COBEM 2023
27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGENEERING
ABCM Brazilian Society of Mechanical Sciences and Engineering
27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1315
MODEL VALIDATION WITH CLASSICAL AND BAYESIAN HYPOTHESIS TESTING UNDER EPISTEMIC UNCERTAINTY.

João Henrique de Camargo Costa
Hélio Fiori de Castro
Unicamp, Rua Mendeleyev 200, BR-13083860 Campinas - Brazil
j265792@dac.unicamp.br, heliofc@fem.unicamp.br

Abstract. *As the systems used in engineering are increasing in complexity, physics-based simulations are being used to replace expensive physical tests for understanding system behavior. The computational model formulated to represent a physical system has an intrinsic uncertainty that needs to be accounted for. Model validation and verification is the processes of determining the degree to which a model is an accurate representation of the real world for its intended use. It involves comparison of model prediction with experimental data, expert intuition, theoretical results, or any combination of them. Epistemic uncertainty is expressed as a lack of knowledge of the physics, variations of the system and environment. A common approach to model validation is the graphical validation, a visually comparison between prediction and experimental observations, is still used to measure the accuracy of the model but this approach does not count for the uncertainty of the model and of system's properties. In this paper a statistic-based quantitative approach will be used to account for these uncertainties, both in model prediction and experimental observation. The classical hypothesis testing is a well-developed approach method of rejecting or accepting a model based on an error statistic. The Bayesian hypothesis testing will compare two hypotheses based on the amount of information available, minimizing the risk of model selection by choosing properly the model acceptance threshold and an avoidance of making type I/II error. Two system responses will be analyzed: model analysis of a rotor test bench and free vibration of the beam. The uncertainty of the test bench are those intrinsic of the measurement, environment and epistemic uncertainty; the parameters of the beam model will be obtained by Monte Carlo Method to assure uncertainty. The expected results are the assessment of model validation and limits of acceptance for both analyses.*

Keywords: *Model Validation, Hypothesis Testing, Rotordynamics*

1. INTRODUCTION

Verification and Validation (V&V) of computational models is necessary to provide confidence that the results of these models, used to solve complex problems, are sufficiently accurate and really represents the solution of the intended problem. Processes includes assessment activities that are performed during the creation and application of the model that approach technical issue of physical systems. The verification process is defined as the assessment of a computational model in representing the underlying mathematical equations and their solutions (Roache, 1998). Validation is defined as the processes of determining the degree to which a computational model is an accurate representation of the real world for its intended use (ASME,2012). The V&V subject is largely researched nowadays, Jiang and Mahadevan (2007) developed a method for risk-based model validation; Rebba *et a*l (2006) worked on a methodology for model confidence assessment. This paper will focus on the validation of a computational model, and the processes of validation can be resumed by Figure 1.

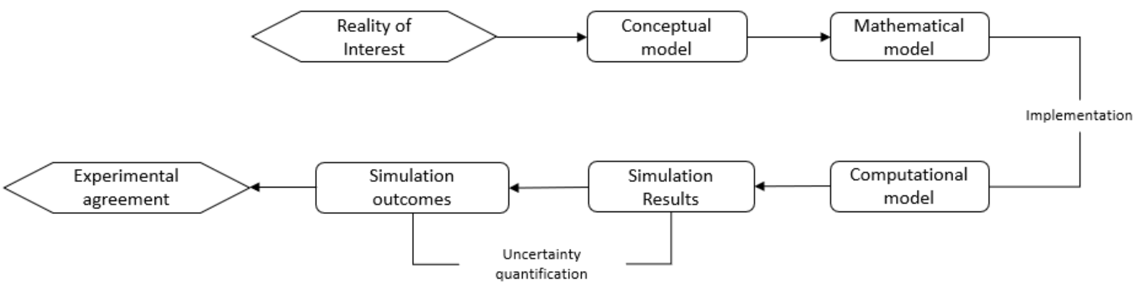

Figure 1. Validation of computational model processes.

Two key challenges of the V&V processes are the preparation of tests and definition of validation metrics. The validation metrics is a measure of agreement between model and experimental observation. According to Mahadevan and

Sankararaman (2011), one of the widely used methods is the graphical validation, a visual comparison between prediction data and results. This is a qualitative approach and do not account for numerical error or uncertainty in prediction or observation, in addition, this comparison provides little information about agreement range between experiment and model. A quantitative approach of the validation metrics are the hypothesis tests.

