



HAL
open science

Méthodes statistiques pour l'exploitation d'une description probabiliste de l'erreur de mesure : estimation des valeurs vraies et amélioration la gestion de la production

Marc Gille

► **To cite this version:**

Marc Gille. Méthodes statistiques pour l'exploitation d'une description probabiliste de l'erreur de mesure : estimation des valeurs vraies et amélioration la gestion de la production. Génie mécanique [physics.class-ph]. Université Clermont Auvergne, 2023. Français. NNT : 2023UCFA0064 . tel-04440441

HAL Id: tel-04440441

<https://theses.hal.science/tel-04440441v1>

Submitted on 6 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse en vue d'obtenir le grade de Docteur d'Université
Spécialité "Génie Mécanique"

**Méthodes statistiques pour l'exploitation d'une
description probabiliste de l'erreur de mesure :
estimation des valeurs vraies et amélioration de la
gestion de production**

présentée par
Marc Gille

Université Clermont Auvergne, Sciences pour l'Ingénieur de
Clermont-Ferrand, 20 juillet 2023

Membres du jury :

Jean-Yves Dantan	Professeur, ENSAM	Président
Pascal Lafon	Professeur, Université de technologie de Troyes	Rapporteur
Francois Thiebaut	Maître de conférence, Université de Paris-Saclay	Rapporteur
Zohra Cherfi-Boulanger	Professeur, Université de Technologie de Compiègne	Examinateur
Nicolas Gayton	Professeur, SIGMA-Clermont	Co-Directeur
Pierre Beaurepaire	Maître de conférence, SIGMA-Clermont	Co-Encadrant

Remerciements

Je remercie Thierry Yalamas et PHIMECA Engineering, pour m'avoir proposé et confié ce sujet de recherche. Je remercie tous les collègues de l'entreprise avec qui j'ai eu le plaisir d'échanger et de vivre ces trois belles années. Je remercie mon équipe de travail, toute l'équipe FIC, et particulièrement Antoine Dumas, Fabien Taghon, Julien Schueller et Guillaume Causses, experts généreux qui m'ont volontiers partagé de leurs expériences et chacun beaucoup appris. Merci encore à Antoine, mon encadrant, qui s'est lancé avec moi dans cette expérience inédite.

Remerciements chaleureux à Pierre Beaurepaire, mon encadrant universitaire, à qui je dois énormément, par son investissement tout au long de cette thèse, constant et décisif. C'est tout aussi chaleureusement que je remercie mon directeur universitaire, Nicolas Gayton, qui m'a fait faire mes premiers pas dans la recherche et m'a supervisé en toute confiance : je suis heureux de pouvoir lui présenter aujourd'hui ce mémoire.

Je remercie l'entreprise DeltaMu et particulièrement J.-M. Pou pour la motivation qu'inspire son investissement sans faille et pour le partage de ses idées innovantes. Remerciements sincères aussi à Christophe Dubois, dont la période de direction sur le projet fut très appréciée.

Je remercie tous les autres membres du consortium, entre autres Mamadou, Richard, Axel et Serge, avec qui nous aurons partagé avec enthousiasme ce projet dans nos parties respectives.

Remerciements à tous les collègues de l'UCA, entre autres Jean-Charles, Antoine, Kristof et Guillaume au bureau, Clément, Alexis et Benoît à l'atelier, Benoît Blaysat, François Auslender et Benoît Thuilot pour m'avoir permis de vivre cette riche expérience d'enseignement via l'ATER.

Je remercie le Fonds Européen de Développement Régional ainsi que la Région Auvergne-Rhône-Alpes pour le financement de ces travaux de recherche.

A Dieu ; à la famille dont il m'a comblé.

Résumé

Dans l'industrie manufacturière, la métrologie est indissociable de la prise de décision. Les pièces à produire font l'objet de cotations fonctionnelles qui garantissent, par exemple, la bonne réalisation d'un futur assemblage. Comme la production présente inévitablement une dispersion, il est nécessaire de définir un intervalle de tolérance (un ensemble de valeurs acceptables) et de mesurer les cotes pour garantir qu'elles répondent bien à ces exigences de qualité, c'est-à-dire que les cotes produites appartiennent bien à l'intervalle fixé.

Or, la mesure présente elle aussi une inévitable dispersion. La métrologie est la science du bien-mesurer, mais rien n'est parfait. Les tests inter-opérateurs en témoignent : non seulement la dispersion des mesures est admise, mais l'on cherche à identifier la dispersion propre à chaque opérateur. Les incertitudes qui affectent la mesure proviennent aussi des outils de mesures (résolution, date d'étalonnage), des méthodes de mesures (directe, indirecte) et de l'environnement de la mesure (température, luminosité, poussières, hygrométrie, champ magnétique...).

La mesure obtenue en métrologie n'est donc pas la lecture parfaite de la cote : la mesure n'est pas la valeur recherchée, celle que l'on nomme "valeur vraie". La présence d'erreurs de mesure compromet les décisions fondées sur la mesure : certaines valeurs vraies conformes seront mesurées non-conformes (c'est-à-dire hors intervalle de tolérance), et vice-versa. La qualité de la production ne peut plus être mesurée de façon fiable : les erreurs de chacune des mesures se propagent à l'échantillon et faussent les indicateurs statistiques usuels, tels que les indices de capabilité.

Pour améliorer la qualité des mesures, une première approche est de déployer les meilleurs moyens à disposition pour réduire autant que possible l'erreur de mesure : cet investissement concerne tant les outils de mesure, que la maîtrise de l'environnement de la mesure. Cette approche intuitive tend à obtenir la mesure parfaite. Une autre approche, compatible avec la première et qui est l'objet de cette thèse, est de bien connaître l'erreur de mesure pour mieux en atténuer l'effet.

La recherche de la valeur vraie par l'utilisation de la connaissance de l'erreur de mesure est prévue dans les standards : la mesure est substituée par un intervalle qui contient la valeur vraie. Cet intervalle, dit "de recouvrement", est construit centré sur la mesure avec soit l'écart-type, soit les bornes minimale et maximale de l'erreur de mesure. Or il y a une erreur intrinsèque à cette méthode, au caractère conservatif et qui a donc un coût. Cette thèse propose d'aborder la recherche de la valeur vraie autrement.

Ces travaux de recherche sont portés par un consortium industriel et académique. Le projet est mené et motivé par une entreprise experte en métrologie qui relève le défi de fournir la densité de probabilité de l'erreur de mesure. L'objectif est d'utiliser la loi de l'erreur de mesure pour améliorer les mesures obtenues en métrologie - *a fortiori* la gestion de la production.

Le premier axe d'amélioration traite d'un échantillon de mesures dans son ensemble. Trois méthodes de déconvolution ont été développées spécifiquement pour répondre à la problématique métrologique. Les trois méthodes répondent à trois hypothèses de modélisation concernant la nature de la loi de la valeur vraie recherchée. La déconvolution utilise la connaissance de la loi de l'erreur de mesure pour produire une loi grandement épurée des effets de l'erreur de mesure. Cette nouvelle loi est une meilleure estimation de la loi de la valeur vraie de l'échantillon, donc une meilleure information pour juger la qualité de la production.

Le second axe de l'amélioration traite individuellement chaque mesure de l'échantillon. L'inférence bayésienne est utilisée pour intégrer deux connaissances à la mesure : la loi de l'erreur de mesure et la loi qui caractérise la production. Cette dernière peut être définie arbitrairement (par expertise) ou être le résultat de la déconvolution. Pour chaque mesure de l'échantillon, l'inférence bayésienne débouche sur une loi dite *a posteriori*. Cette nouvelle loi améliore significativement la prise de décision et la gestion de la production, car elle permet d'obtenir une estimation de la valeur vraie de chaque mesure et de connaître la probabilité pour cette mesure d'être hors tolérance.

