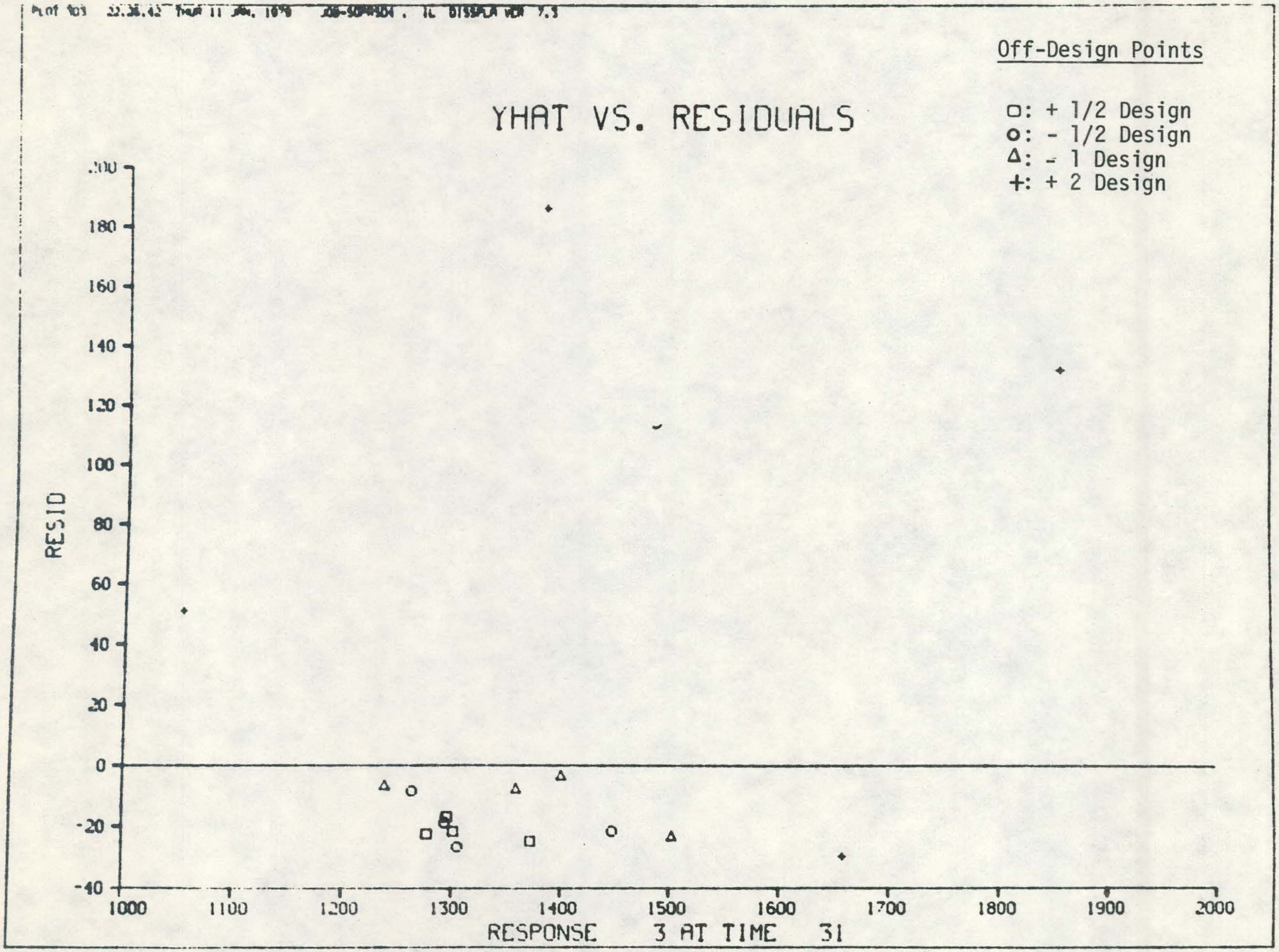

Fig. 6 Fuel centerline temperature (K) off design residuals at 15 seconds.

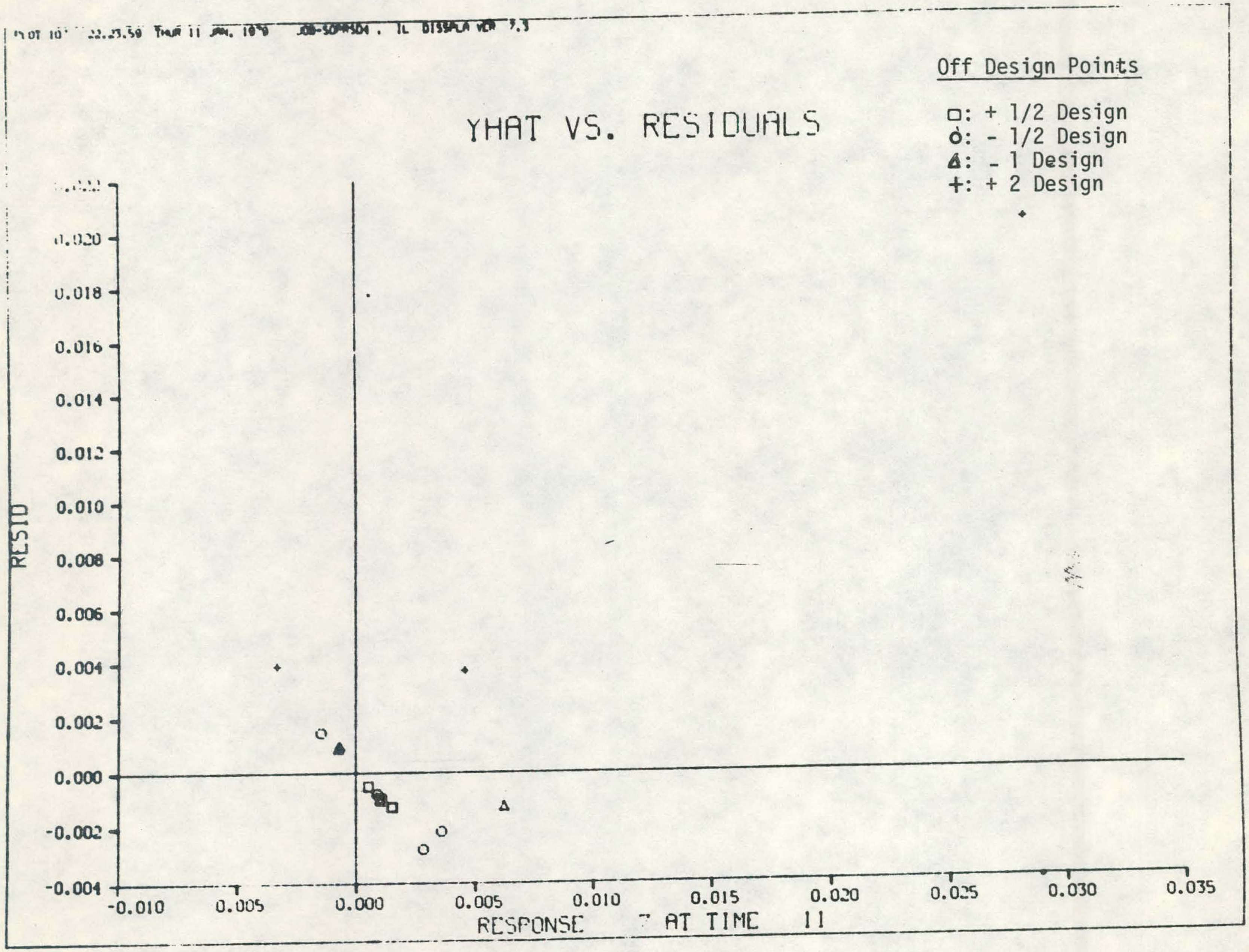

Fig. 7 Permanent cladding hoop strain off design residuals at 5 seconds.

