LA-UR- 10-07534

Approved for public release; distribution is unlimited.

Title: Feature Extraction for Structural Dynamics Model Validation

Author(s):
Mayuko Nishio, University of Tokyo
Francois Hemez, LANL XTD-3
Keith Worden, University of Sheffield
Gyuhae Park, LANL INST-OFF
Nobuo Takeda, University of Tokyo
Charles Farrar, LANL INST-OFF

Intended for: IMAC XXIX A Conference and Exposition on Structural Dynamics
Jacksonville, FL
Jan.31 - Feb 3, 2011

Los Alamos NATIONAL LABORATORY
EST.1943

Feature Extraction for Structural Dynamics Model Validation

Mayuko Nishio[1], Francois Hemez[2], Keith Worden[3], Gyuhae Park[4], Nobuo Takeda[5], Charles Farrar[6]

[1] University of Tokyo, Dept. Civil Engineering, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan, nishio@bridge.t.u-tokyo.ac.jp
[2] Los Alamos National Laboratory, XCP-1, Los Alamos, NM 87545, hemez@lanl.gov
[3] University of Sheffield, Dept. Mechanical Engineering, Mappin St., Sheffield S1 3JD, UK, k.worden@sheffield.ac.uk
[4] Los Alamos National Laboratory, INST-OFF, Los Alamos, NM 87545, gpark@lanl.gov
[5] University of Tokyo, Dept. Advanced Energy, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan, takeda@smart.k.u-tokyo.ac.jp
[6] Los Alamos National Laboratory, INST-OFF, Los Alamos, NM 87545, farrar@lanl.gov

ABSTRACT This study focuses on defining and comparing response features that can be used for structural dynamics model validation studies. Features extracted from dynamic responses obtained analytically or experimentally, such as basic signal statistics, frequency spectra, and estimated time-series models, can be used to compare characteristics of structural system dynamics. By comparing those response features extracted from experimental data and numerical outputs, validation and uncertainty quantification of numerical model containing uncertain parameters can be realized. In this study, the applicability of some response features to model validation is first discussed using measured data from a simple test-bed structure and the associated numerical simulations of these experiments. Issues that must be considered were sensitivity, dimensionality, type of response, and presence or absence of measurement noise in the response. Furthermore, we illustrate a comparison method of multivariate feature vectors for statistical model validation. Results show that the outlier detection technique using the Mahalanobis distance metric can be used as an effective and quantifiable technique for selecting appropriate model parameters. However, in this process, one must not only consider the sensitivity of the features being used, but also correlation of the parameters being compared.

1. Introduction

The purpose of structural model validation is to assess whether a numerical model, such as a finite element model, has adequate predictive capability for the model's intended purpose by comparing analytical predicted and experimentally observed structural responses quantities. In constructing numerical models of structures, many quantities are assigned based on incomplete and/or unavailable knowledge of their true value. Uncertainties can result from measurement error, environmental variability, allowable manufacturing tolerances and variability associated with assembly procedures; while others are due to lack-of-knowledge about the actual structural condition; i.e., materials, loads, friction, energy dissipation (damping), and boundary condition. Therefore, it is important to assess whether the assumptions used in the modeling process provide accurate simulations on the intended purpose of the model.

In this study, the model validation is more generally defined that it should be validating statistically accurate models. Some previous studies were carried out on the basis of this definition according to the uncertainty quantification of the numerical model, e.g., [1]. In those works, it was indicated that it was so important to use appropriate response features to compare the numerical output and the measurement data. A response feature is the quantity extracted from the dynamic response that is used to compare the structural system's experimentally observed response characteristic to those predicted by its numerical model. Fundamentally, the feature extraction process is based on processing the data waveforms or spectra of waveforms, or fitting some model to the data. Many feature extraction techniques have been developed for structural health monitoring. Those include finding indications of nonlinear response and identifying system changes due to damages on structures, e.g., [2] and [3]. Many of these features can potentially be used for the model validation applications.

This paper focuses on defining and illustrating some response features that can be used for structural dynamics model validation studies. After a brief general discussion regarding response features, the application of these features for dynamic model validation are studied using experimental and numerical response data from a test-bed structure. Then, a comparison procedure for multivariate feature vectors based on Mahalanobis distance analysis is presented for the statistical model validation. This paper concludes with an additional discussion regarding the importance of appropriate feature selections.

