The Use of FE Model Updating and Probabilistic Analysis for Dealing with Uncertainty in Structural Dynamics Simulation

Eddy Dascotte

Dynamic Design Solutions N.V.
Interleuvenlaan 64
B-3001 Leuven
Belgium
Eddy.Dascotte@dds.be - www.femtools.com

ABSTRACT

Finite element analysis has become an essential tool to support virtual product development. However, a lot of uncertainty exists in the modelling of physical structures, and in the effects on their performance of variability in manufacturing and usage of products. To successfully make the move to digital prototyping, and thereby reduce the number of physical prototypes, predictions of performance should be provided with a measure of confidence and validated against experimental data. This requires quantifying the physical and numerical uncertainty.

Current FE model validation and updating practice is based on nominal values for input parameters and test results. The consequences of using random updating parameters that represent uncertainty and statistical test data are described. Scenarios for how to incorporate probabilistic simulation tools like Monte Carlo Simulation in model validation and updating procedures are outlined.

Keywords: *simulation model validation, model updating, uncertainty analysis, probabilistic analysis, structural dynamics*.

1. INTRODUCTION

The integration of CAD and CAE solvers, automatic meshing algorithms, parallel and distributed computing and the availability of relatively low cost engineering workstations has led to an inflation of finite element model sizes.

However, although the complexity of engineering problems that can be addressed nowadays has become impressive, model validation should not be overlooked. A numerical model is worth only as much as the level of confidence in the results that the analyst is able to guarantee. Without this guarantee, numerical analysis cannot become the trustworthy engineering simulation tool on which the virtual product development revolution will depend. Validation methodologies have to consider the increasing complexity of numerical simulations when applied to realistic industrial problems.

Comparison of numerical analysis results with experimental data is an accepted way of validating a model, and in the absence of theoretical solutions may be the only possible one. In many situations, the outcome of this test-analysis correlation is not satisfactory and needs to be followed by a model updating phase. However, although most analysts will recognize that uncertainty exists in both the numerical simulations and the experimental data that serves as reference data, this aspect has not really yet been taken into account by the current correlation and updating techniques [1].

Classical deterministic approaches may lead to false and misleading conclusions if incomplete, uncertain and noisy data is used. It is clear that this danger can only be overcome by adopting statistical techniques. Instead of comparing single analysis or test results, statistical correlation compares sets of data that reflect the variability of the structural responses as the result of scatter, uncertainty and noisiness. The result of this working with fuzzy data will be that good or bad correlation will be quantified with a given probability and thus provide the analyst a measure of confidence on the correlation metrics.

Model updating technology has matured over the past decades and today commercial software solutions are available that assist analysts not only with data management and graphical visualization, but provide all the essential analysis tools like pretest analysis, correlation analysis, error localization, data expansion and/or reduction, sensitivity analysis and dedicated model updating optimization solvers [2]. Extension of these tools with probabilistic analysis capabilities now offers a solution to also incorporate uncertainty in the model validation and updating process. This holds a promise to solve some of the remaining difficulties like selection of updating parameters, definition of targets (i.e. when is correlation satisfactory) and interpretation of results. Without the additional insight that can be gained from probabilistic analysis, these decisions have to be made mainly based on engineering judgment.

Probabilistic analysis is here defined as the collection of tools that produce and interpret point clouds rather than the single point results produced by deterministic analysis. Together with model updating (and its supporting tools like correlation analysis, and sensitivity analysis), probabilistic techniques provide the framework to enhance the validity of realistic simulation models by dealing with uncertainty.

Although the techniques described hereafter are generally applicable, emphasis will be on linear structural dynamics. This is no surprise because structural dynamics today continues to be one of the more challenging types of engineering analysis. The many influencing variables, often badly known like damping, and their interactions, seldom allow analysts to produce satisfactory simulation models from their first version. Luckily, however, the technology and instruments are available to measure dynamic responses of structures in a relative fast and reliable way. In addition, testing can be done under controlled laboratory or real-life operating conditions resulting in potentially abundant collections of data that contain valuable statistical information for analysts.