Gap heat transfer coefficient (5 seconds) off-design residuals, Figure 8, show the same trends as the original residuals. This equation contained the effects of one extreme data point caused by a threshold behavior in the fuel rod physical response. The majority of off-design residuals are reasonably well behaved and don't exceed roughly 10% of the response. Two points at twice the design appear to cross the threshold and are widely dispersed. Coincidentally, one point was twice the design of the same point that caused the original bias. This point was still fit rather well. The other point obviously didn't do as well. The original conclusion reached was that the data should be grouped depending on whether the threshold had been reacted or not. This still appears to be true with the added comment that extrapolation of any kind might be risky since the threshold could inadvertently be crossed.

4. SAMPLE PROBLEM

4.1 Introduction

The sample problem used in the developmental verification studies was a Zion fuel rod subjected to the blowdown portion of a double ended cold leg break LOCA. Beginning of life conditious were assumed and the only deviation from nominal conditions was artifically raising the average fuel rod power to 42.1 kW/m. A RELAP4 (MOD5)[6] analysis was used for the deterministic thermal hydraulic boundary conditions and the uncertainty analysis option itself used FRAP-T5 for the fuel rod calculations. Complete results of the analysis are included on microfiche at the back of this report.

The complete uncertainty analysis used 30876 system seconds on the CDC 7600 and cost about $1390. Sixteen individual FRAP-T5 runs would have used 29800 system seconds ($1341), so the uncertainty analysis subcode used about 1076 system seconds and cost $49, which includes the cost of the nominal FRAP-T5 execution.

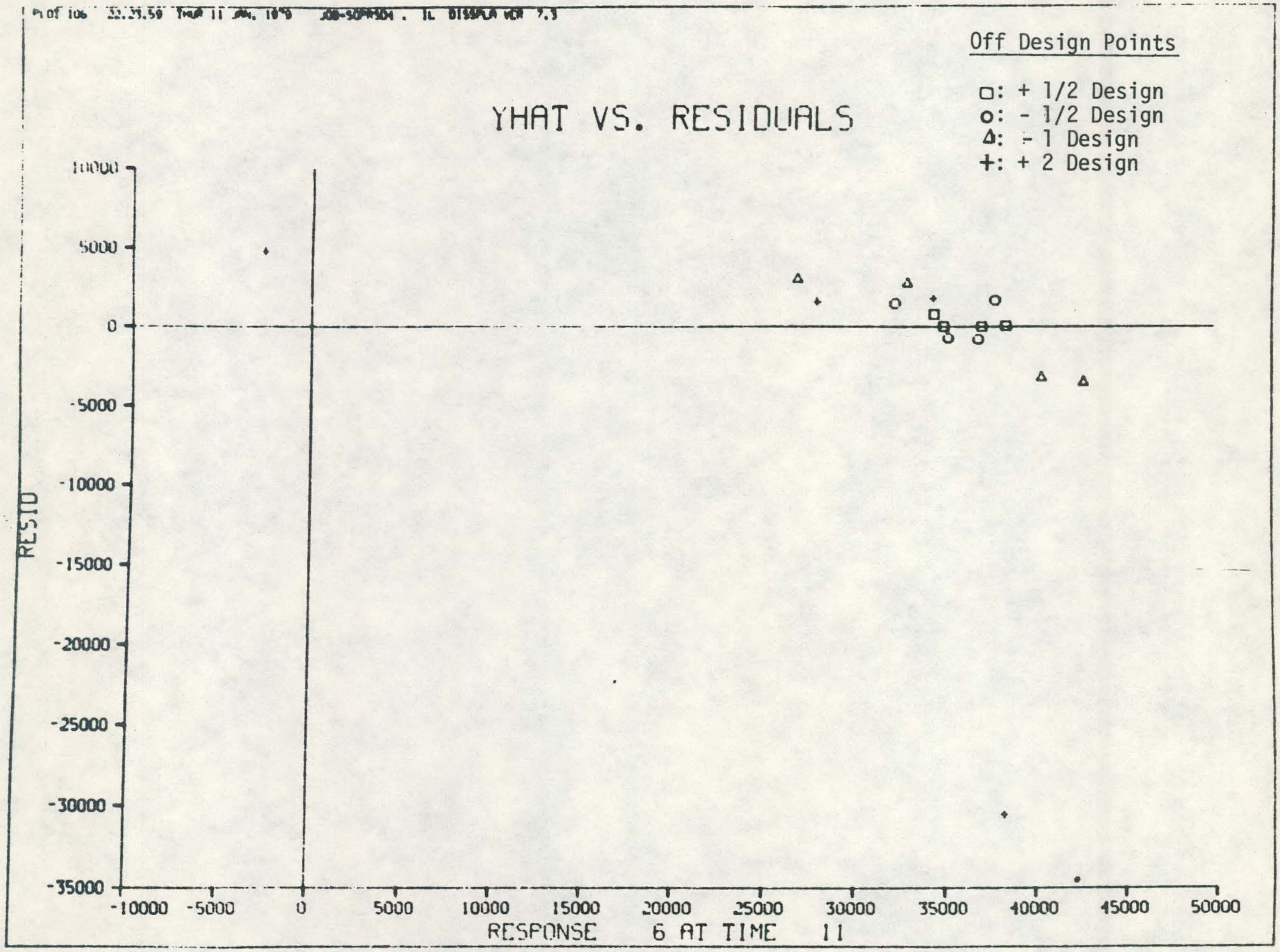

Fig. 8 Gap heat transfer coefficient (W/m^2-K) off design residuals at 5 seconds.

4.2 Design

Results of a previous LOCA uncertainty analysis[7] indicated that 10 input variables accounted for a large majority of the uncertainty in responses for this problem. These 10 variables and their uncertainties are shown in Table I. The fractional factorial design for 10 factors is made in 16 problem executions and so allows the estimation of 5 interaction (crossproduct) coefficients as well as 10 main effects and a constant. Ten factors have 45 two factor interactions so the factors were carefully ordered so that the five interactions thought most likely to be influential could be estimated. The estimable interactions included the products of fuel thermal conductivity and pellet radius, gas thermal conductivity, and power, respectively, and the products of pellet radius and cladding diametral thermal expansion, and cladding radius, respectively. A total of 10 responses from a variety of thermal and mechanical code outputs were chosen as representative of the fuel rod behavior during a LOCA. These responses are listed in Table II.

4.3 Results

The 10 calculated responses were sampled at half second intervals throughout the 29 second blowdown history of the LOCA, generating 590 sets of 16 data points. A response equation was fit to each data set and used to estimate the mean and variance of each response at each point in time. Finally, fractional contributions to the variance of each term were determined.