2. Test-bed structure and numerical model description

The LANL three-story share building structure shown in Fig.1 (a) was used as a test-bed structure in this study [2]. The structure consists of aluminum plates and columns assembled using bolted joints. The structure slides on rails that allow movement only in the x-direction. The input force was applied in the x-direction by an electromagnetic shaker connected to the base floor. A force transducer was attached at the end of a stinger to measure the input force, and four accelerometers were attached at the centerline of each floor on the opposite side from the excitation to measure the system response at each floor.

This structure was modeled as a 4-DOF lumped-mass system as shown in Fig.1 (b). Mass, stiffness and damping of i-th story was defined as m_i, k_i and c_i (i= 1~4), respectively. The values of m_i and k_i (except k_1), were determined from the measured sizes of structural members and nominal values for Young's Modulus, and the mass density of aluminum. The stiffness k_i (i= 2~4) was the summation of the bending stiffness of four columns treated as beams that are constrained against rotation at their ends. For stiffness k_1, a relatively low numerical value was assigned because the friction between the rails and the structure was negligible. Notice that the model includes the base mass that slides on the rails. The equation of motion of the structure was described in a matrix notation using mass, stiffness, and damping matrixes; $[M]$, $[K]$, and $[C]$, as follows

$$[M]\{\ddot{x}\}+[C]\{\dot{x}\}+[K]\{x\}=\{F(t)\}\,, \tag{1}$$

where vector $\{F(t)\}$ is the input force vector, and $\{x\}$ is the displacement vector of x_1~x_4 in Fig.1 (b). In the application of time-response analysis, the displacement vector $\{x(t)\}$ was calculated by a Runge-kutta numerical integration scheme. A proportional damping was adopted to assign a damping matrix $[C]$ that could be discribed as

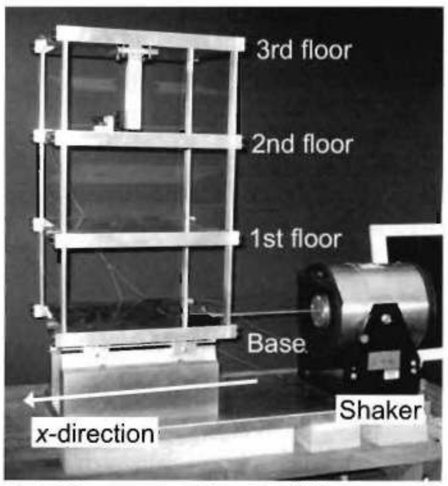
$$[C]=\alpha[M]+\beta[K]. \tag{2}$$

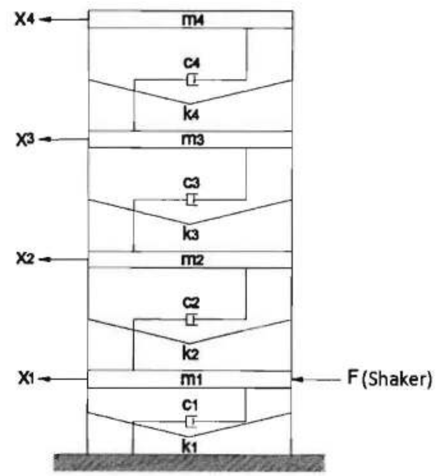
Theoretically, coefficients α and β can be related to k-th mode resonant frequency ω_k and modal damping ratio ζ_k as in

$$\zeta_k=\frac{1}{2}\left(\frac{\alpha}{\omega_k}+\beta\omega_k\right). \tag{3}$$

Therefore, α and β can be assigned when two modal parameter sets (ω_k, ζ_k) are given.

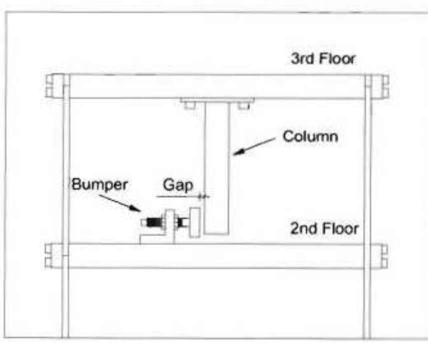
Additionally, this structure included a center column suspended from the 3rd floor, and a bumper mechanism attached on the 2nd floor as shown in Fig.2 (a). This system produces nonlinearity in the system response when the column impacts the bumper mechanism. The clearance (gap) between the bumper and the column can be adjusted to vary the extent of impacting that occurs during a particular excitation. This nonlinearity is intended to produce a small perturbation to an essentially stationary process, causing a nonlinear phenomenon. In the modeling, this nonlinearity can be described by adopting a kind of bilinear model in stiffness k_4 as shown in Fig.2 (b), where k_c is the summation of the bending stiffness of four assembled columns, and the k_b is the bending stiffness of the suspended column, corresponding to a cantilever beam. In the modeling here, distance Δ was defined to make a smooth transition between two stiffness states by a quadric function to satisfy C_1-continuity. This transition area indicates that the actual structural condition should not show a complete discontinuous point at X=*Gap* because, in reality, some kinds of compliance should exist in the actual assembling system.

