

LA-UR- - 01 - 5619

Approved for public release;
distribution is unlimited.

C.1

Title: Variable Screening Methods in Metamodel Design for a
Large Structural Dynamics Simulation

Author(s): Amanda L. Cundy, ESA-WR
John F. Schultze, ESA-WR
Francois M. Hemez, ESA-WR
Scott W. Doebling, ESA-WR
Derek Bingham, ESA-WR

Submitted to: The 20TH International Modal Analysis Conference
Los Angeles, CA
February 4-7, 2002



Los Alamos

NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

VARIABLE SCREENING IN METAMODEL DESIGN FOR A LARGE STRUCTURAL DYNAMICS SIMULATION

A.L. Cundy, J.F. Schultze, F.M. Hemez, S.W. Doebling, D. Bingham*
 Los Alamos National Laboratory, ESA-WR
 MS P946
 Los Alamos, New Mexico 87545

ABSTRACT

This research addresses the problem of analyzing the nonlinear transient response of a large degree-of-freedom structural dynamics simulation. A threaded joint assembly's response to impulse loading has been studied. Twelve parameters relating to the input level, preloads of the joint and friction between components are thought to influence the acceleration response of the structure. Due to the high cost of physical testing and large amount of computation time to run analytical models (three hours on 504 processors of LANL's ASCI supercomputer, Blue Mountain to simulate three milliseconds of response) a fast-running metamodel is being developed. In this case, a metamodel is a statistically developed surrogate to the physics-based finite element model and can be evaluated in minutes on a single processor desktop computer. An unreasonable number of runs is required ($3^{12} > 500,000$) to generate a three level full factorial design with 12 parameters for metamodel creation. Some manner of down-selecting or variable screening is needed in order to determine which of the parameters most affect the response and should be retained in subsequent models. A significant effects methodology has been used, involving a design of experiments technique, in which all parameters were first included in the model and then eliminated on the basis of statistical contributions associated with each parameter. Bayesian variable screening techniques, in which probabilities of effects are generated and updated, have also been explored. Encouraging results have been obtained, as the two methods yield similar sets of statistically significant parameters. Both methods have been compared to general sensitivity analysis (GSA). The resulting compact metamodel can then be explored at more levels to appropriately capture the underlying physics of the threaded assembly with a much smaller set of simulations (only 64 runs required for a three level model with quadratic and some third order terms).

NOMENCLATURE

Y	response feature modeled
X'	parameter
B	parameter coefficient
σ	error coefficient
ϵ	error

INTRODUCTION

A threaded joint assembly (Figure 1) has been studied to determine its response to impulse loading from the side. It has steel upper and lower mass simulators and aluminum upper and lower shells. The upper mass simulator is bolted to a titanium mount. The upper shell and retaining nut (which holds the lower shell in place) are threaded onto the mount. It also has a unique steel "tape joint" which is located at the bottom of the assembly, wedged between the lower mass simulator and the titanium mount. It is the friction between the interfaces of the lower mass simulator,

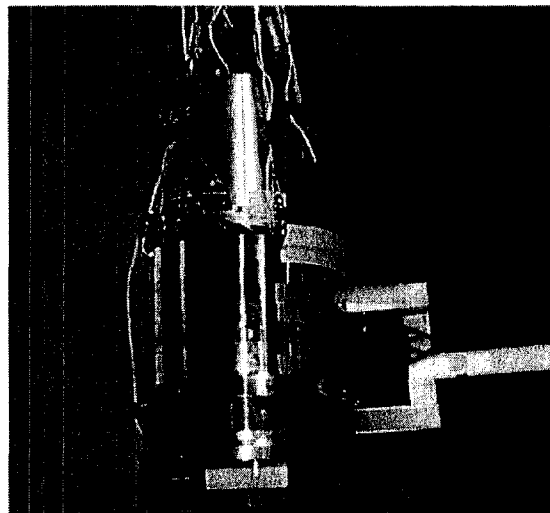


Figure 1: Physical testing set up of the threaded assembly. The charge array is held in place with foam blocks at the lower right of the photo.

* University of Michigan, dbingham@umich.edu

the titanium mount,, and the tape joint that hold these pieces in place.

To date four physical tests have been conducted, in which shell tolerances and "gaps" between the shell and mount were varied [1, 5]. These parameters were chosen based on initial engineering judgement. They had to be isolated before the finite element runs were executed due to scheduling requirements of the physical tests. A charge array was attached to one side of the assembly and detonated. Acceleration responses were monitored at the top and bottom of the assembly, as well as between the upper and lower mass simulators (see Figure 1). While much information was gained from the tests, it is economically prohibitive (>\$60,000/test) to conduct a sufficient number to generate a response surface.