Developmental verification studies, described in detail in preceding sections, analyzed the fit to the data of response equations chosen by the PRESS criterion. Then, each of the full model equations was analyzed for its ability to predict responses at off design points. Results of these studies determined which equations reasonably approximated the true response surface and which required the addition of higher order terms or other independent variables. The results for each response may be summarized as follows. Responses 1 and 2, cladding surface temperatures, are well approximated by a linear function throughout most of the sample

TABLE I

UNCERTAINTY FACTORS

	Factor	Value (σ)
1.	Fuel thermal conductivity	0.4 W/m-K
2.	Cladding roughness	10%
3.	Fuel density	0.67%
4.	Cladding diametral thermal expansion	10% T >1073 (k) 50% T <1073 (k)
5.	Pellet shoulder radius	3.4%
6.	Fuel roughness	10%
7.	Cladding inner radius	1%
8.	Pellet outer radius	1%
9.	Gas thermal conductivity	$-.00068 + 1.61 \times 10^{-6} T$ W/m-K
10.	Transient power level	5%

TABLE II

RESPONSES
(11 equally spaced axial nodes)

	Response	Axial Node
1.	Cladding surface temperature (K)	6
2.	Cladding surface temperature (K)	7
3.	Fuel centerline temperature (K)	6
4.	Fraction of failed fuel rods	
5.	Gap heat transfer coefficient (W/m^2-K)	6
6.	Gap heat transfer coefficient (W/m^2-K)	7
7.	Permanent cladding hoop strain	6
8.	Permanent cladding hoop strain	7
9.	Cladding hoop strain	6
10.	Gap pressure (N/m^2)	6

problem. Response 3, fuel centerline temperature, and response 10, gap pressure, are fit acceptably well although some bias due to local curvature in the response surface is indicated. Response 4, fraction of failed rods, and responses 7, 8 and 9, cladding strain, are not fit well by a linear approximation. Cladding strains, particularly permanent strains, are very sensitive functions of cladding conditions and this result is not unexpected. Response 4 is a function of cladding strain and so follows in behavior and response 9, total hoop strain, shows similar, although not as pronounced, behavior due to contributions from thermal expansion and elastic strains. Responses 5 and 6, gap heat transfer coefficients, are fit well except during threshold behavior. The recommendation for these particular responses was to divide the data into groups and reanalyze. Figures 9-18 show the means plus and minus one standard deviation for all responses for completeness. Only those for which a linear fit is justified should be interpreted in a meaningful manner. Note, for example, that minus one standard deviation for response 4, fraction of failed rods, is less than zero. This is clearly erroneous. At the very least the output distribution is not normally distributed as the figure might imply. More likely the original equation is inadequate, as stated before, and thus produces erroneous results.

One of the objectives of this sample problem was to look at the contributions of two factor interactions to the variance of estimated responses. The results showed that the crossproduct of fuel thermal conductivity and power level was a significant contributor to the calculated uncertainties in cladding strains and particularly, gap heat transfer. Unfortunately, the uncertainties in cladding strains are probably not valid since the model approximation was inadequate. The true uncertainties are likely to be larger and therefore, the relative contribution of a particular term like a crossproduct may diminish. Nevertheless, an orthogonal design estimates coefficients independently of one another and the contribution is significant.

In summary, the results of developmental verification studies applied to this sample problem showed that a linear approximation for cladding surface temperature, gap pressure, and heat transfer coefficients produced meaningful uncertainty estimates. If these were the primary responses of interest to the user, the analysis could stop here. Cladding strain and related responses were not fit well and meaningful results were not obtained. If these responses were of particular interest, the analysis should continue by generating more data for estimating higher order terms.

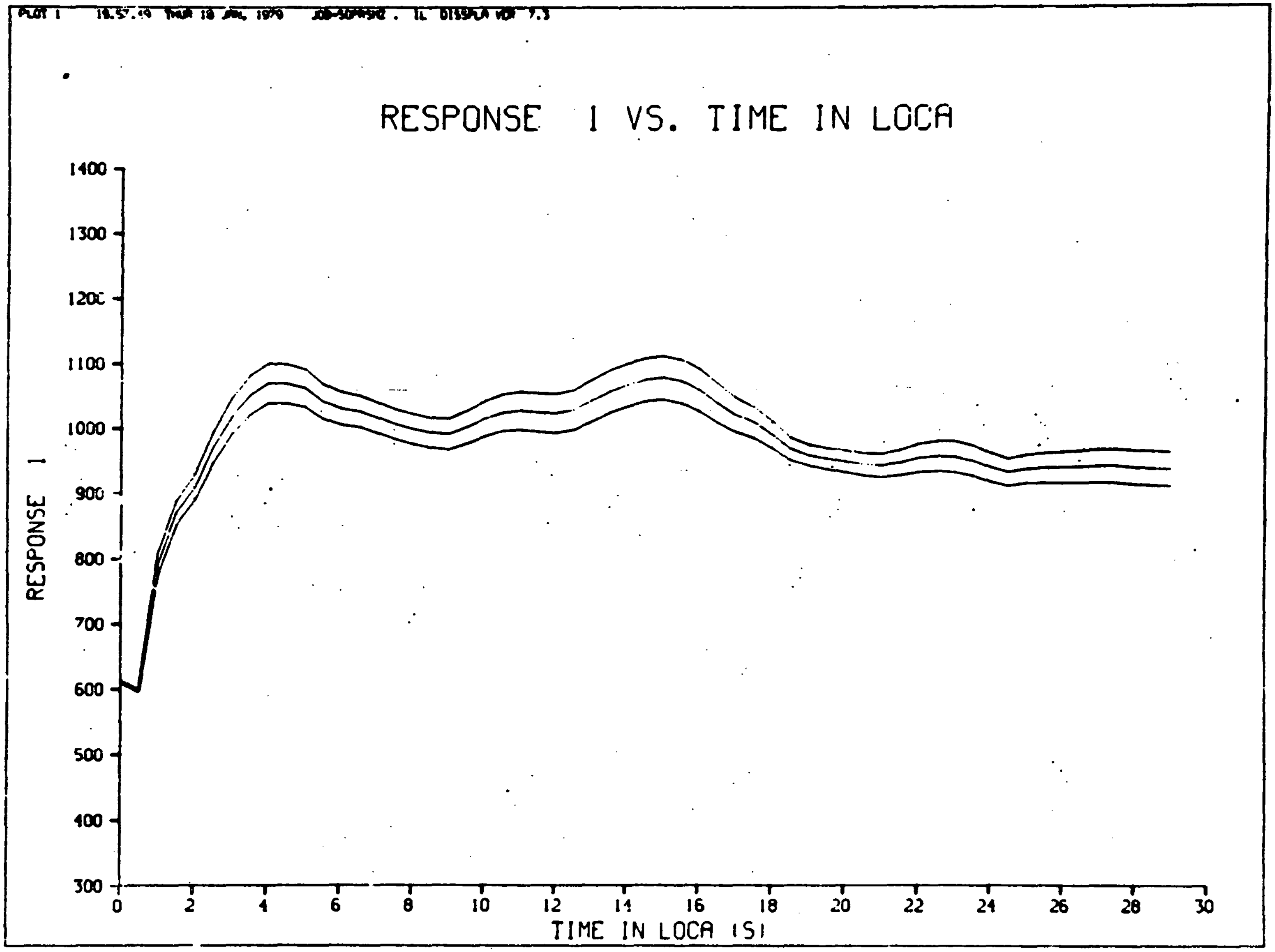

Fig. 9 Cladding surface temperature (K) at Node 6, mean response ± 1 standard deviation.

