SAND2007-7585C

Response Measures for Validation of Structural Dynamic Systems

Prasad Raghavendra
Nokia Inc.
12278 Scripps Summit Drive
San Diego, CA 92131

Thomas L. Paez
Sandia National Laboratories
Dept. 1544, MS 0557, PO Box 1500
Albuquerque, NM 87185-0557

Abstract

Advances in computation and modeling capabilities have led to the potential for creating extremely high resolution finite element models of structural dynamic systems. Concurrently, the hope to replace, in some sense, at least some of the activities previously accomplished through structural dynamic experimentation has led to increasing interest in the field of model validation. Model validation is the activity wherein mathematical models are tested for their accuracy (or adequacy) for the prediction of physical system response in certain limited behavior regimes. However, it has been shown that, except under certain very restrictive conditions, the raw response time histories predicted by mathematical models provide a poor means for comparing model predictions to experimentally measured responses. The reasons for this will be explored, with reference to experiments performed on an ensemble of cell phones and the corresponding mathematical model. Further, some response measures that overcome the difficulties encountered in the use of raw time histories for validation will be recommended.

Nomenclature

η Windowed measure of temporal RMS response
$W(f)$ Window function
$Y(f)$ DFT of experimental response of structure
$y(t)$ Experimental response of structure
y Windowed measure of DFT modulus of structural response

Introduction

Finite element models of structural dynamic systems that include up to millions of degrees of freedom are in common use today. They are built to simulate the behavior and responses of large, small, and intermediate-sized structures. High resolution models of relatively small structures may be so detailed as to simulate the behaviors of parts that have dimensions on the order of one millimeter.

Along with progress in model size and resolution, advances in modeling accuracy have led to the hope that models might be constructed to simulate the behavior of a physical system with enough precision to permit system design, or design changes. The assumption is that time and money can be saved through the use of mathematical models in place of physical experiments to establish the behaviors of structures. But in order for decision-makers to accept the predictions of structural responses made using mathematical models, the prediction accuracy (or, at least, adequacy) of the models must be shown. The process for demonstrating the accuracy or adequacy of a mathematical model for making predictions in a specific framework is known as model validation. The validation process is described in many publications, for example [1,2].

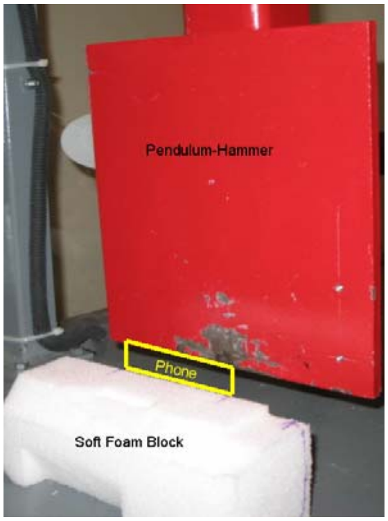
Model validations compare the predictions of mathematical models to the results of physical experiments, and one of the requirements of a validation is that a quantity of interest (QOI) be specified. A QOI is a measure of structural response that the model is intended to predict, or a measure of structural behavior that must be adequately represented in order for model predictions to be satisfactory. For example, when structural strength is at issue the strain responses at some critical locations might be QOIs. Or for example, when an equipment item is known to malfunction when subjected to accelerations that surpass a pre-established level then the QOI may be maximum acceleration at a point on a structure where the equipment item is mounted. As a third example, it may be required that a frequency response function (FRF) of a finite element model match, at some level of accuracy, the corresponding FRF of a physical system.

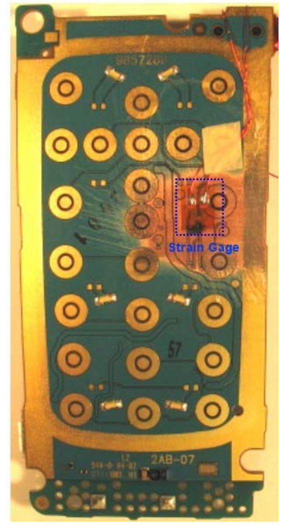
Structural dynamic analyses with finite element models normally predict response time histories. But it has been widely observed that time histories are not well-suited for use, in raw form, as QOIs in validation comparisons. This paper reports on an investigation into when and why response time histories are not well-suited for use as QOIs, and it suggests some measures that are well-suited to this purpose. The study is carried out in terms of the experimental shock responses of cell phones, in the hope that the similarities of this system and its excitation to other structural dynamic systems may shed light on the validation of many types of structural mathematical models.

Section 1 describes some experiments in which an ensemble of cell phones is subjected to a sequence of nominally identical mechanical shock environments. The testing procedure is described and measured response time histories are shown. It is demonstrated using the measured response time histories that apparent experiment-to-experiment variations exist, even when an experiment is repeated on a single member of the ensemble (a single cell phone), and that variabilities also exist among different members of the ensemble of nominally identical structures. Section 2 characterizes the variabilities in both the time and frequency domains and explores the reasons for the variabilities. Behavior and response measures that extract similarities from apparently dissimilar data are also developed in that section. Section 3 briefly considers prediction of experimental response using a finite element model, and shows how predictions from a deterministic model might be compared to experimental results (which are random in nature) in a validation activity. We conclude with comments on some practical aspects of mathematical model validation.