2. MODEL PREDICTION: QUANTITATIVE APPROACH

Assuming a computational model constructed to predict a physical phenomenon, the quantitative model validation involves the comparison between experimental observation and model prediction. Based on Ling and Mahadevan(2013), the notation is as it follows:

- Y represents the "true value" of the system response.
- Y_m is the model prediction of the true response Y
- Y_D represents the experimental observation of Y

Quantitative validation metrics consider assumptions stablished in Y, Y_m and Y_D, relating it to different sources and types of uncertainty regarding the type of validation data. In the case that the validation data is from a fully characterized experiment, and inputs $\boldsymbol{x}$ are measured as point values, other uncertainty sources can contribute to Y and Y_m be stochastic than the input uncertainty, such as variations on micro-structure that leads in a stochastic Young's module. Ling and Mahadevan (2013) stated that, in the Bayesian approach, the epistemic uncertainty (lack of knowledge) about $\boldsymbol{x}$ is represented through a subjective probability distribution. Thus, Y_D is the result of error measurement in addition to the true value Y, by definition $Y_D = Y + \varepsilon_D$, where ε_D is the measurement error.

3. HYPOTHESIS TEST METHODS

According to Montgomery (2003), statistical hypothesis tests and evaluation of confidence intervals are fundamental methods used in data analysis of a comparative experiment. It is important to remember that hypothesis is a statement about the population or distribution under study, not about the sample.

3.1 Classical Hypothesis Test

It is a well stablished method, first a statistical test is formulated and the probability distributions under the null and alternative hypothesis are evaluated theoretically or by approximations. After that, the test statistic value can be computed based on validation data and calculate the p-value, which is a metric that defines the probability of the test statistic falls outside a defined range set by the test statistic under the null hypothesis (H_0). The p-value is an important indicator of quality of the null hypothesis is, since a better H_0 gives a narrower range defined by the calculated value of the test statistic therefore a higher probability of the test statistic escapes the defined range.

Model validation is proven to be useful for decision making in terms of model usage. Probabilities of making errors type I and type II can help decision making in whether to accept or reject the model. The significance level α defines the maximum probability of committing a type I error, and the probability of committing a type II error can be estimated based on α and the probability distributions of the test under H_1 . If the p-value is less than α or if β exceeds the acceptable value the null hypothesis will be rejected. Therefore, the model will be accepted if H_0 is not rejected and will be rejected if H_0 is rejected. As an alternative to compare the significance level α and the p-value, a confidence interval can be used. The level of confidence can be set at γ=1-α, thus, H_0 will be rejected if the confidence interval does not include the predicted value. If the computed p-value is equal or less than α, then the null hypothesis is rejected and the result is said statistically relevant (Rebba, et al., 2006). Failing to reject H_0 indicates that the model's accuracy is acceptable, but does not prove that the null hypothesis is true. Also, comparisons between the p-value and the significance level α are insignificant if experimental data are abundant. As almost no null hypothesis H_0 is true, the p-value decays as the number of samples increases given a significance level α, tending to reject H_0 (Sankararaman & Mahadevan, 2011).

There are several statistical tests developed to correspond different kinds of hypothesis. The *t*-test, based on Student's *t*-distribution, is well fitted when model prediction meets the mean of a random variable. Suppose that the quantity of interest Y is a normal random variable with mean μ and standard deviation σ, both unknown, and the distribution error ε_D is a normal random variable with mean zero and standard deviation ε_D. Therefore, the experimental observation $Y_D = Y + \varepsilon_D \sim N(\mu, \sigma^2 + \sigma_D)$, hence, Y_D also follows a normal distribution with mean μ and standard deviation $\sigma + \sigma_D^2$. Let $\sigma_{YD} = \sqrt{(\sigma^2 + \sigma_D^2\ )}$. A set of random samples of Y_D with size n are the validation data (i.e., $y_{D1}, y_{D2}, \dots, y_{Dn}$) and the sample mean is $\overline{Y_D}$ and the sample standard deviation is S_D. With a t distribution of (n-1) degrees of freedom, the variable $(\overline{(Y_D\ )} - \mu)/\ (S_D/n\ )$ is modeled. Assuming that the mean prediction μ_m is the mean of Y, then, the statistic test t, under the null hypothesis $H_0\colon \mu = \mu_m$ is represented by the following (Ling & Mahadevan, 2013):

$$t = \frac{\bar{Y}_D - \mu_m}{s_D/\sqrt{n}}. \tag{1}$$

The two-tailed t-test p-value can be given by:

$$p = 2F_{T,n-1}\ (-|t|). \tag{2}$$

Where $2F_{\mathrm{T,n-1}}$ it's the cumulative distribution function (CDF) of a t distribution with $(n-1)$ degree of freedom (DOF). With significance level α chosen, the null hypothesis will be rejected if $p < \alpha$, or fail to reject if $p > \alpha$.