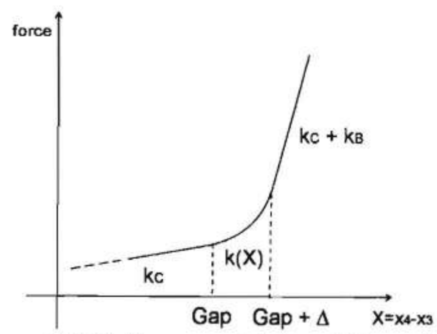




(a) Three-story share building set-up (b) 4DOF lumped-mass model

Fig.1. Test-bed structure setup and its numerical model





(a) Bumper and suspended column mechanism (b) Bi-linear model for stiffness k_4

Fig.2. Nonlinear behaviour mechanism between 2nd and 3rd floor

Some uncertain parameters were then recognized in the modeling of this lumped-mass model. Notice that the application of this model was defined to be the accurate prediction of the system's time-history response. From this viewpoint, not only m_i and k_i, but also damping parameters related to c_i will influence this prediction. The damping parameters were thus considered to be uncertain parameters in the calculation. Furthermore, the parameters, *Gap* and Δ, were also considered to have uncertain values in the nonlinear system calculation.

3. Response feature for structural model validation

3.1. General about feature selection

Response features are quantities that can be used to compare the measured and calculated system response. When used for model validation, the extracted response features should be sensitive to the target uncertain parameters of the numerical model. It should be noted that the intended purpose of the numerical simulation should also be considered when performing feature selection because the parameters that most influence the response may be different depending on in the intended purpose of the analysis. Dimensionality is another important consideration in the feature selection process. The feature dimension is the number of independent scalar quantities that are necessary to describe the feature. Low-dimensional features are preferable to high-dimensional features because this makes it easier to compare values and to statistically analyze their trends. Furthermore, the feature selection should reflect the type of response that is being considered, such as linearity or nonlinearity. Some response features that have been suggested for dynamic response calculations are [4]:

- Linear, stationary, Gaussian vibrations: Direct and inverse Fourier transforms, Power spectral density, Input-output transfer functions, Frequency responses, Modal parameter.
- Transient dynamics and mechanical shock response: Peak values, Energy content, Decrement and exponential damping, Shock response spectrum, temporal moments.
- General-purpose time-series analysis: AR, ARMA, ARX, AR-ARX models, Time-frequency transforms, Wavelet transform, Principal component decomposition.
- Unstable, chaotic, multiple-scale dynamics: Holder exponent, State-space maps, Time-frequency and higher-order transforms, Symmetric dot pattern, Fractal analysis.

It should be emphasized that there is no one feature that will be applicable to all structural dynamics predictive modeling scenarios. If multiple aspects of the system response are of interest, the model validation process may require different features to be extracted from the data in an effort to validate different aspects of the modeling process.