Les méthodes développées sont appliquées à trois jeux de données construits par simulation. Les trois exemples présentent une difficulté croissante du point de vue de la modélisation. La déconvolution puis l'inférence bayésienne sont appliquées pour chaque exemple. Dans tous les cas, les méthodes proposées permettent une nette amélioration de la mesure et de la maîtrise de la production. Ce gain est évalué par différents indices de qualité et illustré à travers divers outils graphiques, dont deux provenant de la *data science*. Enfin, des tests de sensibilité aux paramètres de simulation (variance, taille de l'échantillon, auto-corrélation et connaissances imparfaites) montrent la robustesse des méthodes proposées.

Mots clés : métrologie, erreur de mesure, déconvolution, inférence bayésienne, conformité.

Abstract

In the manufacturing industry, the decision making regarding compliance relies on metrology. It is necessary to measure the key dimensions to ensure that the produced parts meet the production requirements for use, or for a future assembly. As it is impossible to produce parts with exactly the same dimensions, tolerance intervals are set so that the measure must be inside a given range of values rather than being equal to a single value.

However, despite all the efforts invested in metrology, the measurements are not perfect. There is always a dispersion. The study of inter-operator effects is an example of investigations carried out to identify the causes and with the aim of reducing such errors. The dispersion also comes from the measuring tools (resolution, calibration date), measuring methods (direct, indirect), measurement environment (temperature, luminosity, dust, hygrometry, magnetic fields).

Therefore, the measurements obtained in metrology are riddled with errors. They do not correspond to the exact reading of the key dimensions, referred to as "true values". This uncertainty, associated with the measurement, disturbs the decision making: some of the compliant true values will be read non-compliant (*i.e.* out of the tolerance interval) and, conversely, off-tolerance true values may be read as compliant. It also reduces the accuracy of quality metrics, such as the well-known capability indexes.

To improve the measurement quality, the first approach consists in reducing the measurement error, *i.e.* investing in the best measuring tools and in a better control of the measurement environment. This strategy aims at achieve perfect measurement. A second, compatible and complementary approach, is to get a better knowledge of the measurement error. The focus of this thesis lies on the use of this knowledge to improve the measurements.

The standards provide a method for using measurement error knowledge: it is used to construct a recovery interval, which contains the true value. This interval is used instead of the measurement to establish the compliance of the part. However, this method could be improved. The recovery interval is build using statistical measures (standard deviation, range) and does not use all of the information available; it is centred on the measurement, which is not correct. This thesis proposes another approach to characterize the true value using the available information.

This research is part of an industrial and academic consortium. The overall project was initiated and lead by a company expert in metrology which can carry out the gathering of all the data necessary to build the measurement error distribution. The objective of this thesis is to use this distribution for the improvement of the measures, to provide the best information for decision making.

The first focus of improvement is about entire batches of measurements. Three methods of deconvolution have been specifically developed to answer to the metrological issue. Each method corresponds to a specific modelling of the data. The deconvolution uses the measurement error distribution to filter the batch of measurement. The outcome is a new distribution to characterize the batch of measurements, which is better for decision making.

The second focus of improvement is about each single measurement of the batch. The Bayesian inference is used to integrate two knowledges: the measurement error distribution and the production distribution. The latter can be defined arbitrarily (by expertise) or the deconvolution outcome can be used. For each measurement of the distribution, the Bayesian inference yields the so-called *posterior* distribution. This new distribution can be used to estimate the true value and to calculate the probability for the measurement to be out of the tolerance interval.

The methods developed here are applied to three datasets that offer an increasing difficulty. The data is simulated in order to confront the result to the truth. The deconvolution then the Bayesian inference are applied to each of the three examples. In every cases, the methods appears to clearly improve the measurement and the decision making. Various measures, including graphical tools from the data science, are used to highlights this improvement. Finally, tests are made to show the robustness of the methods to imperfect knowledges, modelling flaws and asses the influence of various values of simulation parameters.

Keywords: metrology, measurement error, deconvolution, Bayesian inference, compliance.

Nomenclature

mesure	Pour améliorer la fluidité de la lecture, le terme "mesure" est utilisé ici pour désigner le "résultat de mesure".
$m^{(i)}$	i-ème mesure de l'échantillon de mesure
$t^{(i)}$	Valeur vraie de l'i-ème mesure
$e^{(i)}$	Erreur de mesure de l'i-ème mesure
$r^{(i)}$	Valeur révisée de l'i-ème mesure
M	Variable aléatoire de la mesure
T	Variable aléatoire de la valeur vraie
E	Variable aléatoire de l'erreur de mesure
IT	Intervalle de tolérance
$\vec{\theta}$	Vecteur des paramètres d'une distribution
L	Score de la vraisemblance
l	Score de log-vraisemblance
π	Densité de probabilité a priori
\mathcal{L}	Densité de probabilité de la vraisemblance
p	Densité de probabilité a posteriori
R^2	Coefficient de détermination
p_{HT}	Probabilité de produire hors-tolérance
FN	Faux-négatifs
FP	Faux-positifs
TP	Vrais-positifs
TN	Vrais-négatifs

Table des matières

Table des figures	15
Liste des tableaux	19
Introduction	21
1 Application de l'approche probabiliste en métrologie	27
Introduction	28
1.1 Définition de la mesure et de ses composantes	29
1.1.1 Standards internationaux et national en métrologie	29
1.1.2 Description de la mesure et de son erreur	30
1.1.3 Approche standard pour l'identification de la valeur vraie	34
1.2 La modélisation probabiliste pour intégrer l'incertitude	36
1.2.1 Les failles et critiques du standard	36
1.2.2 Modèles probabilistes en métrologie	38
1.2.3 Modélisation probabiliste des données métrologiques	40
1.3 Méthodes probabilistes pour améliorer la mesure	42
1.3.1 Recherche de la loi de la valeur vraie	42
1.3.2 Recherche de la valeur vraie d'une mesure	44
Conclusion	47
2 Méthodes pour la correction des erreurs de mesure	49
Introduction	50
2.1 Estimation de la loi de la valeur vraie	51
2.1.1 Cas de variables aléatoires gaussiennes	51
2.1.2 Identification de la loi de la valeur vraie	53
2.1.3 Reconstitution par approximations locales	58
2.2 Recherche de la valeur vraie d'une mesure	61
2.2.1 Application de l'inférence bayésienne	61
2.2.2 Le cas particulier des variables aléatoires gaussiennes	65

2.2.3	Quantification de l'amélioration	67
2.3	Étude de la conformité et des risques associés	69
2.3.1	Lecture de la conformité	69
2.3.2	Probabilités et risques de mal-classifier	74
Conclusion	82
3	Applications et évaluation des méthodes probabilistes	85
Paramétrage de la simulation	86
3.1	Application à des variables aléatoires gaussiennes	87
3.1.1	Paramétrage et données d'entrée	87
3.1.2	Application et résultats de la déconvolution	87
3.1.3	Application et résultat de l'inférence bayésienne	91
3.2	Application des méthodes pour une loi de la valeur vraie identifiable	95
3.2.1	Paramétrage et données d'entrée	95
3.2.2	Application de la déconvolution	95
3.2.3	Application et résultat de l'inférence bayésienne	100
3.3	Application des méthodes pour des données quelconques	103
3.3.1	Paramétrage et données d'entrées	103
3.3.2	Application et résultat de la déconvolution	104
3.3.3	Application et résultat de l'inférence bayésienne	107
Conclusion	112
4	Sensibilité des résultats aux paramètres de l'étude	113
Introduction	114
4.1	Impact d'une loi de la valeur vraie non identifiable	115
4.2	Impact d'une loi de l'erreur de mesure mal identifiée	123
4.3	Impact de mesures auto-corrélées	129
4.4	Étude de l'impact de la taille de l'échantillon de mesures	132
4.4.1	Cas de variables aléatoires gaussiennes	132
4.4.2	Déconvolution de données quelconques	135
4.4.3	Discrétisation de la courbe ROC	137
4.5	Impact du ratio des écarts-types de la valeur vraie et de l'erreur de mesure	138
Conclusion	140
Conclusion & Perspectives		141
Annexes		148