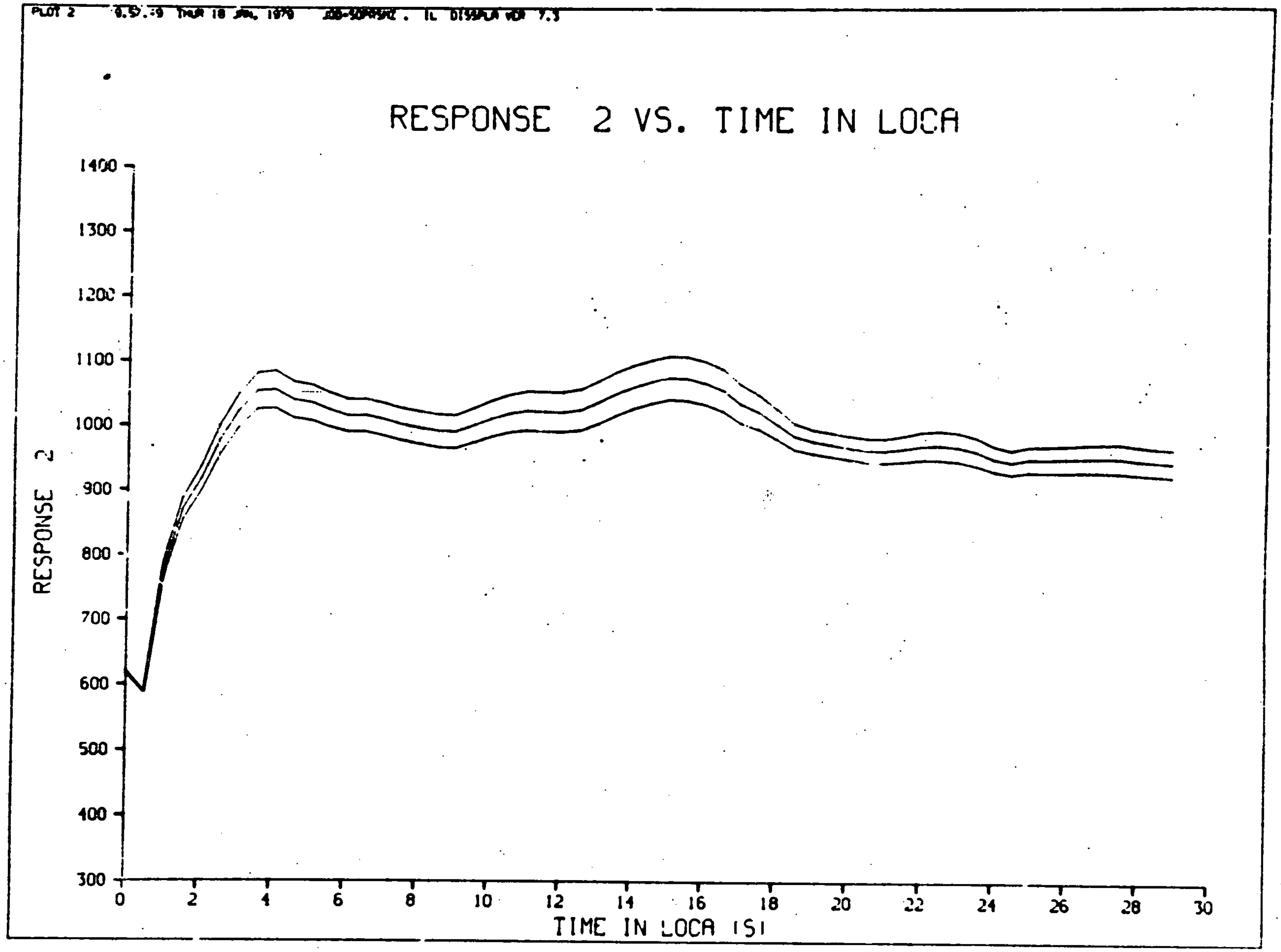

Fig. 10 Cladding surface temperature (K) at Node 7, mean response ± 1 standard deviation.

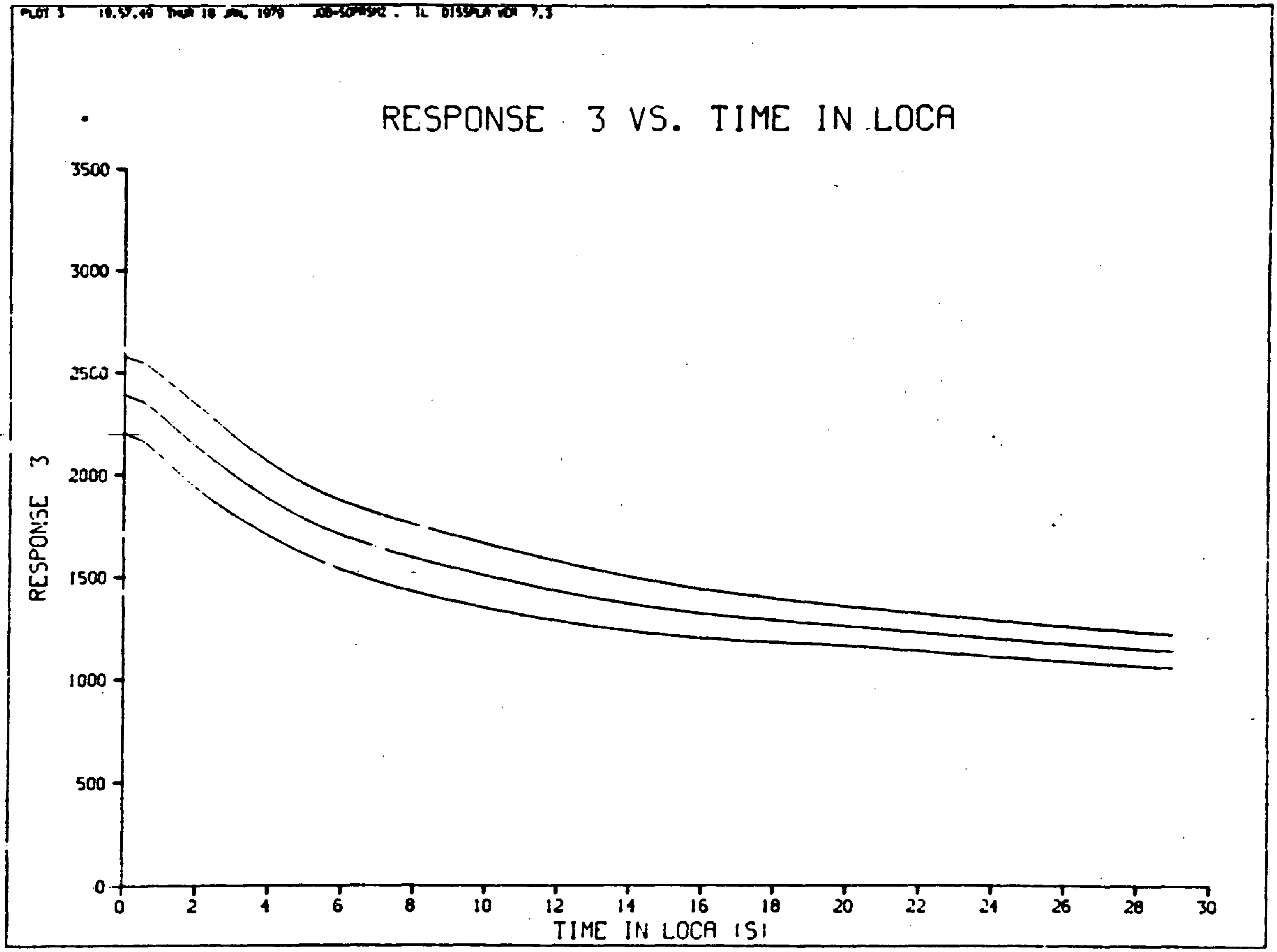

Fig. 11 Fuel centerline temperature (K) at Node 6, mean response ± standard deviation.