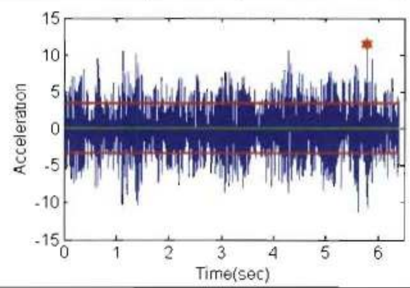
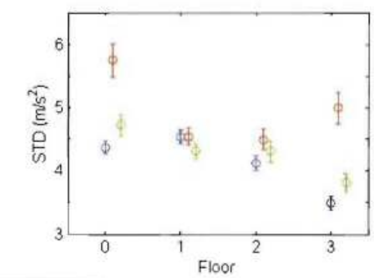
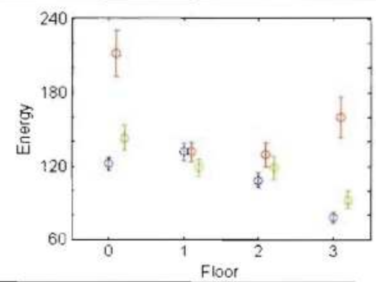
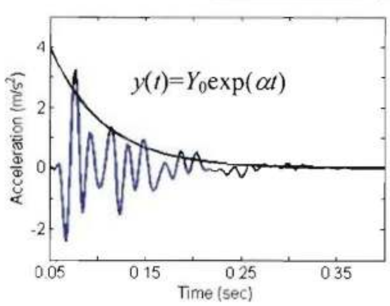
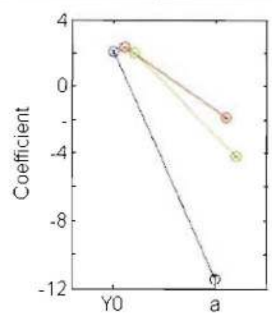
3.2. Discussion: feature extraction for the structural dynamics model validation

In this study, the damping parameters were defined to be the most influential uncertain parameters in the application of dynamic response analysis of the test-bed structure. Table 1 presents some extracted response features expected to be used in the damping parameters validation. Blue results are all from experimental data, and red and green results are from numerical outputs. Notice that the numerical output used in deriving green results was more accurate than red one because assigned damping ratios in the green output were values obtained form an experimental modal analysis. Considering the physical meaning of damping, the parameters should influence on the amplitude behaviour of the response, which is directly related to the energy dissipating behaviour. In the use of a random and stationary response from the linear system; i.e., w/o the bumper-column mechanism, the standard deviation, which was one of basic statistics, showed sensitivity to the accuracies of damping parameters. The feature that had the same physical meaning for a transient response was the energy value of temporal moments, and it also showed almost similar sensitivity as the standard deviation.

On the other hand, the features that were not extracted directly from the waveform, such as features derived by a subsequent model fitting process, were not as sensitive as previous two features. The parameters set of a decaying function fit to the data is an example of such features (see Table 1). It was considered that this low-sensitivity was caused by the difference of error components between experimental and numerical responses. The experimental data generally includes the measurement noise and environmental variability. However, numerical responses are free from such noise components. The result of fitting process was greatly influenced by variability in the experimental data; therefore, it was not considered to be appropriate to use such features for comparing experimental and numerical outputs. High dimensional features such as frequency response function were not considered because of difficulties in making quantified comparisons of such high-dimensional quantities.

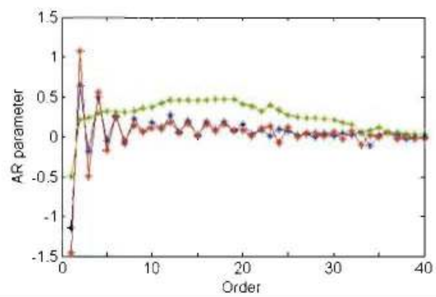
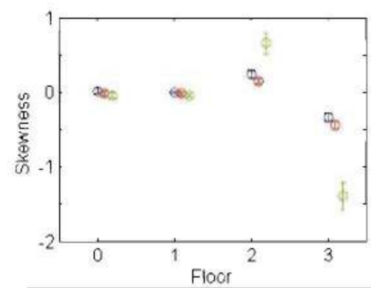
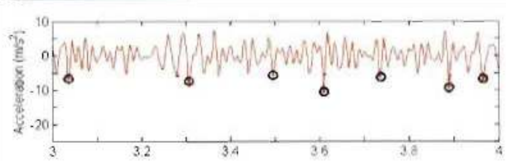
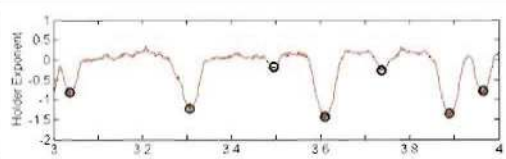
For validating nonlinearity modelling parameters; *Gap* and Δ, we found that some time-series model parameters; e.g., AR model parameters as shown in Table 2, had sensitivity to the accuracy of response. This point was also indicated in the previous work that summarized response features for analysing nonlinearity in dynamic responses for the application of structural monitoring [2].