To account the probability of making a type II error, an alternative hypothesis H_1 is needed, then, $H_1: \mu = \mu_m + \varepsilon_\mu$ it's a usual choice. Under the alternative hypothesis, in t test, the statistic T follows a non-central t-distribution with noncentrality parameter $\delta = \sqrt{n}\epsilon_\mu/\sigma_{YD}$, the probability of making type II error defined as β is estimated as:

$$\beta = 1 - P(|t| > t_{1-\frac{\alpha}{2},n-1}|\delta) \tag{3}$$

3.2 Bayesian Hyphotesis Test

Bayes theorem reveals the relation between two conditional probabilities, e.g., the probability occurring of an event A given the occurrence of an event C (known as $Pr(A \mid C)$), and the probability of occurring the event C given the occurrence of event A (known as $\mathrm{Pr}(\mathrm{C} \mid \mathrm{A})$) (Haldar, A. Mahadevan,S. 2000). Then:

$$Pr(C|A) = \frac{Pr(A|C)Pr(C)}{Pr(A)} \tag{4}$$

Assume event A as the observation of a single validation data point y_D and event C is the hypothesis H_0 being true. Applying the Bayes theorem, the ratio between the posterior probability of the two hypotheses considering the validation data y_D can be given by:

$$\frac{Pr(H_0|y_D)}{Pr(H_1|y_D)} = \left[\frac{Pr(y_D|H_0)}{Pr(y_D|H_1)}\right] * \frac{Pr(H_0)}{Pr(H_1)} \tag{5}$$

where $Pr(H_0)$ and Pr (H_1) are the prior probabilities of the hypotheses, representing the knowledge on the validity of H_0 and H_1 before gathering experimental data. The term $P(H_i|y_D), i = 0\ or\ 1$ is the posterior probability of H_0 and H_1, representing the updated knowledge given validation data point y_D. The ratio $Pr(y_d|H_0)/Pr(yd|H_1)$ is called as Bayes Factor, and its used as a validation metric (Ling & Mahadevan, 2013).

Two formulations for the Bayesian hypothesis test were proposed by Rebba e Mahadevan (2006,2008), the equality-based formulation ($H_0: y_m = y_D, H_1: y_m \neq y_D$) and the interval-based formulation (H_0: $|y_m - y_D| < \epsilon$ and H_1: $|y_m - y_D| > \epsilon$), where the model prediction for a particular $\boldsymbol{x}$ is y_m. Consider the case when model prediction Y_m and the quantity to be predicted Y are both random variables. Two null hypotheses can be formulated: 1) hypothesis that the difference between the means and between the standard deviation of Y_m and Y are within a desired range 2) hypothesis that the PDF of Y_m is equal to the PDF of Y. The second null hypothesis does not require an interval the first one and results in a direct test on probability distributions instead of distribution parameters. In the case that Y_m and Y are deterministic, the first null hypothesis can be applied by setting the standard deviation to zero on both quantities; however, the second hypothesis only applies to the case that both quantities are stochastic.

The interval hypothesis can be stated as:

$$\begin{aligned} &H_0: \epsilon_{\mu 1} \leq \mu_m - \mu \leq \epsilon_{\mu 2}, \qquad \epsilon_{\sigma 1} \leq \sigma_m - \sigma \leq \epsilon_{\sigma 2} \\ &H_1: \mu_m - \mu > \epsilon_{\mu 2}\ ou\ \mu_m - \mu < \epsilon_1, \qquad \sigma_m - \sigma > \epsilon_{\sigma 2}\ ou\ \ \sigma_m - \sigma < \epsilon_1 \end{aligned} \tag{6}$$

Where μ_m and μ are the model Y_m and variable mean Y, respectively, σ_m and σ are the standard deviation of Y_m and Y. Under the interval hypothesis H_0, μ can assume any value between $[\mu_m - \epsilon_{\mu 2}, \mu_m - \epsilon_{\mu 1}]$, thus, is intuitive to consider $\mu \sim Unif(\mu_m - \epsilon_{\mu 2}, \mu_m - \epsilon_{\mu 1})$, then the PDF $\pi_0(\mu|\mu_m) = 1/(\epsilon_{\mu 2} - \epsilon_{\mu 1})$. Likewise, $\sim Unif(\sigma_m - \epsilon_{\sigma 2}, \sigma_m - \epsilon_{\sigma 1})$ and $\pi_0(\sigma|\sigma_m) = 1/(\epsilon_{\sigma 2} - \epsilon_{\sigma 1})$.

$$\pi_0(y|\mu_m,\sigma_m) = \iint \pi(y|\mu,\sigma)\pi_0(\mu|\mu_m)\pi_0(\sigma|\sigma_m)d\mu\, d\sigma = \frac{1}{(\epsilon_{\mu2}-\epsilon_{\mu1})(\epsilon_{\sigma2}-\epsilon_{\sigma1})}\int_{\sigma_m-\epsilon_{\sigma2}}^{\sigma_m-\epsilon_{\sigma1}}\left\{\int_{\mu_m-\epsilon_{\mu2}}^{\mu_m-\epsilon_{\mu1}}\pi(y|\mu,\sigma)\, d\mu\right\}d\sigma \quad (7)$$

