

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTÉRATURE	7
1.1 Démarche de <i>Vérification & Validation</i> proposée par l'ASME	7
1.1.1 Terminologie adoptée par l'ASME	7
1.1.2 Aperçu de la norme ASME V&V 10	7
1.1.3 Vérification	10
1.1.4 Validation	11
1.2 Outils de validation	12
1.2.1 Régression linéaire simple (RLS)	13
1.2.2 Régression linéaire pondérée (RLP)	16
1.2.3 Régression linéaire orthogonale (RLO)	17
1.2.4 Tests d'hypothèse	19
1.2.5 Métrique de surface - <i>Area metric</i>	21
CHAPITRE 2 MÉTHODOLOGIE	23
2.1 Instrumentation	23
2.2 Analyse par ÉF	23
2.3 Incertitudes de simulation	24
2.4 Validation	24
2.5 Structure du mémoire	25
CHAPITRE 3 ARTICLE #1 : NUMERICAL MODEL VALIDATION USING EXPERIMENTAL DATA : APPLICATION OF THE AREA METRIC ON A FRANCIS RUNNER	27
3.1 Résumé	27
3.2 Abstract	28
3.3 Introduction	29
3.4 Area metric	31
3.5 Case study	32
3.5.1 Probability distribution of simulation uncertainty	34
3.5.2 Numerical models	34
3.5.3 Experimental process	36
3.6 Results	36
3.6.1 Evolution of the area metric as a function of various operating conditions.	37
3.6.2 Comparison with classical methods	38
3.7 Discussion	39
3.8 Conclusion	40

CHAPITRE 4 ARTICLE #2 : ON THE APPLICATION OF THE AREA METRIC : A CASE STUDY ON HYDROELECTRIC TURBINE 43

4.1 Résumé 43

4.2 Abstract 44

4.3 Introduction 44

4.4 Overview of the area metric 46

4.5 Overview of u-pooling method 48

4.6 Proposed methodology 51

4.7 Study case 53

4.8 Results 54

4.9 Discussions 55

4.10 Conclusions 57

CONCLUSION ET RECOMMANDATIONS 59

BIBLIOGRAPHIE 63

LISTE DES TABLEAUX

	Page
Tableau 3.1	Positioning uncertainty distribution laws used for simulation 35
Tableau 3.2	Area metric results..... 36
Tableau 3.3	Evolution of the area metric as a function of wicket-gate opening..... 38
Tableau 4.1	Area metric results - model #1 54
Tableau 4.2	Area metric results - model #2 55

LISTE DES FIGURES

		Page
Figure 0.1	Solution d'un problème par une approche analytique - Tirée et adaptée de Coleman et Steele (2009)	3
Figure 0.2	Cycle d'élaboration d'une théorie ou d'un modèle.....	4
Figure 1.1	Quartet d'Anscombe	15
Figure 1.2	Comparaison entre régression linéaire simple et pondérée	17
Figure 1.3	Comparaison entre régression linéaire simple et orthogonale	19
Figure 1.4	Représentation de la zone de rejet pour une distribution donnée	20
Figure 2.1	Structure du mémoire	25
Figure 3.1	Verification & Validation activities (ASME V&V 10).	30
Figure 3.2	Example data set ($n = 6$).....	33
Figure 3.3	Comparison of two different data sets with $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$	33
Figure 3.4	Definition of uncertainty sources : welded gauge behavior, location uncertainty (including positioning and alignment errors) and integration effect Arpin-Pont <i>et al.</i> (2012).	34
Figure 3.5	Position of the gauges placed on the hydraulic turbine blade during measurements.	35
Figure 3.6	Site 3 Area metric comparison of model #1 vs. model #2.	37
Figure 3.7	Comparison of computed stress and experimental stress (from strain measures) at 75% (left) and 100% (right) opening.	39
Figure 4.1	Application of the area metric - Example data sets ($n = 4, k = 10$).	45
Figure 4.2	Illustration of <i>u-pooling</i> method and <i>back-transformation</i>	47
Figure 4.3	Influence of the <i>G</i> function choice in back-transformation process.	49
Figure 4.4	Example of u_i calculation. Only one observation is distributed according to its simulation distribution.	50

Figure 4.5	Position of the strain gages placed on the hydraulic turbine blade during measurement.	51
Figure 4.6	Example of the area metric applied on aggregated distribution $H(x)$	52
Figure 4.7	Area metric applied on the entire distribution H for two different models.	55
Figure 4.8	Area metric computed on u -pooling graph	56

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AM	Area Metric
ASME	American Society of Mechanicals Engineers
CFD	Mécanique des fluides numérique (<i>Computational Fluid Dynamics</i>)
CAO	Conception Assistée par Ordinateur
PS	Face intérieure (<i>Pressure Side</i>)
SS	Face extérieure (<i>Suction Side</i>)
SRQ	Métrique utilisée (<i>System Response Quantity</i>)
PDF	Fonction de densité de probabilité (<i>Probability Density Function</i>)
CDF	Fonction de distribution cumulative (<i>Cumulative Density Function</i>)
FEA	Analyse par éléments finis (<i>Finite Elements Analysis</i>)
ÉF	Éléments Finis
V&V	Vérification et Validation
RLS	Régression Linéaire Simple
RLP	Régression Linéaire Pondérée
RLO	Régression Linéaire Orthogonale
MNEIM	Méthode Numérique d'Estimation des Incertitudes de Mesure

LISTE DES SYMBOLES

x	Variable de la simulation
$F(x)$	Fonction de distribution cumulative
$S_n(x)$	Fonction par morceaux des données expérimentales
$d(F, S_n)$	Fonction intégrale - Aire entre les courbes des fonctions F et S_n
n	Taille de l'échantillon
μ	Moyenne de la distribution
σ	Écart type de la distribution
Δx_i	Incertitudes selon l'axe x
Δy_i	Incertitudes selon l'axe y
w_i	Pondération des points
$\widehat{\beta}_1$	Pente de la régression linéaire
$\widehat{\beta}_0$	Ordonnée à l'origine de la régression linéaire
M^{SRQ}	<i>Area metric</i> entre les observation et la <i>CDF</i>
u_i	Données <i>u-values</i> issues de la transformation
H	Fonction de distribution cumulative des données agrégées
y_i	Données issues de la transformation inverse
G	Fonction de distribution de la transformation inverse

INTRODUCTION

D'après les données de 2013, les centrales hydroélectriques représentent 12% de la production électrique mondiale et 72% des sources d'énergie renouvelable (IEA, 2011) (IEA, 2015). Avec les enjeux environnementaux actuels et les engagements pris par les plus grandes nations, la transition énergétique qui s'opère se fera notamment avec l'exploitation des centrales hydroélectriques existantes et le développement de nouveaux projets. Il est donc primordial de maintenir ces installations dans un bon état de fonctionnement.

Comme dans beaucoup de domaines, les phases de conception, fabrication et installation d'une roue de turbine hydroélectrique requièrent la validation des performances obtenues et plus globalement la confirmation du respect contractuel. En effet, lorsqu'une telle machine est conçue, on s'attache à développer un modèle pour un rendement maximal en fonction notamment de caractéristiques de base telles que le débit traversant la machine et la hauteur de chute. Néanmoins, il convient aussi de s'assurer de la pérennité de la machine durant toute la durée de son exploitation. Par conséquent, à titre d'exemple, en retenant la propagation de fissure comme mode de défaillance, une estimation précise des contraintes mécaniques sur la roue est essentielle pour une prédiction rationnelle de sa durée de vie.

La maintenance industrielle, nécessaire au bon fonctionnement des machines, représente depuis des décennies un réel enjeu économique pour les entreprises responsables de l'exploitation des centrales hydroélectriques. En effet, ces installations très coûteuses à réaliser sont conçues pour avoir des durées de vie de plus d'un siècle. De plus, pour des raisons concurrentielles et économiques entre autres, les équipements sont exploités jusqu'à la limite de leurs spécifications pouvant engendrer plusieurs cycles de départ et arrêts au cours d'une même journée¹. On prend dès lors conscience de l'importance de réaliser des opérations de maintenance régulières en tenant compte des opérations passées (historique de fonctionnement) et à venir (planning

1. La plupart des turbines hydroélectriques ont été pourtant conçues pour subir des cycles de démarrages/arrêts moins fréquents.

d'opération). Ainsi, afin de conserver un maximum de productivité, la maintenance préventive (entretien) est privilégiée par rapport à la maintenance curative (réparation en cas de bris matériel) qui entraîne dans la plupart des cas des pertes d'exploitation bien plus importantes. Par ailleurs, on doit la dégradation du matériel dans une grande majorité des cas au phénomène de rupture par fatigue. Ce phénomène complexe à caractériser est fonction de nombreux paramètres et nécessite une connaissance fine des conditions d'opération de la machine pour formuler des prédictions de détérioration.

L'étude réalisée dans ce mémoire porte sur une turbine hydroélectrique de type Francis installée au Québec (d'une hauteur de chute moyenne et une puissance nominale supérieure à 100 MW). Pour évaluer les risques et estimer le niveau d'endommagement des roues de turbine, il est nécessaire de connaître l'état initial de celle-ci (c.-à-d. dimensions des défauts internes et/ou de fissures), les amplitudes de sollicitation mécanique auxquelles elle est soumise ainsi que son profil d'utilisation (Szczoła, 2012). L'ensemble de ces paramètres sont ensuite étudiés afin de formuler un modèle probabiliste de propagation de fissures, tel que développé par Gagnon *et al.* (2013) qui s'appuie sur le diagramme de Kitagawa et Takahashi (1976). On doit dès lors estimer quelles sont les parties de l'aube de turbine subissant les plus grands niveaux de contraintes et quelles sont les étendues de ces dernières (Gagnon *et al.*, 2014). Pour cela, on peut faire appel à deux méthodes distinctes. La première est empirique ; elle repose sur l'utilisation de quelques jauges de déformation bien disposées pour capter et enregistrer les déformations expérimentales. L'estimation des contraintes réelles sur les aubes ne peut être réalisée qu'en des points stratégiques à cause du coût élevé des mesures *in situ* qui nécessitent un arrêt complet de la turbine, engendrant des pertes d'exploitation conséquentes. De plus, la préparation et la mise en œuvre du matériel de mesure sont délicates en raison des conditions extrêmes auxquelles est soumis celui-ci durant la phase d'acquisition. La deuxième méthode fait appel aux résultats d'analyse par Éléments Finis (ÉF) qui permettent d'estimer des déformations et donc des niveaux des contraintes en tout point du modèle 3D. L'analyse ÉF est

elle-même paramétrée à partir de données issues d'une analyse d'écoulement fluide (CFD). Il est alors nécessaire de proposer un modèle fluide et mécanique pour réaliser la simulation.

Comme dans le cas des aubes de turbine, il est souvent nécessaire de modéliser un problème physique à l'aide d'un formalisme mathématique issu de la mécanique des milieux continus et d'avoir recours à certaines simplifications comme l'illustrent Coleman et Steele (2009) dans la Figure 0.1. Sans cela, la complexité de certains systèmes ou configurations rend impossible la résolution du problème. De nos jours, les simulations numériques sont réalisées grâce à ces modèles et aux outils informatiques. Avec l'augmentation des puissances de calcul des ordinateurs, il est désormais possible de réaliser des calculs mécaniques avec des résolutions de maillage très importantes.

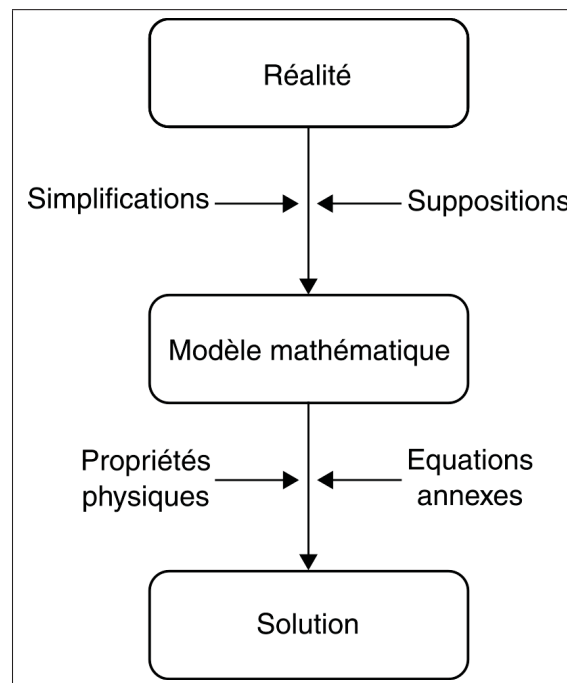


Figure 0.1 Solution d'un problème par une approche analytique - Tirée et adaptée de Coleman et Steele (2009)

Certains peuvent alors se demander pourquoi les ingénieurs sont encore soucieux de réaliser des expériences alors qu'ils disposent pourtant de résultats numériques avec un grand niveau de discrétisation. Il convient alors de se rappeler que les lois physiques et les modèles mathématiques servent à expliquer des phénomènes observés sur un large groupe et à prédire le résultat de futures observations à l'aide d'approximations (Shortley *et al.*, 1961). Ces approximations sont alors vérifiées à l'aide d'expériences et ne sont acceptées qu'à partir du moment où elles expliquent de façon cohérente les faits observés. Ce processus cyclique est illustré par la Figure 0.2.

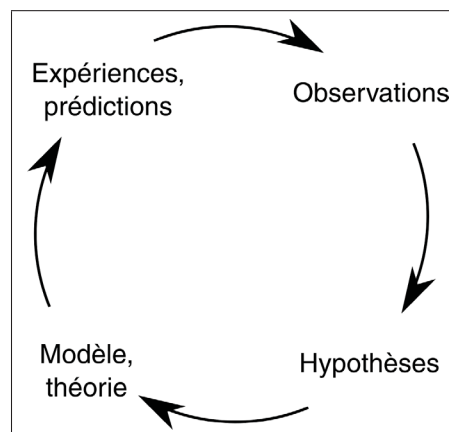


Figure 0.2 Cycle
d'élaboration d'une théorie
ou d'un modèle

Dans le cas des roues de turbines hydroélectriques, on s'intéresse à déterminer dans une collection de modèles, quel est le meilleur pour une utilisation donnée. En effet, dépendamment du comportement que l'on cherche à prédire, l'ingénieur peut préférer un modèle à un autre. Il faut cependant garder en tête que l'exploitation des résultats doit répondre à une question fondamentale : les résultats de simulation reflètent-ils la réalité de la pièce physique ?

Ainsi, afin de pouvoir valider des modèles mécaniques ou encore fluidiques employés, nous proposons d'utiliser une méthodologie rigoureuse de validation basée sur le respect des re-

commandations faites par l'ASME (2006) et sur l'utilisation d'une métrique d'évaluation du modèle particulièrement adaptée. Le but est de proposer aux concepteurs et aux exploitants de turbines hydrauliques une méthodologie rigoureuse et reconnue par la communauté scientifique permettant de valider, dans un sens probabiliste et selon un risque connu et acceptable *a priori*, la qualité des estimateurs par les éléments ÉF.

CHAPITRE 1

REVUE DE LITTÉRATURE

1.1 Démarche de *Vérification & Validation* proposée par l'ASME

1.1.1 Terminologie adoptée par l'ASME

La suite du mémoire utilise les termes suivants et leur signification associée, telle que décrite dans le guide de l'ASME (2006).