Table 1. Examples of response features for validating damping parameters
(Blue: experiment, Red: numerical with all 1% damping ratios,
Green (accurate); numerical with damping ratios from experimental modal analysis)

	Feature description and definition: output x_i	Response feature plot
Standard deviation (Basic statistics)		
Energy (Temporal moments)	$M_k = \sum_{i=1}^{N} t_i^k x_i^2$ in k=0	
Decaying function fitting		

The time-resolution was also recognized as an important factor for investigating the nonlinear phenomena. The parameters; *Gap* and Δ, greatly influenced the frequency and the amplitude of impact events. Therefore, the feature that is sensitive to the number of impact events (i.e., the skewness for random response), had great sensitivity to the accuracy of numerical output as shown in the table. However, this feature alone could not be used in the detail validation of *Gap* and Δ because the dynamic amplitude behaviour during the impact events was not well captured by this feature. The response feature that provided a time-resolution measure was the Holder exponent [5]; however, there was trade-off between the feature dimension and the time-resolution.

Table 2. Examples of response features for validating nonlinearity modeling
(Blue: experiment, Red (accurate): numerical with Δ=0.5mm,
Green; numerical with Δ=0.1mm)

	Feature description and definition: output x_i	Response feature plot
AR model parameter	$\hat{x}_i = \sum_{j=1}^{p} \alpha_j x_{i-j} + \varepsilon_i$	
Skewness (Basic statistics)	$S_x = \frac{1}{\sigma_x^3} \cdot \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_x)^3$	
Holder exponent		

4. Statistical model validation using Mahalanobis distance comparison method

As mentioned in previous section, if multiple aspects of the system response are of interest, the model validation process may require comparing several features that have sensitivity to each uncertain parameter. A multivariate analysis technique was then expected to be useful in this process. Worden et al. had proposed a Mahalanobis distance outlier detection method for comparing multivariate feature vectors for the application of statistical structural damage detection in their previous study [6]. Applicability of the same approach for the statistical model validation was investigated in this study.

4.1. Mahalanobis distance outlier detection method

Mahalanobis distance is a distance measure used for multivariate statistics defined as

$$D_k = (\mathbf{y}_k - \overline{\mathbf{y}})^{\mathrm{T}} \mathbf{S}^{-1} (\mathbf{y}_k - \overline{\mathbf{y}}), \tag{4}$$

where $\mathbf{y}_k$ is a multi-dimensional feature vector for which the normalized distance from the mean is being calculated, $\overline{\mathbf{y}}$ and $\mathbf{S}$ are a mean vector and a covariance matrix based on all acquired vectors defining the nominal condition, respectively. The low distance value D_k can be obtained if $\mathbf{y}_k$ is similar to the set of feature vectors that defines the nominal condition. The outlier detection procedure is then summarized as follows.

(1) Acquire experimental responses (experimental data set) from the target structure.

(2) Calculate a feature vector from each measured set.

(3) Derive a Mahalanobis distance each experimental data set k, given as

$$D_k^{\mathrm{E}} = \left(\mathbf{y}_k^{\mathrm{E}} - \overline{\mathbf{y}}^{\mathrm{E}}\right)^{\mathrm{T}} \mathbf{S}_{\mathrm{E}}^{-1} \left(\mathbf{y}_k^{\mathrm{E}} - \overline{\mathbf{y}}^{\mathrm{E}}\right). \tag{5}$$

Notice that mean feature vector $\overline{\mathbf{y}}^{\mathrm{E}}$ and covariance matrix $\mathbf{S}_{\mathrm{E}}$ are calculated from all acquired experimental data. A set of D_k^{E} is then the experimental baseline distribution of Mahalanobis distance.

(4) Create a feature vector from a numerical output, and calculate a Mahalanobis distance D^{N} using the mean vector and the covariance matrix from the experimental data set; $\overline{\mathbf{y}}^{\mathrm{E}}$ and $\mathbf{S}_{\mathrm{E}}$,

$$D^{\mathrm{N}} = \left(\mathbf{y}^{\mathrm{N}} - \overline{\mathbf{y}}^{\mathrm{E}}\right)^{\mathrm{T}} \mathbf{S}_{\mathrm{E}}^{-1} \left(\mathbf{y}^{\mathrm{N}} - \overline{\mathbf{y}}^{\mathrm{E}}\right). \tag{6}$$

(5) Compare D^{N} with the experimental baseline distribution D^{E}.

If the feature vector form numerical output were similar to those from the experimental data set, D^{N} would show a lower value. The validity of a numerical model was then expected to be investigated statistically by comparing D^{N} with the experimental baseline distribution.