Considering measurement errors, experimental observation is a random variable with conditional probability $P(y_D|y)$. Therefore, the likelihood function under the null hypothesis H_0 is given by:

$$P(y_D|H_0) = \int P(y_D|y)\pi_0(y|\mu_m,\sigma_m)\, dy \quad (8)$$

Under the alternative hypothesis, μ can assume any value outside the interval $[\mu_m-\epsilon_{\mu2},\mu_m-\epsilon_{\mu1}]$, but the uniform distribution is not applicable for infinite spaces in practical cases. To avoid this problem, a finite interval $[\mu_l,\mu_u]$ can be assumed based on the underlying physics. Thus, $\mu\sim Unif(\mu_l,\mu_m-\epsilon_{\mu2})\cup(\mu_m-\epsilon_{\mu1},\mu_u)$ and its PDF $\pi_1(\mu|\mu_m) = 1/(\mu_u-\mu_l+\epsilon_{\mu2}-\epsilon_{\mu1})$. Likewise, $\sigma\sim Unif(\sigma_l,\sigma_m-\epsilon_{\sigma2})\cup(\sigma_m-\epsilon_{\sigma1},\sigma_u)$, and the PDF for the alternative hypothesis $\pi_1(\sigma|\sigma_m) = 1/(\sigma_u-\sigma_l+\epsilon_{\sigma2}-\epsilon_{\sigma1})$, then:

$$\pi_1(y|\mu_m,\sigma_m) = \iint \pi(y|\mu,\sigma)\pi_1(\mu|\mu_m)\pi_1(\sigma|\sigma_m)d\mu\, d\sigma = \frac{A}{(\mu_u-\mu_l+\epsilon_{\mu2}-\epsilon_{\mu1})(\sigma_u-\sigma_l+\epsilon_{\sigma2}-\epsilon_{\sigma1})} \quad (9)$$

Where A is given by:

$$A = \int_{\sigma_l}^{\sigma_m-\epsilon_{\sigma2}}\left\{\int_{\mu_l}^{\mu_m-\epsilon_{\mu2}}\pi(y|\mu,\sigma)\, d\mu + \int_{\mu_m-\epsilon_{\mu1}}^{\mu_u}\pi(y|\mu,\sigma)\, d\mu\right\}d\sigma + \int_{\sigma_m-\epsilon_{\sigma1}}^{\sigma_u}\left\{\int_{\mu_l}^{\mu_m-\epsilon_{\mu2}}\pi(y|\mu,\sigma)\, d\mu + \int_{\mu_m-\epsilon_{\mu1}}^{\mu_u}\pi(y|\mu,\sigma)\, d\mu\right\}d\sigma \quad (10)$$

The likelihood function under the alternative hypothesis can be obtain:

$$P(y_D|H_1) = \int P(y_D|y)\pi_1(y|\mu_m,\sigma_m)\, dy \quad (11)$$

and the Bayes Factor can be calculated by dividing $P(y_D|H_0)$ by $P(y_D|H_1)$. The Bayes Factor can be used to accept a model if $BF < 1$ or reject a model if $BF > 1$.

In the case that Y_m and Y are deterministic, it is straightforward to apply this method. With Y_m being deterministic, let σ_m zero, and the rest of the equations remains the same. For the case that Y is deterministic, the interval is assumed only in μ and μ_m, it is known that σ is zero and the other steps are the same.

A directional bias can be found applying two Bayesian hypothesis test. In the first, let $\epsilon_{\mu1} = -\epsilon_\mu$ and $\epsilon_{\mu2} = 0$, and then under the null hypothesis $-\epsilon_\mu = -\mu_m - \mu \leq 0$. In the second test, let $\epsilon_{\mu1} = 0$ e $\epsilon_{\mu2} = \epsilon_\mu$, and then under the null hypothesis $0 = -\mu_m - \mu \leq \epsilon_\mu$. The model is rejected if any of these hypothesis fail in its respective test.

4. MODEL DESCRIPTION

The model validation methods are applied to a prediction model of the first natural frequency of an undamped rotor in free vibration. The rotor is composed by shaft, rotor and bearing trunnion. In the free vibration case, and natural frequency is treated as a random variable, since it is influenced by Young's module and mass variations. Young's module, or modulus of elasticity, is the mechanical property that measures the stiffness of a solid material, in this case, acting directly in the stiffness of the beam. The finite element method (FEM) was used to construct the model and beam's behavior, both considering a rigid beam (Euler-Bernoulli's beam element) and a flexible beam (Timoshenko's beam element. The model was designed to assess a test bench, in which was applied a free vibration experiment and extracting the validation data. The model was described in a 10 element FEM model staggered shaft rotor, as shown in Figure 2. It contains 8 are from a solid shaft and two are the elements of the disc. The staggering is used to adjust the natural frequencies of the system.