- **code** : implémentation informatique d'algorithmes développés pour simplifier la formulation et l'approximation d'une famille de problèmes donnée.
- **modèle** : représentation conceptuelle, mathématique et numérique d'un phénomène physique expliquant des conditions ou un scénario de la réalité. Le modèle intègre la représentation géométrique, les équations principales, les conditions initiales et aux limites, le chargement, les modèles constitutifs, les paramètres liés aux matériaux, les approximations spatiales et temporelles et enfin les algorithmes de solution numérique.
- **résultats de simulation** : le produit généré par le modèle numérique.
- **vérification** : processus visant à déterminer que le modèle informatique représente correctement le modèle mathématique sous-jacent ainsi que ses solutions.
- **validation** : processus visant à déterminer dans quelle proportion un modèle est une "bonne" représentation de la réalité (dans la perspective d'une utilisation donnée).

1.1.2 Aperçu de la norme ASME V&V 10

Depuis toujours, les ingénieurs s'interrogent sur la validité des modèles qu'ils utilisent pour réaliser des simulations. En effet, même si on explique bien de nombreux phénomènes physiques grâce à des modèles mathématiques, des incertitudes persistent sur la validité des postu-

lats (par exemple : homogénéité de la matière, exactitude des constantes, etc.) et sur le renseignement des conditions aux limites (ou encore initiales) ; de plus, ces modèles ne sont applicables que sous certaines hypothèses. De plus, les simulations numériques sont réalisées grâce à l'implémentation des modèles dans des programmes informatiques. De ce fait, le résultat de ces simulations est alors sujet aux possibles erreurs de programmation ou encore aux artefacts informatiques. Pour démocratiser l'usage d'un vocabulaire lié aux activités de Vérification et Validation (V&V, cf. Sec. 1.1.1), l'ASME a publié plusieurs guides appliqués à la mécanique des fluides ou encore à la mécanique du solide (ASME, 2006) et (ASME, 2009). Ces guides constituent un outil de travail dans lesquels sont proposés des conseils et bonnes pratiques à adopter. Ainsi, ils permettent entre autres d'établir de façon appropriée un plan dans lequel sont décrites les différentes étapes de V&V.

La démarche V&V telle que décrite par l'ASME V&V 10 (ASME, 2006) repose sur les points clés suivants :

- La vérification (concerne les erreurs de programmation et de calculs numériques) doit précéder la validation (évalue la précision du modèle physique en comparant les prédictions avec des valeurs expérimentales).
- Les expériences nécessaires à l'évaluation de la précision du modèle doivent être cohérentes par rapport à l'utilisation *in fine* de ce dernier et doivent faire partie intégrante des activités du plan de V&V.
- La validation d'un système complexe doit être opérée de façon hiérarchique : en commençant du composant pour aller au niveau du système.
- La validation est propre au modèle numérique choisi et à l'utilisation que l'on compte en faire.
- Les résultats expérimentaux et de simulation doivent prendre en compte les incertitudes pour être en mesure d'évaluer leurs probabilités.

La documentation des différentes étapes du processus de V&V est primordiale pour assurer un travail de qualité et la pérennité de l'information. Pour ce faire, l'établissement d'un plan de

V&V est particulièrement recommandé. Ce dernier doit être de préférence réalisé avant même d'avoir entamé les différentes étapes de V&V.

Les points clés à retrouver dans le plan de V&V tels que décrits par le guide de l'ASME (2012) sont les suivants :

- a. une description du composant de plus haut niveau hiérarchique ;
- b. une présentation de l'utilisation *in fine* du modèle de plus haut niveau hiérarchique ;
- c. pour chaque élément de la hiérarchie, on doit retrouver :
 - une sélection des caractéristiques à étudier (*i.e.* : ce qui est mesuré, simulé, comparé) ;
 - une présentation des vérifications nécessaires, y compris les logiciels utilisés et les tests de convergence réalisés ;
 - la métrique utilisée pour comparer les valeurs simulées à celles mesurées ;
 - la précision attendue du modèle ;
 - les actions correctives à mener si le modèle ne respecte pas le niveau de précision ;
 - une liste des expériences à mener ;
 - le coût estimé et un planning des tâches.

Un renseignement correct des points précédemment cités permet aux projets subséquents ou même futurs de bénéficier des connaissances acquises lors du processus de V&V.

Bien souvent les composants de niveaux hiérarchiques proches peuvent présenter des modélisations similaires les unes des autres. Ainsi, la consultation de la documentation associée à un projet donné pourra être mise à profit dans un projet connexe. Enfin l'investissement en termes de coût et de temps réalisé dans la constitution d'une documentation de qualité permet à terme de réduire les coûts liés aux activités de V&V.

1.1.3 Vérification

L'étape de vérification vise à déterminer la fidélité du modèle numérique (algorithme informatique) par rapport au modèle mathématique. Pour cela, on scinde cette étape en deux (2) sous-processus : ① la vérification du code informatique et ② la vérification des solutions numériques. La première vise à s'assurer que le code informatique est exempt d'erreur et que les algorithmes pour résoudre les équations de discrétisation donnent des solutions proches de celles des équations à dérivées partielles (qui sont considérées comme exactes). La deuxième vise à estimer (quantifier) les erreurs de résolution dues à la discrétisation, aux itérations successives et aux arrondis. En effet, lors de l'implémentation d'un modèle mathématique dans un code informatique, l'ingénieur est souvent obligé de passer par des processus de discrétisation (par éléments finis par exemple), introduisant alors des notions de convergence de la méthode. En effet, il convient alors de prouver que plus la discrétisation est fine (taille du maillage en ÉF), plus les erreurs de discrétisation tendent vers 0. Il est alors nécessaire, lors des simulations, de choisir judicieusement un niveau de discrétisation suffisant pour converger vers des solutions stables. Cette étape doit être renseignée dans le plan de V&V pour assurer une bonne documentation des simulations.

Pour vérifier qu'un algorithme est correctement implémenté, on commence par résoudre des problèmes disposant d'une solution analytique. Ainsi, il est simple de comparer les résultats issus de l'algorithme à ceux des solutions exactes connues des équations à dérivées partielles. Enfin, pour estimer les erreurs de résolution, dans la plupart des cas le modèle numérique est testé au travers de plusieurs types de maillage (en nombre et par type d'éléments).

Il faut noter que bien souvent, en raison du fait que les ingénieurs sont amenés à utiliser des logiciels commerciaux, le processus de vérification n'est pas (ou peu) réalisable par ces derniers. En effet, l'accès au code source est la plupart du temps protégé par l'éditeur du logiciel et la vérification est censée être réalisée par ce dernier. Pour cette raison, l'étape de vérification ne sera pas plus développée dans ce mémoire. Néanmoins, même avec une solution commer-

ciale, des erreurs peuvent persister, d'où l'attention particulière que l'ingénieur doit porter aux résultats de simulation.

1.1.4 Validation

Le but de l'étape de validation est de déterminer la capacité de prédiction d'un modèle numérique pour une utilisation donnée. Ceci est réalisé en comparant les prédictions (produits de la simulation) aux observations (mesures *in situ*) obtenues à partir d'expériences appropriées. Comme le soulignent Oberkampf et Trucano (2008), évaluer la validité d'un modèle à partir seulement d'une représentation graphique ne constitue pas une méthode rigoureuse et la conclusion peut varier d'une personne à l'autre. Afin de rendre l'étape de validation significative, l'ASME souligne les trois prérequis suivants que doit suivre l'ingénieur :

- avoir une idée précise de l'utilisation *in fine* du modèle ;
- avoir réalisé les étapes de vérification de code et de calcul numérique de façon rigoureuse afin d'être en mesure d'isoler les erreurs rencontrées lors de l'étape de validation ;
- avoir quantifié les incertitudes rencontrées dans les simulations et dans les expériences.

Souvent, des expériences préalables à l'étape de validation ont été réalisées afin de calibrer le modèle numérique. Le guide de l'ASME précise que l'étape de validation ne doit pas réutiliser les données issues des expériences précédentes dans la mesure où celles-ci ne reflètent pas l'étendue totale du spectre d'utilisation du modèle.

Par ailleurs, le plan de V&V doit stipuler le niveau d'adéquation à partir duquel le modèle est considéré comme suffisamment précis pour l'utilisation donnée. Cette étape doit être réalisée impérativement avant de conduire les expériences de validation afin que le niveau d'adéquation imposé ne soit pas biaisé par les constats tirés de ces dernières.

1.2 Outils de validation

Dans aucun de ses guides l'ASME, n'impose de métrique de validation spécifique. En effet, étant donné la complexité et la diversité des requis de validation, il est impossible de généraliser l'usage d'un outil de validation spécifique. L'ASME décrit dans son guide la métrique de validation comme "*une méthode permettant de comparer les résultats de simulation aux résultats expérimentaux*". Néanmoins, il est ajouté qu'une métrique est une mesure (au sens mathématique du terme) de la différence entre les deux résultats. Celle-ci est égale à zéro, si et seulement si, les deux résultats sont identiques.

Même si le comité normatif de l'ASME ne préconise pas l'utilisation d'une métrique en particulier, un certain nombre de caractéristiques sont recherchées dans celle-ci, comme le résume les travaux de Liu *et al.* (2011) et Ferson *et al.* (2008).

Les principales caractéristiques attendues sont les suivantes :

- La métrique doit exprimer une **mesure quantitative** de la différence entre les valeurs simulées et expérimentales. La métrique doit être par ailleurs objective, dans le sens où le résultat ne doit pas dépendre de l'ingénieur qui mène l'analyse.
- La métrique doit être en mesure de **capturer les différences** entre simulation et observations sur l'ensemble des distributions statistiques et pas seulement sur la moyenne et la variance de celles-ci. Cela signifie que la forme de la distribution doit être prise en considération dans la mesure.
- Les critères d'acceptation du modèle doivent être indépendants du résultat retourné par la métrique utilisée. Elle doit servir uniquement à quantifier la mesure de la différence.
- Les **incertitudes** à la fois sur la simulation et les observations doivent être prises en compte.
- La métrique doit être en mesure de discriminer les modèles qui n'ont pas les mêmes niveaux d'incertitudes. Cela signifie que si l'analyste augmente le niveau d'incertitude d'un modèle en élargissant sa fonction de distribution de probabilité, la valeur de la métrique ne doit pas être améliorée de telle façon que le modèle soit plus facilement accepté.

- La métrique doit être en mesure de quantifier la concordance des observations par rapport à une simulation aussi bien au niveau local que global d'un système.

Dans les sections suivantes sont présentés les principaux outils de validation qu'on peut retrouver dans la littérature ou dans les bureaux d'étude de l'industrie. Il conviendra alors de sélectionner un outil de validation regroupant le maximum de caractéristiques de la liste précédente, mais aussi adapté au type de simulation et au nombre d'observations que nous disposons dans notre cas d'étude.

1.2.1 Régression linéaire simple (RLS)

Il n'est pas nécessaire de rappeler la définition mathématique d'une régression linéaire par la méthode des moindres carrés tant elle est couramment employée. Il s'agit en effet d'un moyen simple d'observer et de quantifier s'il y a une corrélation ou non entre des valeurs observées et simulées. Même si elle ne produit pas une mesure à proprement dit de l'écart entre les valeurs simulées et mesurées, elle est régulièrement employée pour caractériser des relations de proportionnalité, en plus de capturer les éventuels biais (Montgomery *et al.*, 2015). La qualité d'une régression linéaire est souvent analysée à l'aide du coefficient de corrélation linéaire de Pearson, r , qui évalue l'adéquation entre le modèle et les données observées. Celui-ci se détermine à l'aide de l'équation 1.1. On définit \bar{x} et \bar{y} comme moyennes respectives sur x et y .

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2} \cdot \sqrt{\sum (y - \bar{y})^2}} \quad (1.1)$$

On utilise plus souvent (dans les logiciels de calcul statistique notamment) le coefficient de détermination R^2 , qui est dans le cas d'une régression linéaire simple, égal au carré du coefficient de corrélation r . Celui-ci est alors compris entre 0 et 1 (0 indiquant que les deux variables sont linéairement indépendantes et 1 indiquant qu'elles sont parfaitement corrélées). Bien que simple à utiliser et interpréter, le coefficient de détermination présente un inconvénient majeur,

comme souligné par Fréchet (1934). En effet, celui-ci croît lorsque le nombre de variables explicatives augmente, que les valeurs soient bien corrélées ou non. C'est pour cela qu'on ne peut se fier à cet indicateur pour quantifier l'erreur entre valeurs simulées et mesurées. Par ailleurs, en statistique on préfère bien souvent utiliser le coefficient de détermination ajusté, $R_{ajusté}^2$ (cf. équation 1.2) qui prend en compte le nombre de variables total n , le nombre de variables explicatives k et qui pallie les problèmes liés au coefficient de détermination classique.

$$R_{ajusté}^2 = R^2 - \frac{k(1 - R^2)}{n - k - 1}, \quad \text{où } k \text{ est le nombre de variables explicatives.} \quad (1.2)$$

Tufte et Graves-Morris (1983) ouvrent le premier chapitre de leur livre en présentant les incohérences que l'on peut rencontrer lorsque des outils statistiques sont utilisés sans qu'une représentation graphique soit préalablement réalisée. C'est avec l'exemple du quartet d'Anscombe, étudié à de nombreuses reprises dans la littérature (Chong *et al.*, 2012), qu'ils illustrent leur constat. Ce dernier présente quatre cas de régressions linéaires partageant les caractéristiques suivantes :

- Droite de régression : $y = 0.5 \cdot x + 3$;
- Coefficient de corrélation : $r = 0.816$;
- Moyenne sur les x : $\bar{x} = 9.0$;
- Variance sur les x : $s_x = 10.0$;
- Moyenne sur les y : $\bar{y} = 7.5$;
- Variance sur les y : $s_y = 3.75$.

On peut remarquer qu'en dépit de l'hétérogénéité des quatre séries de données, les régressions linéaires associées ont la même droite (pente et ordonnée à l'origine). Seule la série 1 exploite de façon cohérente la régression linéaire pour modéliser le nuage de points. Dans les cas 2 et 3, il est clair, avec la représentation graphique, que le modèle linéaire n'est pas adéquat pour modéliser les différentes séries. Par ailleurs, on peut remarquer dans l'exemple 2 que seul

l'avant dernier point n'est pas aligné avec les précédents, ce qui se traduit par une déviation de la pente de la droite de régression (ici une modélisation linéaire est d'ailleurs particulièrement adaptée).