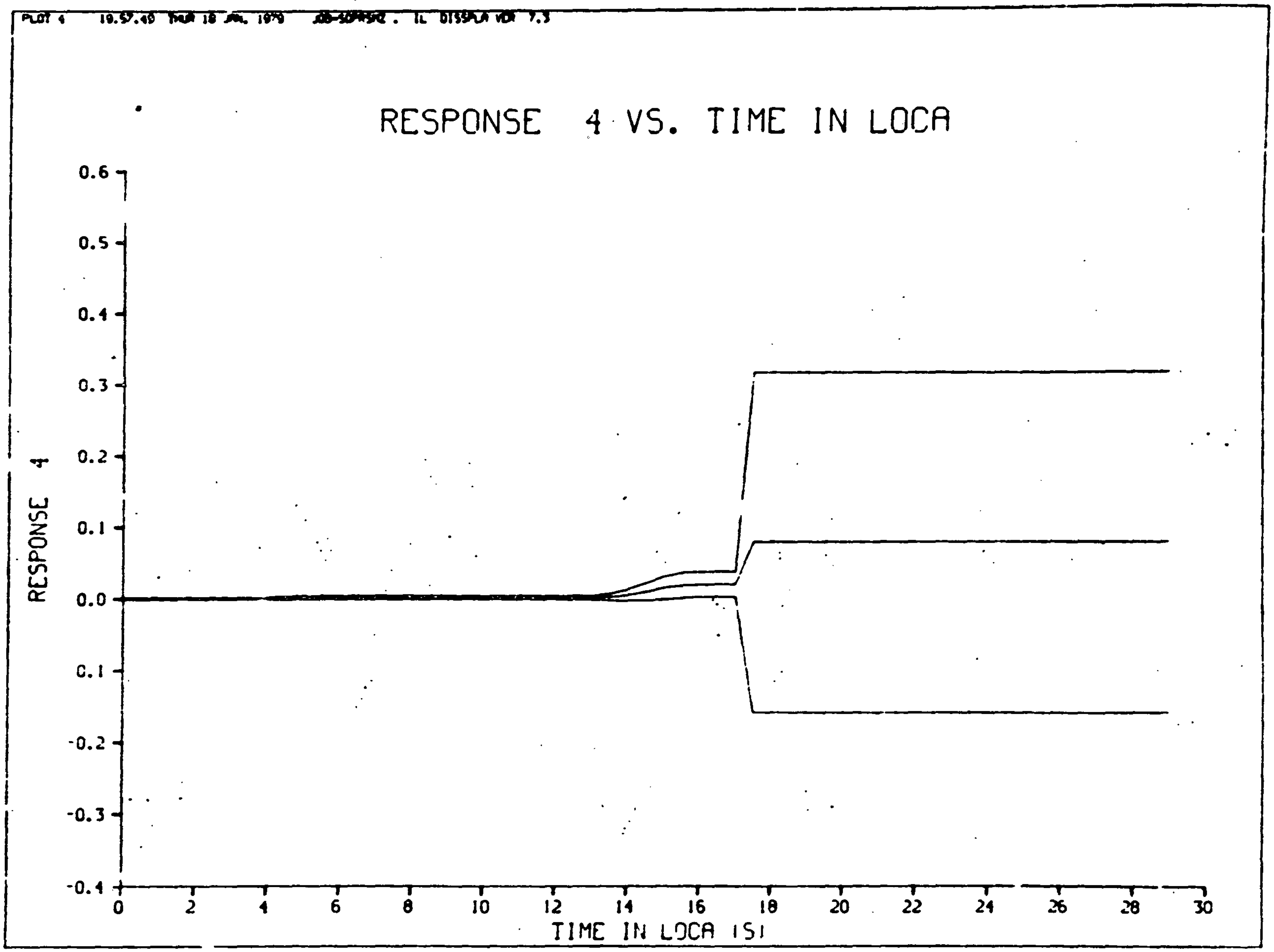

Fig. 12 Fraction of failed fuel rods, mean response ± 1 standard deviation.

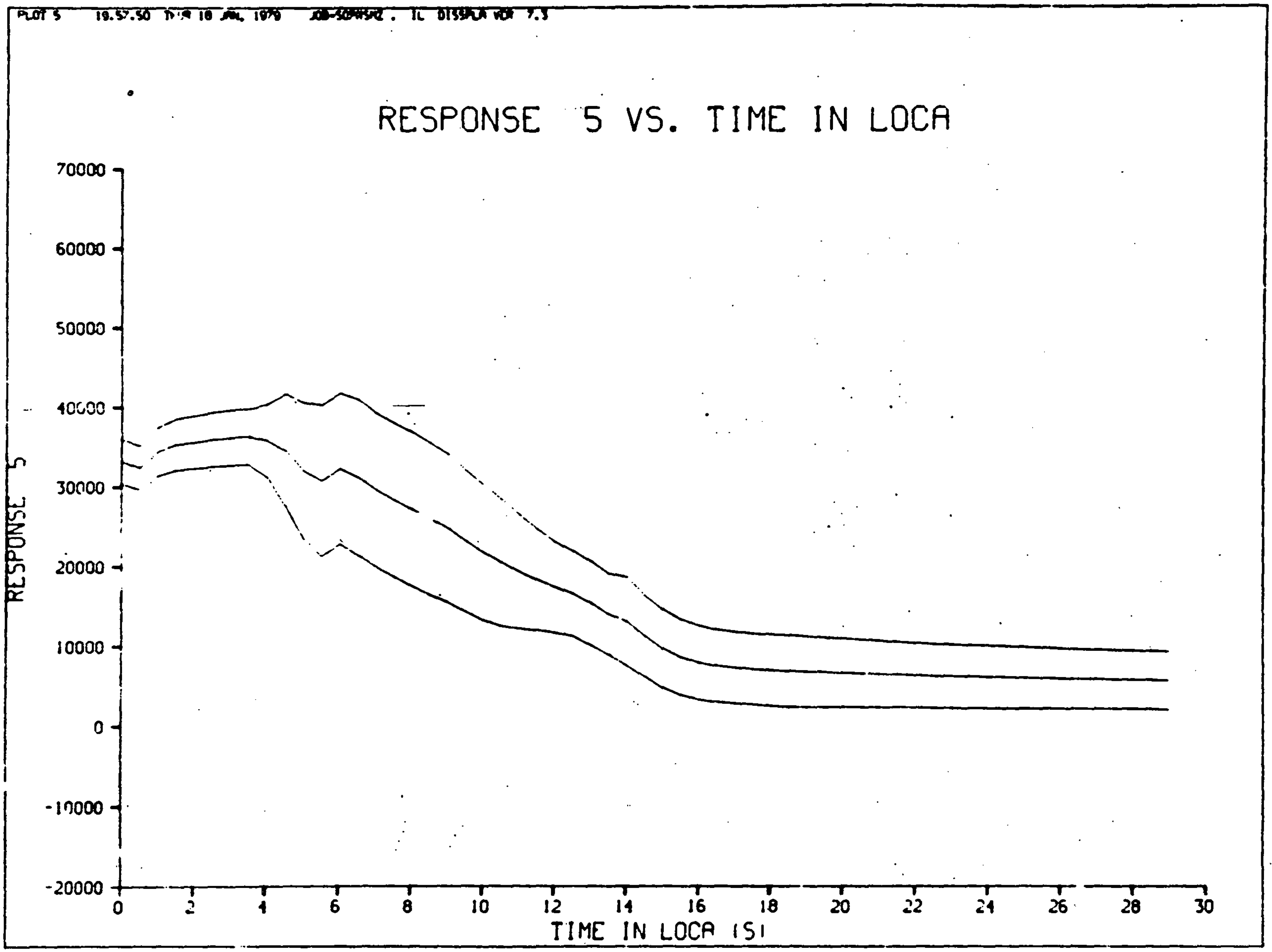

Fig. 13 Gap heat transfer coefficient (W/m^2-K) at Node 6, mean response ± one standard deviation.