2. SOURCES OF UNCERTAINTY

Uncertainty in numerical simulation results manifests itself in 2 main classes: physical uncertainty and numerical uncertainty.

There exist four main levels at which *physical uncertainty*, or scatter, becomes visible:

- Boundary and initial conditions - *impact velocity, impact angle, mass of vehicle, characteristics of barrier, etc.*
- Material properties – elasticity, *yield stress, density, damping, local imperfections, etc*
- Geometry - *shape, thickness, joint stiffness, manufacturing and assembly tolerances, etc.*
- Loads - *earthquakes, wind gusts, sea waves, blasts, shocks, impacts, etc.*

Uncertainty is further increased because many of these properties may vary substantially with temperature, frequency, or load level.
Information on these forms of scatter can be obtained by measurement. A sufficient large number of samples need to be evaluated to distinguish the natural and intrinsic scatter from the (often high) scatter that may be attributed to a small number of statistical samples. Probability distribution functions and their associated properties can be obtained from the statistical analysis of the test data. For example, the elastic modulus of isotropic materials can be described using a normal (Gauss) distribution that is characterized by a mean value and standard deviation.

The following types of *numerical uncertainty* can be identified:

- Conceptual modeling uncertainty - *lack of data on the physical process involved, lack of system knowledge.*
- Mathematical modeling uncertainty - *accuracy of the mathematical model validity.*
- Discretization error uncertainties - *mesh density*.
- Numerical solution uncertainty - *rounding-off, convergence tolerances, integration step.*
- Human mistakes - *programming errors in the code, data and units mistakes, wrong utilization of the software.*

These types may or may not exist regardless of the physics involved. Another example of the exhibit of numerical uncertainty is the different results that may be obtained by two finite element codes, using the same finite element model. Indeed changing solver, computing platform, and/or element formulation can be possible causes.

It is clear that uncertainty also exists in testing. Possible causes of *physical* uncertainty are related to:

- Test definition – *fixture, mounting procedure, excitation method, transducer location, sensor weight, etc.*
- Dynamic loading
- Instrumentation – *calibration, distortions, cabling noise, etc.*
- Data acquisition – *digital signal processing, measurement and filtering error.*

Techniques like Experimental Modal Analysis, are also subject to *numerical* uncertainty in the mathematical models that are used for modal parameter estimation.

3. HOW TO DEAL WITH UNCERTAINTY

Increasing model size (i.e. mesh density) is not automatically leading to better simulation models. Instead, correctly representing the physics that govern the performance of a product is much more rewarding. Increasing complexity of engineering systems and the existence of scatter result in uncertainty, a layer of physics that in the past has often been overlooked in CAE. When incorporating uncertainty in the analysis of large and complex systems, the performance is described in terms of trends, patterns, averages or most likely performance. Classical deterministic analysis, on the other hand, attempts to improve the nominal performance, which practically never coincides with the most likely performance.

This is illustrated in the following example [3]. The expected value is a measure of central tendency and is commonly referred to as the mean value of a random variable. Now let g(X) be a continuous function of the continuous random variable X. For this case the expected value E of g(X) is given by

$$E(g(X)) = \int_{-\infty}^{+\infty} g(X)\, f(X)\, dX \qquad (1)$$

with f(X) the Probability Density Function (PDF), which defines the distribution of the probability density associated with X.

It can be shown that for the general case,

$$E(g(X)) \neq g(E(X)) \qquad (2)$$

For example, if $g(X) = X^2$, then

$$E(g(X)) = (E(X))^2 + \sigma_X^2 \qquad (3)$$

Where σ_X is the Standard Deviation on variable X, while

$$g(E(X)) = (E(X))^2 \qquad (4)$$

The importance of this result is that, in general, the expected value of a structural response cannot be obtained by simply calculating the response associated with the nominal values of the random variables. It is for this reason that specific probabilistic analysis techniques are required to estimate the probability distribution of responses from the known (or assumed) distributions on variables. Statistical methods, and Monte Carlo Simulation (MCS) in particular, provide a solution to incorporate uncertainty in numerical simulation and thus compare simulation with test in terms of probability and confidence rather than in classical one-to-one deterministic correlation.