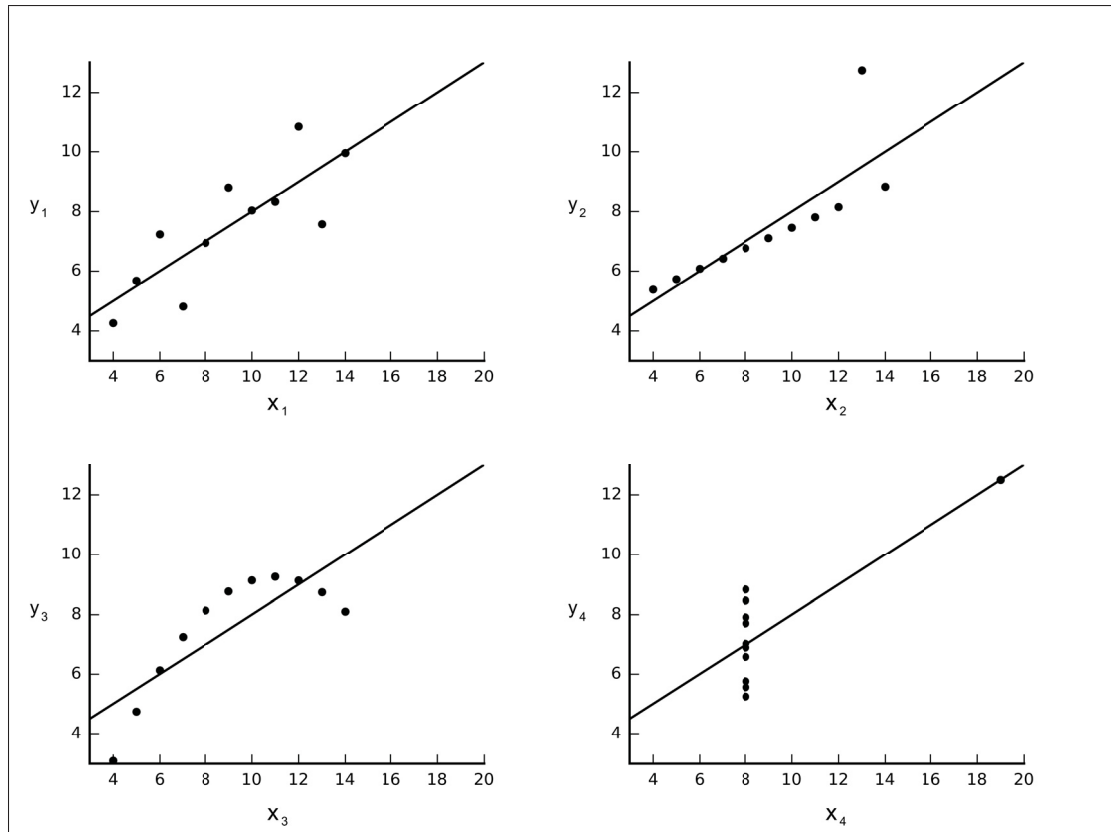


Figure 1.1 Quartet d'Anscombe

Bien que largement utilisée dans différents domaines, la régression linéaire simple présente aussi l'inconvénient de ne pas pouvoir tenir compte des incertitudes de mesure sur chacun des axes (Minitab, 2016). En effet, on intègre, dans le meilleur des cas, les erreurs sur la réponse uniquement. Afin de tenir compte dans la régression des erreurs de mesure à la fois sur la valeur prédite et celle mesurée, il faut se tourner vers les modèles de régression pondérée ou orthogonale, comme développées dans les Sec. 1.2.2 et 1.2.3.

1.2.2 Régression linéaire pondérée (RLP)

Afin de prendre en compte les incertitudes à la fois sur les prédictions et les réponses, il est possible d'introduire une pondération des points dans le modèle de régression linéaire. Le but est alors d'accorder plus d'importances à certains points de la régression linéaire car ils présentent le moins d'incertitude (selon chacun des axes). Cette méthode de pondération a été traitée à plusieurs reprises dans la littérature, dans laquelle on peut trouver différentes méthodes de pondération des données (Chatterjee et Hadi, 2015) et (Rouaud, 2014). Elle est semblable à celle utilisant les moindres carrés d'où les similarités dans l'équation 1.4. On y retrouve l'ordonnée à l'origine et le coefficient de la droite de régression notés respectivement β_0 et β_1 . Les moyennes pondérées sur les deux axes sont définis par \bar{x}_w et \bar{y}_w . Le poids accordé à chaque point dans la régression est inversement proportionnel à la somme des carrés des incertitudes selon chacun des axes (Δx_i et Δy_i) comme le détaille l'équation 1.3.

$$w_i = \frac{1}{(\Delta x_i)^2 + (\Delta y_i)^2} \quad (1.3)$$

$$\begin{cases} \widehat{\beta}_0 = \bar{y}_w - \widehat{\beta}_1 \bar{x}_w \\ \widehat{\beta}_1 = \frac{\sum w_i (x_i - \bar{x}_w)(y_i - \bar{y}_w)}{\sum w_i (x_i - \bar{x}_w)^2} \end{cases}, \quad \text{avec} \quad \begin{cases} \bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} \\ \bar{y}_w = \frac{\sum w_i y_i}{\sum w_i} \end{cases} \quad (1.4)$$

Les deux méthodes de régression linéaires (simple et pondérée) sont comparées dans la figure 1.2 avec quatre points présentant plus ou moins d'incertitude sur les deux axes de coordonnées. La droite continue grise correspond à la régression simple tandis que la droite en tirets noirs représente la régression pondérée. On peut remarquer qualitativement que les points de coordonnées (2 ; 5) et (4 ; 10) présentent le moins d'incertitude (rectangle formé par les barres d'erreur) et "attirent" le plus la droite de régression pondérée.

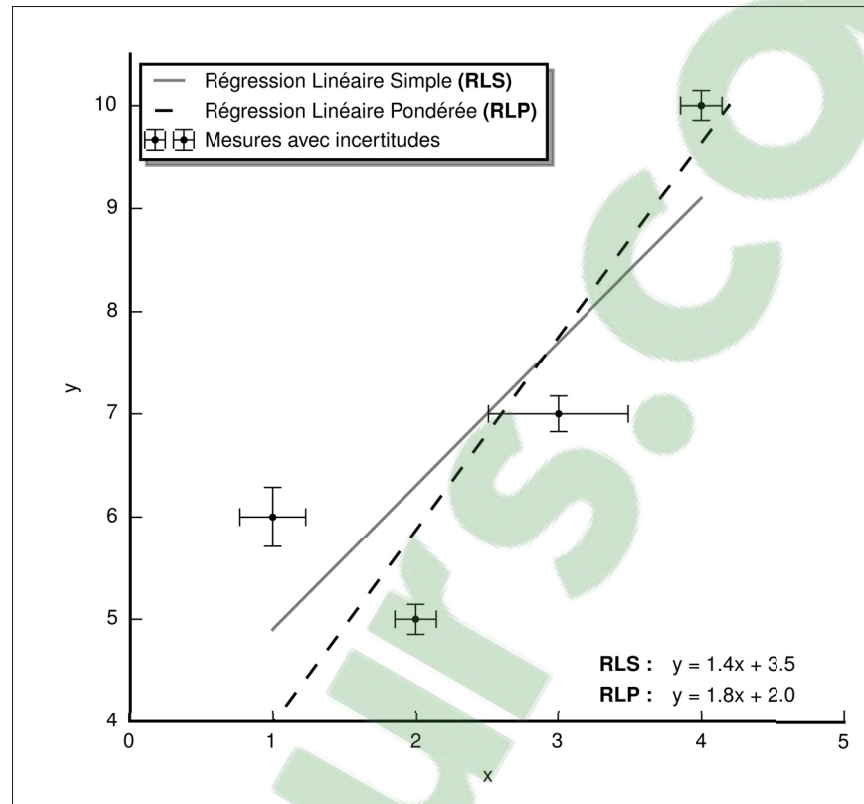


Figure 1.2 Comparaison entre régression linéaire simple et pondérée

1.2.3 Régression linéaire orthogonale (RLO)

Afin de pallier l'incapacité de la régression linéaire simple à prendre en compte les incertitudes de mesure sur les deux axes, la régression linéaire orthogonale peut être employée. En effet, les erreurs de mesure sur la prédiction et la réponse sont intégrées à la détermination de la droite de régression (Minitab, 2016). Néanmoins, son utilisation est soumise au respect des conditions suivantes :

- la prédiction et la réponse contiennent une valeur notée x et y respectivement ainsi qu'une composante d'erreur associée ;
- les erreurs sont indépendantes ;
- les erreurs sont de moyenne nulle et de variance constante ;

- la prédiction et la réponse sont linéairement corrélées.

Ainsi, si les conditions précédentes sont vérifiées, Fuller (2009) a démontré que si le rapport des variances est connu et peut être défini tel que $\eta = \sigma_y^2 / \sigma_x^2$, alors la meilleure estimation des coefficients de régression (au sens des moindres carrés) est définie par les équations 1.5 et 1.6

$$\begin{cases} \widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x} \\ \widehat{\beta}_1 = \frac{m_{yy} - \eta \cdot m_{xx} + \sqrt{(m_{yy} - \eta \cdot m_{xx})^2 + 4 \cdot \eta \cdot m_{xy}^2}}{2 \cdot m_{xy}} \end{cases} \quad (1.5)$$

$$\text{avec les variances,} \quad \begin{cases} m_{xx} = \frac{1}{N} \sum (x_i - \bar{x})^2 \\ m_{yy} = \frac{1}{N} \sum (y_i - \bar{y})^2 \\ m_{xy} = \frac{1}{N} \sum (x_i - \bar{x})(y_i - \bar{y}) \end{cases} \quad (1.6)$$

Les régressions linéaires simples sont basées sur la minimisation de la distance verticale (ou horizontale, dans certains cas) des différents points à la droite de régression par la méthode des moindres carrés. La régression orthogonale, elle, vise à minimiser la projection orthogonale des points sur la droite de régression en prenant en compte les erreurs associées. Les différentes projections ont été représentées sur la figure 1.3. La détermination des coefficients de la droite de régression orthogonale a été réalisée à l'aide du code de Boggs *et al.* (1987) dans l'environnement Python. La droite de régression orthogonale y est représentée par des tirets noirs. On peut constater que les coefficients des droites de régression orthogonale et pondérée sont proches par rapport à ceux de la régression linéaire simple (qui n'incorpore pas les incertitudes selon les deux axes).

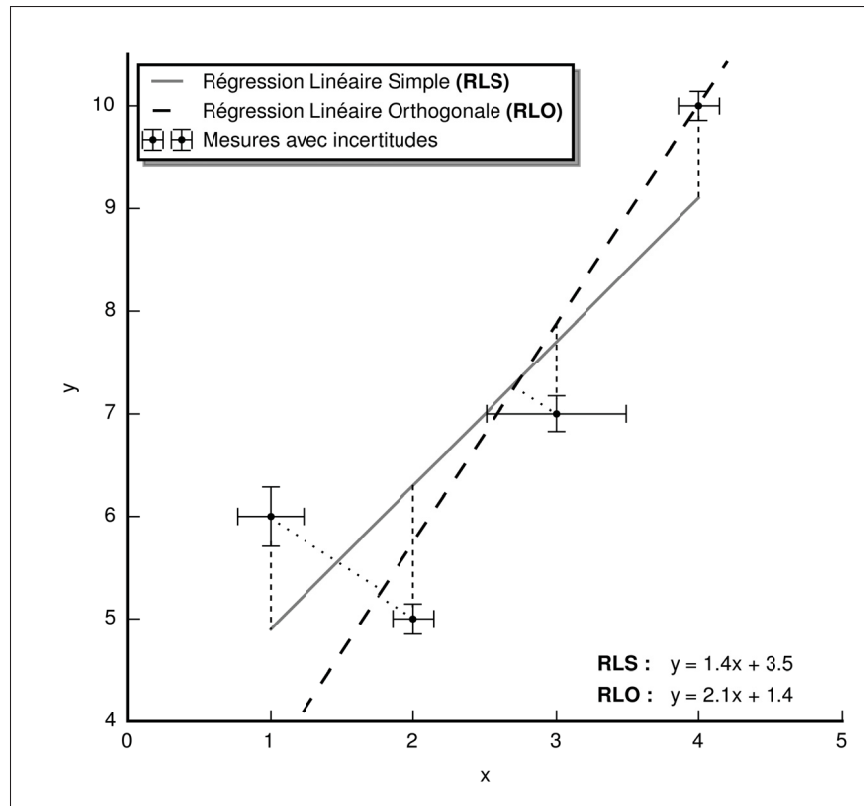


Figure 1.3 Comparaison entre régression linéaire simple et orthogonale

1.2.4 Tests d'hypothèse

Les tests d'hypothèse reposent sur une démarche visant à vérifier une hypothèse statistique pour un échantillon donné (Montgomery *et al.*, 2011). Dans le cas du processus de validation, on cherche à déterminer si l'hypothèse nulle H_0 est vraie. Dans le cas contraire on privilégie l'hypothèse alternative H_1 . On définit alors les hypothèses suivantes :

$$\begin{cases} H_0 : \text{les valeurs observées proviennent de la population simulée} \\ H_1 : \text{les valeurs observées ne proviennent pas de la population simulée} \end{cases} \quad (1.7)$$

Sous l'hypothèse que H_0 est vraie, un test statistique T qui suit une certaine distribution (par exemple : test du χ^2 , de Student ou de Kolmogorov-Smirnov) peut être construit. Ainsi si on considère que les valeurs \widehat{T} du test statistique issues des valeurs mesurées n'appartiennent pas à l'intervalle $[-T_{crit}; T_{crit}]$ du test statistique T , alors l'hypothèse nulle H_0 doit être rejetée.

L'intervalle précédent est déterminé à l'aide du seuil de signification α qui reflète la probabilité de rejet de l'hypothèse nulle H_0 lorsque celle-ci est vraie (appelée erreur de type I). On choisit typiquement $\alpha = 5\%$ ou 10% et l'erreur de type I se traduit par l'équation 1.8. Cette zone de rejet de H_0 est représentée dans la figure 1.4.

$$P_I = P(|\widehat{T}| > T_{crit} | H_0) \quad (1.8)$$

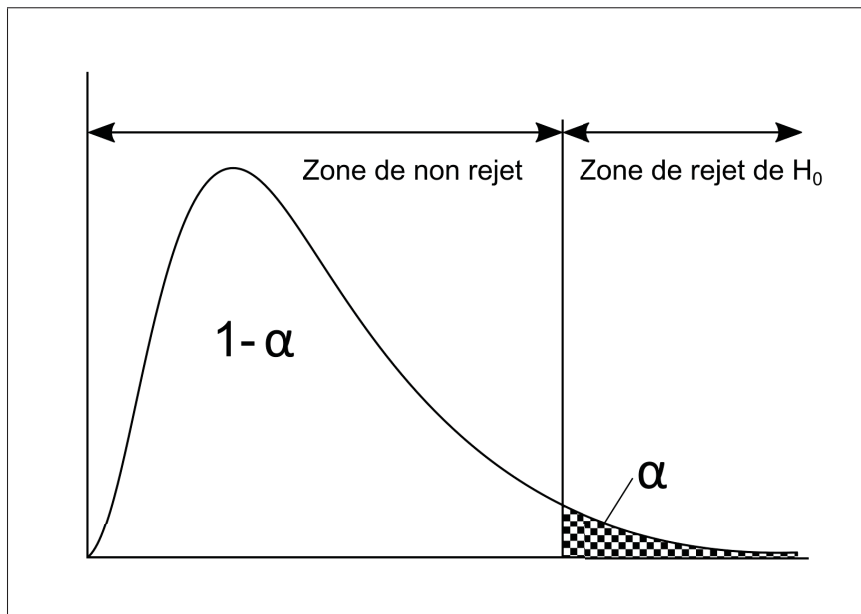


Figure 1.4 Représentation de la zone de rejet pour une distribution donnée

De la même manière, on peut rencontrer des erreurs de type II (défini par l'équation 1.9) qui consiste à accepter l'hypothèse nulle H_0 alors que celle-ci est fautive.

$$P_{II} = P(|\widehat{T}| < T_{crit} | H_1) \quad (1.9)$$

Le choix du test statistique T est dépendant de paramètres tels que l'hypothèse testée, les formulations sous-jacentes ou encore la taille de l'échantillon disponible. Ainsi, on ne peut pas privilégier un test en particulier.

Dans le cas de la validation, on choisit souvent ① de réaliser la comparaison de la moyenne des simulations avec la moyenne des observations et ② de comparer les distributions statistiques des simulations et des observations. Par exemple, pour le cas ②, il peut être intéressant de comparer les courbes de probabilité cumulée des simulations et des observations au lieu de s'intéresser uniquement à l'étude de la moyenne et de l'écart-type. Pour cela, le test de Kolmogorov-Smirnov est particulièrement adapté (Haldar et Mahadevan, 2000). Un des inconvénients de cette méthode réside sur la nécessité d'avoir plusieurs répétitions de mesure. Même si d'autres tests peuvent être utilisés avec une seule observation (Hills et Trucano, 1999), des cas d'application montrent que la méthode a tendance à ne pas rejeter les modèles inadéquats (Liu *et al.*, 2011).

