Numerical model validation using experimental data: Application of the area metric on a Francis runner

Q Chatenet¹, A Tahan¹, M Gagnon² and J Chamberland-Lauzon³

¹ Mechanical Engineering Department, École de Technologie Supérieure (ÉTS), 1100 Rue Notre-Dame O, Montréal, QC, Canada H3C1K3
² Institut de Recherche d’Hydro-Québec (IREQ), 1800 boul. Lionel-Boulet, Varennes, QC, Canada J3X1S1
³ Andritz Hydro Canada Inc., 6100 aut. Transcanadienne, Pointe Claire, QC, Canada H9R1B9

E-mail: quentin.chatenet@gadz.org

Abstract. Nowadays, engineers are able to solve complex equations thanks to the increase of computing capacity. Thus, finite elements software is widely used, especially in the field of mechanics to predict part behavior such as strain, stress and natural frequency. However, it can be difficult to determine how a model might be right or wrong, or whether a model is better than another one. Nevertheless, during the design phase, it is very important to estimate how the hydroelectric turbine blades will behave according to the stress to which it is subjected. Indeed, the static and dynamic stress levels will influence the blade’s fatigue resistance and thus its lifetime, which is a significant feature. In the industry, engineers generally use either graphic representation, hypothesis tests such as the Student test, or linear regressions in order to compare experimental to estimated data from the numerical model. Due to the variability in personal interpretation (reproducibility), graphical validation is not considered objective. For an objective assessment, it is essential to use a robust validation metric to measure the conformity of predictions against data. We propose to use the area metric in the case of a turbine blade that meets the key points of the ASME Standards and produces a quantitative measure of agreement between simulations and empirical data. This validation metric excludes any belief and criterion of accepting a model which increases robustness. The present work is aimed at applying a validation method, according to ASME V&V 10 recommendations. Firstly, the area metric is applied on the case of a real Francis runner whose geometry and boundaries conditions are complex. Secondly, the area metric will be compared to classical regression methods to evaluate the performance of the method. Finally, we will discuss the use of the area metric as a tool to correct simulations.

1. Introduction

Given the complexity of the needs required to perform the validation process, the ASME V&V 10 committee [1] prefers not to recommend specific metrics to evaluate the accuracy or the goodness-to-fit of the model, but rather suggest to specify which metric is used, and what is the chosen acceptance level during the validation plan design. This work agrees within the scope of the Verification & Validation process as shown in figure 1. Nevertheless, we decided to focus on the validation part of the process because most of the difficulties occur during this phase (the validation phase is shown in...
the grey rectangle in Fig. 1). Besides, it is often difficult to go through the verification process as most software used are not open source; thus, only the developer has access to the code.

Specifically, in the hydropower field, it is important for manufacturers or plant operators to estimate the life expectancy of their hydraulic power plants and plan when to perform maintenance including inspections. In particular, for hydroelectric turbine runners, they need to rely on estimation of the stress level at the blade’s hot spots. One of the challenges facing today’s design engineers is determining the suitable mechanical models (good enough for the intended purpose) compared to experimental data, and be able to choose the best one. To ensure this, engineers have to use an adequate metric to quantify the discrepancy between the predicted results and actual data. This metric should take into account the uncertainties in both: the simulation outcomes and the experimental outcomes [1]. The metric used to perform model validation should also satisfy these conditions according to ASME V&V 10 recommendations [1]: “A validation metric provides a method by which the simulation outcomes and the experimental outcomes can be compared. A metric is a mathematical
measure of the difference between the two outcomes such that the measure is zero only if the two outcomes are identical.” The proposed validation metric, in this paper, is the area metric developed by Ferson et al. [2] to integrate these recommendations and excludes any belief and acceptance criterion contrary to classical hypothesis tests [3].

This paper is structured as follows. First, a review of the area metric and its characteristics is presented in Sec 2. Next, in Sec. 3 we apply the method on a real case of a hydraulic turbine runner blade in order to determine which model is more accurate. Finally, a discussion on the results of the study case and a comparison with classical methods concludes the paper.

2. Area metric

The area metric displays several advantages compared to other validation methods (i.e.: hypothesis tests, Bayes factor or frequentist’s metric) [3]. One of these is objectiveness, since the conclusion given by the area metric does not depend on how an engineer conducting the analysis, or on the assumptions he makes.

The area metric allows engineers to conduct the validation process based on a quantitative measure of the discrepancy between predictions and data. At the same time, it gives a graphical representation (in the physical units used) of this discrepancy. Moreover, this metric allows one to evaluate the differences across the full range of prediction distribution while taking into account simulation uncertainties.