The Monte Carlo method is a numerical method of solving mathematical problems through random sampling. As a universal numerical technique, the method became possible only with the advent of computers, and its application continues to expand with each new computer generation. Although the focus nowadays is still on reducing the number of samples (like pseudo-random sampling, or Latin Hypercube Design Of Experiments), the Monte Carlo method is a general applicable probabilistic method that works in all cases and is therefore considered as a reference tool.

The concept of a probabilistic analysis based on Monte Carlo Simulation is rather simple. A set of design parameters is specified with their probability distribution functions (PDF) and the objective is to obtain a description of the performance of the system on a statistical basis, i.e. via histograms. This can be done by repeatedly generating a random selection of possible parameter values (based on their PDF) and run the solver for each selection. Each random selection will lead to a different analysis result. All results are statistically post-processed to obtain the PDF of the responses or to obtain confidence intervals. If a sufficient large number of analyses are run, then it can be shown that PDF of the responses approaches their real PDF. With only a limited number of runs, the response PDF can only be known approximately. However, for most engineering analysis an extremely high precision is not required (unless the application is reliability studies) and in practice modern and efficient sampling schemes enable to work with between 100-150 random samples. This number is independent of the number of random parameters used.

It is essential to understand that this type of repetitive random sampling and re-analysis is required because it is not possible to know the most likely value of the structural responses simply by calculating the response associated with the nominal values of the input random variables, as was shown previously. In the same way, the practical minima and maxima of structural responses cannot be obtained by calculating the responses associated with the minimum and maximum values of the input parameters respectively (interval analysis).

This is illustrated in the following simple example [4]: imagine that the norm of a N-dimensional vector is the performance value that needs to be minimized. Suppose that the N components of the vector are random normal (Gauss) variables with mean 0 and standard deviation 0.3. Samples of these components show that all variables are in the range [-1, +1] for N = 10. Solving this problem in a classical deterministic way leads to the trivial solution that the minimum norm of the vector is zero, requiring that all components also be equal to zero. This is true for any value of N. The norm computed for the extreme values of the variables is 3.16. However, solving this problem with Monte Carlo Simulation leads to the following surprising results:

N	Min. Nom	Max. Norm	Mean Norm
10	0.2	1.5	0.9
50	1.5	2.7	2.1
100	2.3	3.6	3.0

Neither the minimum nor the mean norm are zero. The table shows additional results for higher values of N, showing that the values increase if the dimension of the vector increases.

The results in the table have been obtained simulating the vectors 1000 times. The minimum and maximum values correspond with a cumulative probability value of 0.001 and 0.999 respectively. However, using much less Monte Carlo samples allows finding approximate but indicative numbers for the norms. These values are an intrinsic property of the vector norm and not of the number of Monte Carlo samples.

The above results show that model validation should be concerned about the most likely response of the structure when comparing results from simulation and testing. For systems that exhibit significant scatter, the most likely response cannot be readily obtained with deterministic analysis and Monte Carlo Simulation is required.

4. PROBABILISTIC VALIDATION AND UPDATING

In [4-5], an approach to statistical validation and updating of simulation models based on the concept of meta-models is presented. Monte Carlo analysis produces histograms and scatter plots. Histograms show the number of times a response value (or a range) is obtained and reflects the underlying probability density. Scatter plots are orthogonal projections of an N-dimensional point cloud into 2D or 3D plots. Each point cloud represents the relation between a random input variable and a corresponding output value. There is a point for each state (also called sample) of the input variable. On statistical grounds, the collection of all point clouds, one for each input-output variable combination, constitute a new concept of model, often referred to as Meta-Model in literature. This model can be complemented by input-input and output-output relationships. An example of point clouds is shown in figure 1.