Enfin, les tests d'hypothèse ne visent pas à quantifier l'adéquation d'un modèle mais formulent une décision (binaire) quant à la validité du modèle. Cet aspect n'est pas conforme aux recommandations formulées par l'ASME qui stipule que la quantification de l'adéquation et le critère de décision doivent être découplés.

1.2.5 Métrique de surface - *Area metric*

L'*Area Metric* (AM), développée par Ferson *et al.* (2008) est une méthode qui vise à quantifier les différences entre deux distributions de probabilité cumulée (CDF). Celle-ci mesure l'adéquation entre les valeurs simulées et mesurées en estimant l'aire comprise entre les deux représentations des distributions. Notre choix s'est porté sur l'utilisation de cette métrique en considérant les avantages suivants :

- Donne une mesure quantitative de la différence entre les valeurs simulées et mesurées.
- Ne donne pas de critère d'acceptation ou de rejet du modèle.
- Renvoie un résultat exprimé dans la même unité que la variable étudiée.
- Prend en compte les incertitudes à la fois sur la simulation et l'expérimentation.
- Donne un résultat intuitif, objectif et reproductible (ne dépend pas des opinions de l'analyste).
- Permet de valider aussi bien localement que globalement un modèle.

Afin d'éviter des redondances avec les articles qui suivent, la revue de littérature à propos de l'AM ne sera pas plus longuement développée. Le lecteur est invité à lire les sections dédiées dans le premier article notamment.

CHAPITRE 2

MÉTHODOLOGIE

Les données expérimentales de cette étude sont issues de tests réalisés en 2005 sur les groupes d'une centrale hydroélectrique d'Hydro Québec (d'une hauteur de chute moyenne et une puissance nominale supérieure à 100 MW). Celle-ci est équipée de deux groupes turbine-alternateurs. Les turbines hydroélectriques sont de type Francis, particulièrement adapté aux moyennes hauteurs de chute.

2.1 Instrumentation

Pour étudier le comportement en opération et les régimes transitoires de la turbine hydroélectrique, la roue de la turbine a été instrumentée avec de nombreux capteurs installés sur les parties fixes et rotatives du groupe. Parmi eux, des jauges de déformations, au nombre de dix, ont été posées à la surface de deux pales de la roue. Le but de ce type de capteur est d'estimer la déformation de la matière au niveau de la surface des pièces d'un système pour en déduire les contraintes mécaniques subies. Pour des raisons pratiques, le choix des jauges s'est porté sur une technologie de type "soudée par points". Ces jauges présentent en effet une facilité de mise en œuvre comparativement aux jauges "collées" avec pour principal défaut d'introduire une incertitude non négligeable sur les mesures de déformation en raison de l'effet de sa plaque porteuse notamment (Arpin-Pont, 2012). En plus de ces jauges, on peut souligner la présence de capteurs d'ouverture des directrices de la turbine qui traduisent entre autres à quelle condition elle est opérée.

2.2 Analyse par ÉF

Le but des présents travaux est de développer une méthode de validation des modèles numériques. Par conséquent, différentes analyses par ÉF sont réalisées. En effet, l'estimation des déformations et des contraintes des pales de la turbine est exécutée à partir d'un modèle CAO

numérique 3D. Ce dernier présente la géométrie nominale d'une des pales de la turbine. Préalablement à l'analyse mécanique des milieux continus, une analyse en mécanique des fluides (CFD) a été nécessaire pour générer le profil de chargement à partir de l'écoulement de l'eau. Deux modèles d'écoulement différents sont utilisés pour paramétrer les analyses précédentes. L'objectif *in fine* était de comparer la validité et l'adéquation des différents modèles fluidiques à la réalité physique. Pour cela, les différents champs de pressions issus de l'analyse CFD sont intégrés au modèle de mécanique des milieux continus.

2.3 Incertitudes de simulation

Une fois que les résultats en déplacement et déformation sont obtenus pour chaque site d'intérêt, une méthode de propagation des incertitudes est utilisée. Cette dernière, appelée, Méthode Numérique d'Estimation des Incertitudes de Mesure (MNEIM) (Arpin-Pont *et al.*, 2012), permet d'intégrer aux résultats ÉF les incertitudes de positionnement et d'orientation des jauges de déformation ainsi que l'effet d'intégration introduit par la technologie de la jauge (effet de la plaque porteuse dans le cas présent). La méthode de propagation des incertitudes consiste à reproduire numériquement le comportement de milliers de jauges virtuelles disposées autour du point cible avec des positions et orientations différentes. La MNEIM s'appuie sur la méthode de Monte-Carlo pour générer une distribution de contraintes possibles à chaque site (1000 répétitions sont générées afin d'atteindre un niveau de convergence satisfaisant).

2.4 Validation

La distribution de contraintes obtenue grâce à la MNEIM représente le comportement statistique des valeurs simulées. Les valeurs observées sont issues des mesures expérimentales après qu'une correction de déviation soit appliquée.

Pour chaque site, l'AM est ensuite calculée pour quantifier le niveau d'adéquation des valeurs simulées par rapport à celles observées. Il est alors possible d'estimer l'adéquation globale

d'un modèle en concaténant les résultats obtenus localement. Dans les articles suivants, trois méthodes sont présentées pour juger de la validité globale d'un modèle.

2.5 Structure du mémoire

La suite du mémoire est organisée tel que décrit par la figure 2.1. Le Chapitre 3 présente l'article #1 : "Numerical model validation using experimental data : Application of the area metric on a Francis runner", qui a fait l'objet d'une publication et d'une présentation lors de la conférence IAHR de Grenoble en juillet 2016. Le Chapitre 4 présente l'article #2 : "On the application of the area metric : a case study on hydroelectric turbine", qui a fait l'objet d'une soumission au Journal of Verification, Validation and Uncertainty Quantification ASME en août 2016.

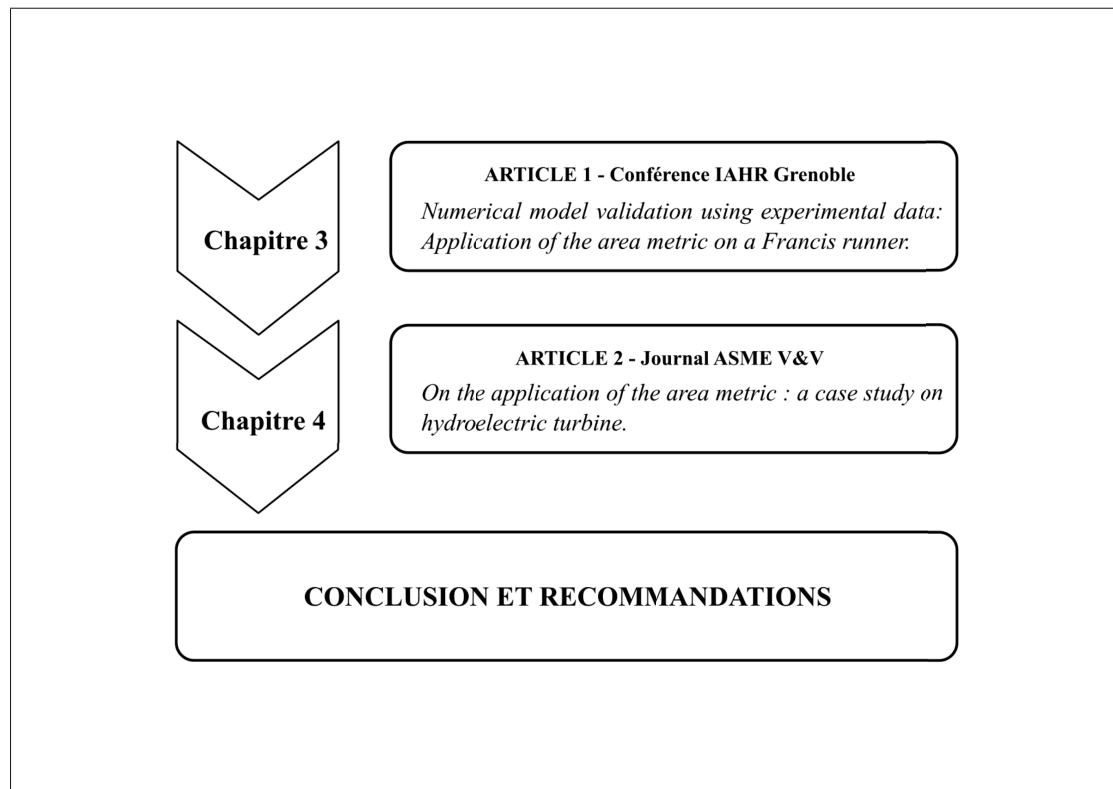


Figure 2.1 Structure du mémoire

CHAPITRE 3

ARTICLE #1 : NUMERICAL MODEL VALIDATION USING EXPERIMENTAL DATA : APPLICATION OF THE AREA METRIC ON A FRANCIS RUNNER

Quentin Chatenet¹, Tahan Antoine¹, Martin Gagnon², Joël Chamberland-Lauzon³

¹ Département de Génie Mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

² Institut de Recherche d'Hydro Québec (IREQ),
1800 boul. Lionel-Boulet, Varennes, Québec, Canada J3X 1S1

³ Andritz Hydro Canada,
6100 aut. Transcanadienne, Pointe Claire, Québec, Canada H9R 1B9

Article accepté à « International Association for Hydro-Environment Engineering and Research (IAHR) » à titre de « Conference Paper » en mai 2016.

3.1 Résumé

De nos jours, les ingénieurs sont capables de résoudre des équations complexes grâce aux capacités accrues des ordinateurs. De ce fait, les logiciels de calcul par éléments finis sont très répandus, et ce particulièrement dans le domaine mécanique afin de prédire le comportement d'une pièce tels que les déformations, les niveaux de contraintes ou encore sa fréquence propre. Cependant, il peut s'avérer difficile de déterminer si un modèle est adapté ou non, ou encore si un modèle est meilleur qu'un autre. Toutefois, lors de la phase de conception, il est très important d'évaluer le comportement des pales de turbines hydroélectriques en fonction des contraintes qu'elles subissent. En effet, les niveaux de contraintes statiques et dynamiques vont influencer la tenue en fatigue de la pale et ainsi sa durée de vie, ce qui est une caractéristique primordiale. Dans l'industrie, les ingénieurs utilisent communément soit les représentations graphiques, les tests d'hypothèses comme celui de Student ou enfin les régressions linéaires afin de comparer les valeurs expérimentales à celles prédites numériquement. À cause des différentes interprétations d'une personne à l'autre (reproductibilité), les validations par méthodes graphiques ne sont pas considérées comme étant objectives. Pour mener une analyse objective, il est essentiel d'employer une métrique de validation pour mesurer la concordance des prédictions par rapport aux données expérimentales. Nous proposons d'utiliser la métrique

de surface (*Area Metric*) dans le cas d'une pale de turbine qui vérifie les points clés de la norme américaine ASME et génère une mesure quantitative de la concordance entre la simulation et les valeurs expérimentales. Cette métrique de validation exclut toute supposition de l'analyste et ne donne pas de critère d'acceptation ce qui renforce sa robustesse. Le but de ce travail est d'appliquer une méthode de validation selon les recommandations de l'ASME Verification & Validation 10. Premièrement, la métrique de surface est appliquée à un cas réel de roue de turbine Francis dont la géométrie et les conditions aux limites sont complexes. Ensuite, elle va être comparée aux méthodes classiques de régression afin d'évaluer sa performance. Enfin, nous allons discuter de l'utilisation de la métrique de surface dans un but de correction des simulations.

3.2 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.

3.3 Introduction

Given the complexity of the needs required to perform the validation process, the ASME V&V 10 committee (ASME, 2006) 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 3.1. 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 figure 3.1). 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 hydro-power 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 (ASME, 2006). The metric used to perform model validation should also satisfy these conditions according to ASME V&V 10 recommendations :*“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*

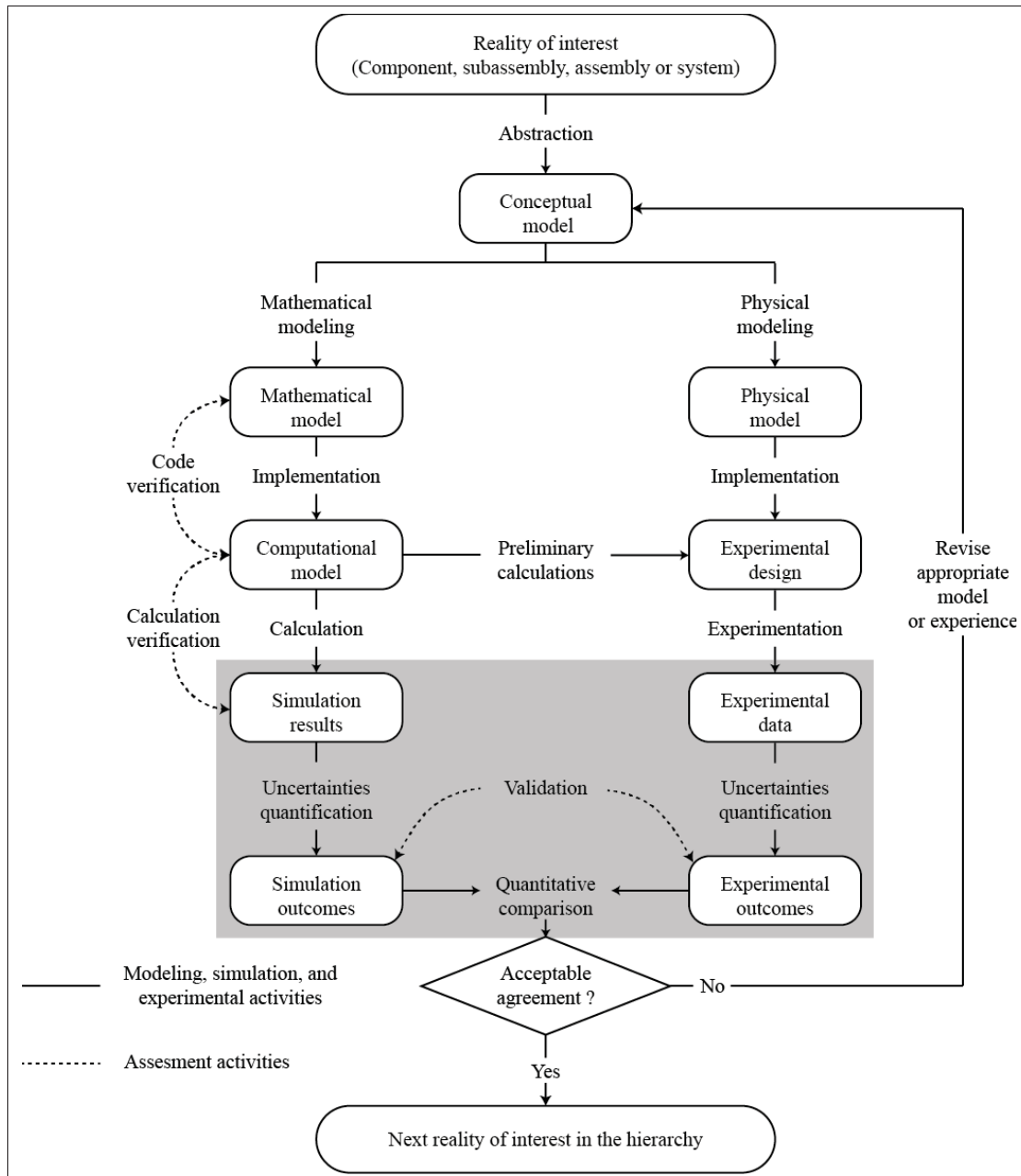


Figure 3.1 Verification & Validation activities (ASME V&V 10).

identical". The proposed validation metric, in this paper, is the area metric developed by Fer-son *et al.* (2008) to integrate these recommendations and excludes any belief and acceptance criterion contrary to classical hypothesis tests (Liu *et al.*, 2011).