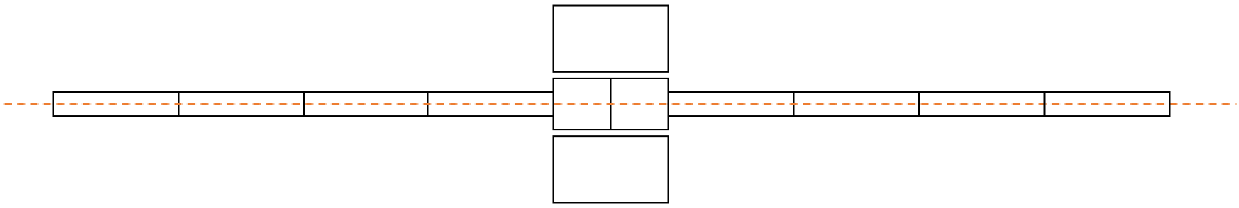

Figure 2. Finite element method model of the rotor.

Nelson (1980) stated that the system equation for a monorotor is as follows:

$$\boldsymbol{M}\ddot{\delta} + \boldsymbol{C}(\Omega)\dot{\delta} + \boldsymbol{K}\delta = \boldsymbol{F}(t) \tag{12}$$

Where $\boldsymbol{M}$ is the symmetric mass matrix, $\boldsymbol{C}$ is the asymmetric matrix, including and antisymmetric gyroscopic matrix function of Ω, and a frequently asymmetric matrix representing the characteristics of the bearing, $\boldsymbol{K}$ is the stiffness matrix, $\boldsymbol{F}(t)$ is the force vector and δ is the vector containing all the nodal displacements.

The natural frequencies and the vibration modes of the rotor are derived from the homogeneous equations of the undamped system (Pereira, 2005):

$$\boldsymbol{M}\ddot{\delta} + \boldsymbol{K}\delta = \boldsymbol{0} \tag{13}$$

The solution of the equation is:

$$\delta = \phi e^{j\omega t} \tag{14}$$

Where ϕ is the amplitude and ω is the natural frequency. Applying eq 14 in eq 13, we obtain a typical eigenvalue-eigenvector problem:

$$(-\omega^2\boldsymbol{M} + \boldsymbol{K})\phi = 0 \tag{15}$$

Hence, as the trivial solution is not of interest, the matrix determinant must be zero.

$$det(-\omega^2\boldsymbol{M} + \boldsymbol{K}) = 0 \tag{16}$$

The histogram of each model natural frequency is presented by Figure 3 and Figure 4 and the convergence by Table 1 and Table 2.

Table 1. ***Timoshenko convergence.***

Model	Sample Size	Mean (Hz)	Standard Deviation (Hz)
5%	500	60.45	1,87
	1000	60.52	1.89
	1500	60.55	1.89
	2000	60.59	1.90
1%	500	60.47	0.37
	1000	60.47	0.37
	1500	60.48	0.37
	2000	60.48	0.38
0.5%	500	60.49	0.19
	1000	60.48	0.18
	1500	60.48	0.18
	2000	60.48	0.18

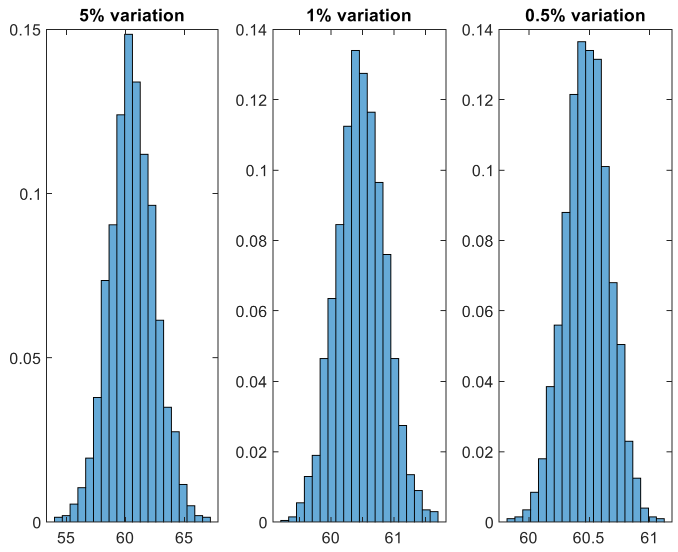

Figure 3. Histogram of Timoshenko's model.

Table 2. ***Euler-Bernoulli convergence***

Model	Sample Size	Mean (Hz)	Standard Deviation (Hz)
5%	500	60.54	1.97
	1000	60.52	1.97
	1500	60.55	2.02
	2000	60.52	2.0
1%	500	60.44	0.4
	1000	60.43	0.4
	1500	60.42	0.39
	2000	60.42	0.39
0.5%	500	60.45	0.19
	1000	60.44	0.2
	1500	60.44	0.2
	2000	60.44	0.19