1 Experiments – Cell Phones Subjected to Mechanical Shock

Experiments were performed to develop a data base that is useful for assessing system characteristics and variations in measures of response. There are two parts to the experiments. Part I tests a single cell phone from a collection of nominally identical units. A typical test setup – picturing a schematic element instead of a cell phone - is shown in Figure 1a. The phone is placed on a soft foam block that allows it to be held in the required drop orientation in a repeatable manner. After collision (impact) with the pendulum-hammer the phone is free to rebound unfettered. At the start of the test the phone is struck by a one and one-half foot square and one-half inch thick steel plate so that the flat surface of the plate impacts the flat front face of the phone (Figure 1a). The impact is designed to occur with the plate moving at a velocity of 5.4 meters per second. The cell phone was subjected to a sequence of nine impacts that were arranged to create an environment as nearly repeatable as possible. Part II of the experiment tests five cell phones from the ensemble of nominally identical units that yielded the cell phone for the Part I tests. Each cell phone was subjected to the same impact as the Part I unit. Every cell phone in the Part I and Part II tests was instrumented with a strain gage located on the printed wiring board (PWB) behind a relatively large chip as shown in Figure 1b. The strain gage is oriented so as to measure the longitudinal strain along the greater dimension of the PWB. Of course, there is a level of randomness associated with location and orientation of strain gages, but the degree of this type of randomness is thought to be small, and so this source of randomness is neglected, in this study.





Figures 1a (left) and 1b (right). Test configuration and strain gage on PWB of cell phone.

The objective of the Part I experiments was to characterize repeatability in a sequence of nominally identical shock tests. Figure 2 shows the strains measured during the Part I tests. In this set of tests and the next, the signals were low-pass filtered at 10 kHz, and the peak strains were approximately aligned. The measured strain responses indicate substantial variation. Because in this part of the tests a single structural realization from the cell phone ensemble is excited, it must be concluded that the variations occur due to test conditions. Among other things, we suppose that variations are related to (1) boundary conditions of the test item at the start of the experiment, (2) the test geometry (e.g., the angle and velocity of impact of the exciting hammer, and the location and orientation of a cell phone within its fixture), and other sources. Another contributor to the observed variations is changes in the tested cell phone between experiments, but that contribution is thought to be small. The sample mean of the shock response time histories and mean minus/plus two sample standard deviations are also shown in Figure 2. The lower and upper sample standard deviation bounds and the placement of the time histories within those bounds show that much of the variation must be explained in terms of the signal frequencies and phases.

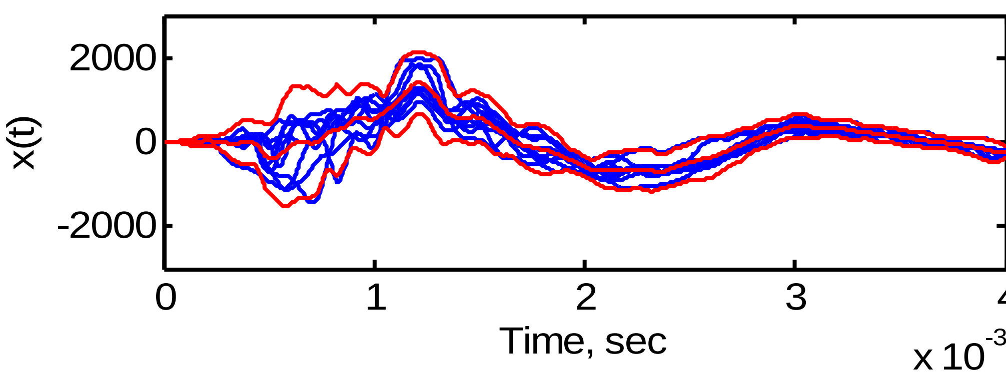


Figure 2. Measured strain responses from the Part I experiments – a single cell phone subjected to nine nominally identical impacts (blue). Estimated mean and mean minus/plus two standard deviations (red).

The objective of the Part II experiments was to characterize experiment repeatability associated with the testing of separate, nominally identical units from a single ensemble. Figure 3 shows the strain measured during each of the Part II tests. The measured strain responses indicate substantial variation. The variation realized here comes from two main sources. First, the variation associated with experiment-to-experiment changes – the variations witnessed in the Part I experiments - are also present here. Second, part of the variation witnessed in this set of experiments is due to differences in the units tested. Though nominally identical, the units must vary because they cannot be perfect replicates of a prototype. Specifically, nominally identical units from a single ensemble actually differ, to some extent, because: (1) geometries vary within (or perhaps, outside) a specified tolerance, (2) mechanical joints vary, particularly bolted connections, (3) material properties are not uniform and vary, even slightly, from one structure to the next, etc. The sample mean of the shock response time histories and mean minus/plus two sample standard deviations are also shown in Figure 3. The two standard deviation limits are near the two standard deviation limits of the Part I test responses (shown in Figure 2) and this indicates that in the present tests the variability associated with unit-to-unit differences is relatively small compared to the setup-to-setup differences. (This may not be true for strain and other measurements of responses made at other locations in the system, and typically, it is not true.)