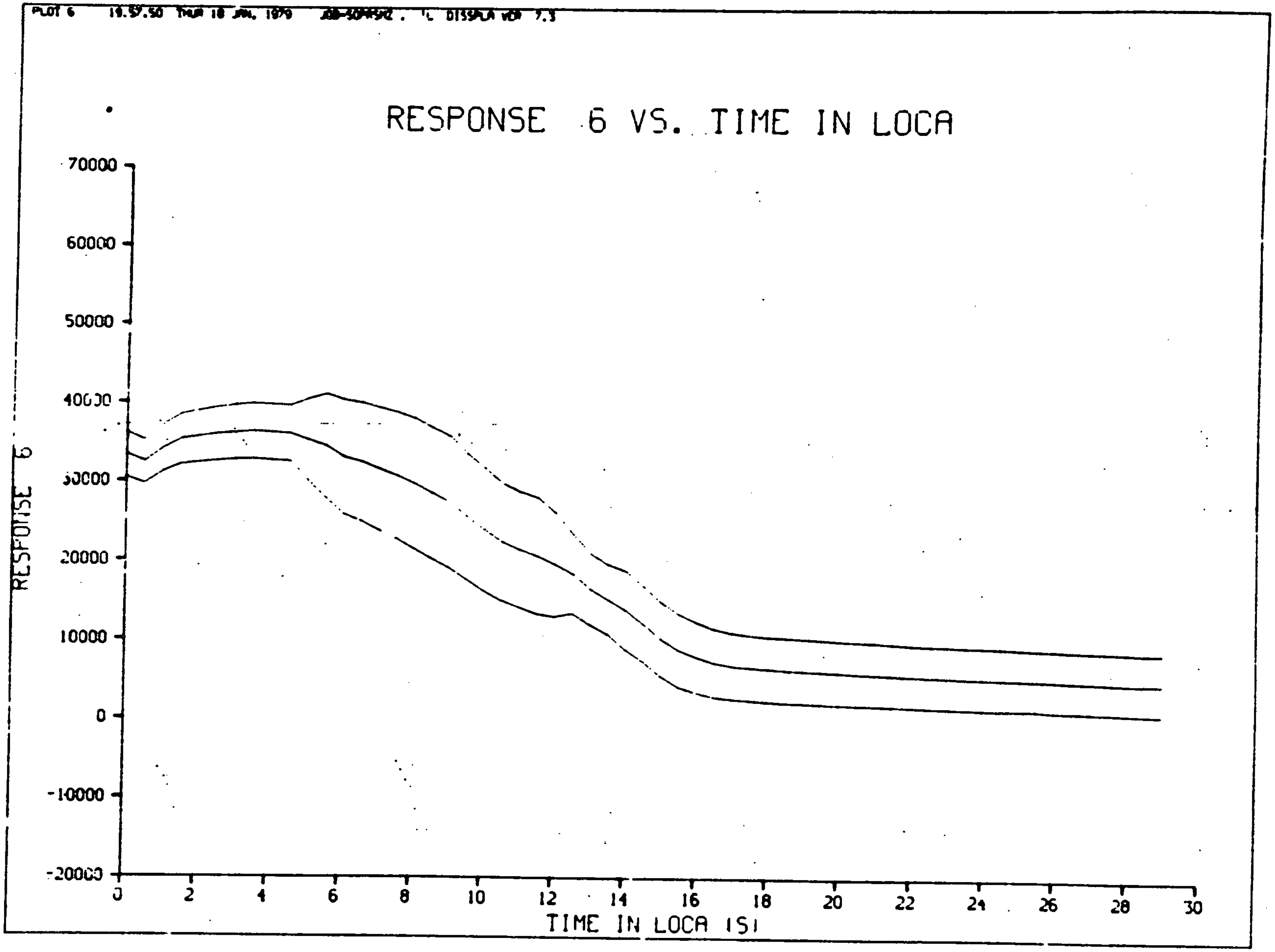

Fig. 14 Gap heat transfer coefficient (W/m^2-K) at Node 7, mean response $\pm$ 1 standard deviation.

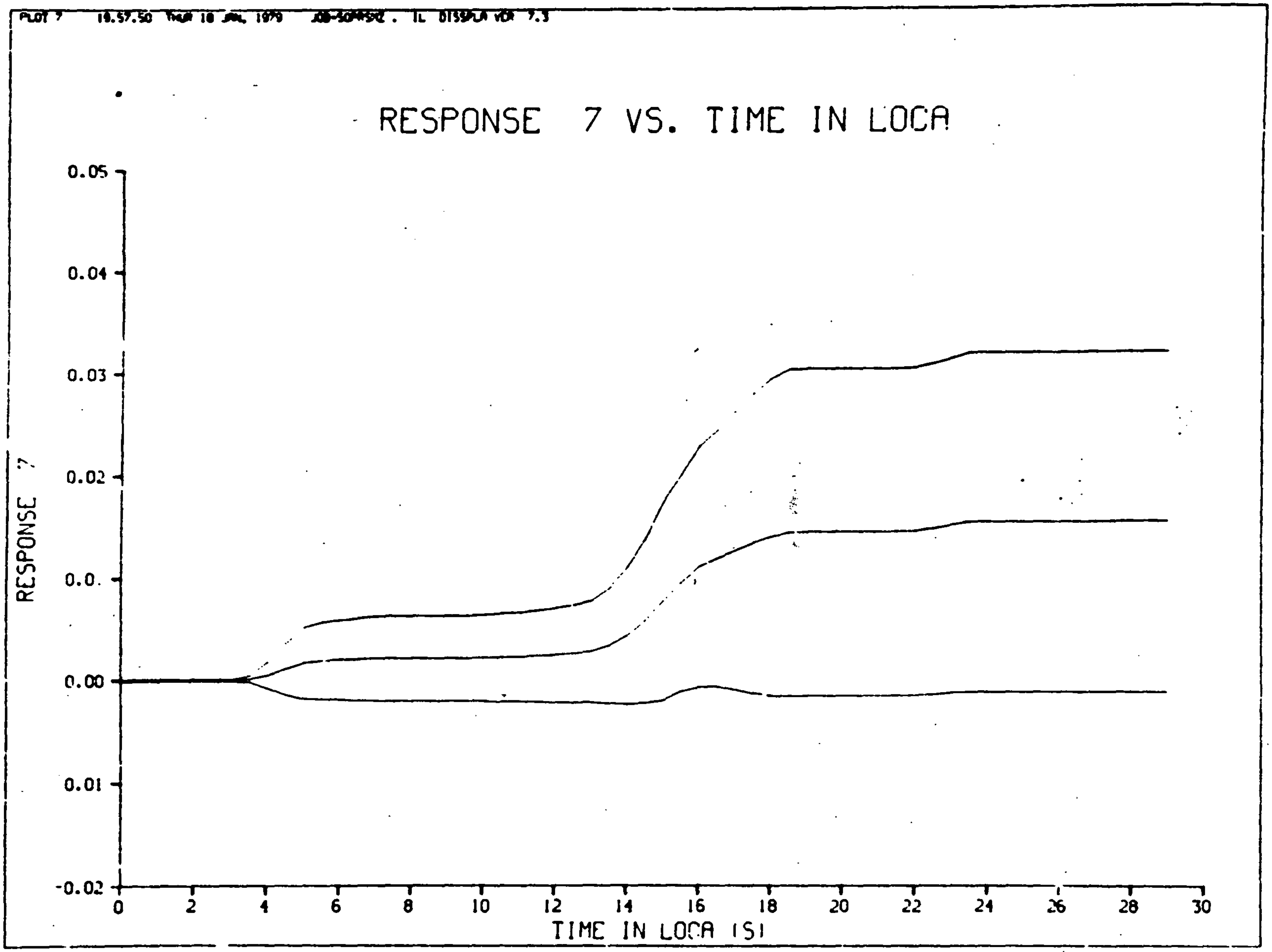

Fig. 15 Permanent cladding hoop strain at Node 6, mean response ± 1 standard deviation.