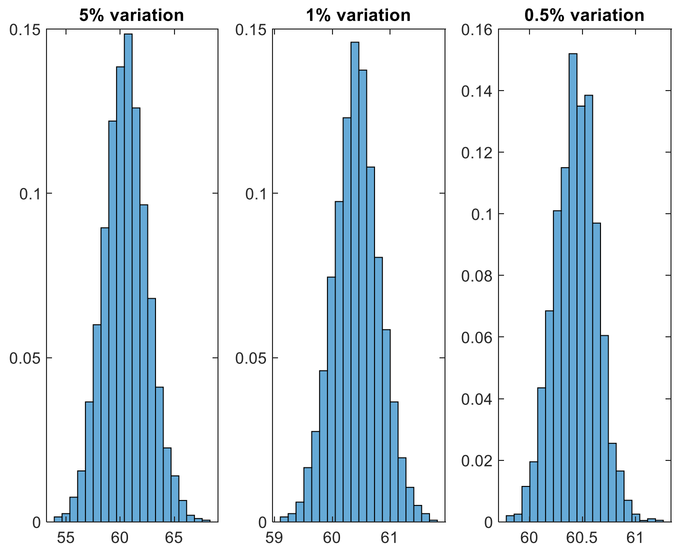

Figure 4. Histogram of Euler-Bernoulli's model.

5. EXPERIMENTAL DATA VALIDATION

To gather the validation data for the finite element method model of the rotor, an experimental test for the rotor was executed to determine the modal parameters, the excitation was caused by a shaker, Figure 5 represents the experiment test bench;

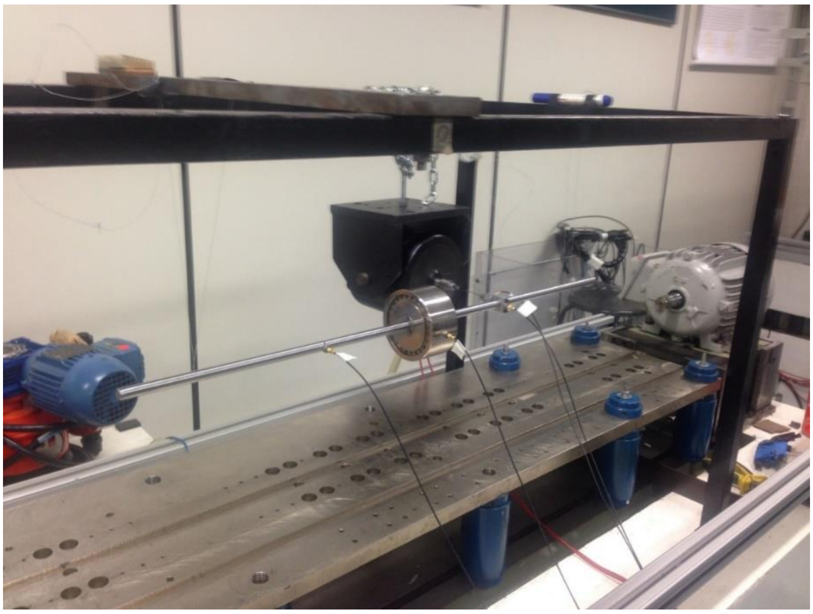

Figure 5. Test bench of the experiment.

The masses of the system were obtained by a precision balance. Three small accelerometers were used, choosing a weight that would not interfere in the results of the modal parameters. Table 3 shows the results obtained in the precision balance.

Table 3. Rotor component masses.

Component	Mass [g]
Disc	2345,80
Shaft	705,90
Accelerometer	2,80

The experimental test resulted in a frequency response function (FRF), Figure 6, in which one can recognize the natural frequencies by the simultaneous peaks in both accelerometers, thus, the first three natural frequencies of the system are around 60Hz, 180Hz and 360Hz.

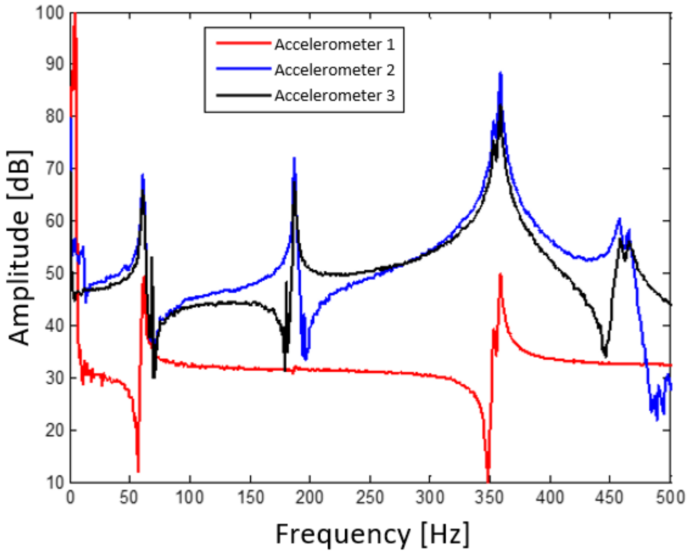

Figure 6. FRF of experiment**.**

This test managed to gather 104 observations, which were compared to the model prediction. The sample generated by the FEM model contains 2000 observations obtained by a Monte Carlo Sampling was used to compute this sample. To validate the experiment data, three cases were considered in the Monte Carlo: 1) coefficient of variation equal to 5% in the Young's module and density of the material. 2) coefficient of variation of 1% variation in both variables and, 3) finally, coefficient of variation equal to 0.5% for both variables. Figure 7 compares the distribution of the three cases with the experimental data.