It is clear now that in the presence of scatter, the single deterministic response represents only one point of the cloud and therefore carries little information on likeliness and trends. The point clouds on the other hand can be interpreted in term of probability of a response value lying below or above a prescribed level. At a minimum, all responses are now defined as an interval with information on the confidence an analyst can have that the true response value will be within this interval. Additional statistical information can be derived if necessary.

Another fundamental concept is that of distance between meta-models. Statistical theory provides a simple measure of distance between 2 Meta Models, namely the Mahalonobis distance:

$$d_M = (\mu_1 - \mu_2)^t \, COV_p^{-1} \, (\mu_1 - \mu_2) \qquad (5)$$

where the vectors μ_1 and μ_2 represent the centres of gravity of each meta model and COV_p the pooled covariance matrix. Whereas a deterministic measure of correlation, like for example the average relative error on resonance frequencies, provides only a snapshot measure that could be good or bad depending on coincidence, the Mahalonobis distance is clearly a much safer measure because it is based on position and shape of point clouds. Coincidence, good or bad luck with parameter estimations, or variable measurement conditions can hardly influence this result. Figure 2 shows illustrates the Mahalonobis distance between 2 point clouds.

The concept of meta models, both for numerical simulation and testing, together with the Mahalonobis metric, enables to compare responses in a statistically sound and rigorous manner. Position, shape and size of point clouds should be compared with the test meta model being the reference. For example, consider the scatter plots shown in figures 3 and 4. Differences in the principal axes of the two ellipses suggest either a major shortcoming in the discretization of structures geometry, a physical discrepancy between the two models or simply modelling errors. It should be clear that relative translation and overall size of point clouds are easier to correct than relative rotations. The former merely indicate systematic or global errors whereas the latter usually indicate (local) physical errors.

Secondly, the level of scatter in the two models is clearly different. Although this may be desirable in some cases, it is in general preferable to obtain a simulation model that exhibits a level of scatter that is in balance with the scatter on the test data.

A fundamental contribution of meta model analysis towards model updating is the possibility of pinpointing the dominating parameters of a system and to quantify the correlations between the input and output variables. This is the equivalent of sensitivity analysis in deterministic analysis. However, the concept of sensitivities, or gradients, no longer exists in the presence of scatter. So unless scatter is very low and can be neglected, other procedures to identify the dominant parameters need to be applied. In a similar way, not all available responses may be of equal relevance. Indeed, statistical postprocessing may reveal hidden relations and identify dependent and independent responses. As a result, the analyst can reduce the order of the system to include the most dominant parameters and independent responses. Using regression analysis, relations between the dominant parameters and independent responses are established. The objective of model updating is then to solve the system of equations for unknown parameter properties that change the centre of gravity, the principal directions and the density of point clouds resulting from probabilistic analysis to match the corresponding test point clouds. In fact this comes down to ‘updating‘ the Probability Density Function (PDF) of input parameters such that the PDF of the outputs correspond with the PDF of the experimental reference responses. In its simplest form, assuming a normal probability distribution, this means that in addition to the nominal value (like in deterministic model updating), also the standard deviation of model parameters should be adjusted.

For a more profound discussion of these concepts and possible procedures, the reader is referred to [4] which is based on initial experiences in automotive and aerospace industry.

It should be noted that the ranking of input parameters based on how much they influence the performance of the system, offers additional benefits in the subsequent design improvement phase. Indeed, a designer or engineer does not need to spend time with input parameters that have only minor influence. Instead the functional performance of the design can be modified most efficiently by working with the most dominant parameters only. Reducing the scatter on these parameters (for example by specifying more severe manufacturing tolerances) is the most rewarding in terms of robustness of the design. On the other hand, the engineer should relax tolerances on the parameters that do not significantly influence the performance, and in the process save money on manufacturing costs. An example of parameter ranking using pie-charts is shown in figure 5.