In order to study the assembly in greater detail, a large finite element model was created (>1 million elements and 4 million degrees of freedom) [1] (Figure 2). The model can be run on the Laboratory's ASCI super-computer, Blue Mountain. However, because Blue Mountain requires about three hours using 504 processors to run three milliseconds of simulation time it is still impractical to execute a large number of runs for response surface generation.

From the finite element run results, twelve parameters were identified as being potentially important to the response of the assembly. They relate to dynamic and static coefficients of friction between the different parts, the preloads, and the input level (see Table 1). Parameters, or factors, were set at discrete levels for each experiment and coded to unitless values. Multiple features were derived from the acceleration time history data collected. Time histories were transformed to yield power spectral densities (PSDs) and shock response spectrums (SRSs). Then moments were taken of each of these functions, with the lowest order being the energy (E) plots and the second and third order moments called Tau and D [7].

Table 1: Potential parameters to be included in metamodel

Parameter	Phenomenon Described
A	Tape Preload
B	Nut Preload
C	Upper Shell Preload
D	Aluminum-Aluminum Static Friction Coefficient
E	Titanium-Titanium Static Friction Coefficient
F	Aluminum-Titanium Static Friction Coefficient
G	Steel-Titanium Static Friction Coefficient
H	Aluminum-Aluminum Kinetic Friction Coefficient
J	Titanium-Titanium Kinetic Friction Coefficient
K	Aluminum-Titanium Kinetic Friction Coefficient
L	Steel-Titanium Kinetic Friction Coefficient
M	Input Load Level

Design of experiments (DOE) methods are being explored to create a fast-running, or metamodel, that is based only on model parameters that affect the response [6]. A Taguchi

orthogonal array, alias free for quadratic and some cubic terms, was used [4]. Using DOE, an Analysis of Variance (ANOVA) is generated using data from finite element runs, in order to estimate a polynomial (which can be nth order, depending on the number of runs available) that relates parameters to responses.

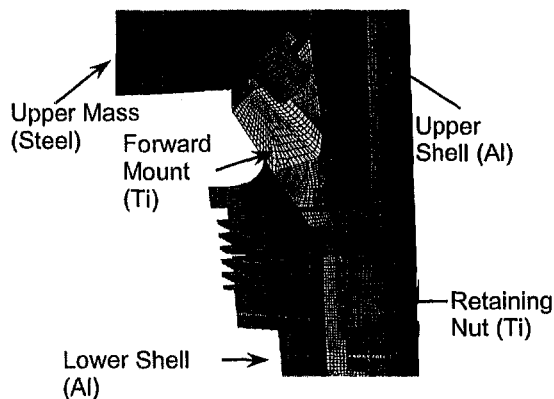


Figure 2: Detail of threaded joint FEM

A concern with developing a metamodel is that as more parameters are included in the model, more FEM runs are required to generate the model. A three level design incorporating all twelve parameters would require over 500,000 runs for a full factorial design, thus defeating the simplifying purpose of the metamodel. A fractional factorial design can be implemented to reduce the number of runs required for a given model, while increasing the model order. However, the use of fewer runs for the creation of a higher order model leads to aliasing of some terms which must then be omitted; therefore it is still desirable to keep as many runs as possible. A design with twelve parameters might be run with as few as 6000 runs, but would be heavily aliased. A variable screening process must be implemented in order to reduce the dimension of the response space so that a model may be designed without too many aliased terms.

GENERAL SENSITIVITY ANALYSIS

General sensitivity analysis (GSA) was conducted for a quick way of looking at important parameters. Parameters A, B and C were not included in the general sensitivity analysis. They were held constant at their nominal values because of difficulty in setting preloads (each preload requires ten extra runs to set because they cannot be set directly; thermal stress methods must be employed). A set of finite element runs were executed with each parameter set at its high and low value and one run with all parameters set at their nominal values (19 runs total). Parameter importance was defined as any parameter causing a relatively large difference in model response between its high and low value (as calculated using simple finite differencing). This method does not account for higher order effects and interactions. Results from this exercise showed that parameters L and M (St-Ti kinetic friction, input load level) were important to model features, with H, K and J (Al-Al kinetic friction, Ti-Ti kinetic friction, Al-Ti kinetic friction) appearing also, though at a lesser magnitude.