4.2. Damping parameters validation using linear system responses

4.2.1. Experimental baseline distribution and numerical run set

An experimental baseline distribution was created by acquiring twenty acceleration data sets (Exp. data #1~#20) under the same linear condition (w/o the bumper-column mechanism) of the test-bed structure. The input force was the band-limited random excitation with frequency range of 20-200Hz. Accelerations from each floor was acquired; the length of data was 16384 points with sampling frequency of 640Hz.

A numerical run set consisted of 200 calculation outputs was created using variable damping parameter sets; variable 2nd and 4th modal damping ratios. 200 parameter sets were sampled from the ranges of the two damping ratios indicated in Table 3 using Latin hypercube sampling method. 200 time-response accelerations (Run #1~#200) were then calculated by using each parameter set. Notice that the input force in all calculations was the same time-history that was measured in one of the experimental data set acquisitions; Exp. data #10.

Table 3. Parameter ranges for Latin hypercube sampling

	Minimum	Maximum
2nd mode damping ratio ζ_2	0.01	0.08
4th mode damping ratio ζ_4	0.001	0.02

The accuracy of all 200 numerical outputs against the experimental data #10 was plotted in Fig.3. The RMS error values were calculated using the 3rd floor acceleration outputs. The highest accuracy was shown in Run #114, and the lowest one was in Run #174 corresponding to damping values of (ζ_2, ζ_4) = (0.046, 0.070) and (ζ_2, ζ_4) = (0.0122, 0.0013), respectively. By examining these two responses when overlaid on the experimental data they were attempting to predict as shown in Fig.4, it can be seen that the different damping parameters mainly influence the amplitude of response as mentioned in the previous chapter.

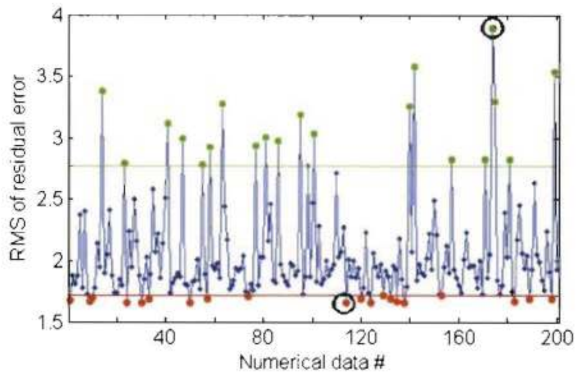
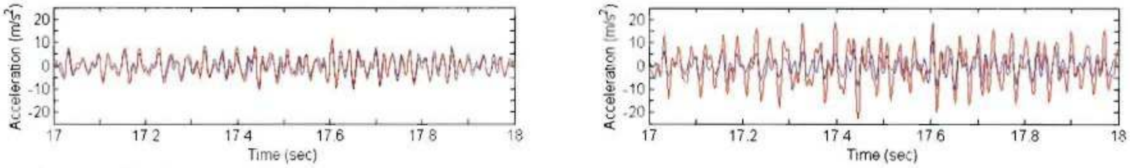


Fig.3. RMSE plots of all 200 numerical data



(a) The highest RMSE numerical run #114 (b) The lowest RMSE numerical run #174

Fig.4. Overlays of numerical (Red) and experimental (Blue) time-histories from linear system

4.2.2. Response feature extraction and Mahalanobis distance comparison

Response features selected for validating the damping parameters from random and linear responses here were then the peak amplitude and the standard deviation. The two values were extracted from outputs in the 2nd and 3rd floors producing a four-dimensional feature vector. Figure 5 is the corresponding Mahalanobis distance plot. Notice that the blue dots are the experimental baseline distribution, and the black dots are distances of the 200 numerical runs. Accurate and inaccurate numerical runs, which showed 10% lowest and highest RMSE values in Fig.3, are indicated by red and green circles, respectively. Seeing this figure, the accurate numerical runs have Mahalanobis distance values that predominantly fall within the experimental baseline distribution. In addition, the numerical runs that show the minimum/maximum Mahalanobis distance agree with Run #114 and #174, respectively as identified by an arrow. It can be said that the Mahalanobis distance derived by using the peak amplitude and standard deviation response features, has appropriate sensitivity to the time-histories; to be appropriated features for validation of the damping parameters.