In order to apply the area metric method, we need to compute predictions which can be represented as a cumulative distribution function \( F(x) \) (CDF), where \( x \) is the variable of the prediction. Observations are represented by a non-decreasing step function with a constant vertical step of \( 1/n \) (\( n \) represents the size of the data set). The \( x \)-axis value of the steps corresponds to the data points. Using equation (1), we can construct this function for data \( x_i \).

\[
S_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(x_i, x), \quad \text{where} \quad I(x_i, x) = \begin{cases} 
1 & x_i \leq x, \\
0 & x_i > x,
\end{cases}
\]

The measure of the mismatch between prediction and observation is thus the area between the prediction function \( F \) and the data distribution \( S_n \) and can be evaluated using equation (2).

\[
d(F, S_n) = \int_{-\infty}^{\infty} |F(x) - S_n(x)| \, dx
\]

It can be noted that if the functions \( F \) and \( S_n \) are identical, then the area metric will be equal to zero, which satisfies one of the ASME V&V recommendations. The figure 2 illustrates the mismatch between the prediction distribution (shown as the gray curve) and the experimental data (shown as the dark step function with the metric being the grey shaded area). The latter \( d(F, S_n) \) is computed using equation (2) on the entire range on both the prediction uncertainty distribution and empirical data. In this example, a normal distribution has been chosen to describe numerical model. Insofar as the \( x \)-axis unit is expressed in MPa and the \( y \)-axis is unit less, the area is expressed in MPa too, which is the unit of interest for stress levels. Moreover, as the metric is not normalized, the engineer is able to evaluate the impact of different values of the area metric according to the level of stress he is dealing with. As an example, if the result of the area metric is 5 MPa for a set of measurements whose mean is about 10 MPa, the engineer will not draw the same conclusions as if the mean value was 100 MPa.
As the area metric is applied on the entire uncertainty distribution of the predictions and the data, the metric is particularly sensitive, while at the same time remaining robust. Indeed, figure 3 illustrates examples with two sets of data, both of which have strictly identical means and standard deviations (i.e.: $\mu_1 = \mu_2 = 180$ MPa and $\sigma_1 = \sigma_2 = 40$ MPa). The grey curve represents same normal distribution on both figures. The results of the metric are respectively 25.14 MPa and 28.88 MPa for those two sets of data. The area metric allows one to get more information from the validation analysis, whereas dealing only with the first two statistical parameters would yield no noticeable difference.

**Figure 2.** Example data set ($n = 6$)

**Figure 3.** Comparison of two different data sets with $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. 

*empirical distribution function*  
*prediction distribution*  

**Area metric = 27.04 MPa**

**Area metric = 28.88 MPa**

**Area metric = 25.14 MPa**
3. Case study

The area metric method has been applied to identify the best mechanical model against experimental data. For this purpose, two sets of finite elements analysis (FEA) have been computed according to two different mechanical models on a real and running Francis hydraulic turbine runner (medium head and nominal power >100MW).

3.1. Probability distribution of simulation uncertainty

In order to perform the validation process, it is better to focus on probability distribution rather than on a single predicted value. Indeed, a distribution better reflects the variability of the model outcomes insofar as uncertainties can be integrated during computation (such as epistemic uncertainties).

In his work, J. Arpin-Pont [4] developed a method of accounting position and orientation uncertainties, integration effect, and uncertainties due to the welded gauge technology. Actually, it is impossible to accurately position a welded gauge at a target location as shown in figure 4, which results in position uncertainties. The method consists in simulating virtual strain gauges located around the finite elements target node, for numerous different positions and orientations. The purpose of this is to reflect the replicability of the measurement using random distribution model. Thus, with a Monte Carlo simulation, a distribution of the strain measured by the virtual gauge is generated and can be compared to the empirical values.

![Figure 4. Definition of uncertainty sources: welded gauge behavior, location uncertainty (including positioning and alignment errors) and integration effect [4].](image4)

3.2. Numerical models

Firstly, computational fluid dynamics (CFD) analysis at maximum power were performed using two different CFD settings giving two pressure-loading cases for this turbine. Thus, in terms of FEA software, different pressure fields were imposed to the blade depending on the model used.