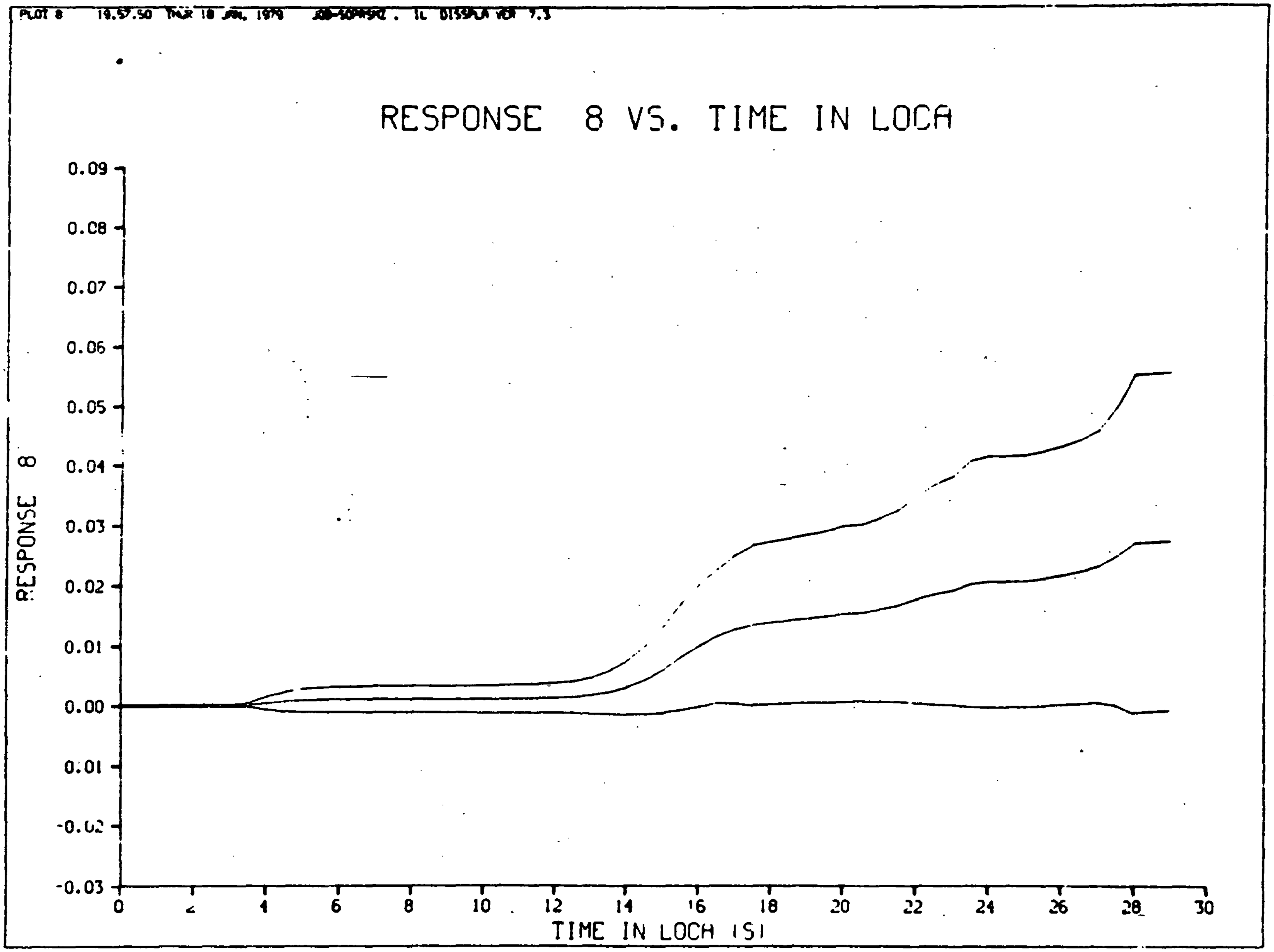

Fig. 16 Permanent cladding hoop strain at Node 7, mean response ± 1 standard deviation.

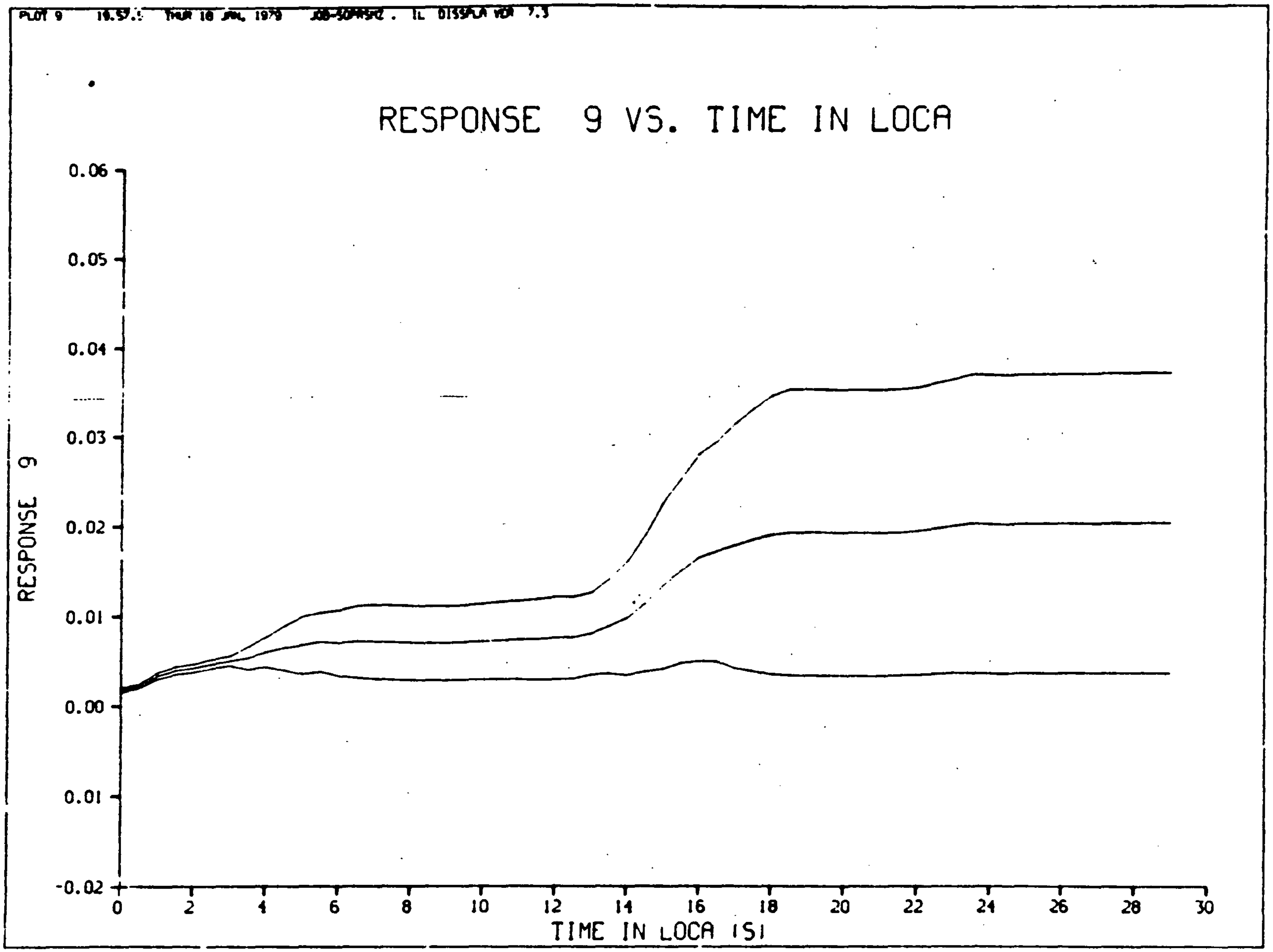

Fig. 17 Total cladding hoop strain at Node 6, mean response ± 1 standard deviation.