This paper is structured as follows. First, a review of the area metric and its characteristics is presented in Sec. 3.4. Next, in Sec. 3.5 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.

3.4 Area metric

The area metric displays several advantages compared to other validation methods (*i.e.* : hypothesis tests, Bayes factor or frequentist's metric) (Liu *et al.*, 2011). 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 4.1, we can construct this function for data x_i .

$$S_n(x) = \frac{\sum_{i=1}^n I(x_i, x)}{n}, \quad \text{where} \quad I(x_i, x) = \begin{cases} 1 & x_i \leq x \\ 0 & x_i > x \end{cases} \quad (3.1)$$

The measure of the mismatch between prediction and observation is thus the area between the prediction function and the data distribution and can be evaluated using equation 4.2.

$$d(F, S_n) = \int_{-\infty}^{\infty} |F(x) - S_n(x)| dx \quad (3.2)$$

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 3.2 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 3.2 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 3.3 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.

3.5 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

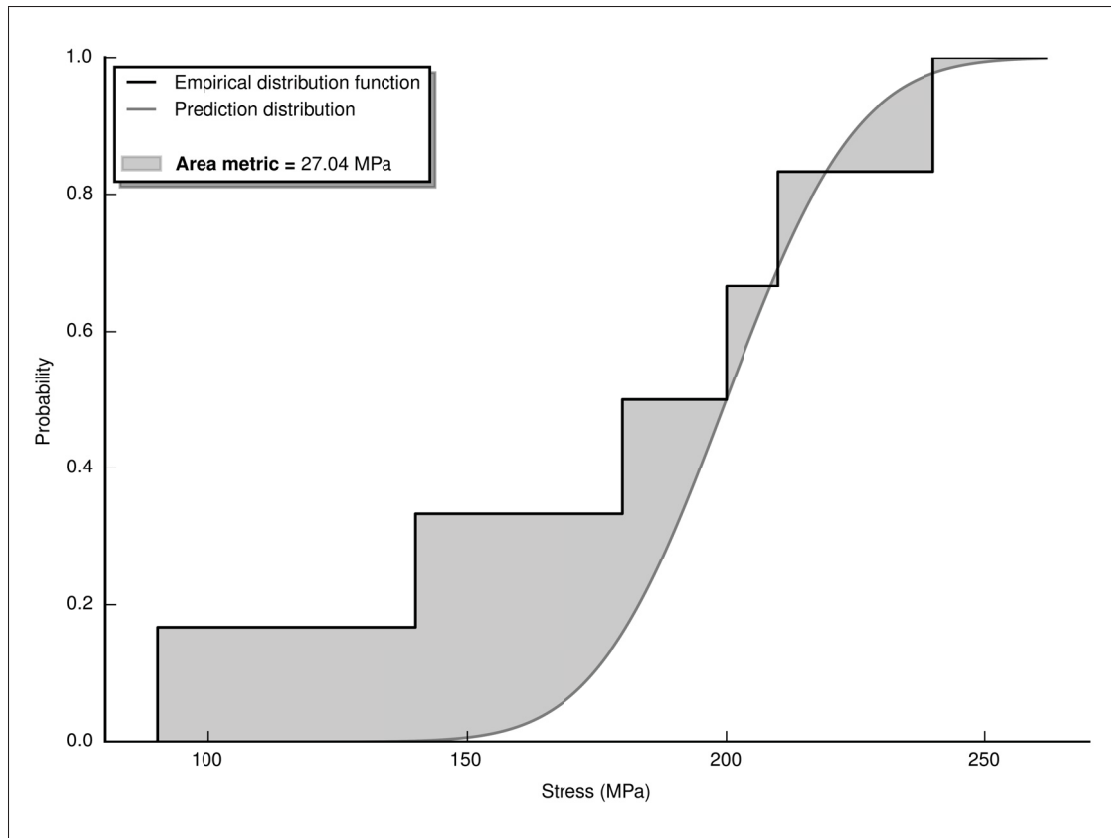


Figure 3.2 Example data set ($n = 6$).

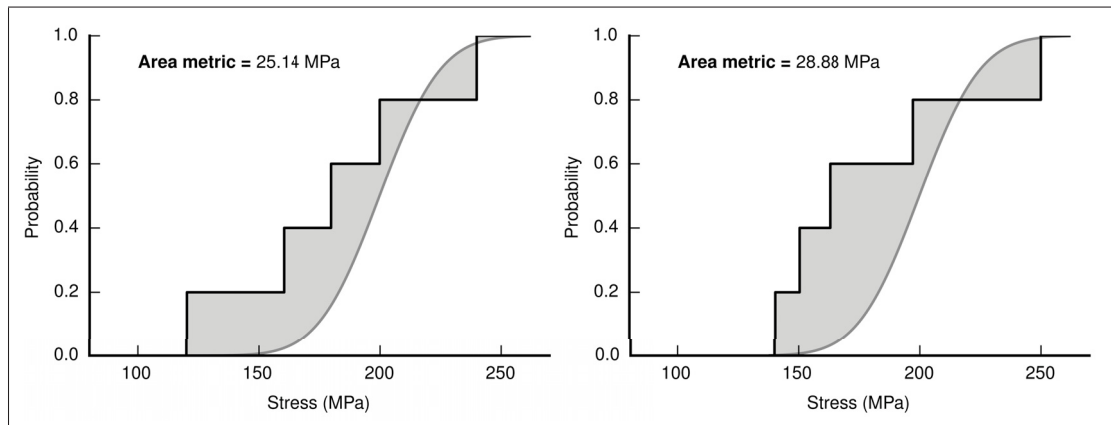


Figure 3.3 Comparison of two different data sets with $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$.

according to two different mechanical models on a real and running Francis hydraulic turbine runner (medium head and nominal power $> 100\text{MW}$).

3.5.1 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 their work, Arpin-Pont *et al.* (2012) 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 3.4, 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.

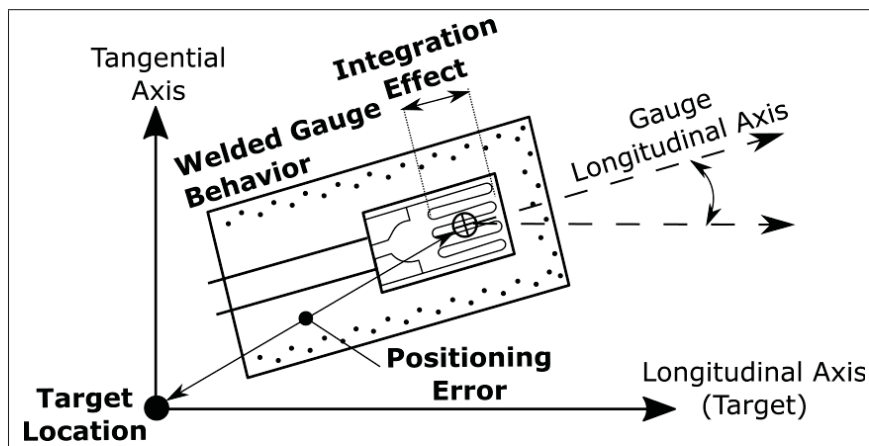


Figure 3.4 Definition of uncertainty sources : welded gauge behavior, location uncertainty (including positioning and alignment errors) and integration effect Arpin-Pont *et al.* (2012).

3.5.2 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.

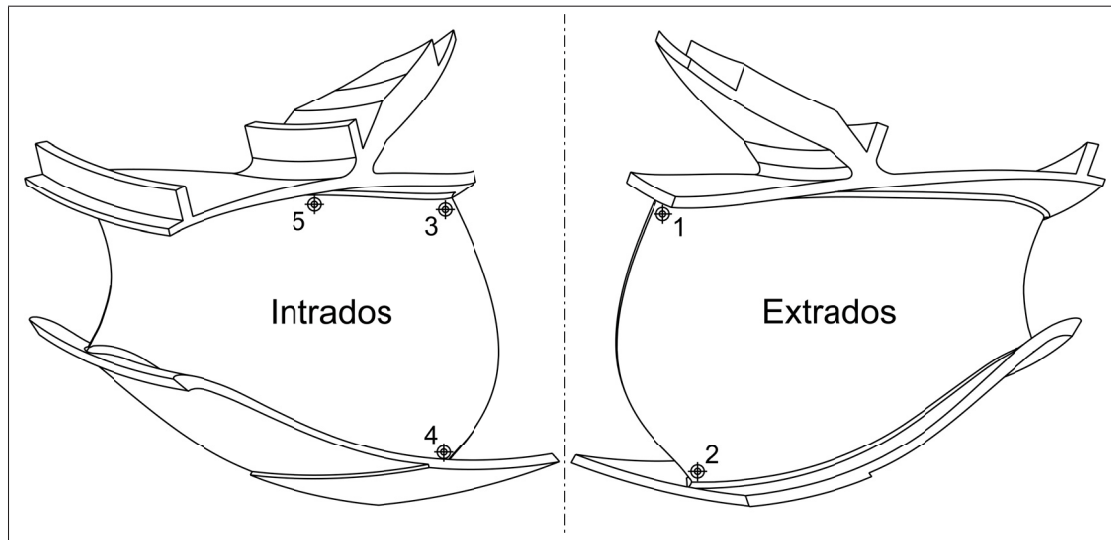


Figure 3.5 Position of the gauges placed on the hydraulic turbine blade during measurements.

According to previous work done by Arpin-Pont (2012), 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 3.1. The simulation distribution obtained for FEA models #1 and #2 are obtained from strain gauge position uncertainty on FEA strain results.

Tableau 3.1 Positioning uncertainty distribution laws used for simulation

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

3.5.3 Experimental process

According to FEA results, strain gauges have been welded at specific locations (as shown in figure 3.5) 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.

3.6 Results

Area metric results comparing simulation distribution and measurements are shown in table 3.2, with detailed graph for site 3 (see figure 3.6).

Tableau 3.2 Area metric results.

Location	Model #1 (MPa)	Model #2 (MPa)
Sites 1 - Crown outflow SS ^a	37.88	255.4
Sites 2 - Band outflow SS ^a	10.53	10.72
Sites 3 - Crown outflow PS ^b	21.69	26.06
Sites 4 - Band outflow PS ^b	51.13	46.97
Sites 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 3.6, 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.

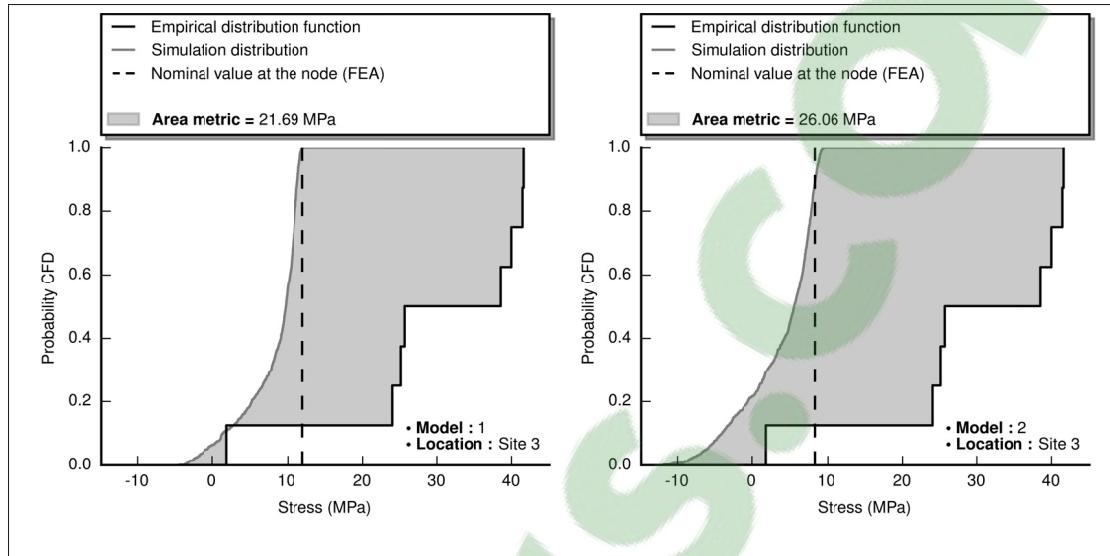


Figure 3.6 Site 3 Area metric comparison of model #1 vs. model #2.

The dashed line in figure 3.6 represents the nominal computed stress value at the node under the strain gauge.

3.6.1 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 3.3. 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.

Tableau 3.3 Evolution of the area metric as a function of wicket-gate opening.

Location	Wicket gate opening			
	12%	40%	75%	100%
Sites 1 - Crown outflow SS ^a	170.80	97.81	8.26	37.88
Sites 2 - Band outflow SS ^a	44.39	18.40	13.86	10.53
Sites 3 - Crown outflow PS ^b	56.68	18.55	23.48	21.69
Sites 4 - Band outflow PS ^b	174.80	104.10	97.89	51.13
Sites 5 - Crown center PS ^b	187.50	147.40	82.72	48.89
Sum	634.17	386.26	226.21	170.12

3.6.2 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 (Chatterjee et Hadi, 2015) (Rouaud, 2014) for example. The weighted least squares estimates ($\widehat{\beta}_0$ and $\widehat{\beta}_1$) are then given by equations 3.4 where \overline{x}_w and \overline{y}_w are the weighted means.

$$w_i = \frac{1}{(\Delta x_i)^2 + (\Delta y_i)^2} \quad (3.3)$$

$$\begin{cases} \widehat{\beta}_0 = \overline{y}_w - \widehat{\beta}_1 \overline{x}_w \\ \widehat{\beta}_1 = \frac{\sum w_i (x_i - \overline{x}_w)(y_i - \overline{y}_w)}{\sum w_i (x_i - \overline{x}_w)^2} \end{cases}, \quad \text{where} \quad \begin{cases} \overline{x}_w = \frac{\sum w_i x_i}{\sum w_i} \\ \overline{y}_w = \frac{\sum w_i y_i}{\sum w_i} \end{cases} \quad (3.4)$$

For each point, weights are assigned as described by equation 3.3. Uncertainties on each axis Δx_i and Δy_i are determined by applying recommendations (ENV13005, 1999), modeled by equation 3.5.

$$\Delta x_i = \frac{\text{range}(x_i)}{2\sqrt{3}} \quad (3.5)$$

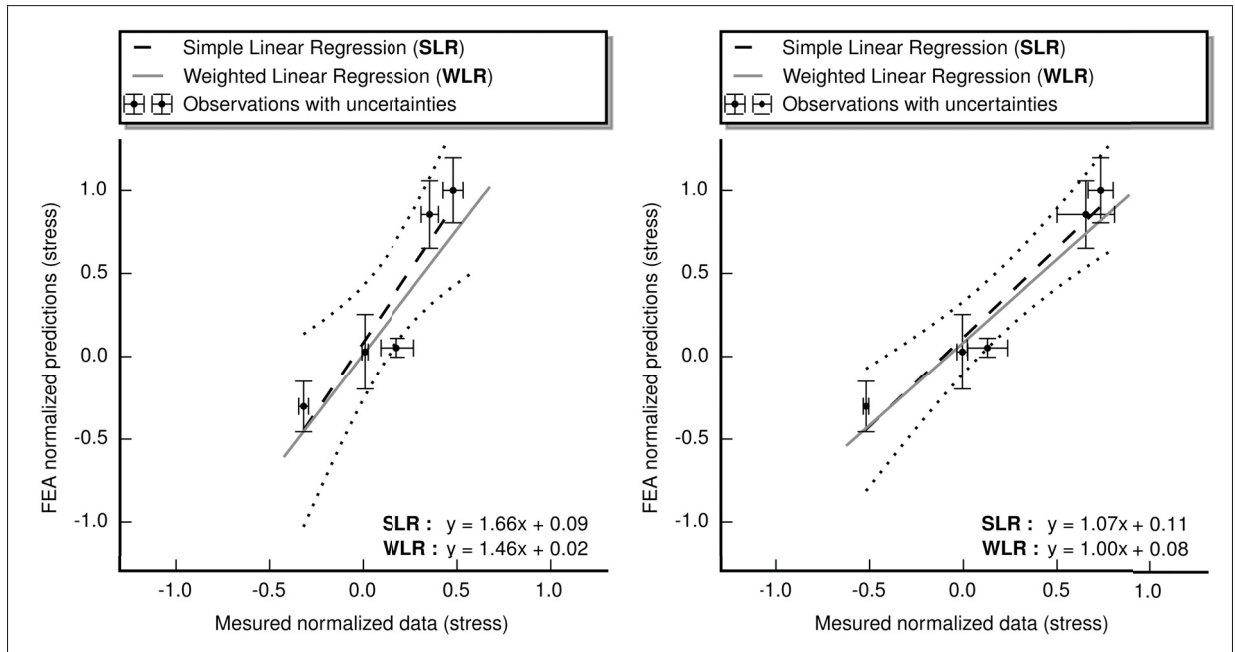


Figure 3.7 Comparison of computed stress and experimental stress (from strain measures) at 75% (left) and 100% (right) opening.