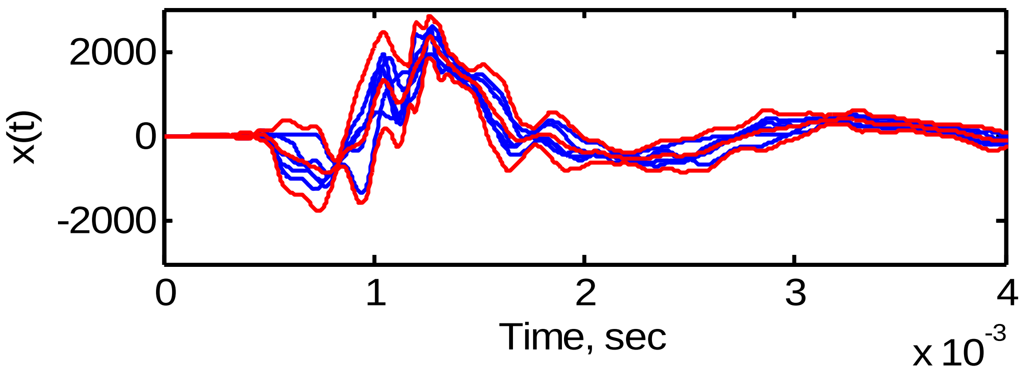


Figure 3. Measured strain responses from the Part II experiments – five cell phones subjected to nominally identical impacts (blue). Estimated mean and mean minus/plus two standard deviations (red).

The degree of variation in the measured responses caused by the random variations listed above is typically not known, a priori. When the degree of variation from all sources is small, then the experimental responses will be close to one another; otherwise, they will not.

A question to be addressed in a later section is: When we have a deterministic finite element model of the cell phone which is expected to simulate the behavior of the physical system, how do we compare the responses predicted by the model to the responses measured on the physical system during experiments? Some reasonable answers to the question will be proposed in the following sections. In this section we have pointed out that the difference between the raw time history of the model-predicted response and the measured response from a physical structure may be a poor means for comparison. This is due to the large test-to-test variation in the measured responses of physical structures.

2 Variabilities in Shock Response Time Histories and some of their Measures

In order to demonstrate the difficulties associated with the use of time history comparisons for model validation, we consider a metric of the Part I test responses. (The results for the Part II tests are similar.) Figure 4 shows the differences between all pairs of measured strain time histories for the Part I tests. There are nine measured strain time histories, therefore, there are 36 signal differences. The maximum and minimum values of the differences are plotted in red, and have magnitudes near – or greater than - the peak magnitudes of the strain time histories, themselves. (Under some circumstances the maximum and minimum values of the signal differences are double the magnitudes of the strain time histories.) The reason why the differences are so great relates to the frequency content and phasing of the strain time histories, and will be clarified, below. The differences between pairs of experimental strain realizations is so great that it appears that it would be difficult to use the difference between a finite element model-predicted response and an experimental response to judge the validity of the model predictions. Moreover, when multiple experimental responses are available for use as the basis of a comparison, it may be difficult to establish how to use them all in one systematic comparison.

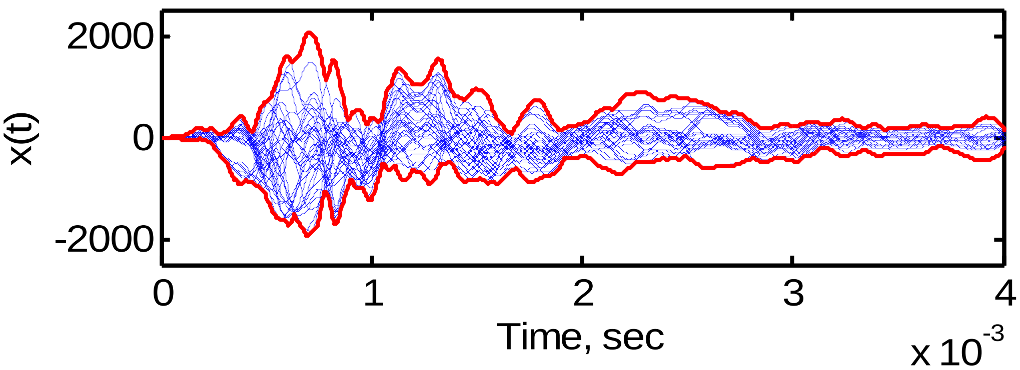


Figure 4. Differences between pairs of measured strain time histories from the Part I tests.

It must be noted that an intuitively appealing approach to assessing the adequacy of a model prediction is to plot the model prediction along with the measured strain response time histories of Figure 2 or 3, and simply observe whether or not the model prediction lies among the measurements. The approach is not unreasonable; however, it is difficult to rigorously judge the adequacy of a model on such a basis because a model may be quite adequate even if it passes outside bounds like the ones shown in Figure 2 or 3.