SIGNIFICANT EFFECTS METHOD OF VARIABLE SCREENING

Another method of variable screening implemented was analysis of significant effects, or linear variable screening. A particular main effect's contribution (no higher order effects) to the total model variance was analyzed [1]. Significant effects method provides an advantage over GSA because a probabilistic assessment of variable importance is obtained through the analysis of variance. Screening was done using a two level fractional factorial design. Features which did not result in a high enough total variance contribution (<30%) were discarded because they did not sufficiently discriminate between runs. Results from this screening are shown in Figure 3. In many of the models, effects A, K, L and M are important parameters. These variables correspond to the tape joint preload, kinematic frictions between parts (Al-Ti, Steel-Ti), and input scaling. Results agreed with the GSA. The results were not intuitive; the thread preloads (B and C) were expected to have more impact on the response, because they appear to be more directly tied to the upper and lower mass simulators, but neither of these effects were screened as important. This lead to a desire to corroborate results with another screening method.

BAYESIAN VARIABLE SCREENING

The second method used was a Bayesian variable screening technique [2,3]. An advantage it provides over both of the previous two methods is that it samples the entire response space instead of just the high and low edges, as well as providing a probabilistic assessment of parameter importance. A simple Markov Chain Monte Carlo (MCMC) method, called the Gibbs Sampler, was used to sample different models. The most probable models are visited most often, with the most probable effects occurring in the models more frequently than the rest of the effects [6].

Bayesian methods make use of prior and posterior probabilities and distributions. Priors are assigned by the analyst; the sampling method then updates these values to posterior values. In an MCMC method, posteriors from one iteration become the priors of the next. In this case, two sets of prior probabilities were established. The first relates to whether main effects and two factor interactions occur in the model. Because comparison to the method of significant effects was desired, only main effects were given a probability (0.25). The rest of the prior probabilities, which describe two factor interactions, were set to zero.

Priors relating to the probability of a particular model given a set of inputs must also be set. A simple linear regression model is used,

$$Y = X' \beta + \sigma \epsilon \quad (\text{Equation 1})$$

Where Y is the response feature vector, X is a vector of parameters and ϵ is the error. The parameter coefficient, β is assumed to have a multivariate normal distribution and the error coefficient, σ an inverse gamma distribution. Each of these distributions has associated parameters for which priors must be set [3].

A Gibbs Sampler was then used to obtain a posterior set of screened variables, derived from their associated probabilities. The Gibbs Sampler repeatedly draws samples from the posterior distributions of each parameter, resulting in a sample that is approximately one from the joint distribution of the above parameters. Chipman, Hamada, and Wu [3] note the chain length can be an implementation issue, as samples close together may be correlated. After trying a longer chain (10,000 samples) with a lag of 10 and discovering that results were nearly the same as using a chain 1000 samples long storing every sample, the shorter chain was chosen, with a 100 sample long "burn-in" (to initialize the chain).