![Figure 5. Position of the gauges placed on the hydraulic turbine blade during measurements.](image5)
According to previous work done by J. Arpin Pont [5], we were able to evaluate position and orientation errors for each gauge location on the blade. Notice that the same Francis runner is used in this study. The blade is instrumented evenly on both sides, position and orientation uncertainties are the same for sites 1 and 3 and for sites 2 and 4 respectively. The parameters used to perform the simulation are presented in table 1. The simulation distribution obtained for FEA models #1 and #2 are obtained from strain gauge position uncertainty on FEA strain results.

| Location     | Uncertainties (95% confidence level) |
|--------------|--------------------------------------|
|              | Position along target axis $\bar{X}$ (mm) | Position along target axis $\bar{Y}$ (mm) | Alignment angle (°) |
| Sites 1 - 3  | [-8; 8]                               | [-2; 5]                               | [-2; 2]                |
| Sites 2 - 4  | [-12; 6]                              | [-9; 3]                               | [-6; 6]                |
| Site 5       | [-14; 14]                             | [-3; 6]                               | [-7; 7]                |

3.3. Experimental process

According to FEA results, strain gauges have been welded at specific locations (as shown in figure 4) on each of the two blades along the principal strain direction given by an initial FEA analysis. One of the assumptions is the strain measured by the gauges is along the principal direction, which cannot be verified since unidirectional welded gauges were used. The use of strain rosette gauges would enable the direction of the strain to be obtained and improve the assessment of the maximum stress at hot spots.

Measured strains were established from at least two independent measurements made at similar operating conditions. At each strain gauge site, the empirical distribution function then include measurements at both instrumented blade and at several independent measurements in time.

4. Results

Area metric results comparing simulation distribution and measurements are shown in table 2, with detailed graph for site 3 (see figure 6).

| Location     | Model #1 (MPa) | Model #2 (MPa) |
|--------------|----------------|----------------|
| Site 1 – Crown outflow SS\(^a\) | 37.88          | 255.4          |
| Site 2 – Band outflow SS\(^a\)  | 10.53          | 10.72          |
| Site 3 – Crown outflow PS\(^b\) | 21.69          | 26.06          |
| Site 4 – Band outflow PS\(^b\)  | 51.13          | 46.97          |
| Site 5 – Crown center PS\(^b\)  | 48.89          | 158.1          |
| **Sum**      | **170.12**     | **497.25**     |

\(^a\) Suction Side
\(^b\) Pressure Side

As we can see in figure 6, the area metric method provides both a numerical result of the mismatch between prediction and data and a graphical representation of the result. Because of the smaller area metric value, we can conclude in this case that the model #1 predictions are closer to experimental data than model #2.
The dashed line in figure 6 represents the nominal computed stress value at the node under the strain gauge.

4.1. Evolution of the area metric as a function of various operating conditions.

The CFD simulation was done for the maximum nominal power of the turbine, which corresponds to a wicket gates opening of 100%. However, measures were performed for several openings (12% to 100%) in order to study transient states in particular. Then, it is possible to compare area metric results depending on the turbine operating condition, as shown in table 3. The aim is to perform validation process while looking for operating conditions which better match with simulation results. As expected, FEA performed with maximum power CFD pressures better predicts measurements at 100% wicket gate opening. The area metric increases significantly as wicket gate opening decreases which suggests that FEA performed with maximum power CFD pressures does not represent low and part load measured strains.

Table 3. Evolution of the area metric as a function of wicket-gate opening.

| Location                        | Wicket gate opening |
|---------------------------------|---------------------|
| Site 1 – Crown outflow SS\(^a\) | 170.80              |
| Site 2 – Band outflow SS\(^a\) | 44.39               |
| Site 3 – Crown outflow PS\(^b\) | 56.68               |
| Site 4 – Band outflow PS\(^b\) | 174.80              |
| Site 5 – Crown center PS\(^b\) | 187.50              |
| **Sum**                         | **634.17**          |

4.2. Comparison with classical methods

The least squares linear regression is one of the most-used methods in model validation because it is easy to compute and to understand. Yet this method has severe limitations. The result of least squares regression is improved by increasing sample size and does not take into account uncertainties on both
axes, unless one uses weighted least squares [6, 7] for example. The weighted least squares estimates ($\hat{\beta}_0$ and $\hat{\beta}_1$) are then given by equations (4) where $\bar{x}_w$ and $\bar{y}_w$ are the weighted means.

\begin{equation}
\begin{aligned}
\sum w_i (x_i - \bar{x}_w)(y_i - \bar{y}_w),
\end{aligned}
\end{equation}

For each point, weights are assigned as described by equation (4). Uncertainties on each axis $\Delta_x_i$ and $\Delta_y_i$ are determined by applying recommendations [8], modeled by equation (5).

\begin{equation}
\Delta x_i = \frac{\text{range}(x_i) - 1}{2\sqrt{3}}
\end{equation}

Figure 7. Comparison of computed stress and experimental stress (from strain measures) at 75% (left) and 100% (right) opening.

Points represent the mean of sample data and probability distribution; bars associated with points correspond to uncertainties according to each axis. In figure 7, the dashed line represents the simple linear regression, which does not take into account uncertainties, unlike the weighted least squares with a solid line. The weighted linear regression curve is given by: $y = \hat{\beta}_1 x + \hat{\beta}_0$.