This discussion raises another important issue. A model may be adequate if it predicts the peak dynamic response to within some pre-established accuracy. This is possible when the response is excited by a mechanical shock, because even fairly accurate - and fully adequate – models tend to predict peak responses early in the response time history, and the peak value of response is not the result of combinations of randomly phased input components. This sort of test might be applied to a model prediction, but it is not normally sufficient to require that a dynamic model simply predict peak responses accurately. It is normally hoped that a dynamic model will make accurate predictions because it contains the appropriate frequency content.

The issue of frequency content in measured and predicted responses can be addressed by computing discrete Fourier transforms (DFTs) of strain response time histories. The moduli (also known as complex magnitudes, or absolute values) of the DFTs of the Part I tests are shown in Figure 5. (The results from the Part II tests are similar, and are not shown for the sake of brevity.) The results consistently indicate that the output contains substantial signal content in the frequency range up to at least 7 kHz. The signal content between 200 and 800 Hz is probably associated with the impact. We reach this conclusion for the following reason. If the structure were linear (and, it is not strongly nonlinear), its response DFT would equal the DFT of its excitation times the structure's frequency response function (FRF). The structure model indicates no modes in the frequency band below about 1000 Hz, therefore, the FRF is relatively flat below 800 Hz. So the feature in the DFT modulus below 800 Hz must be caused by the excitation.

Several system modes appear consistently in the one through ten kHz range, but change values, at least slightly, from one test to the next. This is an indication that the excitation and system character vary from one test to the next. Therefore, one answer to the question posed at the end of the previous section may be that if the response predicted by a mathematical model has frequency domain amplitude characteristics that resemble the physical structure's frequency domain amplitude characteristics, then the model may be an adequate representation of the structural system ensemble. But, based on Figure 5, the DFT amplitude spectrum of the model prediction need not replicate the amplitude spectra of the physical system – the amplitude spectra of the nominally identical structures excited by nominally identical impacts do not even replicate one another. Rather, another criterion is needed.

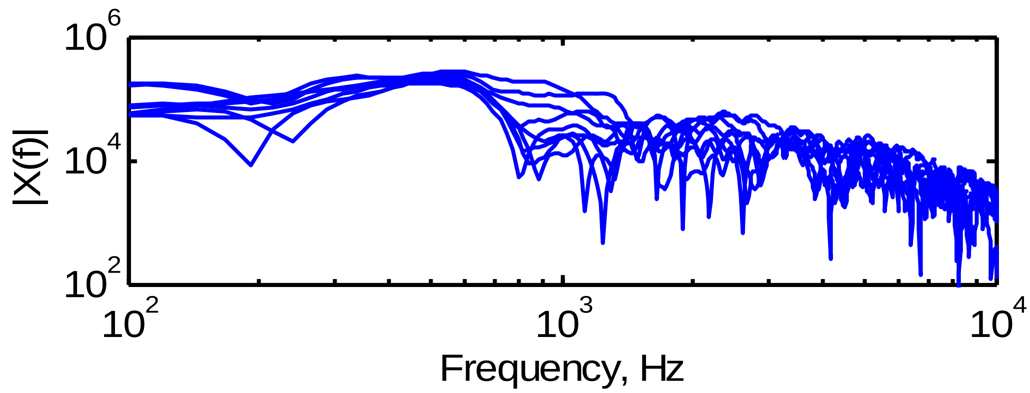


Figure 5. DFT moduli of strain time histories from Part I tests.

Before developing such a criterion, we note that the differences in time domain responses arise, to a great extent, from phase differences among structural responses. Figure 6 shows the complex phases of the strain time histories from the Part I tests. There is some consistency among the complex phases of the shock response DFTs below and slightly beyond 1 kHz, but at higher frequencies the phases do not coincide with one another. When the DFT phases of harmonic components differ, then components of the time domain signals, typically, cannot match.

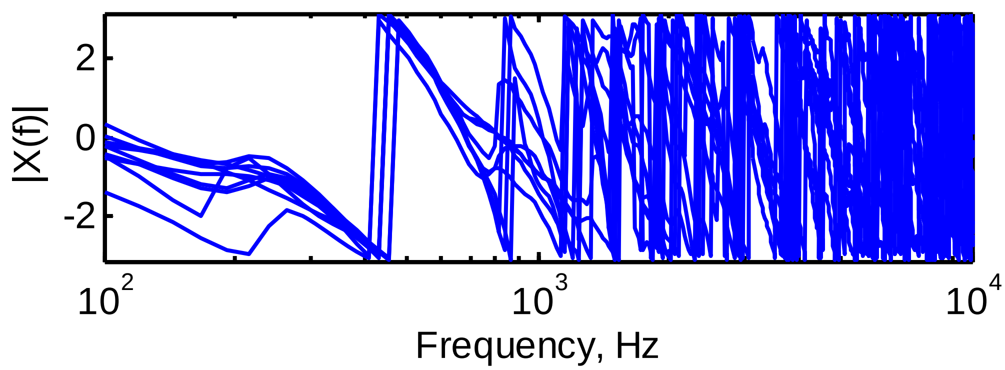


Figure 6. DFT phases of strain time histories from Part I tests.