A Etudes analytiques	149
A.1 Expression analytique de la distribution de la valeur vraie dans le cas gaussien . . .	150
A.2 Expression de la valeur révisée dans le cas gaussien	150
A.3 Extension symétrique de l'intervalle de sélection et corrélation négative entre FN et FP	151
A.4 L'inférence bayésienne dans le cas gaussien est une extension de l'intervalle de tolérance	153
A.5 La valeur vraie est toujours comprise entre la moyenne de la valeur vraie et la mesure quand la moyenne de l'erreur est nulle	154
A.6 Détails concernant la transformation de Nataf et l'identification des points d'inté- gration et des poids de la quadrature de Gauss-Hermite	155
A.7 L'augmentation de la moyenne de l'erreur de mesure ne détériore que le coefficient de détermination des mesures	155
B Etudes graphiques	157
B.1 Comportement en zone non-homogène	158
B.2 Population avantagée par la révision	163
C Descriptions des lois	167
Bibliographie	171

Table des figures

1	Logos des acteurs du projet.	24
1.1	Logos des différentes autorités nationales et internationales en métrologie. Celui du BIPM apparaît en fond des premières pages des documents GUM et VIM.	30
1.2	Illustrations des valeurs vraies d'une mesurande. A cause des imperfections que présente une pièce issue de la production, toutes les valeurs vraies ne sont pas représentatives de la grandeur recherchée.	33
1.3	Illustration de l'application de l'intervalle de recouvrement.	35
1.4	Illustrations de l'inversion du concept dans la construction de l'intervalle de recouvrement.	37
1.5	Illustrations de la mauvaise identification de l'intervalle de recouvrement.	38
1.6	L'erreur de mesure est décrite par une variable aléatoire.	39
1.7	Les valeurs vraies issues de la production sont décrites par une variable aléatoire.	40
1.8	Schéma des mesures à risques. Les mesures dont la conformité est brouillée par l'erreur de mesure sont celles proches des bornes de l'intervalle de tolérance.	41
1.9	Schéma du concept de bande de garde.	41
2.1	Exemples de fonctions de répartition de différentes lois. A gauche, trois exemples de lois de moyenne 0 et d'écart-type 1. A droite, leurs fonctions des répartitions respectives.	57
2.2	Schéma de la seconde méthode de déconvolution, pour deux familles candidates.	58
2.3	Illustration d'un mélange de lois pour le cas du moulage multi-empreintes.	59
2.4	Schéma de la construction de la loi de vraisemblance.	62
2.5	Trois exemples d'application.	64
2.6	Le risque maximum (dans le cas gaussien)	66
2.7	Exemple de matrice de confusion de pièces triées par une machine.	70
2.8	Disposition des termes de la matrice de confusion (classification binaire).	71
2.9	Schéma de la lecture graphique de la conformité.	71
2.10	Illustration des ratios utilisés pour tracer la courbe ROC.	72
2.11	Différents axes et pôles du graphique de la courbe ROC.	74
2.12	Schéma des combinaisons possibles.	76

2.13	Exemple de courbes de probabilité de prendre une mauvaise décision (lois uniformes).	77
2.14	Exemple d'une mesure dans la zone non homogène.	78
2.15	Exemple de courbes des risques d'erreur (lois uniformes).	79
2.16	Illustration du risque total.	81
3.1	Déconvolution de la loi des mesures obtenues en métrologie.	87
3.2	Probabilités qu'une mesure soit mal classifiée.	89
3.3	Risque qu'une mesure soit mal classifiée.	90
3.4	Comparaison des résidus des mesures et des valeurs révisées.	91
3.5	Classification de l'échantillon.	93
3.6	Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas gaussien).	93
3.7	Tracé de la courbe ROC (cas gaussien).	94
3.8	Score BIC du représentant de chaque famille candidate (cas log-normal).	96
3.9	Déconvolution de la loi des mesures obtenues en métrologie (cas log-normal).	97
3.10	Probabilités qu'une mesure soit mal classifiée (cas log-normal).	98
3.11	Risque qu'une mesure soit mal classifiée (cas log-normal).	99
3.12	Comparaison des résidus des mesures et des valeurs révisées (cas log-normal).	100
3.13	Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas log-normal).	101
3.14	Classification de l'échantillon (cas log-normal).	102
3.15	Tracé de la courbe ROC (cas log-normal).	103
3.16	Déconvolution de la loi des mesures obtenues en métrologie (cas multimodal).	104
3.17	Probabilités qu'une mesure soit mal classifiée (cas multimodal).	106
3.18	Risque qu'une mesure soit mal classifiée (cas multimodal).	107
3.19	Comparaison des résidus des mesures et des valeurs révisées (cas multimodal).	108
3.20	Classification de l'échantillon (cas multimodal).	109
3.21	Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas multimodal).	110
3.22	Tracé de la courbe ROC (cas multimodal).	111
4.1	Score BIC du représentant de chaque famille candidate (cas test).	116
4.2	Déconvolution de la loi des mesures obtenues en métrologie (f_T non identifiable).	116
4.3	Probabilités qu'une mesure soit mal classifiée (f_T non identifiable).	118
4.4	Risque qu'une mesure soit mal classifiée (f_T non identifiable).	119
4.5	Comparaison des résidus des mesures et des valeurs révisées (f_T non identifiable).	120
4.6	Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (f_T non identifiable).	121

4.7	Classification de l'échantillon (f_T non identifiable).	122
4.8	Tracé de la courbe ROC (f_T non identifiable).	123
4.9	Tracé des différentes lois erronées de l'erreur de mesure.	125
4.10	Impact d'une mauvaise identification de la loi de l'erreur de mesure lors de la déconvolution de données gaussiennes.	127
4.11	Nuage de points des résidus de valeurs révisées prises parmi les 50 lots générés.	128
4.12	Impact de l'auto-corrélation de l'échantillon de mesures sur la déconvolution de lois normales.	131
4.13	Impact de la taille de l'échantillon de mesures sur la déconvolution de lois normales.	133
4.14	Distribution des scores R^2 des mesures et valeurs révisées en fonction de la taille d'échantillon.	135
4.15	Impact de la taille de l'échantillon de mesures sur la déconvolution de lois normales.	136
4.16	Courbe ROC pour un échantillon de taille 10^4 (cas gaussien).	138
4.17	Évolution du score R^2 en fonction du rapport des écarts-types (application au cas gaussien).	139
B.1	Illustration des effets de la zone non homogène (1/4).	159
B.2	Illustration des effets de la zone non homogène (2/4).	160
B.3	Illustration des effets de la zone non homogène (3/4).	161
B.4	Illustration des effets de la zone non homogène (4/4).	162
B.5	Illustration de l'influence de l'inférence bayésienne (cas gaussien).	164
B.6	Isolement des différentes catégories de données.	165
C.1	Exemple d'une loi de moyenne 0 et d'écart-type 1 pour chacune des fonctions candidates.	169

Liste des tableaux

1	Membres du consortium.	23
2.1	Paramétrage de l'exemple.	76
2.2	Synthèse des méthodes probabilistes utilisées.	84
3.1	Paramétrage des lois d'entrées.	87
3.2	Précision de l'identification des paramètres.	88
3.3	Comparaison des probabilités de produire hors tolérance (p_{HT}).	88
3.4	Risque total des différentes sélections.	90
3.5	Indices de qualité des mesures et des valeurs révisées.	92
3.6	Bornes de l'intervalle de tolérance équivalent.	92
3.7	Paramétrage des lois d'entrées (cas log-normal).	95
3.8	Évaluation de la précision de l'identification (cas log-normal).	97
3.9	Comparaison des probabilités de produire hors tolérance (p_{HT}) (cas log-normal).	97
3.10	Risque total des différentes sélections (cas log-normal).	99
3.11	Indices de qualité des mesures et des valeurs révisées (cas non gaussien).	101
3.12	Définitions des lois (cas multimodal).	104
3.13	Évaluation de la précision de la reconstruction (cas multimodal).	105
3.14	Comparaison des probabilités de produire hors tolérance (p_{HT}) (cas multimodal).	105
3.15	Risque total des différentes sélections (cas multimodal).	107
3.16	Indices de qualité des mesures et des valeurs révisées (cas multimodal).	108
4.1	Paramètres des lois des données d'entrées.	115
4.2	Évaluation de la précision de l'identification (f_T non identifiable).	117
4.3	Comparaison des probabilités de produire hors tolérance (p_{HT}) (f_T non identifiable).	117
4.4	Risque total des différentes sélections (f_T non identifiable).	119
4.5	Indices de qualité des mesures et des valeurs révisées (f_T non identifiable).	120
4.6	Synthèse des scores R^2 pour les différentes valeurs de F	129
4.7	Synthèse des scores R^2 obtenus avec les échantillons auto-corrélés.	132
4.8	Moyenne et écart-type des R^2 en fonction de la taille d'échantillon.	134