Points represent the mean of sample data and probability distribution; bars associated with points correspond to uncertainties according to each axis. In figure 3.7, 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 : $y = \widehat{\beta}_1 x + \widehat{\beta}_0$.

3.7 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 3.3. 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 3.7, 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.

3.8 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.

CHAPITRE 4

ARTICLE #2 : ON THE APPLICATION OF THE AREA METRIC : A CASE STUDY ON HYDROELECTRIC TURBINE

Quentin Chatenet¹, Martin Gagnon², Tahan Antoine¹

¹ Département de Génie Mécanique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

² Institut de Recherche d'Hydro Québec (IREQ),
1800 boul. Lionel-Boulet, Varennes, Québec, Canada J3X 1S1

Article soumis à «Journal of Verification, Validation and Uncertainty Quantification ASME »
à titre de « Research Paper » en août 2016.

4.1 Résumé

Avec les améliorations des capacités de simulation, un besoin croissant en validation de modèles se fait sentir dans les milieux des sciences et de l'ingénierie. En effet, les analystes doivent s'appuyer sur des métriques de validation pour évaluer les performances des modèles afin de les comparer entre eux. Bien qu'il n'y ait pas de métrique de validation standardisée ou largement acceptée, la métrique de surface (*Area Metric*) semble se démarquer grâce à ses caractéristiques intéressantes. Même si la métrique de surface est particulièrement efficace pour générer une mesure quantitative de la concordance entre les valeurs prédites et mesurées, on remarque que l'extension de cette méthode (appelée u-pooling) utilisée pour déterminer la validité globale d'un modèle, présente certains inconvénients et limites. Dans cet article, nous présentons une approche alternative s'appuyant également sur la métrique de surface qui permet de résoudre les problèmes liés à la méthode u-pooling. D'abord, un cas d'application est utilisé pour montrer les limites de cette dernière. Puis, nous présentons notre approche. Enfin, différents modèles de simulation numérique sont comparés et analysés.

Mots clés : validation, métrique de surface, u-pooling, adéquation globale de modèle, turbine hydroélectrique

4.2 Abstract

With improvement in simulation capacity, there is an increasing need for model validation in scientific fields and engineering. Indeed, analysts must rely on validation metrics to evaluate model performance in order to compare different models. Although there are no generalized or widely accepted validation metrics, the area metric seems to stand out because of its distinctive features. Even if the area metric is particularly efficient at producing a quantitative measurement of the agreement between simulation predictions and observations, we observe that its extension for global model validity assessment (using the u-pooling method) presents some drawbacks and limitations. In this paper, we present an alternative approach also based on the area metric which solve some of the u-pooling method issues. First, a study case is used to highlight the limitations of the u-pooling method. Next, our alternative approach is presented. Finally, different numerical simulation models are compared and discussed.

Keywords : validation, area metric, u-pooling, global model adequacy, hydroelectric turbine.

4.3 Introduction

Nowadays, it has become easier to compute numerical simulations due to the increase of computing capacity. This allows a more accurate prediction of real system behavior using numerical model. Meanwhile, the validation plan, as was outlined in the ASME Verification and Validation Guide ASME (2006), is not well defined or does not specify validation requirements to accept a model. Thus it is difficult agree on how to quantify models accuracy. In order to solve this issue, the ASME proposed a terminology alongside with a complete verification and validation methodology combined with some best practices.

According to ASME's guide ASME (2006), uncertainties should be taken into account in both simulation results and simulation data during the validation process. Even though, a specific metric is not imposed, the guide recommends using one to quantify the difference between simulation and experimental data. This metric should be equal to zero if both are identical. In order to illustrate this, in ASME (2012), the ASME Committee provided two different vali-

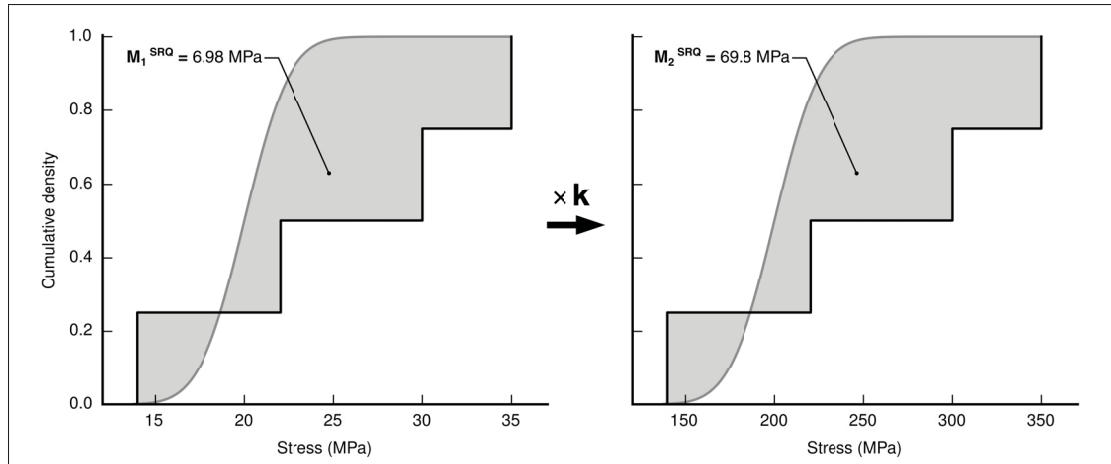


Figure 4.1 Application of the area metric - Example data sets ($n = 4$, $k = 10$).

dition approaches : the first one deals with case where only a single experiment and a single simulation are available, the second, the area metric, is useful when experimental replications and probabilistic simulation distribution are available. We chose to focus on the area metric which was proposed by Ferson *et al.* (2008) because uncertainty estimation does not rely on experts (ASME, 2012). This metric enable the comparison of the entire distributions of simulation results with experimental data. Moreover, this metric meets all the ASME's mathematical requirements and consideration for uncertainties. The requirements are as follows :

- the metric must measure the difference between experimental outcomes and simulation outcomes.
- the metric is zero if and only if the two outcomes are identical.
- the metric does not conclude on the model validity but provides a coherent numerical value expressing the model accuracy/adequacy.

According to the comparison made by Liu *et al.* (2011), the area metric also features some advantages compared to more classical ones (*i.e.* : hypothesis tests, Bayes factor or frequentist's metric) such as excluding any belief and criterion of accepting model.

Since most numerical simulations are complex, it is preferable to identify beforehand the experiments needed to ensure proper validation. Indeed, even if numerical results are available at each location in the model. Because of costs or technical restrictions, only a few strategic measurement locations can be considered during the experimental testing. Thus, the analyst must conclude on the accuracy of a model by relying on limited observed data. Furthermore, it is important to note that we can distinguish between local and global consideration in terms of model accuracy. A model can adequately predict the reality for some local subset of the domain, but may not be good enough for the entire domain. The area metrics enable one to quantify the local accuracy of a numerical model by comparing individually one or several experimental measurement locations with predicted results. However, the area metric method can be extended in order to conclude on the global accuracy of a model using *u-pooling* methodology which consists in concatenating local results in a single metric to assess global model performance (Ferson et Oberkampf, 2009).

The remainder of the paper is organized as follows : First the area metric is presented in Sec. 4.4. Then, the *u-pooling* method and its drawbacks are presented in Sec. 4.5. Next an alternative approach is suggested in Sec. 4.6 in order to solve *u-pooling* issues. This is followed by a study case in Sec. 4.7. Finally, we discuss the results and conclude on the applicability of the methodology.

4.4 Overview of the area metric

The area metric has numerous advantages compared to classical validation methods (hypothesis tests, Bayes factor, frequentist's metric, least squares regression). Because its result does not depend on who conduct the analysis or the assumptions made (since it relies only on the data distribution shape). It produces a quantitative measure of the discrepancy between predictions and data which is considered objective and robust (Ferson *et al.*, 2008). Moreover, it enable one to obtain a graphical representation of this discrepancy (in the same physical units as the data used). The metric allows one to evaluate the differences across the full range of prediction results taking into account simulation uncertainties.

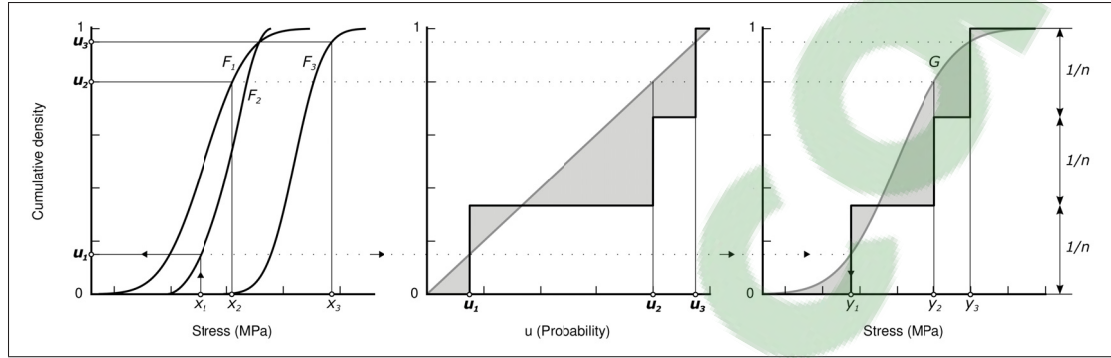


Figure 4.2 Illustration of u -pooling method and $back$ -transformation.

In order to apply the area metric method, we need prediction uncertainties which can be represented as a cumulative distribution function $F(x)$ (CDF), where x is the variable of the prediction. In the case where multiple observations are available, observation uncertainties 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 Eqn. 4.1, we can construct this function for data x_i .

$$S_n(x) = \frac{\sum_{i=1}^n I(x_i, x)}{n}, \quad \text{where} \quad I(x_i, x) = \begin{cases} 1 & x_i \leq x, \\ 0 & x_i > x \end{cases} \quad (4.1)$$

The mismatch between prediction and observation is thus the area between the prediction function and the observed data distribution and can be evaluated using Eqn. 4.2.

$$M^{SRQ}(F, S_n) = \int_{-\infty}^{\infty} |F(x) - S_n(x)| dx \quad (4.2)$$

Figure 4.1 (left) illustrates the mismatch between the prediction distribution (shown as the grey curve) and the experimental data (shown as the dark step function with the metric being the grey shaded area). The metric $M^{SRQ}(F, S_n)$ is obtained using Eqn. 4.2 using the entire range of both the prediction uncertainty distribution and empirical data. In this example, a normal distribution has been chosen for numerical simulation uncertainties but any parametric

or empiric distribution could be used. Finally the area metric is function of the magnitude of the values studied. Hence, a magnitude effect should be expected given the units used for the data. Thus, for two similar distribution shapes with data of different magnitude will produce area metric values proportional to the scale factor k as shown in Fig. 4.1.

4.5 Overview of u-pooling method

The *u-pooling* method, as an extension of the area metric, allows to aggregate validation results over the entire validation domain using a single metric. It is possible to evaluate the global performance of different models and thus compare them even if the resulting distributions of each simulation model are different. The method consists in calculating a *u-value*, u_i , for each experimental measurements x_i , by evaluating its corresponding CFD, F_i , as

$$u_i = F_i(x_i) \quad (4.3)$$

Figure 4.2 (left graph) shows the u_i transformation. Given the fact that u_i expresses the percentile (between 0 and 1) in which the measurement is with respect to simulation uncertainty distribution. Then if x_i is larger than the maximum value of simulation uncertainty distribution, $F_i(x_i) = 1$ and $F_i(x_i) = 0$ for any value smaller than the minimum of the distribution. Thus, if the x_i are perfectly distributed, the set of u_i will be uniformly distributed on $[0; 1]$ according to the probability integral transform theorem (Angus, 1994). Thus, *u-values* can be compared against a standard uniform distribution over $[0, 1]$ using the area metric to quantify the model accuracy in a global sense. Figure 4.2 (center graph) shows the area metric using the uniform distribution. Meanwhile, Ferson *et al.* underlines the fact that transforming values into probability results in the loss of the original units used for comparison. Therefore, he suggested solving this “*by back-transforming the u-values through a suitable distribution G that restores the scale and its interpretation*”. The aim of this procedure is to keep the intuitive features of the area metric. Similarly, a non-decreasing step function must be generated from a chosen distribution G and previous *u-values*. The new empirical step function is constructed using the

back-transformed data values (through the inverse of the G function) :

$$y_i = G^{-1}(u_i) = G^{-1}(F_i(x_i)) \quad (4.4)$$

Figure 4.2 (right graph) shows the *back-transformation* of u -values for a chosen G function.

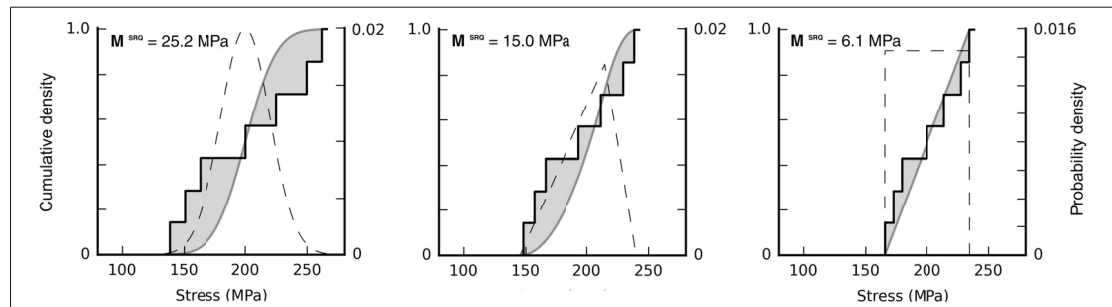


Figure 4.3 Influence of the G function choice in back-transformation process.