A measure of structural response that captures the essential similarities of the signals in Figure 5 is one that weights and averages each DFT modulus over a range of frequencies. An approach to accomplishing this sort of weighting and averaging is to multiply each DFT modulus by a weighting function that is centered at a particular frequency and has finite value over a band of frequencies, then integrate to obtain a scalar measure of the DFT modulus. The weighted measure described here is defined

$$y_{ik} = \int_{-\infty}^{\infty} W(f - f_k)|Y_i(f)|df \qquad i = 1,...,n, k = 1,...,N \tag{1}$$

where $Y_i(f), i = 1,...,n,$ is the Fourier transform of the i^{th} measured experimental response, and $|Y_i(f)|$ is its complex modulus, and $W(f), -\infty < f < \infty$, is a non-negative, symmetric (about the origin), absolutely integrable function. The function $W(f - f_k)$ is centered at the frequency f_k.

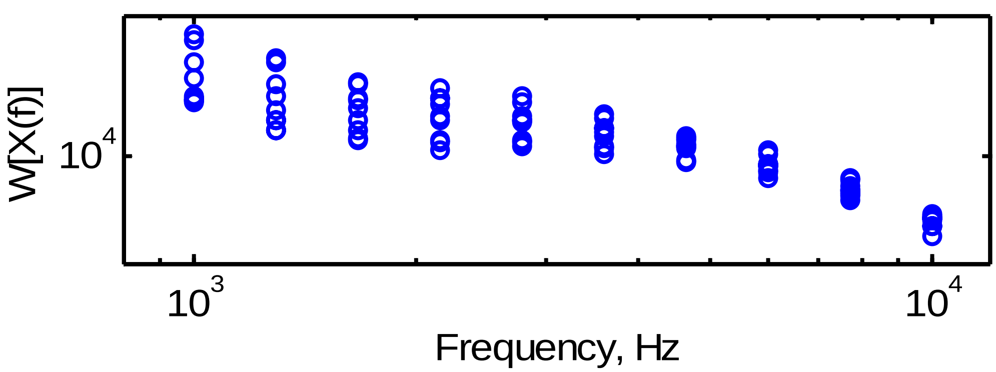
In order to use this windowed measure to compare model to experimental data, choose a set of center frequencies and window widths and compute the windowed measure of DFT modulus for all the experimental DFTs and for the model-predicted DFT. Such an analysis provides the weighted and averaged DFT moduli at multiple frequencies for both the model and the experiments. Normally the scalar quantities that are the windowed measures of DFT moduli are plotted above the window center frequencies. These windowed measures can be used as the basis for comparison of model-predicted results to experimental results. An example is given in the following section.

For any given center frequency f_k, the measured experimental response strains will have windowed response measures y_{ik} that span a range of magnitudes. If the analog to y_{ik} computed using the model-predicted response falls among the experimental y_{ik}, then the model might be inferred to be valid – even accurate - with respect to the averaged measure of response. If the model-predicted response measures are conservative, in some sense, with respect to the y_{ik}, then the model might be said to be adequate, but not necessarily accurate.

Figure 7a shows the windowed measures of the modulus of the DFT of the experimental strain responses to shock excitations corresponding to the DFT moduli shown in Figure 5. The window used here is the truncated Gaussian window given by

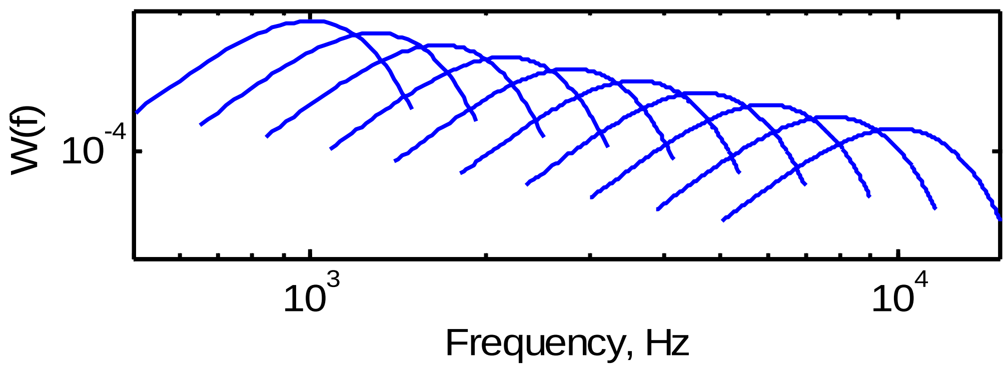
$$W(f)=\frac{1}{\sqrt{2\pi}\sigma_W}exp\left[-\frac{f^2}{2\sigma_W^2}\right] \qquad -2\sigma_W \le f \le 2\sigma_W \tag{2}$$

The parameter σ_W usually defines the standard deviation of a Gaussian distribution, but here, it scales the width of the truncated window. The total window width is $4\sigma_W$. The frequency centers vary and are equal to $f_k = 1000(1.292)^{(k-1)} Hz, k = 1,...,10$. The window widths all equal their corresponding center frequencies. The window centers and widths are arbitrary, but of course, their choice affects the validation comparison. The center frequencies normally span the frequency band where the model predictions are to be compared to the experimental results. The window widths are normally chosen to weight and average frequency domain information over a range of frequencies across which random variation in the physical systems is anticipated. The windowed measures of the DFT moduli, shown in Figure 7a, vary, depending on the center frequencies about which they were computed, but they appear to provide a rational basis for assessing the validity of mathematical model predictions.