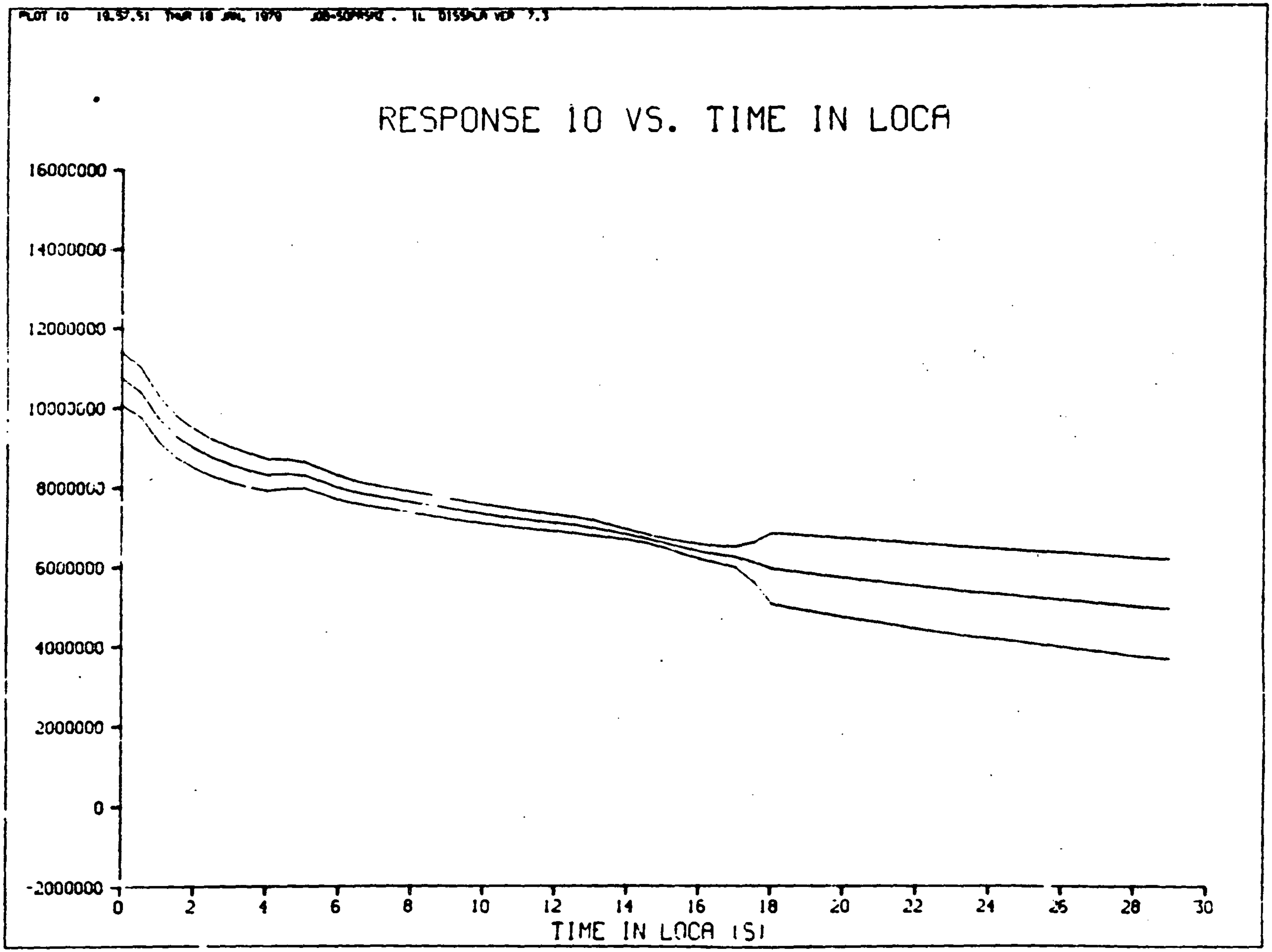

Fig. 18 Gap pressure (Pa), mean response ± 1 standard deviation.

5. CONCLUSIONS AND RECOMMENDATIONS

The developmental verification of the Uncertainty Analysis Option incorporated in the FRAP-T5 code included a comparison of the code predictions with a hand solution in order to verify the correctness of the subcode programming, and a study of the adequacy of the fit of the approximate response surfaces generated for a hypothetical LOCA example problem. This study showed that the majority of the code output responses were adequately modeled by a simple linear response surface equation. For those responses for which the linear response surface fit was less than adequate, the fit could potentially be improved by continuing the analysis to include quadratic terms in the response surface equations. The adequacy of fit is highly problem dependent, and a study of response surface requirements for the typical hypothetical accidents important to Light Water Reactor Safety Analysis is underway. This study will result in specific recommendations for the choice of response surface equation forms for the accident scenarios in question.

In the current version of the Uncertainty Analysis Option (as incorporated in FRAP-T5) the correct choice of the completeness of the response surface equation is left to the user. In the developmental verification, the ANYOLS code was used to automatically ascertain the best form of the response equations. In addition, it was found that the response equation selected by the ANYOLS code was sometimes more accurate than using the full model response equation. Thus it is recommended that calculations such as performed in ANYOLS be incorporated in future versions of the FRAP-T Uncertainty Analysis Option. In addition, it is recommended that future versions should include a method for graphically plotting the nominal responses and comparing them to the estimated mean and variance. An alternative method would be to overlay plots of all the data together with a plot of the mean and nominal. In this way the user could visualize the data spread and the location of the mean and nominal relative to the data.

Finally, it should be noted that the evaluation of response equation validity relies heavily on visual techniques. Ideally, a quantitative measure should be found that could be used in the same manner as a confidence limit. That is, response equation X can be used with Y% confidence. The possibility of incorporating such confidence limits will be studied as part of the on-going code optimization tasks to be performed during the remainder of FY-1979.

6. REFERENCES

1. FRAP T-5 A Computer Code for the Transient Analysis of Oxide Fuel Rods, EG&G, Idaho, Inc. Report CDAP-TR-79-043, to be published in March, 1979.

2. S. O. Peck, FRAP Uncertainty Analysis Option, EG&G, Idaho, Inc., Report CDAP-TR-78-024, July, 1978.

3. N. D. Cox, An Investigation of Methods for Uncertainty Analysis of the Best-Estimate RELAP4 Models, Aerojet Nuclear Company Report No. SRD-95-76, May 1976.

4. N. D. Cox, A Report on a Sensitivity Study of the Response Surface Method of Uncertainty Analysis of a PWR Model, EG&G Idaho, Inc. Report No. RE-S-77-7, January, 1977.

5. C. L. Atwood, User's Guide to ANYOLS-A Regression Program Allowing Many Ways to Select a Model, EG&G Idaho, Inc. Internal Technical Report No. RE-A-78-049, March, 1978.

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

7. S. O. Peck, Uncertainty Analysis of the FRAP Code, ENS/ANS International Meeting on Nuclear Power Reactor Safety, October 16-19, 1978, Brussels, Belgium.