C.1 Définition des lois choisies. Les densités de probabilité prennent une valeur nulle en dehors du support de la variable.	168
--------------------------------------------------------------------------------------------------------------------------------------	-----

Introduction

Problématique

Dans l'industrie manufacturière, les pièces produites sont pensées et dessinées par des ingénieurs qui utilisent aujourd'hui des logiciels de conception assistée par ordinateur (CAO). La description géométrique de la pièce est établie théoriquement, les valeurs des cotes dimensionnelles sont définies avec exactitude. Or, il y a nécessairement une dispersion de production. Il est impossible de produire une pièce qui corresponde parfaitement aux cotes dimensionnelles espérées. Il est aussi impossible de produire exactement deux fois la même pièce. Le dessin des pièces réalisé sur ordinateur s'accompagne ainsi de la définition d'intervalles de tolérance. Un intervalle de tolérance est associé à chaque cote. L'exigence de ces intervalles de tolérance varie : les cotes les plus importantes sont dites fonctionnelles et requièrent une grande précision (elles seront par exemple sollicitées dans un assemblage, ou soumises à une stricte réglementation de sécurité). L'industriel qui s'engage à fabriquer la pièce en question doit garantir que sa production répondra à ces exigences.

C'est la mesure qui est gage de la conformité de la pièce par rapport aux exigences de production. Le résultat de mesure - que l'on désignera ici sobrement par mesure, est la lecture de la cote. Or, le bien-mesurer est toute une science : c'est la métrologie. Il faut maîtriser l'environnement de la mesure, en sachant que le domaine d'action de la métrologie est très vaste (de la métallurgie à la cosmétique). Il faut une connaissance précise des différents outils de mesures déployés, qui sont eux-mêmes de plus en plus précis et complexes (robots multi-axiaux, mesures 3D, mesures optiques...) [139], et suivre avec précaution leurs fréquences d'utilisation et d'étalonnage. La précision et la fiabilité du résultat demande généralement plusieurs mesures pour s'assurer de la valeur d'une cote. Enfin, le résultat attendu est parfois la combinaison de plusieurs mesures, comme pour la mesure d'un volume ou d'une surface. Les travaux de cette thèse se limitent à la métrologie unidimensionnelle et uni-variée (telle la mesure d'une tige droite dont seule la longueur importe), et une seule mesure est associée à seule cote (dont la valeur peut être associée à la moyenne de trois mesures, par exemple). En pratique, toutes les cotes ne font pas nécessairement l'objet d'une mesure, et toutes les pièces d'un lot de pièces ne sont pas nécessairement contrôlées.

Or, il y a nécessairement une dispersion des mesures obtenues en métrologie. La mesure est toujours entachée d'une erreur de mesure. La complexité croissante des outils déployés en métrologie témoigne de ce constat inéluctable. Il existe tout simplement des limites technologiques et humaines aux processus et méthodes de mesures qui ne permettent évidemment pas de viser la perfection. Ainsi, les erreurs de mesure ne peuvent qu'être admises.

Les ingénieurs et métrologues s'efforcent de réduire au maximum les erreurs de mesure. Ces dernières se propagent à d'autres indicateurs de la maîtrise statistique des procédés [138, 106] - comme les capacités C_p et C_{pk} ([82, 35, 122, 88]), et à d'autres méthodes d'évaluation de qualité de la production ([135]) qui se fondent sur les mesures pour juger de la qualité de la production. Les plus grands investissements permettront de déployer les solutions techniques de mesure les plus précises de la métrologie. Mais la capacité à mesurer un grand nombre de pièces d'un lot et celle de réaliser plusieurs mesures par cote demandent des moyens importants, qui sont parfois impossibles à mettre en place. Les mesures sont parfois coûteuses (obtenues par contrôles destructifs ou limitées pour des questions de cadence). Il se peut aussi que la cote à mesurer ne soit pas accessible aux outils de mesures les plus précis et qu'il faille en combiner d'autres pour l'estimer (ce qui détériore la qualité des mesures puisque la combinaison de mesures implique l'accumulation de leurs incertitudes). La minimisation de l'erreur de mesure est indubitablement un besoin, mais n'est cependant pas suffisante en soi : même si l'ordre de grandeur des erreurs de mesure est faible devant celui des mesures, il existe toujours un risque que la lecture de la conformité soit faussée (spécifiquement pour les pièces dont la mesure est à la frontière de la conformité). Une mauvaise mesure peut discréditer la pièce alors qu'elle est opérationnelle et la reconduire vers des usinages supplémentaires voire à être rebutée ; ou bien faire croire à tort que la pièce respecte les exigences de production, et les conséquences seront alors subies par le client qui la réceptionnera.

Les ingénieurs et métrologues cherchent à maîtriser les erreurs de mesure : le problème serait en quelque sorte résolu si la part de l'erreur de mesure dans la mesure était parfaitement connue. A défaut de pouvoir connaître exactement l'erreur de mesure pour une mesure donnée, il s'agit de l'estimer le mieux possible pour en maîtriser les conséquences induites, ce qui demande d'identifier au mieux les différents facteurs et lois phénoménologiques qui interviennent [64]. Le diagramme Ishikawa par exemple [27], propose un cadre structuré pour un ciblage pertinent des potentielles causes d'erreurs de mesure. Pour certaines composantes de l'erreur de mesure, cette connaissance aboutit à des indicateurs statistiques : la moyenne, les valeurs maximales ou encore l'écart-type. Les causes humaines (l'effet inter-opérateur par exemple) introduisent une dispersion dont on ne peut guère observer que la variance. Pour les causes physiques, il est parfois possible d'obtenir des informations plus précises sur l'environnement de la mesure, par des capteurs : la température et/ou l'humidité par exemple. Mais encore, de prédire leur impact sur la mesure : les formules qui lient la dilatation des matériaux à la température et à l'hygrométrie sont connues. Un graphique des différentes causes d'incertitudes a été dressé par Pananias [111]. A l'heure de l'"Industrie 4.0", du Big Data et des plates-formes connectées [109], où les algorithmes prennent une importance croissante [38], où la puissance des ordinateurs et le débit de données font rarement défaut, où les capacités de stockage sont toujours plus grandes, et où les parcs industriels sont inondés par les données de leurs nombreux capteurs, la description de l'environnement, bien que difficile, semble à portée de main. Pour cette seconde approche, il est possible d'admettre qu'une bonne connaissance de l'erreur de mesure est accessible, et la question est alors de savoir comment l'utiliser.

L'objectif global de cette thèse est d'exploiter la connaissance de la densité de probabilité de l'erreur de mesure pour améliorer les mesures obtenues en métrologie, et par extension la gestion de la production.

Contexte

Cette thèse est financée par le FEDER et la Région Auvergne-Rhône-Alpes et l'encadrement sur le plan universitaire est rattaché à l'école d'ingénierie SIGMA Clermont. Elle s'inscrit dans le projet "We Are Metrology" (WAM) [32] porté par un consortium clermontois dont les membres sont énumérés dans le Tableau 1. L'ensemble de ces acteurs sont représentés dans la Figure 1.

Le consortium vise à fédérer la communauté des métrologues autour d'une plate-forme, qui contient les outils classiques de la métrologie ainsi que d'autres plus innovants. Les travaux présentés dans cette thèse synthétisent spécifiquement la recherche et le développement des méthodes métrologiques inédites qui pourront être proposées sur la plate-forme.