Figures 7a. Windowed measures of the DFT moduli of experimental strain shock responses shown in Figure 5.

The actual windows used to obtain the measures graphed in Figure 7a are shown in Figure 7b. The first window is centered at 1000 Hz and has width of 1000 Hz, therefore it is defined on the frequency interval [500,1500] Hz. The second window is centered at 1292 Hz and has a width of 1292 Hz, therefore, it is defined on the frequency interval [646,1938] Hz, etc.



Figures 7b. Windows used t obtain the measures in Figure 7a. (Keep in mind that the windows are Gaussian, but they are plotted on logarithmic axes.)

The frequency averaging idea used above can be extended to a temporal averaging idea for the time domain response signals. In fact, various window-averaged time domain responses can be defined. One fundamental measure is the windowed root-mean-square (RMS). It may be defined

$$\eta_{ij} = \left[\int_{-\infty}^{\infty} w(t - t_k) y_i^2(t) dt \right]^{1/2} \qquad i = 1, \ldots, n, \; j = 1, \ldots, M \tag{3}$$

where $y_i(t), i = 1, \ldots, n, t \geq 0$, is the i^{th} measured experimental response, and $w(t), -\infty < t < \infty$, is a window that satisfies the requirements imposed on the frequency domain window. The quantity is always non-negative because of the square inside the integral. The absolute value of a single model-predicted response might be required to be comparable to the collection of windowed metrics defined by Eq. (3) in order for the model to be an accurate representation of the ensemble of physical structures. For the model predictions to be adequate, the requirement that the model-predicted responses be comparable to the windowed metrics in Eq. (3) might be relaxed to require, for example, only that the metrics of the model-predicted response be comparable to or greater than the windowed metrics in Eq. (3).

Figure 8 shows the windowed measures of the time domain response RMS of the experimental strain responses to shock excitations corresponding to the response time histories shown in Figure 2 The window used here is the truncated Gaussian window of Eq. (2). The temporal centers are $t_j = 2 \times 10^{-4} + 4 \times 10^{-4} \times (j-1) sec, j = 1, \ldots, 10$. The temporal widths all equal $\Delta t = 4 \times 10^{-4}\, sec$. The windowed measures of the strain response time histories vary, depending on the time centers about which they were computed, but these also appear to provide a rational basis for assessing the validity of mathematical model predictions.

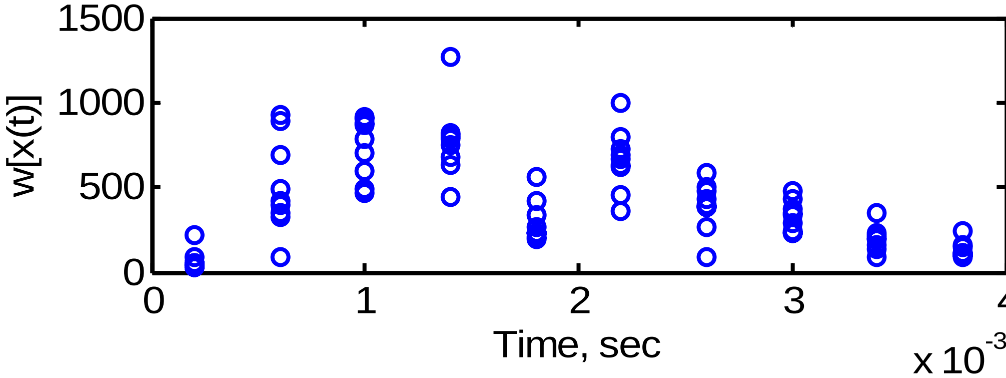


Figure 8. Windowed measures of the RMSs of the experimental strain shock responses shown in Figure 2.

3 Response Prediction using Finite Element Model

Though the focus of this investigation is not the validation of a particular mathematical model, but rather what measures of response might rationally be used to compare model predictions to experimental responses, we briefly present here some finite element model predictions of the response of the cell phones to shock excitation. The reason is that there are some lessons to be learned from the comparisons. A deterministic finite element model of the cell phones tested in Parts I and II was constructed in the framework of the ABAQUS, commercial finite element code. The model uses a mixture of solid and shell first order elements and includes nonlinear material models to represent foams, gaskets, and keymats. The plastic housings and the PWB are modeled using linear elastic material models. The model includes approximately two million degrees of freedom, and, is solved in the time domain using an explicit integration scheme. The structure is excited via an initial condition equal to the impact velocity.