Nonetheless, some desirable features of the area metric are not preserved using the u -pooling method. The area metric is considered to be objective in the sense that the result does not depend on how the validation process is conducted. Once uncertainties have been quantified, no assumption need to be made by the analyst which could influence the discrepancy evaluation between simulation and experimental data. The u -pooling and the *back-transformation* do not preserve this objectiveness. Indeed, insofar as y_i , in Eqn. (4.4), depends on the choice of the distribution function G . Indeed, this function could be chosen in such a way that force the metric M^{SRQ} to meet the validation requirement (e.g. $M^{SRQ} \leq \text{validation requirement}$). In Fig. 4.4, three (3) different functions G have been chosen for a same set of u_i (0, 0.2, 0.5, 0.7, 0.9, 1) in order to illustrate this issue. These functions are respectively normal, triangular and uniform distributions with identical mean ($\mu = 200$ MPa) and standard deviation ($s = 20$)¹. Using this example, we observe that, the area metric is very sensitive to the distribution G and its parameters. In the case where the normal distribution is used, the area metric is 25.2 MPa which is about 1.6 times higher than for the triangular distribution (15.04 MPa) and about 6 times higher than for the uniform one (6.1 MPa). Since the *back-transformation* process, as developed

1. The unit of interest in Sec. 4.7 is expressed in MPa. The same work can be done regardless of the unit used.

by Ferson *et al.*, has no rules for choosing the G function, this choice belongs to the analyst conducting the validation. As a results, a loss of objectiveness is to be expected.

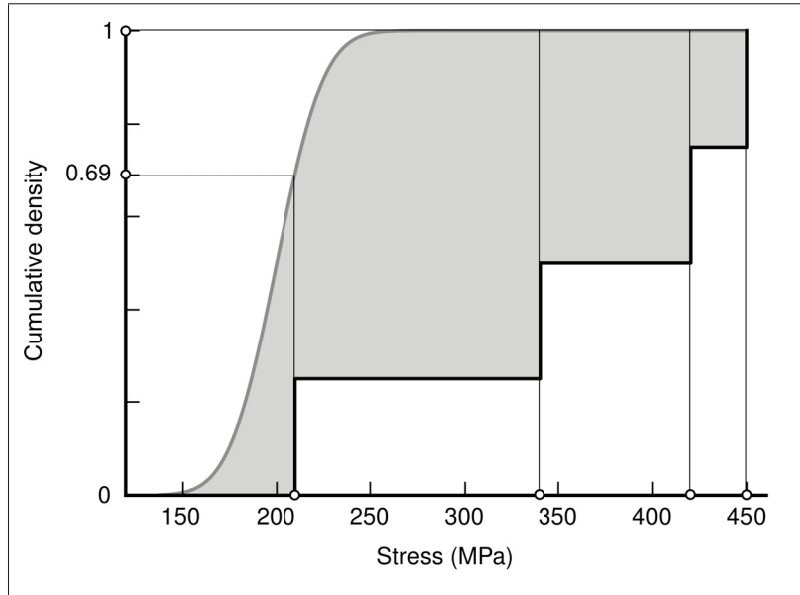


Figure 4.4 Example of u_i calculation. Only one observation is distributed according to its simulation distribution.

For Ferson *et al.* (2008), a validation metric “*should be unbounded in the sense that, if the prediction is completely off the mark of the measurement, the metric characterizing this discrepancy should be able to grow to be an arbitrarily large value.*” The validation metric then stays sensitive even if the prediction is ‘*far*’ away from the measurement. The area metric integrates this desirable property. Meanwhile, this is not the case for the *u-pooling*. Figure 4.4 shows an illustration of calculating the values u_i for four (4) hypothetical transformations. Insofar as only one of four values x_i belongs to its simulation uncertainty distribution, we get the following set of transformations of u_i ’s : [0.69, 1, 1, 1]. Thus, if we consider any value x_i close to the prediction distribution but which is not in the support range, its *u-value* will be the similar to another value which could be ‘*far*’ from the distribution. This situation represents a limitation for the *u-pooling* method. Indeed, as soon as a data x_i does not belong to its simulation uncertainty distribution support or become extreme, it is impossible to determine from *u-values* how ‘*far*’ away the measure is from its prediction and the sensitivity is lost.

Finally, the *u-pooling* method relies on the probability integral transform, “which is to say, so long as the model is predicting the observations well” (Ferson et Oberkampf, 2009), $u_i = F_i(x_i)$ will be uniformly distributed on $[0; 1]$. If observations x_i are within distribution F_i support, the *u-pooling* graph (cf. Fig. 4.2 (center graph)) is similar as the representation of the Kolmogorov-Smirnov statistical test (Massey Jr, 1951). The latter can be used to determine if a sample of data belongs to a given reference distribution ; here, if it belongs to the uniform distribution on $[0; 1]$. Concatenating *u-values* from different prediction distributions F_i allow the analyst to see if there is a leading trend, such as aggregated values in a specific area of the distribution. This can be done only if there is a significant number of observations and if it can be assumed that the sample is representative. As soon as some observations x_i are out the distribution support, it is impossible to compare models side by side at an global level.

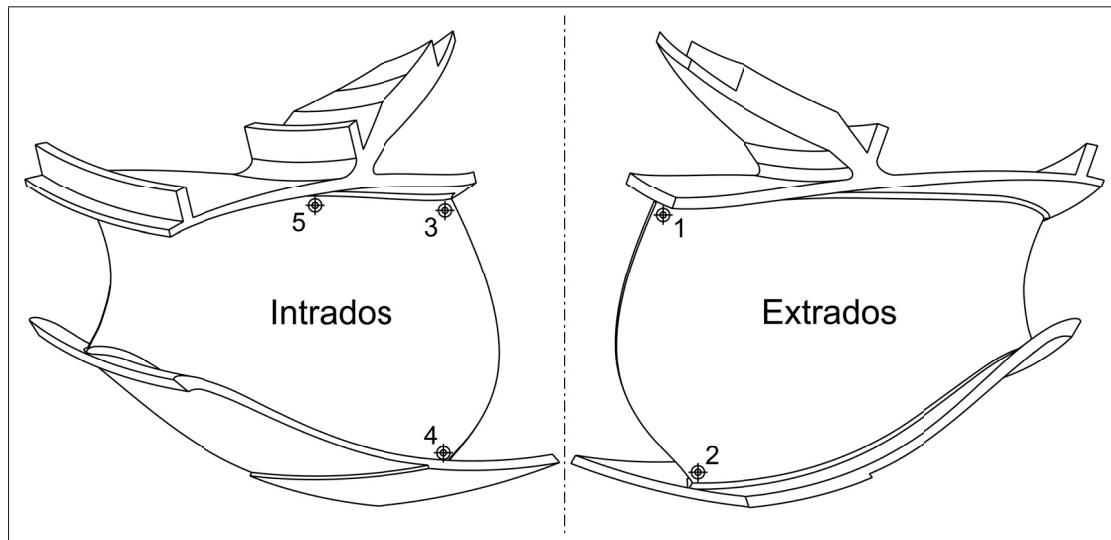


Figure 4.5 Position of the strain gages placed on the hydraulic turbine blade during measurement.

4.6 Proposed methodology

As described in Sec. 4.5, the *u-pooling* method presents several limitations and is consistent under the condition that observations belong to its prediction distribution (*e.g.* when model predicts well reality). To solve these issues, we propose in this section an alternative approach

to evaluate model adequacy on an entire validation domain. We believe that the metric used to evaluate global model adequacy should be as objective as the area metric for local accuracy. Thus, regardless of who is conducting the analysis, the result should be similar and the conclusion reproducible. For each site, the simulation returns a prediction distribution described by $F_i(x)$. Concatenating all prediction uncertainty distributions F_i in a single cumulative function H , as detailed in Eqn. 4.5 (n represents the number of combined distributions F_i), allows to aggregate simulation data from different sites. Once data are concatenated in H function, it is possible to apply the area metric method to assess the global model performance. Therefore in this method the distribution used is a function of the data and does not rely on an arbitrary choice from the analyst.

$$H(x) = F_1(x) \cup F_2(x) \cup \dots \cup F_n(x) \quad (4.5)$$

Likewise, empirical observations are combined in a single data set to construct the non decreasing step function $S_n(x)$. Finally, an area metric is computed to measure discrepancy between prediction distribution and empirical data.

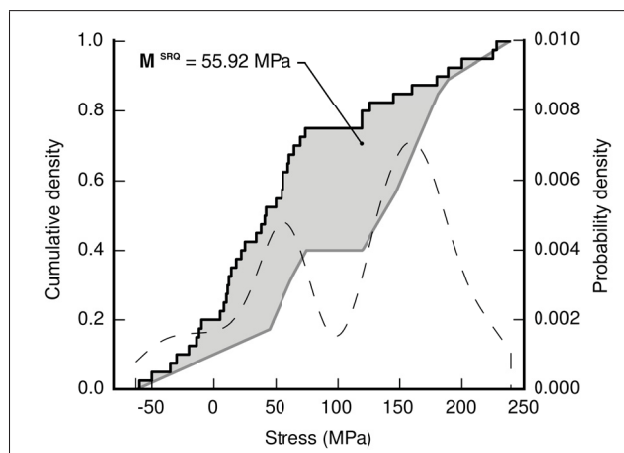


Figure 4.6 Example of the area metric applied on aggregated distribution $H(x)$.

Figure 4.6 illustrates an example of the area metric applied on the entire validation domain. The prediction distribution $H(x)$ shown as the smooth gray curve represents the cumulative density function (CDF), whereas the black dotted line shows the probability density function. The function $H(x)$ has been computed from five (5) different distributions $F_i(x)$ and it is compared to function $S_n(x)$ (black step function) which represents a set of forty (40) observations (eight (8) per simulation).

4.7 Study case

The aim of this study case is to determine between two (2) numerical simulation models which one better estimate stress on hydraulic turbine runner blade. Stress levels are important parameters for life expectancy assessment of hydraulic turbines (Gagnon *et al.*, 2013) (Gagnon *et al.*, 2014). First, the *u-pooling* method has been applied to compare two models against experimental data. For this purpose, two sets of finite element analysis (FEA) have been computed to predict the stress on an in-service Francis hydraulic turbine runner. The finite element numerical simulation computes results at each node of the model. Based on these results, the experimenter decides the location where the measurements will be performed which are shown in Fig. 4.5. Sites 1 to 4 have been chosen because those regions present high stress levels ; site 5 is considered a checkpoint because stress is moderated. In order to experimentally determine the stress at each site, strain gages have been welded at the surface of the blade according to locations presented in Fig. 4.5.

In our cases, given the significant experimental variability, compared to numerical simulation, these were not taken into account in our study. Because of the strain gradient under gages, the position and orientation of the latter have a great influence on the value obtained (Kobayashi et Chui, 2005) (Little *et al.*, 2006). Hence, the sensor technology (welded gages) must be taken into account. Furthermore, with computing capacity, it is possible to refine the element size in such a way that we are mesh independent. Taking into account experimental uncertainties, we used the method developed by Arpin-Pont *et al.* (2012) which consists in the simulation of virtual strain gages located around the finite element target node accounting for positioning

error. Using a Monte Carlo simulations, we then account for the numerous possible positions and orientation. The aim is to reflect the reproducibility of the measurement. Thus, a distribution of the stress measured by the virtual gages is generated. On the experimental side, the data is collected for each site (1 to 5) within similar operating conditions to replicate the measurement.

4.8 Results

Results are presented in Tables 4.1 and 4.2. Data *Overlapping* represents the percentage of measured data which belongs to its prediction uncertainty distribution (as defined in Eqn. 4.6). For the two models, the area metric has been computed at each site of the blade.

$$\text{Overlapping} : \frac{x_i \in F_i}{n_i} \quad (4.6)$$

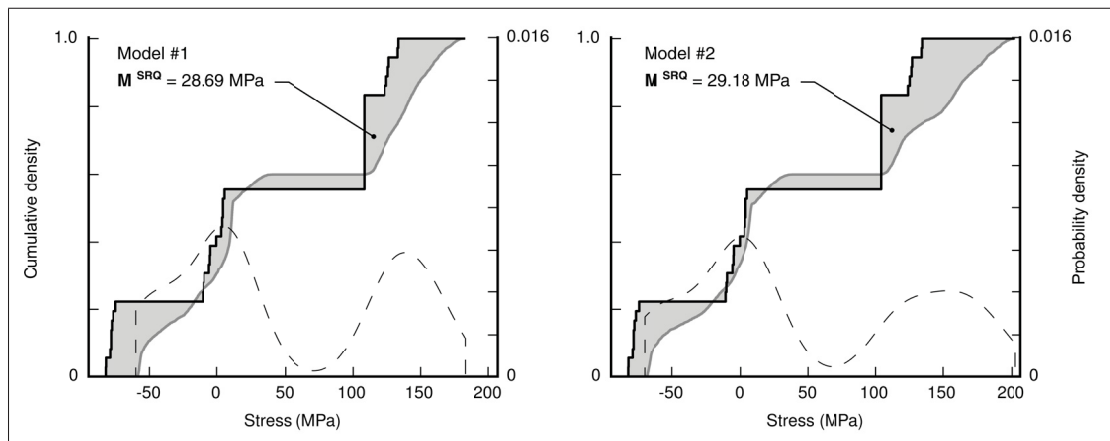
Column '*FEA value*' in Tables 4.1 and 4.2 represents the normalized nominal value computed during numerical simulation using 100 and 0 to represent respectively the maximum and the minimum of stress levels. Signs of values indicate a state of compression (negative) or tension (positive). Figure 4.8 shows area metric results while it is applied to *u-values* against uniform distribution on [0 ; 1] (*u-pooling* method). Finally, Fig. 4.7 illustrates results of the alternative solution proposed in Sec. 4.6. The five (5) distributions F_i have been aggregated to provide the single distribution H . In our case the function S_n is generated from 36 observations.

Tableau 4.1 Area metric results - model #1

Site	Area metric (MPa)	Overlapping	FEA value
1	32.76	0 %	- 41
2	11.23	100 %	0
3	21.47	25 %	4
4	32.08	50 %	72
5	40.71	25 %	100
Sum	138.25		

Tableau 4.2 Area metric results - model #2

Site	Area metric (MPa)	Overlapping	FEA value
1	23.29	0 %	- 27
2	11.43	100 %	0
3	20.47	50 %	13
4	30.00	50 %	79
5	60.31	0 %	100
Sum	146.00		

Figure 4.7 Area metric applied on the entire distribution H for two different models.