L'objectif commun aux différents acteurs est de produire une preuve de concept (Proof Of Concept - POC). Les travaux qui ont été initiés lors de cette thèse sont les premières bases de ce projet ; ils ne s'appuient pas sur de précédents résultats.

TABLEAU 1: Membres du consortium.

Entités	Type	Spécialité particulièrement sollicitée
DeltaMu	Entreprise	Métrologie
Phimeca Engineering	Entreprise	Calculs d'incertitudes
PerfectMemory	Entreprise	Sémantique, Deep Learning
Limos	Laboratoire de recherche	Informatique
Clerma	Laboratoire de recherche	Management

Les besoins d'innovation et d'émulation dans la métrologie ont été identifiés par l'entreprise DeltaMu dont le cœur de métier est de diffuser une métrologie "avancée" [115] (*Smart Metrology*). Les derniers outils statistiques qu'ils suivent de près, à l'instar de Jean-Michel Pou son président fondateur et membre de l'Association Française de NORmalisation (AFNOR) « Métrologie », sont bien souvent absents de la réalité pratique. Actuellement, leur logiciel phare "Optimu" [31] permet entre autres de modéliser et de synthétiser les différentes sources d'erreurs de mesure. L'expérience et le savoir-faire de l'entreprise de métrologie ainsi que certaines méthodes de la littérature [132, 143] assurent la crédibilité de l'hypothèse clef de ces travaux : la connaissance de la loi qui décrit l'erreur de mesure.

Sur tous les enjeux que renferme le projet, l'entreprise Phimeca Engineering recouvre la partie statistique et méthodes numériques. Les variables aléatoires et le calcul d'incertitudes sont l'objet d'une de ses expertises, à l'image du module Python *Openturns* [7] qui est activement développé depuis plus de quinze ans. L'objectif est d'utiliser les outils probabilistes pour l'estimation des valeurs vraies à partir des mesures et de la loi de l'erreur de mesure.



FIGURE 1: Logos des acteurs du projet.

Plan du mémoire

Les travaux s'organisent en trois chapitres :

- Le premier chapitre commence par exposer la notion de mesure, telle que donnée par les organismes faisant autorité aujourd'hui en métrologie. Il explicite aussi les définitions de ses deux composantes : l'erreur de mesure et la valeur vraie. Il introduit ensuite la modélisation de la mesure, et les méthodes qui l'accompagnent pour la gestion de l'erreur de mesure. Dans un second temps, ce premier chapitre montre les critiques qui peuvent être faites aux définitions, à la modélisation et aux méthodes du standard. Il propose de modéliser de façon probabiliste l'erreur de mesure, de la valeur vraie et de la mesure, ainsi que d'utiliser des outils probabilistes pour l'amélioration des mesures. L'amélioration des mesures est déclinée

sous deux facettes : l'amélioration de la lecture statistique d'un échantillon de mesures entier, et l'amélioration de toutes les mesures de l'échantillon prises individuellement. La dernière partie de ce chapitre est dédiée à la présentation de méthodes probabilistes qui répondent à ces deux objectifs. Elle est principalement focalisée sur **les méthodes de déconvolution pour traiter l'échantillon tout entier** et sur **l'inférence bayésienne pour l'amélioration individuelle des mesures**.

- Le deuxième chapitre présente en détails les solutions adoptées et développées pour l'amélioration des mesures, puis les outils utilisés pour exploiter les nouvelles informations pour la gestion des risques de production. La première partie traite de la déconvolution et recouvre trois méthodes répondant à trois modélisations du problème : le cas de variables aléatoires suivant des lois soit normales, soit paramétriques et non gaussiennes, soit quelconques. La seconde partie présente l'inférence bayésienne appliquée dans le contexte métrologique et les outils statistiques qui en quantifient le gain. La dernière partie de ce chapitre montre les bénéfices des deux méthodes d'amélioration sur la lecture de la conformité et la gestion des risques de production. Cette section présente plusieurs outils graphiques, dont la matrice de confusion et la fonction d'efficacité du récepteur qui proviennent de la *data science*.
- Le troisième chapitre a pour but d'illustrer le gain apporté par les méthodes d'amélioration développées dans ces travaux ainsi que leurs robustesses. Les trois premières sections présentent chacune un exemple d'application et en analysent les résultats. La dernière section étudie l'impact de différents paramètres de la simulation. Les différents exemples d'application sont tous construits avec des données simulées pour connaître les valeurs vraies. L'évaluation de l'amélioration met en jeu les différents outils exposés dans le chapitre deux. Chaque exemple des trois premières sections de ce chapitre correspond à une modélisation adéquate pour l'application d'une méthode de déconvolution. L'application de la déconvolution est toujours suivie de l'application de l'inférence bayésienne, utilisant pour *a priori* la loi déconvoluée. La quatrième et dernière section évalue la robustesse des méthodes : elle cible particulièrement l'impact de connaissances imprécises (lois valeur vraie et de l'erreur de mesure), des valeurs relatives des dispersions de l'erreur de mesure et de la valeur vraie, de la taille de l'échantillon de mesures et enfin de données auto-corrélées.

Chapitre 1

Application de l'approche probabiliste en métrologie

Points clefs

- Enjeux autour de la définition de "mesure"
 - Gestion de l'erreur de mesure
 - Les solutions probabilistes au problème posé
-

Sommaire

Introduction	28
1.1 Définition de la mesure et de ses composantes	29
1.1.1 Standards internationaux et national en métrologie	29
1.1.2 Description de la mesure et de son erreur	30
1.1.3 Approche standard pour l'identification de la valeur vraie	34
1.2 La modélisation probabiliste pour intégrer l'incertitude	36
1.2.1 Les failles et critiques du standard	36
1.2.2 Modèles probabilistes en métrologie	38
1.2.3 Modélisation probabiliste des données métrologiques	40
1.3 Méthodes probabilistes pour améliorer la mesure	42
1.3.1 Recherche de la loi de la valeur vraie	42
1.3.2 Recherche de la valeur vraie d'une mesure	44
Conclusion	47

Introduction

La définition de la mesure est aussi complexe qu'intuitive. Une mesure est communément perçue comme la lecture d'une information recherchée mais imparfaite : une valeur vraie mais associée à une erreur de mesure. La définition précise de la mesure renferme les définitions de ces deux notions et leur imbrication. Définir clairement la mesure, c'est amorcer la modélisation et expliciter les enjeux autour de son amélioration. L'amélioration de la mesure peut être décrite comme la recherche de la valeur vraie [11]. En admettant qu'il y aura toujours une certaine erreur de mesure, la recherche de la valeur vraie doit s'appuyer sur l'identification de l'erreur de mesure. La qualité de la recherche est à la fois fonction de la finesse de la modélisation et de ce qui est connu de l'erreur de mesure. Deux approches sont présentées pour améliorer la mesure. La première est l'application des standards actuels, tant pour la modélisation que pour les méthodes. La seconde est une alternative fondée sur une modélisation probabiliste pour laquelle les méthodes relèvent encore du domaine de la recherche. L'approche standard utilise une description statistique de l'erreur de mesure pour la construction d'un intervalle qui contient la valeur vraie. L'approche probabiliste a pour but d'améliorer les performances du standard. Pour cela la notion d'incertitude est élargie à toutes les composantes. Ce chapitre présente des méthodes probabilistes pour la gestion des incertitudes. Les méthodes dépendent fortement des données disponibles pour caractériser l'erreur de mesure. Or, ce qui distingue particulièrement ces travaux est que la densité de probabilité de l'erreur de mesure est connue. C'est une hypothèse exigeante et donc traitée dans la littérature. La dernière section donne une vue globale des méthodes probabilistes de la littérature relatives à l'amélioration de données.

1.1 Définition de la mesure et de ses composantes

1.1.1 Standards internationaux et national en métrologie

La problématique qui motive les travaux de cette thèse concerne la métrologie, et les méthodes statistiques et probabilistes qui sont abordées ici ne pourraient être dissociées de ce domaine. Sans remonter aux origines de la métrologie (et le fait même de mesurer ne saurait être daté), il est important d'en rappeler la définition car c'est aussi là que les controverses commencent. Ce que représente réellement une mesure est une question complexe qui nourrit différents angles de vue. La réponse a évolué et possède aujourd'hui une longue histoire, entretenant la seule certitude qu'elle n'est pas aussi triviale qu'intuitive. La solution apportée par ces travaux s'adresse à des problématiques indissociables des concepts métrologiques contemporains en vigueur.