5. A MIXED DETERMINISTIC-PROBABILISTIC PROCEDURE

The consequences of adopting a probabilistic approach to CAE and simulation model updating in particular are significant. Concepts like gradients (sensitivity analysis), variational analysis and optimization rely on a functional relationship between parameters and responses. These are no longer usable when this relation is described in terms of meta-models. Now does all this mean that we should throw overboard all our existing deterministic updating technology? Certainly not!

The probabilistic procedure that is described in the previous section, represents an idealized process in which an engineer has access to unlimited simulations and tests. In practice the following obstacles arise:

- The amount and quality of test data is limited by tight budgets and time frames available for testing. In practice, only one test may be available and no reliable statistical information on the probability distribution.
- The probability distribution of input parameters can only be known as the result of intensive testing. Although this is certainly encouraged, most of the time only nominal, or minimum and maximum values are available.
- Model size and solver time still restrict the number of re-analysis runs that can be run. Fast but approximate solutions may be available but these also introduce numerical uncertainty.
- Uncertainty on the number of required samples. For example: critical reliability analysis will not accept MCA with only 100 samples if the requirements state that only less than 1 out of 10000 products may fail.
- The 'distance' between FEA and test can be significant and exceed what could be explained by scatter only. In a preliminary analysis phase, it has no use to run expensive probabilistic analysis if the FE model that is used exhibits to many deficiencies (for example, the FE model is simply not capable of representing the true system response, wrong physical properties are used, there exist too many numerical errors etc.)

For these reasons, classical deterministic correlation and model updating will continue to provide good services for a long time to come. However, the benefits of probabilistic analysis should not be neglected. It is therefore proposed to adopt a mixed deterministic-probabilistic procedure in which classical test-analysis correlation and model updating fits in without changes.

Deterministic correlation and updating techniques for structural dynamics applications are based on the functional relationship between the measured modal characteristics and the structural parameters that can be expressed in terms of a Taylor series expansion limited to the linear term. This relationship can be written as:

$$\{R_e\} = \{R_a\} + [S](\{P_u\} - \{P_o\}) \quad (6)$$

or

$$\{\Delta R\} = [S]\{\Delta P\} \quad (7)$$

Where:

$\{R_e\}$ Vector containing the reference system responses (experimental data).
$\{R_a\}$ Vector containing the predicted system responses for a given state $\{P_o\}$ of the parameter values.
$\{P_u\}$ Vector containing the updated parameter values.
$[S]$ Sensitivity matrix.

The discrepancy between the initial model predictions and the test data is resolved by minimizing a weighted error E, given by:

$$Min(E = \{\Delta R\}^t [C_R]\{\Delta R\} + \{\Delta P\}^t [C_P]\{\Delta P\}) \quad (8)$$

and subject to constraints

$$g_i(P) \le 0 \quad ; P_{\min} \le P \le P_{\max} \quad (9)$$

The matrices $[C_R]$ and $[C_P]$ respectively express the confidence of the user in the reference system responses and initial parameter estimates [6]. In case the confidence matrices are derived from statistical postprocessing of multiple tests, then they can be obtained from the covariance matrices. Equivalence with the Mahalonobis distance (equation 5) should be noted here.

Deriving equation (8) and minimizing E with respect to the parameter values, leads to a updated value for the parameter values that simultaneously reduce the distance between the simulation and test results (based on single values for resonance frequencies, modal correlation coefficients etc.), but at the same time keep the distance between the original and updated model minimal (in terms of parameter changes).

The implementation of the above method in commercial software like FEMtools [2], offers the analysts a broad choice of response and parameter types to explore the behaviour of the simulation model and quickly try different strategies to improve correlation (see figure 6). However, validation in dynamics focuses on mode orthogonality (based on standards used in defence and aerospace industry) and mode vector correlation. Engineering judgment remains critical for the selection of updating variables, definition of targets and interpretation of results. Automating these tasks will most benefit from the introduction of probabilistic concepts.