4.9 Discussions

First, we think that we can simply and intuitively evaluate the global accuracy of a given model, by adding area metric results of each site of interest. Then, the smaller the sum is, the more accurate is the model. Hence, in our study case, according to Tables 4.1 and 4.2, the model #1 is the better with a sum of 138.25 MPa. Furthermore, we have to underline the fact that the highest values of the area metric will exert leverage on the sum. Indeed, because the area metric is not normalized, the results exhibit a scale effect which is function of the magnitude of the data values. Thus, for the two similar sets of data (prediction and observation distributions, cf. Fig. 4.1) with different scales (k factor), the set whose magnitude is the higher (right graph) produces the highest area metric (cf. Eqn. 4.7). In Fig. 4.1, we observe that with $k = 10$ that the scaling effect is observed in the area metric result. This means that the higher magnitude

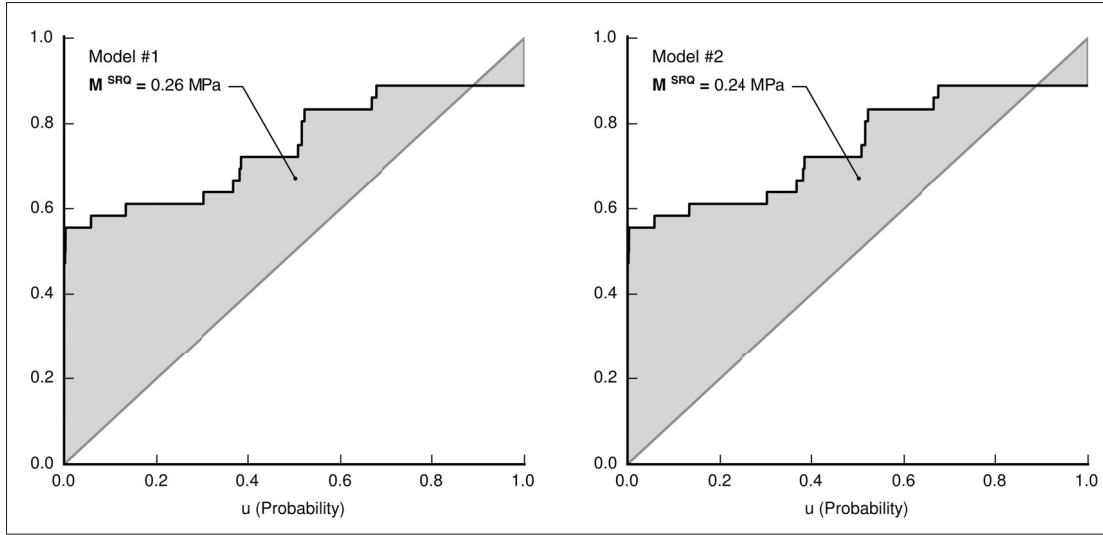


Figure 4.8 Area metric computed on u -pooling graph

is for a similar discrepancy in proportion, the more the area metric will influence the sum. This property of the area metric (which is considered as a desirable feature by Ferson *et al.* Ferson et Oberkampf (2009)) does not represent an issue in our case insofar as we are interested in evaluating specifically the high stress area for fatigue issues. Furthermore, as shown in Tables 4.1 and 4.2, area metric results and FEA values are sorted in the same order (excluding site 1 in model #1) which is consistent with the previous statement.

$$F_2(x) = k * F_1(x) \quad \Rightarrow \quad M_2^{SRQ} = k * M_1^{SRQ} \quad (4.7)$$

Tables 4.1 and 4.2 show that only for site no. 2, all observations belong to prediction distribution. For sites no. 1, 3, 4 and 5, 50 % or less of empirical data overlap with simulation distribution as seen in the overlapping results defined in Eqn. 4.6. Then, when the area metric is applied to the u -values compared against the standard uniform distribution over $[0; 1]$ (*cf.* Fig. 4.8), the area metric is smaller for model #1 than for model #2. This contradicts the conclusion made from the area metric sum. The reason is that the u -pooling method does not discriminate observations out of their respective prediction uncertainty distribution support. When observa-

tions are 'far' from the simulation distribution, u -values are either 0 or 1. In other words, the u -pooling method is not capable of capturing entire information.

Finally, with our alternative approach for the evaluation of global model accuracy illustrated in Fig. 4.7, we conclude that the model #1 is the more accurate (area metric = 28.96 MPa) compared to model #2 (area metric = 29.18 MPa). This conclusion is consistent with results in Tables 4.1 and 4.2. Moreover, the methodology, which is based on the area metric provides a graphical representation which allows the analyst to visually determine sites which exert leverage effect on the metric. In fact, it is visually possible to identify sites whose predictions do not match with observations : the larger the grey area is, the less the model predict correctly the reality.

4.10 Conclusions

In this paper we underlined the restrictions of the u -pooling and back-transformation methods through three examples. The result shows that transforming observations into u -values, can result in some cases in a loss of information. We demonstrated that the back-transformation result can be sensitive to the distribution chosen by the analyst which could be chosen *a posteriori* to meets the validation requirement. This lack of objectiveness is considered not desirable. Instead of using a predefined distribution, we proposed a solution which uses the obtained predicted values distribution itself over the entire validation domain.

The proposed solution has been illustrated with a study case. In this study case, the data comes from measurements carried out on hydraulic turbine runner and a finite element analysis (FEA) of the corresponding numerical model. The results demonstrated that proposed solution provides more consistent results than the u -pooling methodology.

Acknowledgments

This study was carried out with the financial support of the Canadian research internship program, Mitacs-Accelerate. The authors would like to 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.

CONCLUSION ET RECOMMANDATIONS

Nous rappelons au lecteur que l'objectif de ce projet est de proposer et d'appliquer une méthode de validation des modèles numériques pour l'estimation des contraintes mécaniques et que cette méthode soit adaptée en particulier au milieu des turbines hydroélectriques. La principale particularité est que dans ce cas, le modèle de simulation procure une information complète et exhaustive de l'état des contraintes mécaniques sur l'ensemble de la roue de turbine, par contre, le nombre de sites de mesures empiriques est très réduit.

L'expérience et la pratique en milieu industriel montre qu'une grande diversité de méthodes sont employées pour conduire des processus de validation sans pour autant que leur usage soit correctement justifié ou adapté. En effet, la littérature a montré que les processus de vérification et de validation devaient faire appel à des outils quantitatifs spécifiques et à une démarche cohérente, sans quoi, les conclusions portant sur la validité du modèle pourraient être erronées. Aussi, les incertitudes associées au modèle et aux mesures empiriques doivent être intégrées afin d'obtenir des résultats de validation de meilleure qualité et surtout plus rigoureux. Lorsque des jauges de déformation soudées sont employées, la variabilité des mesures *in situ* est prépondérante par rapport à celle du calcul par ÉF (en considérant que l'étape préalable de vérification a été réalisée sur le code de calcul ÉF). Pour cette raison, notre intérêt s'est porté sur l'évaluation de ces erreurs. C'est dans cette démarche que la méthode MNEIM a été employée. Cette dernière nous a permis de générer des distributions statistiques des déformations mécaniques à partir de résultats nominaux par ÉF qui ne renvoient qu'une seule valeur (sans incertitude). Ces distributions simulées sont le résultat de l'intégration des incertitudes de mesure au modèle.

Les essais expérimentaux en centrale sont contraints par des considérations de coût et de temps d'arrêt de production. Ainsi, seulement quelques répétitions de mesure sont réalisées. En plus de respecter les recommandations des experts en V&V, la métrique de surface (*Area Metric* -

AM) permet de quantifier la validité d'un modèle même lorsqu'on ne dispose que d'un nombre réduit de mesures empiriques. On peut par ailleurs rappeler les qualités que présente l'AM :

- Donne une mesure quantitative de la différence entre les valeurs simulées et mesurées sur l'ensemble des distributions de données.
- Permet de réaliser une validation locale et globale (par la méthode de concaténation des données).
- Prend en compte les incertitudes à la fois sur la simulation et sur l'expérimentation.
- N'est pas influencée par les opinions de l'analyste. La méthode donne des résultats objectifs, robustes et reproductibles.
- Respecte les recommandations formulées par l'ASME V&V à propos des caractéristiques de la métrique.

Pour ces raisons, et après l'avoir comparé aux méthodes classiques de validation, l'AM a été retenue afin de quantifier la validité de deux (2) modèles distincts en mécanique des fluides. Le but est de comparer les performances de deux modèles présentant des similarités dans les résultats. Une fois l'AM appliquée, ceci nous a permis de déterminer lequel des deux modèles reflète le mieux la réalité physique. En ce sens, l'objectif du projet a été rempli. Néanmoins, il pourrait s'avérer pertinent de s'intéresser au cas de la validation de modèles en régime transitoire (contraintes dynamiques); nos travaux se limitant à une étude de l'état des contraintes en régime stationnaire (statique). En effet, la connaissance des amplitudes maximales des contraintes mécaniques lors des phénomènes transitoires est essentielle à toute prédiction de la limite de vie et au calcul du dommage induit à la structure (fatigue, mécanique de la rupture). Notre première recommandation va dans ce sens, il sera opportun de tenter une extension des travaux effectués vers le cas du régime transitoire afin de procurer une méthodologie de validation cohérente, qui tient compte du comportement global du modèle et qui procure une aide précieuse pour les concepteurs et les exploitants ; autant sur la validation des étendues des cycles de contraintes que sur les valeurs crêtes de ces dernières.

Un autre volet de l'AM s'intéresse à la validation globale d'un modèle en agrégeant les résultats de différents sites pour donner un niveau de validité commun. Cette méthode, appelée *u-pooling* montre cependant des inconvénients lors de son application. Nous avons par ailleurs démontré, à travers d'exemples théoriques et pratiques, les limites qui pouvaient être rencontrées. Pour pallier cela, nous avons proposé une méthode alternative remplissant le même objectif : *quantifier la validité d'un modèle numérique à un niveau global*. Celle-ci est basée aussi sur l'évaluation d'une surface comprise entre deux fonctions de distribution. Elle présente l'avantage de se baser sur la distribution réelle des valeurs simulées et non pas sur une distribution choisie *a priori* par l'analyste comme c'est le cas pour la méthode de *u-pooling*. À notre sens, cela fait d'elle une solution plus objective, ce qui est un critère important en V&V. Son application sur un cas réel s'est avérée fructueuse et les résultats obtenus sont cohérents avec ceux de méthodes basiques de validation.

Notre deuxième et dernière recommandation concerne l'évaluation des incertitudes sur les paramètres du modèle et ses conditions aux limites. En effet, que ce soit en mécanique des milieux continus ou en mécanique des fluides, l'ingénieur doit renseigner de façon fidèle les multiples paramètres de simulation. Cette tâche est très complexe et toujours entachée d'erreurs à cause des approximations nécessaires. Il serait alors intéressant d'estimer les incertitudes sur, entre autres : les paramètres du mode d'écoulement, les différentes conditions aux limites et les propriétés des matériaux (variabilité de constitution, inhomogénéité de la matière, propriétés mécaniques altérées localement, etc.). Aussi, un intérêt particulier pourrait être porté sur l'influence dans les résultats de simulation de la géométrie réelle des pièces par rapport au modèle numérique CAD idéal.

Nous sommes convaincus que nous avons répondu, peut-être partiellement, à l'objectif principal de notre projet ; une recherche scientifique sera-t-elle terminée un jour ?

"La poursuite des perfectionnements exclut la recherche de la perfection. Perfectionner s'oppose à parfaire" - Paul Valéry

BIBLIOGRAPHIE

- Angus, J. E. 1994. « The probability integral transform and related results ». *SIAM review*, vol. 36, n° 4, p. 652–654.
- Arpin-Pont, J., M. Gagnon, S. Tahan, A. Coutu, et D. Thibault. 2012. « Strain gauge measurement uncertainties on hydraulic turbine runner blade ». In *IOP Conference Series : Earth and Environmental Science*. p. 062042. IOP Publishing.
- Arpin-Pont, J. 2012. « Méthode de détermination des incertitudes de mesures par jauges de déformation ». Master's thesis, École de technologie supérieure, Montréal, 129 p.
- ASME. 2006. *ASME V&V 10-2006 : Guide for Verification and Validation in Computational Solid Mechanics*.
- ASME. 2009. « ASME V&V 10-2012 : Guide on Verification and Validation in Computational Fluid Dynamics and Heat Transfer ». *New York : American Society of Mechanical Engineers (ASME)*.
- ASME. 2012. « ASME V&V 10-2012 : An illustration of the Concepts of Verification and Validation in Computational Solid Mechanics ». *New York : American Society of Mechanical Engineers (ASME)*.
- Boggs, P. T., R. H. Byrd, J. R. Donaldson, et R. B. Schnabel. 1987. *Odrpack software for weighted orthogonal distance regression*. Technical report.
- Chatterjee, S. et A. S. Hadi, 2015. *Regression analysis by example*.
- Chong, J., D. Halcoussis, et G. M. Phillips. 2012. « Misleading betas : An educational example ». *American Journal of Business Education (Online)*, vol. 5, n° 5, p. 617.
- Coleman, H. W. et W. G. Steele, 2009. *Experimentation, validation, and uncertainty analysis for engineers*.
- ENV13005, N. 1999. « Guide pour l'expression de l'incertitude de mesure ».
- Ferson, S. et W. L. Oberkampf. 2009. « Validation of imprecise probability models ». *International Journal of Reliability and Safety*, vol. 3, n° 1-3, p. 3–22.
- Ferson, S., W. L. Oberkampf, et L. Ginzburg. 2008. « Model validation and predictive capability for the thermal challenge problem ». *Computer Methods in Applied Mechanics and Engineering*, vol. 197, n° 29, p. 2408–2430.
- Fréchet, M. 1934. « Sur l'usage du soi-disant coefficient de corrélation ». International Statistical Institute.
- Fuller, W. A., 2009. *Measurement error models*, volume 305.

- Gagnon, M., A. Tahan, P. Bocher, et D. Thibault. 2013. « A probabilistic model for the onset of High Cycle Fatigue (HCF) crack propagation : Application to hydroelectric turbine runner ». *International Journal of Fatigue*, vol. 47, p. 300–307.
- Gagnon, M., A. Tahan, P. Bocher, et D. Thibault. 2014. « Influence of load spectrum assumptions on the expected reliability of hydroelectric turbines : A case study ». *Structural Safety*, vol. 50, p. 1–8.
- Goupy, J. 2005. « La régression PLS 1, cas particulier de la régression linéaire séquentielle orthogonale (RLSO) ». *Revue Modulad*, vol. 33.
- Haldar, A. et S. Mahadevan, 2000. *Probability, reliability, and statistical methods in engineering design*, volume 1.
- Hills, R. G. et T. G. Trucano. 1999. « Statistical validation of engineering and scientific models : Background ». *Sandia National Laboratories, Albuquerque, NM, Report No. SAND99-1256*.
- IEA, I. E. A. 2011. *Key world energy statistics*. Technical report.
- IEA, I. E. A. 2015. *Key world energy statistics*. Technical report.
- Kitagawa, H. et S. Takahashi. 1976. « Applicability of fracture mechanics to very small cracks or the cracks in the early stage ». In *Second International Conference on Mechanical Behavior of Materials*. ASM, Metals Park, Ohio. 1976, 627-631.
- Kobayoshi, M. et Q. S. H. Chui. 2005. « The positioning influence of dial gauges on their calibration results ». *Measurement*, vol. 38, n° 1, p. 67–76.
- Little, E., D. Tocher, D. Colgan, et P. O'Donnell. 2006. « A Probabilistic Evaluation of the Least Squares Strain Tensor Derived from a Three-Dimensional Modular Strain Rosette Using the Monte-Carlo Technique ». *Strain*, vol. 42, n° 3, p. 207–216.
- Liu, Y., W. Chen, P. Arendt, et H.-Z. Huang. 2011. « Toward a better understanding of model validation metrics ». *Journal of Mechanical Design*, vol. 133, n° 7, p. 071005.
- Massey Jr, F. J. 1951. « The Kolmogorov-Smirnov test for goodness of fit ». *Journal of the American statistical Association*, vol. 46, n° 253, p. 68–78.
- Minitab. 2016. « Orthogonal regression ». <<http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/regression-and-correlation/basics/orthogonal-regression/>>.
- Montgomery, D. C., G. C. Runger, et N. F. Hubele, 2011. *Engineering statistics*. éd 5th ed. Hoboken, NJ : John Wiley & Sons.
- Montgomery, D. C., E. A. Peck, et G. G. Vining, 2015. *Introduction to linear regression analysis*.

- Oberkampf, W. L. et T. G. Trucano. 2008. « Verification and validation benchmarks ». *Nuclear engineering and Design*, vol. 238, n° 3, p. 716–743.
- Rouaud, M., 2014. *Calcul d'incertitudes*. Paris, France : Creative Commons.
- Shortley, G., D. Williams, I. Dayton, et M. Heald. 1961. « Elements of physics ». *American Journal of Physics*, vol. 29, n° 10, p. 719–720.
- Szczota, M. 2012. « Modélisation de l'historique d'opération de groupes turbine-alternateur ». PhD thesis, École de technologie supérieure.
- Tufte, E. R. et P. Graves-Morris, 1983. *The visual display of quantitative information*, volume 2.