The cell phone response was computed and the strain predicted at the location of interest is shown in Figure 9 superimposed on the measured responses of Figure 2. The model response appears to correctly predict, in a qualitative sense, the fundamental trends of the measured responses. It is clear that the peak value of the predicted response is on the low end of the range of experimentally measured responses and there are numerous potential reasons for this, but as stated earlier, it would be difficult to quantitatively establish the validity of the model based on the response time histories.

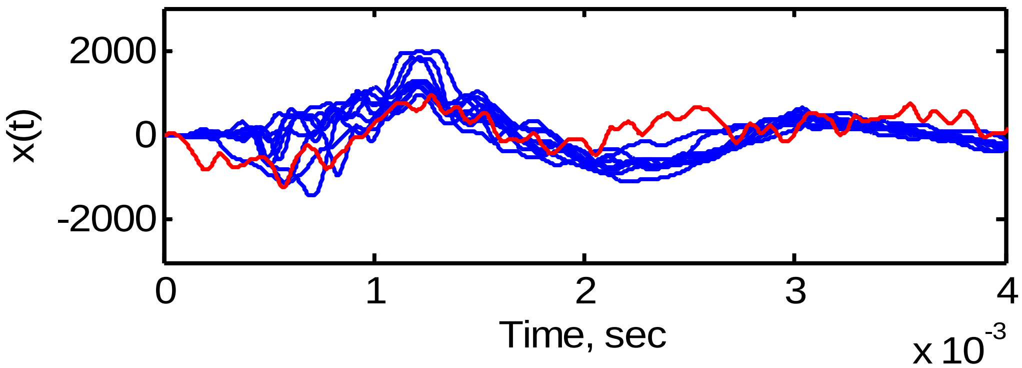


Figure 9. Finite element model-predicted strain response of the cell phone (red) superimposed on the measured strain responses (blue).

The DFT of the model-predicted strain response was computed and is shown superimposed on the moduli of the experimental response DFTs (from Figure 5) in Figure 10. It is clear that the match is good, although the model prediction is slightly low in the [300,600] Hz range, and slightly high in two high frequency ranges. As with the predicted and experimental response time histories, it would be difficult to quantitatively establish the validity of the model based on the response time histories.

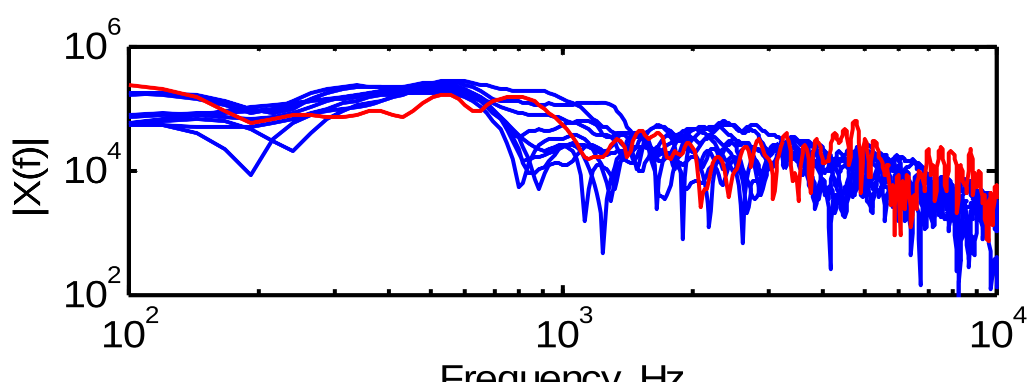


Figure 10. Modulus of the DFT of the model-predicted strain response (red) superimposed on the moduli of the DFTs of the measured strain responses (blue).

In view of these difficulties, we resort to use of the windowed measure of RMS response discussed in the previous section. Figure 11 is a modified version of Figure 8 - windowed measures of the RMSs of the experimental strain shock responses, but with the RMS of the model-predicted strain response added. The figure makes it clear that the model predicted response starts high, reaches its greatest (time averaged) value before the experimental systems, and may decay too slowly. Whether or not the model would be validated depends on the criteria established for validation before the comparisons were commenced, but the comparison shows how many structural model behaviors can be clearly diagnosed using the averaged measure of response.

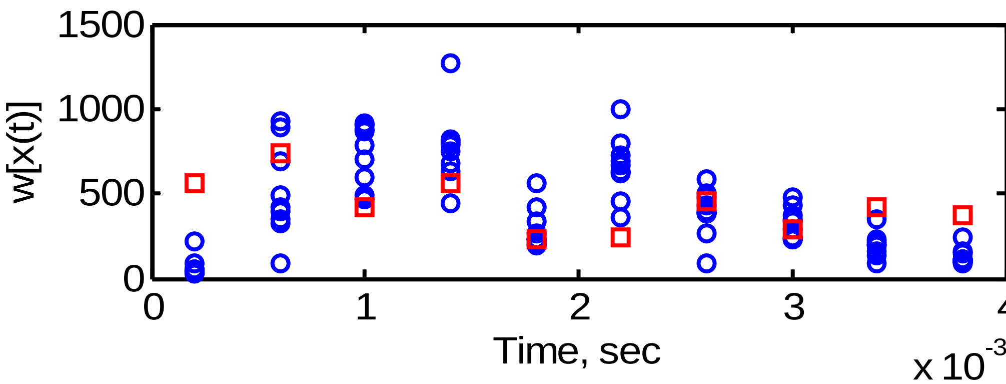


Figure 11. Windowed measures of the RMSs of the experimental strain shock responses (blue circles) and the RMS of the model-predicted shock response (red boxes).