Autour de la notion de mesure, il faut expliciter les concepts et poser les bons mots sur les concepts, mais encore garantir la cohérence du tout. Ces éléments sont toujours aujourd'hui au cœur du débat. Car de même que les mots orientent la pensée, c'est la description d'un problème qui en façonne la réponse. Une fois les mots posés, il faut encore les traduire dans les différentes langues. La cohérence est un besoin évident qui se réalise dans la logique : certes, ce n'est pas encombré du sens des mots, mais c'est un objectif d'autant plus difficile à atteindre qu'il y a un grand nombre d'éléments en jeu.

Quand une réponse scientifique est apportée, il faut encore que les hommes s'accordent et la diffusent pour qu'elle ait réellement un effet ; tout comme une langue, aussi riche, complexe et cohérente qu'elle est, ne serait plus d'aucun intérêt si elle n'est plus connue que d'un seul homme. L'histoire est riche d'une multitude de solutions adoptées uniquement localement, par certains peuples et dans certaines zones géographiques. Dans le contexte actuel, où tout a pris une dimension mondiale, il faut resituer quels sont les acteurs et textes qui font loi.

La convention du mètre signée à Paris en 1875 oriente encore aujourd'hui l'organisation fonctionnelle des institutions internationales de métrologie et s'exerce par le biais de trois organismes : le Bureau International des Poids et Mesures (BIPM), la Conférence générale des poids et mesures (CGPM) et le Comité International des Poids et Mesures (CIPM). Les deux derniers ont une mission plus spécialement dirigée vers la définition des systèmes internationaux d'unités et leur publication, un sujet qui n'affecte pas directement ces travaux. Le BIPM lui héberge le Comité commun pour les guides en métrologie, plus connu sous son acronyme anglais "JCGM", qui "est chargé de maintenir et de promouvoir l'utilisation du Guide pour l'expression de l'Incertitude de Mesure (GUM) [9] et du Vocabulaire International des termes fondamentaux et généraux de Métrologie (VIM) [11]. Le JCGM succède à l'ISO TAG 4 qui avait la responsabilité de publier le GUM et le VIM sous la supervision du BIPM, de l'IEC, de l'IFCC, de l'Organisation Internationale de Normalisation (International Organization for Standardization, ISO), de l'Union internationale de chimie pure et appliquée (International Union of Pure and Applied Chemistry, IUPAC), de l'Union internationale de physique pure et appliquée (International Union of Pure and Applied Physics, IUPAP) et de l'Organisation internationale de métrologie légale (International Organization of Legal Metrology, OIML)". Les états membres et associés au BIPM sont présents sur tous les continents et représentent une très large majorité des pays dans le monde.

Le VIM et le GUM sont les deux documents techniques clefs pour ces travaux, qui apparaissent au bout d'une chaîne dont l'autorité découle des instances majeures de la métrologie mondiale. A l'image du VIM qui est publié dans l'ISO/IEC GUIDE 99:2007 [73], on peut légitimement affirmer que ce sont les documents qui font aujourd'hui loi dans la mise en pratique de la métrologie. Ce

statut est assumé par le VIM et est partie intégrante de son ambition : "Le Guide ISO/CEI 99:2007 se propose d'être une référence commune pour les scientifiques et les ingénieurs" [73]. En étant à l'interface de la théorie et des applications, ces documents doivent répondre à une problématique de pédagogie et d'accessibilité dont ces travaux héritent directement ; ils seront dans tous les cas mis à l'épreuve par des contraintes financières, de savoir ou de temps [54, 129].

Garantir l'application du vocabulaire, des concepts et des méthodes de ces deux documents fait l'objet d'accréditations, comme en témoigne l'activité du Comité français d'accréditation (Cofrac) en France. Le Cofrac est une association à but non lucratif et est l'unique instance nationale reconnue comme "organisme national d'accréditation". La notoriété du VIM et du GUM transparait dans la volonté des laboratoires ou des entreprises d'obtenir une accréditation. On pourra retrouver dans la publication de Mallet [100] le témoignage d'un laboratoire qui en énumère la liste des motivations, de l'exigence pour s'y conformer ainsi que les changements notables que cela a impliqué. Une critique est notamment portée par Coton au sujet de l'accréditation [28].



FIGURE 1.1: Logos des différentes autorités nationales et internationales en métrologie. Celui du BIPM apparaît en fond des premières pages des documents GUM et VIM.

1.1.2 Description de la mesure et de son erreur

Le VIM et GUM servent ainsi de référence pour la définition métrologique du problème. De même qu'il faut assurer une traçabilité métrologique (par exemple que les étalons des étalons découlent bien du prototype international du dépôt souterrain du Pavillon de Breteuil de 1889), il faut aussi se raccorder à une référence commune pour le vocabulaire et les méthodes métrologiques. Les discussions que nourrissent le VIM et le GUM dans la communauté scientifique, même critiques,

attestent leur notoriété. La dernière version du Guide pour l'expression de l'incertitude de mesure est parue en 2008 [9] et sa sixième partie nommée "Élaboration et utilisation des modèles de mesure" a été publiée en 2020. La dernière version du Vocabulaire international des termes fondamentaux et généraux de métrologie est parue de 2012 [11], c'est une révision de la troisième édition parue en 2008 et qui comprend des corrections importantes [14].

Parmi les définitions du VIM, il est intéressant dans le cadre de cette étude de relever celles qui sont en lien avec la notion de mesure. On trouve notamment dans [11] que :

- la métrologie "est la science des mesurages et ses applications". Et le VIM comprend une note, en paratexte : qu'elle "comprend tous les aspects théoriques et pratiques des mesurages, quels que soient l'incertitude de mesure et le domaine d'application." ;
- le mesurande est la "grandeur que l'on veut mesurer" ;
- la mesure est le "processus consistant à obtenir expérimentalement une ou plusieurs valeurs que l'on peut raisonnablement attribuer à une grandeur".

On notera que la notion de résultat de mesure est distincte de celle de mesure :

- "Le résultat d'un mesurage est l'ensemble de[s] valeur[s] attribuée[s] à un mesurande, complété par toute autre information pertinente disponible" ;

Enfin :

- enfin "La valeur vraie d'une grandeur est la valeur d'une grandeur compatible avec la définition de la grandeur". En note, il est écrit que "Dans l'approche « erreur » de description des mesurages, la valeur vraie est considérée comme unique et, en pratique, impossible à connaître. L'approche « incertitude » consiste à reconnaître que, par suite de la quantité intrinsèquement incomplète de détails dans la définition d'une grandeur, il n'y a pas une seule valeur vraie mais plutôt un ensemble de valeurs vraies compatibles avec la définition."

Ces définitions synthétisent efficacement les notions à la base de ce travail de thèse et cristallisent les différentes problématiques qui seront rencontrées. Par exemple, une notion peut être définie rigoureusement mais présenter une incertitude dans son application, comme c'est le cas pour le mesurande (la mesure réalisée peut ne pas être représentative de la grandeur mesurée). Une définition peut contenir en soi une incertitude, comme celle du résultat de mesure qui comporte d'"autres informations pertinentes autour" : il y a de l'incertitude dans l'appréciation des informations pertinentes ou non. Enfin, à l'instar de la note de la valeur vraie, il peut y avoir une incertitude au sujet de la définition même : la proposition n'est pas définitive et comporte plusieurs alternatives, laissées au choix du lecteur selon l'"approche" qu'il adoptera. Toutes ces incertitudes identifiées dans le vocabulaire international des termes fondamentaux et généraux de métrologie [11] réapparaîtront nécessairement dans la mise en équation de la mesure.