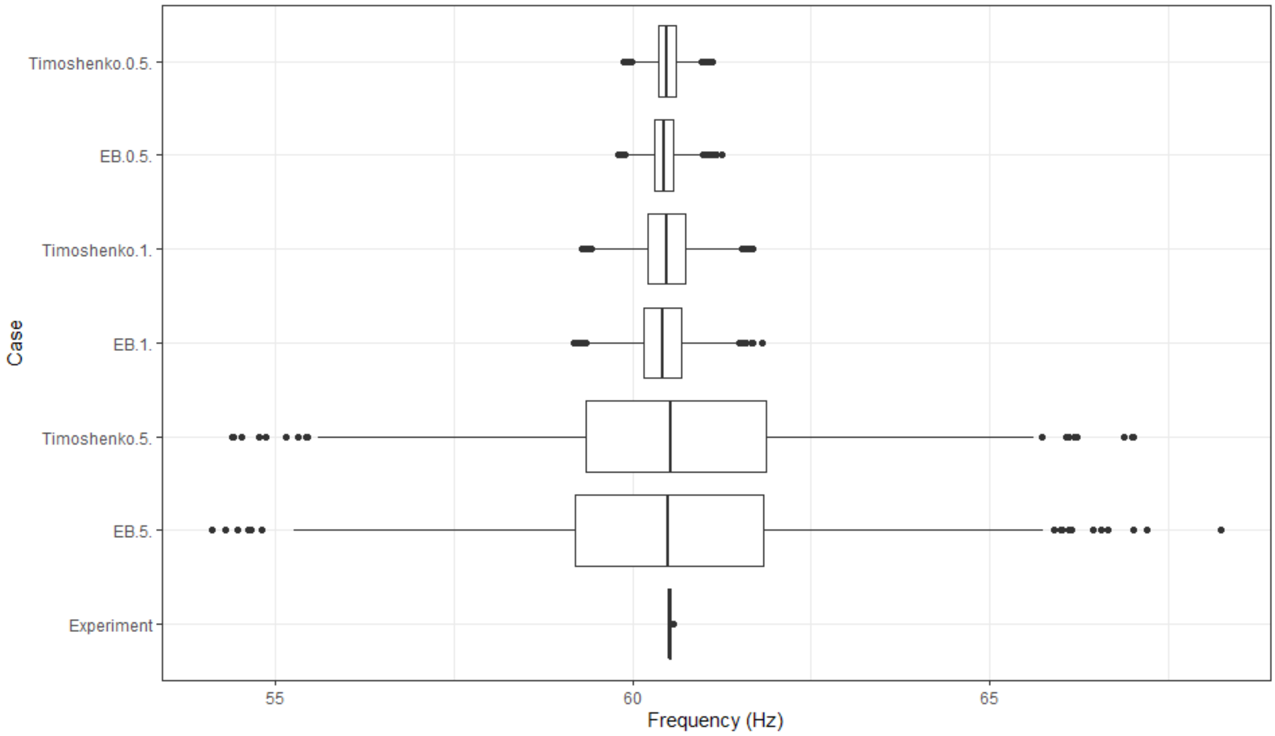

Figure 7. Comparison of experiment and model distributions.

5.1. Validation With Classical Hypothesis Test

As the sample size obtained by the FEM model was large enough, a t-test was applied for each of the cases of study. The significance level $\alpha = 0.05$ was chosen to represent the accuracy of the model, and the null hypothesis $H_0: \mu_m = \mu_D$ was tested. The model is considered to fail when the correspondent p-value is less than the level of significance α. The results are shown in the Table 4:

Table 4. Model validation for Classical Hypothesis Test

Beam Element	Variation	Confidence Interval (Hz)	p-value	Model mean (Hz)	Stardard Error of the Mean
Euler-Bernoulli	5%	-0.38 to 0.38	0.98	60.52	0.2
	1%	-0.18 to -0.02	0.01	60.42	0.04
	0.5%	-0.12 to -0.05	6.e-06	60.43	0.02
Timoshenko	5%	-0.3 to 0.42	0.74	60.58	0.18
	1%	-0.12 to 0.02	0.19	60.47	0.04
	0.5%	-0.07 to -0.01	0.02	60.48	0.02

By definition, the Classic Hypothesis Test fails to reject the both models considering 5% coefficient of variation, rejecting the other.