A mixed deterministic-probabilistic procedure can be used as an intermediate step before the advent of entirely probabilistic procedures like the one described in the previous section. Let's review this procedure in more detail (see figure 7):

- Distinguish model-related uncertainty (partially physical, partially numerical) from uncertainty due to manufacturing and operational uncertainty.
- Apply classical deterministic correlation and updating. The results are improved global correlation levels and information on the minimal parameter changes required to obtain these changes.
- Estimate the probability distribution function (PDF) of input parameters.
- Run Monte Carlo Simulation, possibly with reduced models and/or using smart sampling, to obtain the simulated PDF on output responses.
- If PDF information on output responses is available (from multiple testing), then compare with the simulated PDF. This comparison provides information to adjust the PDF of the input parameters.
- The updated simulation model can be used for subsequent optimisation (deterministic).
- Apply the updated PDF on input parameters to predict output responses PDF of optimised FEM and interpret in terms of quality, cost, robustness and reliability.

A critical part is the numerous re-analysis that is required for Monte Carlo Simulations. Because in general finite element models are used, this means that for every sample of the parameter values, a new FE model needs to be generated and passed through a solver. In addition to using modern MCS techniques that allow to reduce the number of samples, possible ways to accelerate the re-analysis phase are:

- Using a modal solver.
- Using reduced system matrices and perturbation of the reduced matrices.
- Using first-order gradients (for very small levels of scatter) up to higher order sensitivity analysis for larger levels. In a general way, response surface can be used. Standard structural dynamics allows for the efficient computation of first-order gradients and thus the construction of a response plane. Design of experiments techniques or advanced variational analysis methods are required if higher order response surfaces are required.

6. CONCLUSIONS

Validation of a single deterministic model using a single test is leading to a snapshot result. While this may be valuable for roughly calibrating the input parameters, and to gain insight in the behaviour of the model, information on the broader picture is missing. This lack of information must currently be compensated by the engineering judgement of the analyst and prohibits further automation of model validation and updating processes. Probabilistic analysis, together with uncertainty management and knowledge databases, and enabled by massive computing capacity, hold the promise to provide the necessary tools to revolutionize CAE and lead to better, more reliable simulation models, not necessarily bigger models as it seems to be the trend nowadays. The available computing capacity, and automatic meshers linked to CAD models, should be used not to run increasingly bigger FE models but rather to construct meta models that can be validated, enhanced and used as the basis for quality improvement, cost reduction, robust design and reliability analysis.

REFERENCES

[1] Friswell M.I., Mottershead J.E., *Model Updating in Structural Dynamics: A Survey.* Journal of Sound and Vibrations 167:347-375.

[2] FEMtools 3.0, Software for Correlation, Validation and Updating of Simulation Models, Dynamic Design Solutions N.V., www.femtools.com.

[3] Lust R.V., Wu Y.-T.J., *Probabilistic Analysis – an Introduction,* Experimental Techniques, September/ October 1998.

[4] Marczyk J., *Principles of Simulation-Based Computer-Aided Engineering,* FIM Publications, 1999.

[5] Marczyk J., *Statistical Validation of FE Models Using Experimental Data,* Proceedings of the 21th ISMA Conference. 1996.

[6] Collins J., Hart G. , Hasselman T., Kennedy B., *Statistical Identification of Structures*, AIAA journal, 12/2, Feb. 1974, 185-190.

FIGURES

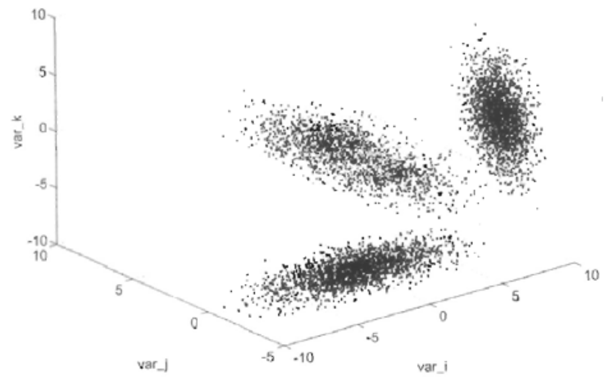

Figure 1: A meta-model represented as a 3D cloud.