L'équation la plus élémentaire du GUM pour la modélisation d'une mesure entachée d'erreurs se retrouve dans la littérature métrologique [101, 130, 63] et est relayée dans les documents de la norme [73, 72] comme dans le livre "Processus de mesure : évaluer les incertitudes - 20 exemples" [27]. Ce dernier est un guide, rédigé par le Collège Français de Métrologie (CFM), ayant pour vocation "de présenter les *bonnes pratiques mises en place ou en cours de développement en entreprise dans des domaines très variés*". L'équation fondamentale sur laquelle repose cette thèse est :

$$m = t + e \tag{1.1}$$

avec m la mesure et t la "valeur vraie" qui aurait été mesurée si la mesure n'était pas faussée par une erreur de mesure e . On remarquera que le nom "mesure" utilisé pour la variable m correspond au "résultat du mesurage" selon le VIM. Cette modélisation retranscrit de façon intuitive le lien entre les concepts en jeu : lire une mesure, c'est lire la grandeur en plus d'une erreur. En contrepartie de la simplicité de cette équation fondamentale, la définition des deux composantes de la mesure est complexe.

La grandeur t est souvent nommée "valeur vraie" [12, 9, 11, 40, 79] et c'est ainsi qu'elle est désignée dans ce document. La mention "vraie" de son nom indique qu'elle est la mesure sans l'erreur de mesure, mais ne garantit pas qu'elle soit la grandeur recherchée. Un micromètre, par exemple, permet de lire la distance entre deux touches de mesure : quand bien même cette mesure serait parfaite (sans erreur de mesure), cela ne garantit pas que la quantité mesurée décrit exactement le mesurande. Cela peut être perçu comme la remontée fiable d'une information dont il reste encore à vérifier la pertinence.

La notion de valeur vraie, pourtant intuitive, est parfois difficile à définir en pratique. Les différentes descriptions géométriques se heurtent aux imperfections de la réalité. Comme schématisé en Figure 1.2, un tube, par exemple, n'est jamais parfaitement cylindrique. L'écart de forme est une des imperfections qui peuvent être observées à différentes échelles : les ondulations, la rugosité et la micro-rugosité. Cette question des imperfections géométriques est traitée avec précision dans les travaux de Goka [58]. Les différents défauts de forme empêchent de définir avec précision le diamètre du tube. Cette réalité fait l'objet de toute une expertise dans la définition des cotations fonctionnelles autour de profils anamorphosés : on substitue à une réalité chaotique un modèle théorique (on se réfère par exemple à l'enveloppe) pour exprimer les spécifications du tolérancement (au minimum ou au maximum de matière par exemple [12]). En conséquence d'un modèle géométriquement imparfait, il faut donc discerner que toutes les valeurs vraies ne représentent pas la grandeur recherchée.

Priel, directeur honoraire de la Métrologie (Laboratoire National de métrologie et d'Essais, LNE), a rapporté que la définition de la valeur vraie était proche d'être changée dans la troisième édition du VIM [116]. Il faut ainsi s'attendre à ce qu'elle évolue dans un avenir proche. La nature radicalement différente de l'alternative proposée en note de la définition de la valeur vraie atteste que le concept est actuellement discuté, et pourrait bien révéler l'orientation des changements à venir.

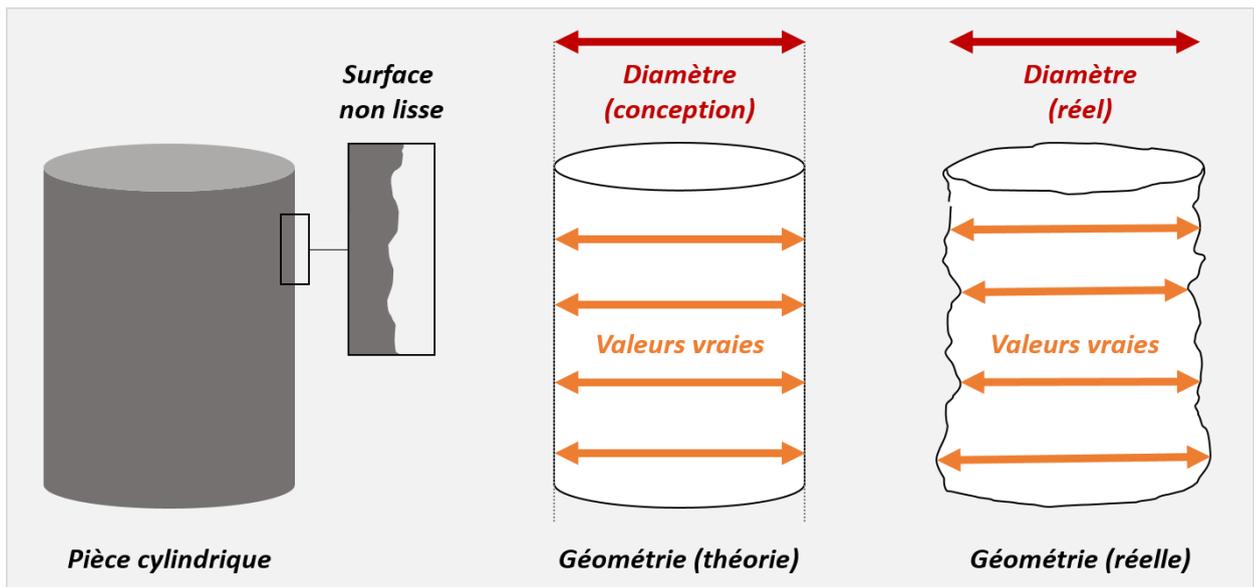


FIGURE 1.2: Illustrations des valeurs vraies d'une mesurande. A cause des imperfections que présente une pièce issue de la production, toutes les valeurs vraies ne sont pas représentatives de la grandeur recherchée.

Le terme e de l'Équation (1.1) est l'erreur de mesure. La définition du VIM indique :

- "L'erreur de mesure [est la] différence entre la valeur mesurée d'une grandeur et une valeur de référence"

L'erreur de mesure est donc l'écart entre la mesure et la valeur vraie. Si un thermomètre affiche une température ambiante de $20,8^{\circ}\text{C}$ et qu'un opérateur retranscrit erronément $21,0^{\circ}\text{C}$, une erreur de mesure de $+0,2^{\circ}\text{C}$ est commise. Si l'instrument avait sous-estimé la température de $0,5^{\circ}\text{C}$, l'erreur de mesure est au final de $0,3^{\circ}\text{C}$. L'erreur de mesure est donc une grandeur scalaire, qui est le résultat de la combinaison de tous les bruitages qui ont eu lieu lors de la mesure de la valeur vraie. Enfin il faut noter qu'il y a une erreur de mesure propre à chaque mesure.

La première composante de l'erreur de mesure qui est distinguée classiquement est dite "systématique". Selon le VIM :

- "**L'erreur systématique** [est une] composante de l'erreur de mesure qui, dans des mesurages répétés, demeure constante ou varie de façon prévisible"

Cette composante d'erreur de mesure est facile à appréhender : c'est le même que tout le monde corrige dans sa cuisine lorsqu'il tare sa balance. Dans cet exemple de vie quotidienne, c'est le réglage de l'outil qui est la cause du problème. Il est facilement concevable qu'un outil mal réglé puisse répéter systématiquement une même erreur. Il peut pourtant aussi y avoir des erreurs systématiques ayant des causes humaines : par exemple un écart de force entre deux opérateurs qui implique en moyenne un couple plus grand et donc plus de précontraintes dans une pièce lors du serrage des vis pour la mise en place avant usinage. Il faut aussi relever que la "grandeur que l'on veut mesurer", le mesurande, est dépendante de la matière qui n'est pas infiniment rigide et peut donc s'effriter ou se déformer lors de la mesure. Le matage local de la pièce est particulièrement remarquable si deux opérateurs mesurent avec un pied à coulisse et avec une force nettement différente : l'outil est capable de transmettre des forces importantes à travers des mâchoires qui présentent une fine zone de contact et pouvant ainsi facilement créer des encoches dans la matière et dénaturer systématiquement la valeur vraie mesurée.