The windowed measure of the DFT modulus of strain shock response is the other quantity suggested for use in comparing model predictions to experimental results. Figure 12 is a modified version of Figure 7 - windowed measures of the DFT moduli of experimental strain shock responses, but with the windowed measure of the DFT modulus of the model-predicted shock response added. This figure makes it clear that the model-predicted response in the frequency domain is very accurate in the frequency range $[1000,7000]Hz$, and slightly too high in the frequency range beyond 7000 Hz. This raises a question regarding why the windowed measure of model-predicted temporal RMS response tends to be too low in Figure 11. The answer is that the DFT modulus of the model-predicted response in the frequency range up to 1000 Hz – the frequency range that characterizes the excitation - tends to be too low. (This is confirmed in Figure 10.) As before, whether or not the model would be validated depends on the criteria established for validation before the comparisons were commenced, but this comparison also shows that many structural model behaviors can be clearly diagnosed using the averaged measure of response in the frequency domain.

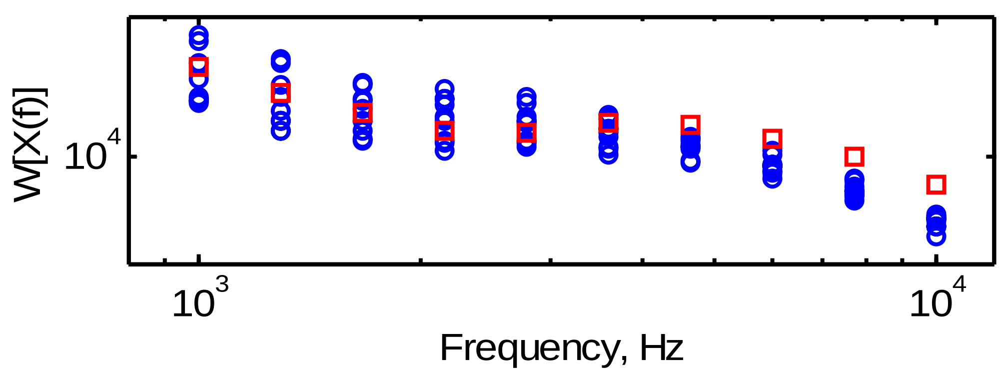


Figure 12. Windowed measures of the DFT moduli of experimental strain shock responses (blue circles) and the DFT modulus of the model-predicted strain shock response (red boxes).

Many references (e.g., [3,4]) develop formal procedures for model validation and, typically, where the data are similar to those available here, probabilistic validation procedures are adopted.

Discussion and Conclusions

This paper shows that, at least under some circumstances, response time histories cannot be easily used to perform validation comparisons, i.e., comparisons of model-predicted response to experimentally measured response. Moreover, it is difficult to use raw Fourier transforms - even the Fourier transform moduli - to perform the comparisons. But there are some measures of structural dynamic response that are well-suited to validation comparisons, measures that characterize the essential behavior of structural dynamic systems. These are measures that average and, perhaps, weight response characteristics in the time domain, frequency domain, or in another space. In this study, a windowed measure of the modulus of the DFT of a structural dynamic response and a windowed measure of the RMS response were considered. It was shown that these measures of response capture sufficient information about behavior of systems to provide a useful basis for judging system validity.

This paper also shows that typical measured structural responses are random because physical systems, themselves are random, as well as their excitations, boundary conditions, etc. Therefore, probabilistic approaches to structural dynamic system validation must be implemented in practical applications.

Many other measures of structural dynamic system response that have the characteristics of the ones defined here can also be defined.

Acknowledgment

Nokia Corporation engages in the manufacture of mobile devices and mobile networks. It also provides equipment, solutions, and services for network operators, service providers, and corporations. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

References

1. ASME, "Guide for Verification and Validation in Computational Solid Mechanics," V&V 10-2006, American Society of Mechanical Engineers, (2006).
2. U.S. Department of Energy, (2000), "Advanced Simulation and Computing (ASCI0 Program Plan," 01-ASCI-Prog-01, Sandia National Laboratories, Albuquerque, New Mexico.
3. Paez, T. L., Massad, J., Hinnerichs, T., O'Gorman, C. C., Hunter, P., (2008), „Validation of Mathematical Models Using Weighted Response Measures," *Proceedings of IMAC XXVI*, Society for Experimental Mechanics, Orlando, FL.
4. Urbina, A., Paez, T., Gregory, D. L., Resor, B., Hinnerichs, T., O'Gorman, C., (2006), "Validation of a Combined Nonlinear Joint and Viscoelastic Encapsulating Foam Model," *Proceedings of the Society for Experimental Mechanics Annual Meeting*, SEM, St. Louis, June.