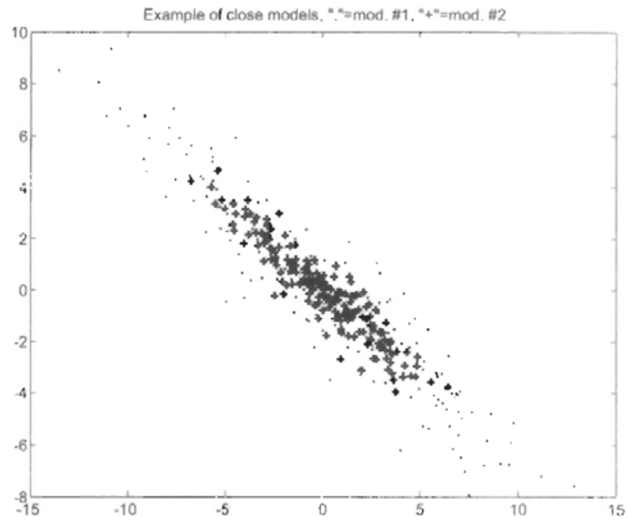

Figure 4: Example of two physically close models.

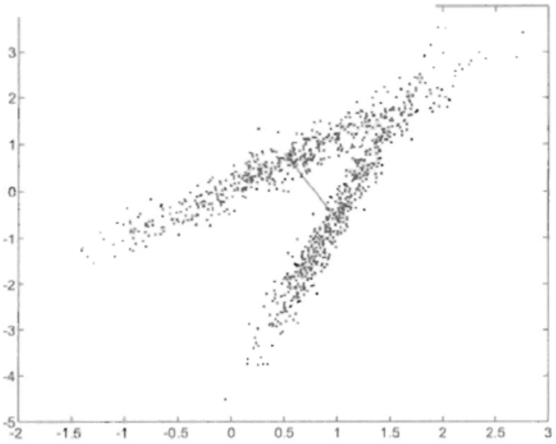

Figure 2: The Mahalonobis distance between two meta-models is measured relative to the means of each cloud.

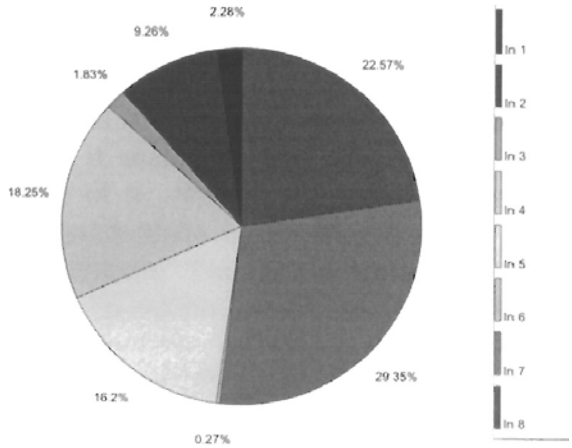

Figure 5: Identification of dominant variables using Spearman's rank correlation technique. The highest ranked variables have the highest relative influence on the response, taking into account the presence of scatter.