5.2 Validation With Bayesian Hypothesis Test

For the two sample Bayesian test, a significance level was set as $\alpha = 0.05$. In this approach, the null hypothesis was set the difference between the mean was lower than an estimated error, as said $H_0: \mu_m - \mu_D < \varepsilon$, the error ε was defined as a decimate in the posterior model mean, $\varepsilon = 0.1$. The results for all the models are shown in the Table 5:

Table 5. Model validation for Bayesian Hypothesis Test

Beam Element	Variation	Confidence Interval (Hz)	Posterior Model Mean	Error of the mean (Hz)	Bayes Factor
Euler-Bernoulli	5%	-0.19 to 0.01	60.52	0.0009	0.11
	1%	-0.18 to -0.03	60.42	-0.1	3.6
	0.5%	-0.12 to -0.05	60.43	-0.08	2190
Timoshenko	5%	-0.12 to -0.04	60.58	0.06	0.117
	1%	-0.12 to 0.02	60.47	-0.05	0.252
	0.5%	-0.07 to -0.01	60.48	-0.5	1.49

In the Bayesian Hypothesis Test case, the model is rejected if Bayes Factor >1. Thus, it failed to reject the Euler-Bernoulli 5% variation and the Timoshenko models with 5% and 1% coefficient of variation, however, as the Euler-Bernoulli 1% variation model and Timoshenko 0.5% variation has presented a Bayes Factor of approximately 1, it still is considered statistically relevant, thus, with more evidence available it could fail to reject these models.

6. CONCLUSION

Verification and Validation of computational models plays a pivotal role in instilling confidence in the accuracy of model results and their representation of real-world problems. In this paper, we focused on the model validation of a rotor experiment, leveraging 104 data points as our basis for assessment, exploring two distinct validation methodologies: classical hypothesis testing and Bayesian hypothesis testing.

The unique advantage that Bayesian hypothesis testing offers is considering the entire PDF of the formulated model prediction, while classical testing focused solely on means of the models. Through these validation approaches, we have arrived at some findings regarding the agreement between the models and the experimental data.

From the perspective of classical hypothesis testing, it is evident that the Timoshenko models with 5% and 1% coefficient of variation in the Young's module, as well as the Euler-Bernoulli model with a 5% coefficient of variation, exhibit a commendable level of agreement with the experimental data.

Furthermore, the Timoshenko model if 0.5% of coefficient of variation in the Young's module and the Euler-Bernoulli model with a 1% coefficient of variation were found to be statistically relevant. These models, though not conclusively accepted by the current hypothesis testis criteria, still maintain Bayes Factor close to 1. This suggests that, with the accumulation of more data and evidence, exists the potential for these models to pass the threshold of statistical relevance.

In sum, this study underscores the importance of employing diverse validation methods and considering the nuanced insight they offer. The Bayesian approach showcases the potential of models that may not initially meet classical hypothesis testing criteria to be deemed relevant with the accumulation of additional information. It is through this approach to validation that we can foster a deeper understanding of the accuracy and applicability of computational models to real-world problem-solving.

7. ACKNOWLEDMENTS

The authors gratefully acknowledge the support by CAPES, Coordination for the Improvement of Higher Education Personnel for the financial support that made this work possible.

8. REFERENCES

American Society of Mechanical Engineers (2012) 'An Illustration of Concepts of Verification and Validation in Computational Solid Mechanics', New York

Haldar A, Mahadevan S. (2000) 'Probability, Reliability, and Statistical Methods in Engineering Design'. *Wiley*.

Jiang, X., Mahadevan, S. 'Bayesian risk-based decision method for model validation under uncertainty' *Reliability Engineering & System Safety*. Elsevier Ltd, Vol. 92, pp 707-718, doi.org/10.1016/j.ress.2006.03.006

Ling, Y., Mahadevan, S. (2013) 'Quantitative model validation techniques: New insights', *Reliability Engineering & System Safety*. Elsevier Ltd, Vol. 111, pp 217-231, doi.org/10.1016/j.ress.2012.11.011.

Montgomery, D. C. & Runger, G.C. (2003) 'Estatística Aplicada e Probabilidade para Engenheiros' LTC.

Nelson, H. D. (1980). "A Finite Rotating Shaft Element Using Timoshenko Beam Theory." ASME. *J. Mech. Des*. October 1980; Vol. 102, pp. 793–803. https://doi.org/10.1115/1.3254824

Pereira, J.C. (2005) 'Introdução à Dinâmica de Rotores', Florianópolis

Rebba, R., Mahadevan, S. & Huang, S. (2006) 'Validation and Error Estimation of Computational Models" *Reliability Engineering & System Safety*. Elsevier Ltd, Vol. 91, pp 1390-1397. doi:10.1016/j.ress.2005.11.035

Roache, P.J., (1998) 'Verification And Validation in Computational Science and Engineering'.: *Hermosa.*

Rebba R., Mahadevan, S. (2008) 'Computational Methods for Model Reliability Assessment' *Reliability Engineering & System Safety*. Elsevier Ltd, Vol. 93, pp 1197-1207, doi.org/10.1016/j.ress.2007.08.001

Sankararaman, S. & Mahadevan, S. (2011) 'Likelihood-Based Representation of Epistemic Uncertainty due to Sparse Point Data and/or Interval Data.' *Reliability Engineering & System Safety*. Elsevier Ltd, Vol. 97, pp 814-824