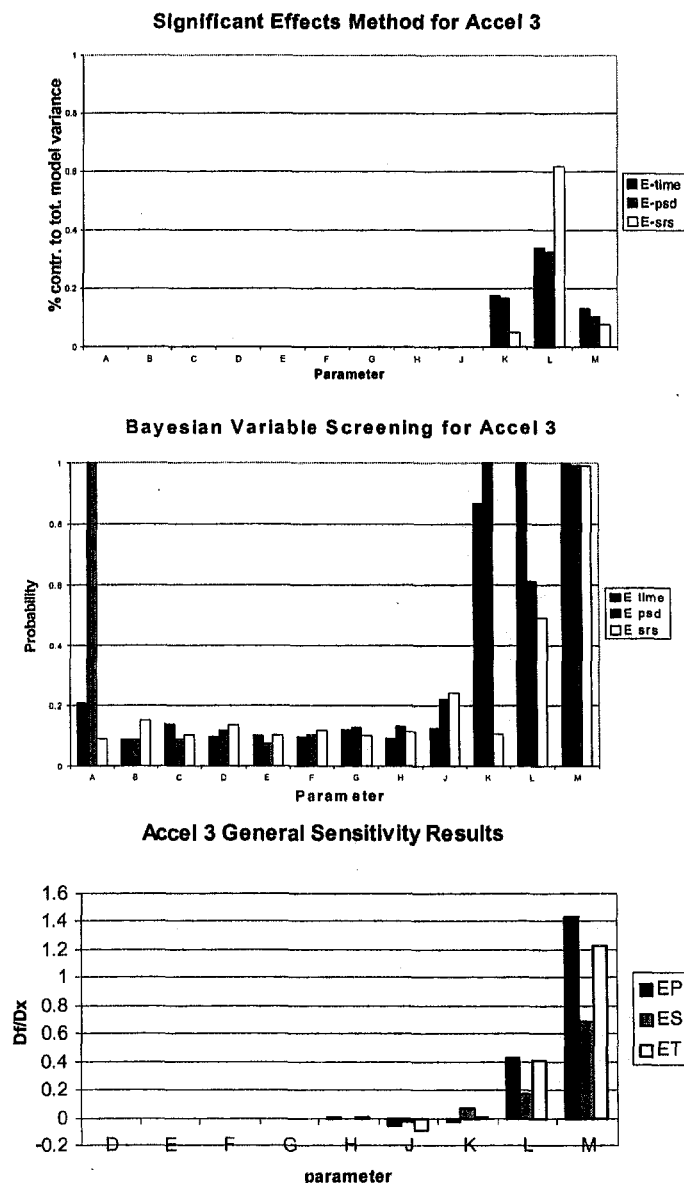


Figure 3: Comparison of Screening Methods for Accelerometer 3 Response Features

Results from the Bayesian variable screening method can be found in Figure 3 for main effects screening. In the Bayesian method, effects are important if their posterior probability is high compared to the rest of the parameters in question.

Again the problem of features which are not discriminating enough arises. Features that have parameters with very low probabilities (generally <0.4), ie, models that are approaching the "mean model," are omitted from comparison. Only results for the first moment features, E-time, E-PSD and E-SRS, are shown. While D and Tau were similar, E features yielded the most conclusive results. For most features, A, K, L, and M are important parameters to the model, matching the results found using the method of significant effects and general sensitivity analysis. Note that in the Bayesian screening method, effects G, H and J occasionally have high probabilities, as well.

Screening was also extended to two factor interactions. Prior probability for two factor interactions was set at .25 for interactions with both parents appearing in the model, .1 for interactions with one parent appearing in the model and .01 for interactions with no parents in the model. Distribution parameter priors used were the same as for main effects analysis. Screened two factor interactions for the most part included combinations of A, K, L, and M with occasional occurrences of B, H and J (nut preload, Al-Al kinetic friction coefficient, and Ti-Ti kinetic friction coefficient).

DISCUSSION

Both methods clearly indicate which effects are "important" and which may be omitted from the model, and these results are comparable. For the set of data that was screened, K, L, and M frequently appear as important model factors in both screening methods, with effects A, G, H and J appearing less frequently. The E features appear to be the most discriminating for nearly all the accelerometers, based on high contribution to total model variance in the significant effects method and its high probability levels in the Bayesian screening. Features used in the final model were A, G, H, K, L, and M based on the results obtained above. General sensitivity analysis provides a good first look at important parameters, however it only accounts for the high and low values of parameters and does not provide a probabilistic assessment of parameter importance.

CONCLUSIONS AND FUTURE WORK

Results achieved are encouraging, showing that the parameters screened are very likely to be at least part of some larger super-set of parameters which have an impact on the response of the threaded assembly. After screening parameters and using fractional factorial design, we now have feature models based on six important parameters, significantly reducing the number of runs needed to define the metamodel (64 runs for a model with third order terms). Effects chosen have been used in design of a higher order metamodel, which more accurately predicts the response of the model.

Future work will concern determining which feature is appropriate to model and the fidelity with which it must be modeled. Additionally, testing of some subsystems will be

conducted so that dynamic friction parameters, which were identified through the screening process as important to model response, may be estimated. To date, because little (no?) literature exists on kinematic coefficients of friction, they have been estimated using engineering judgement and this value checked by using the metamodel to solve the inverse problem. Parameters values are calculated through inverse means and minimization of error between actual and predicted parameter values. Once better estimates of these parameters are obtained, a new metamodel will be generated which more accurately represents the physics of the problem.

REFERENCES:

- [1] Butler, T. A., F. M. Hemez, J. F. Schultze, H. Sohn, S. W. Doebling, "Model Validation for a Complex Jointed Structure," *19th IMAC, International Modal Analysis Conference*, Kissimmee, FL, Feb. 5-8, 2001.
- [2] Bingham, Derek. Matlab software for Bayesian Variable Screening. University of Michigan. 2001.
- [3] Chipman, H., M. Hamada, C. F. J. Wu. "A Bayesian Variable-Selection Approach for Analyzing Designed Experiments with Complex Aliasing." *Technometrics*, Vol. 39, No. 4. American Statistical Association and the American Society for Quality. 1997.
- [4] Design Expert Version 6 User's Guide. Stat-Ease, Inc. 2000.
- [5] Groethe, Mark A. *Manufacturing Joint Validation Tests*. SRI Project 10129. SRI International. Menlo Park, CA. November 1999.
- [6] Wu, C. F. Jeff, Michael Hamada. *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley and Sons, Inc. New York. 2000.
- [7] Smallwood, D.O. Matlab software for calculation of Temporal Moments. Sandia National Labs. 1994.