The consistency check was then also carried out to confirm the effectiveness of this method for the statistical model validation. Consistency here meant that the parameter set that showed the low Mahalanobis distance provided accurate responses to any input forces. Additional twenty numerical runs were created by using each of twenty input force data in Exp. data #1~#20. Notice that the damping parameter set that showed the lowest Mahalanobis distance in Fig.5 was used in all calculations. Figure 6 (a) is the RMSE plot of additional twenty numerical runs, indicated in green points, presented with that of previous 200 runs. It was clearly observed that comparably high accuracies were obtained in all runs. Calculated Mahalanobis distances from the twenty runs are also presented in Fig.6 (b). All of them are thus distributed in the same Mahalanobis distance order as that of the experimental distribution. These results indicate that the parameter set that shows a low Mahalanobis distance can constantly provide accurate numerical outputs and vice versa. It can then be concluded that the statistical validity of uncertain parameters can be evaluated by this Mahalanobis distance comparison method.

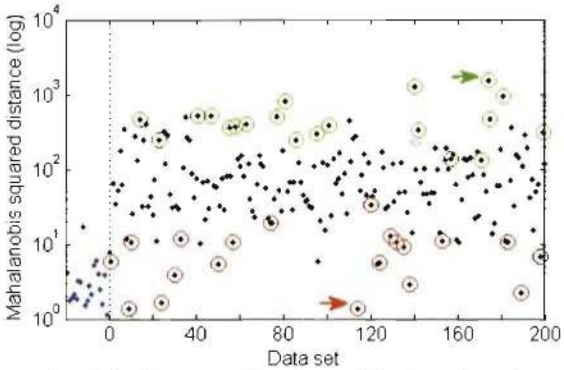
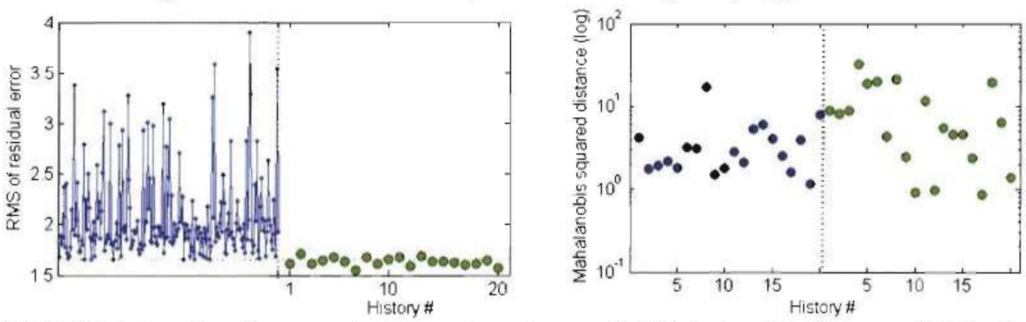


Fig.5. Mahalanobis distance plot for validating damping parameters



(a) RMSE plots using the accurate parameter set (b) Mahalanobis distance distributions

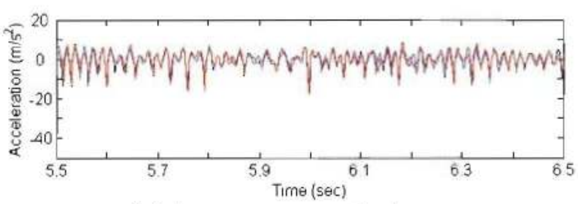
Fig.6. Consistency check of the Mahalanobis distance comparison method

4.2.3. Discussion: difficulty in validating correlated uncertain parameters

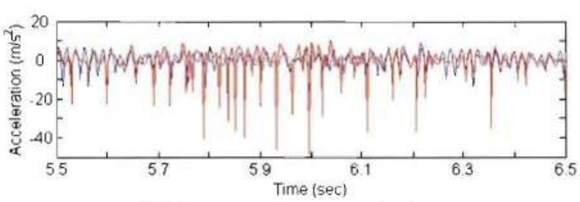
The success in the damping parameters validation presented in the previous section was realized by appropriate response feature selection. This need for appropriate response feature selection was further confirmed in the nonlinearity modeling parameters validation study. The uncertain parameters were *Gap* and Δ; however, it was difficult to find appropriate features that could assess the validity of each parameter independently.