5. Discussion
First, with respect to the model comparison, we observe that the area metric result is lower for model #1, which means this model better reflects empirical observations. However, this is not the case for one of the five sites with an area metric result lower for model #2. It could be explained by the fact that we have only four (4) observations for this specific site whereas we have eight (8) observations for others, resulting in a lack of information. Note that compiling area metric results for all sites, considering the mean, model #2 is about three times higher in terms of area metric than model #1 (34.02 MPa for model #1, 99.45 MPa for model #2).
A second observation can be made from table 3. Indeed, for sites 2, 4 and 5, the further away from a nominal operating point, the more the area metric increases, which is consistent since simulations have been made for maximum wicket gates opening. Conversely, for sites 1 and 3, which are located symmetrically on opposite sides, the area metric value is the lowest for a wicket opening of 75% and 40%, respectively. This result can be used to correct FEA simulations, considering values evolve proportionally as a function of wicket-gate opening. Thus, applying this correction for site 1 predictions, the area metric result drops to 22.50 MPa instead of 37.88 MPa (for a wicket gates opening of 100%).

As shown with linear regressions in figure 7, often used in engineering to compare models, it is rather difficult to evaluate if the model is right or wrong according to regressions parameters. Indeed, even if the slope is close to one (1), we cannot conclude on the model accuracy insofar as points can be symmetrically distributed around the linear regression curve.

Also, the real turbine geometry is not taken into account as an uncertainty source. However, FEA is based on design geometry while strain measures are from the real runner whose geometry varies in the tolerance interval because of manufacturing constraints. This variability added to gauge position uncertainties, might explain the mismatch between model predictions and experimental data in some cases.

6. Conclusion
In this paper, we presented an application of the area metric method and we conducted a validation process according to ASME V&V recommendations using the former metric. This method relies on a calculation area between two probability distribution: one from simulation prediction and the second from experimental data. Thus, according to area metric values, engineer is able to choose which model is the more accurate. When applied to in situ measurements carried out on hydraulic turbine runner blades, area metric results show that model #1 is about three times more accurate (on average) than model #2. However, we observe that in particular locations, model #2 predictions can be closer to experimental values. Likewise, using results for different operating conditions, we show that using the area metric, an extrapolation can be made to correct some prediction simulations.

Finally, one of the most difficult questions to answer concerns the level of acceptance. Indeed, validation plan must predetermine the accuracy requirements to validate a model according to the metric used.

Acknowledgments
This study was carried out with the financial support of the Canadian research internship program, Mitacs-Accelerate. The authors acknowledge Andritz Hydro Canada, the Institut de Recherche d’Hydro-Québec (IREQ) and the École de technologie supérieure (ÉTS) for their support, materials and advice.

Nomenclature

\[ d(F, S_x) \] Integral function

\[ \sigma \] Standard deviation of distribution

\[ S_x(x) \] Step function of empirical data

\[ \Delta x_i \] Uncertainties along x-axe

\[ n \] Sample size

\[ \Delta y_i \] Uncertainties along y-axe

\[ x \] Variable of the prediction

\[ w_i \] Point weight

\[ F(x) \] Cumulative distribution function

\[ \beta_i \] Linear regression slope

\[ \mu \] Mean of distribution

\[ \beta_0 \] Linear regression Y intercept

References
[1] ASME 2006 Guide for Verification and Validation in Computational Solid Mechanics vol PTC 60/V&V 10
[2] Ferson S, Oberkampf W L and Ginzburg L 2008 Model validation and predictive capability for the thermal challenge problem Computer Methods in Applied Mechanics and Engineering \textbf{197} 2408-30

[3] Liu Y, Chen W, Arendt P and Huang H-z 2011 Toward a Better Understanding of Model Validation Metrics Journal of Mechanical Design \textbf{133} 071005

[4] Arpin-Pont J, Gagnon M, Tahan S A, Coutu A and Thibault D 2012 Strain gauge measurement uncertainties on hydraulic turbine runner blade IOP Conference Series: Earth and Environmental Science \textbf{15} 062042

[5] Arpin-Pont J 2012 Méthode de détermination des incertitudes de mesures par jauges de déformation. (Montréal: École de technologie supérieure) pp 1 ressource en ligne (xxiv, 129 p.)

[6] Chatterjee S and Price B 1991 \textit{Regression analysis by example} (New York, N.Y.: J. Wiley and Sons)

[7] Rouaud M 2014 \textit{Calcul d'incertitudes} (Paris, France: Creative Commons)

[8] ENV13005 N 1999 Guide pour l’expression de l’incertitude de mesure. (Paris, France: AFNOR)