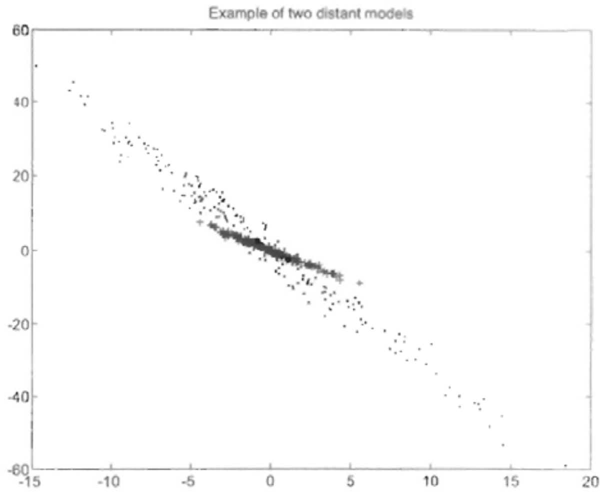

Figure 3: Example of two physically distant models.

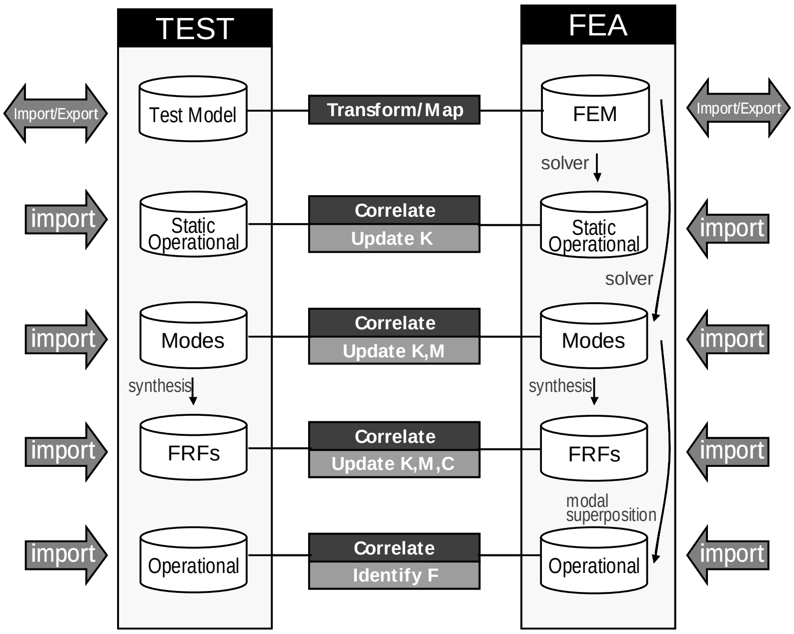

Figure 6: Test-analysis correlation and updating of finite element models for structural dynamics.

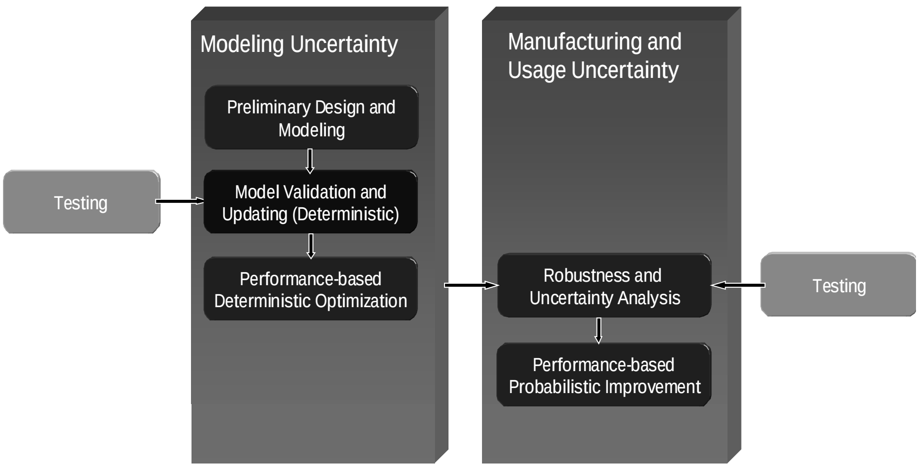

Figure 7: A mixed deterministic-probabilistic methodology to combine classical model validation and probabilistic analysis to introduce uncertainty on the input parameters and in the reference test data.