In the validation, the experimental baseline distribution was created from twenty experimental data acquired under the same random excitation as that in the linear system data acquisition. Notice that the clearance between the bumper and the suspended column (= *Gap*) was set to 0.1mm in all measurements. However, the clearance adjustment was carried out in each data acquisition using a feeler gauge. This measurement method led to variability in actual conditions related to *Gap* and Δ. Four-hundred numerical runs were then created using sampled parameter sets (*Gap*, Δ). The same

Mahalanobis distance comparison procedure was carried out as that used in the damping parameters validation study. Even though some response features that were expected to be sensitive to the impact event; such as skewness, kurtosis, and time-series model parameters, the Mahalanobis distances from the most accurate numerical runs (as assess by RMSE) never distributed in the same order of the experimental baseline distribution. One main reason for these results was that the two parameters; *Gap* and Δ, were strongly correlated and response features that could have sensitivity to each of two parameters independently had to be used. Figure 8 is the parameter set plot, which indicates all parameter sets sampled in the creation of numerical runs. In this figure, the red dots are accurate numerical runs, which show low RMSE values, and the green dots indicate inaccurate numerical runs. Although the gap was set to 0.1mm in the experiment, the runs calculated using smaller *Gap* but with large Δ are also recognized as accurate outputs; i.e., two parameters are correlated. This observation indicates that the correlation of uncertain parameters in the numerical model also is a significant issue for the feature selection process. If such correlation cannot be prevented in the numerical model construction, it will require using response features that are independently sensitive to the validity of each uncertain parameter. Moreover, results here then confirmed that the Mahalanobis distance comparison method could be effectively applied to the statistical model validation only by using appropriate response features.



(a) Accurate numerical run



(b) Inaccurate numerical run

Fig.7. Overlays of numerical (Red) and experimental (Blue) time-histories from nonlinear system

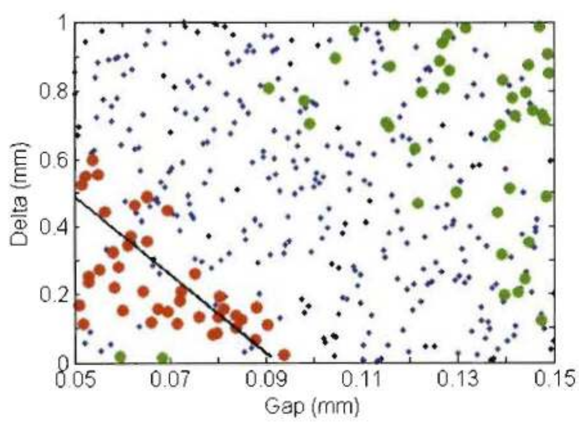


Fig.8. Map of sampled parameter sets with indicating correlation between *Gap* and Δ

5. Conclusions

- When selecting features for the model validation, issues that must be considered are not only dimension of the feature vector and type of response, but also the difference in sources of variability between experimental and numerical outputs. When using some features that require a fitting procedure, this variability can influence the feature's sensitivity to the parameters of interest.

- It was shown that the Mahalanobis distance comparison method was useful for the statistical model validation based on multivariate feature vectors.
- Correlation of uncertain parameters in the numerical model greatly influences the success of Mahalanobis distance comparison. This issue should be considered both in the construction of a numerical model and in the response feature selection.

The first and third issues should be advanced in the future works. They will become important considerations to realize more generalized structural dynamics model validation strategy and accurate uncertainty quantifications.

References

[1] Atamturktur, S., Hemez, F., Unal, C., Calibration under Uncertainty for Finite Element Models Masonry, LA-14414, Los Alamos, NM, Los Alamos National Laboratory, 2010.

[2] Figueiredo, E., Park, G., Figueiras. J., Farrar, C., Worden, K., Structural Health Monitoring Algorithm Comparisons Using Standard Data Sets, LA-14393, Los Alamos, NM: Los Alamos National Laboratory, 2009.

[3] Farrar, C., Worden, K., Todd, M., Park, G., Nichols, J., Adams, D., Bement, M., Farinholt, K., Nonlinear System Identification for Damage Detection, LA-14353, Los Alamos, NM: Los Alamos National Laboratory, 2007.

[4] Hemez, F., Farrar, C., Lecture notes: Finite Element Model Validation, Updating, and Uncertainty Quantification, Los Alamos National Laboratory, 2009.

[5] Robertson, A., Farrar, C., Sohn, H., Singularity Detection for Structural Health Monitoring using Holder exponent, Mechanical Systems and Signal Processing, Vol. 17, Issue 6, 1163-1184, 2003.

[6] Worden, K., Manson, G., Fieller N.R.J., Damage Detection using Outlier Analysis, J. Sound and Vibration, 229 (3), 647-667, 2000.