La seconde composante de l'erreur de mesure que l'on distingue classiquement est dite "aléatoire". Selon le VIM, toujours :

- **L'erreur aléatoire** [est une] composante de l'erreur de mesure qui, dans des mesurages répétés, varie de façon imprévisible

Cette notion englobe toutes les causes qui font que l'on peut obtenir deux mesures différentes pour une unique valeur vraie ; ou bien, de la même façon, que deux mesures identiques peuvent être le résultat de mesurage de deux valeurs vraies différentes. Pour se projeter sur les causes physiques à l'origine de l'erreur de mesure, on peut se référer à la définition du VIM du résultat de mesurage. Il est précisé, qu'en plus d'une valeur numérique, le résultat est complété par "toute autre information pertinente disponible". Cela inclut l'état de l'instrumentation, de l'environnement mais aussi la précision des méthodes et des facteurs humains [139], qui contribuent à l'environnement dans lequel la mesure a été faite. Cet environnement est contrôlé avec précision dans les laboratoires de métrologie, mais ne sera jamais parfaitement identique entre deux mesures.

Lorsque l'erreur aléatoire est dite "imprévisibles", il faut simplement entendre que l'estimation de son impact sur la mesure est hors de portée (ce qui réfère donc à une connaissance dont la limite peut évoluer dans le temps), et non que son impact change à chaque instant et de façon parfaitement aléatoire. Ce point est synthétisé dans les travaux de Gregis : "l'introduction d'un concept d'*erreur aléatoire* est une hypothèse théorique. Celle-ci découle d'un constat : les paramètres influant sur l'issue de l'expérience sont trop complexes pour être appréhendés individuellement." [64].

Bien que l'étude d'un phénomène incontrôlable puisse sembler inutile - et présente de surcroît une difficulté, il est nécessaire d'étudier les causes de l'erreur de mesure aléatoires qui faussent la lecture de la production. Comme exprimé dans la norme ISO de 1998 : "Pour prendre des décisions relatives à la qualité d'un produit, par exemple en ce qui concerne la conformité d'un produit fabriqué à des spécifications données dans le cadre de relations client-fournisseur, il est nécessaire de considérer l'incertitude du résultat de mesure" [71].

1.1.3 Approche standard pour l'identification de la valeur vraie

Si le champ d'action de la métrologie est vaste, son objectif reste de permettre et de garantir l'acquisition de mesures de référence. En se fondant sur les définitions du VIM et la modélisation du GUM, on peut exprimer l'objectif comme étant l'identification de la valeur vraie. Mais il est impossible de savoir ce que vaut exactement la valeur vraie, impossible même de l'étudier spécifiquement. **C'est donc de façon indirecte qu'il faut résoudre la question : en connaissant la mesure et l'erreur de mesure, on connaîtra la valeur vraie.** Par cette approche, la mesure est connue : c'est un nombre réel fixé (la valeur donnée par l'instrument). Mais la valeur exacte de l'erreur de mesure quant à elle, est impossible de connaître. Cette incertitude réduit l'objectif à identifier une zone qui contient la valeur vraie [138, 9].

La construction de la zone qui contient la valeur vraie requiert une certaine connaissance de l'erreur de mesure. Dans le GUM, il est proposé de les classer selon deux types :

- Évaluation de l'incertitude de type A : évaluation d'une composante de l'incertitude de mesure par une analyse statistique des valeurs mesurées obtenues dans des conditions définies de mesurage ;

- Évaluation de l'incertitude de type B : évaluation d'une composante de l'incertitude de mesure par d'autres moyens qu'une évaluation de type A de l'incertitude.

Les descriptions des différentes causes qui contribuent à l'erreur de mesure sont ensuite interprétées statistiquement par des grandeurs telles que la valeur moyenne, la variance ou ses bornes. En pratique, les évaluations de type A font généralement références à toutes les composantes de l'erreur que l'on peut caractériser par leur variance [141, 106], tandis que les évaluations de type B sont caractérisées par des lois de densité de probabilité définies par l'utilisateur. Un exemple de recensement et de combinaison de type A et B est donné dans les travaux de Stolfi [131]. L'objectif de ces évaluations est que les descriptions physiques laissent place à des indicateurs statistiques qu'il est ensuite possible de combiner.

L'intervalle contenant la valeur vraie est ensuite construit de la façon suivante : la moyenne de l'erreur de mesure est soustraite à la mesure et l'erreur de mesure maximale est appliquée de part et d'autre de la valeur obtenue. Une mesure de 250 grammes entachée d'une erreur de mesure estimée en moyenne de $-0,1$ gramme et dispersée au maximum de $0,3$ grammes mènera ainsi à la conclusion que la valeur vraie est comprise entre 249.8 et 250.4 grammes. L'intervalle contenant la valeur vraie est désigné dans le GUM par "intervalle de recouvrement" [9]. Un schéma de l'application de cet intervalle est proposé en Figure 1.3. Du point de vue des exigences de production, il ne s'agit plus que la mesure soit à l'intérieur de l'intervalle de tolérance mais désormais que l'intervalle contenant la valeur vraie soit entièrement inclus.

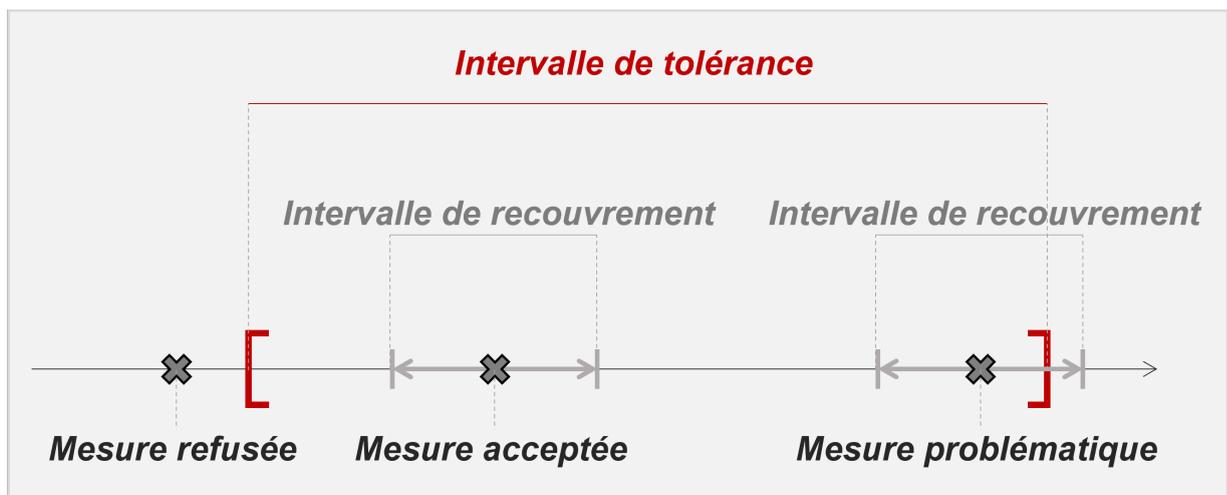


FIGURE 1.3: Illustration de l'application de l'intervalle de recouvrement.

L'intervalle de recouvrement, en remplaçant la mesure, hérite de tous les enjeux décisionnels et notamment de ceux de la conformité. Comme il est le résultat et la combinaison de beaucoup d'estimations, il suscite logiquement un certain scepticisme. Cette méfiance face à l'erreur de mesure qui est difficile à cerner [132], peut se concrétiser par un élargissement arbitraire de l'intervalle de recouvrement. Le GUM lui-même propose l'application d'un coefficient dit "de recouvrement" (K) [9], qui multiplie la variance de l'erreur de mesure pour obtenir l'"incertitude étendue" [18, 131, 106]. La valeur du coefficient K est généralement fixée égale à deux (voire trois), pour garantir à 95% le bon recouvrement [9]. Une

Analyser les mesures avec un œil sceptique, c'est plus communément déployer une approche dite "conservative". Les différentes composantes de l'erreur de mesure sont prises "au pire des cas", afin de ne jamais sous-estimer les erreurs. En effet, surestimer l'erreur de mesure, c'est minimiser

le risque de garantir la conformité d'une pièce à tort. La métrologie étant garante de la qualité de la production, une approche conservatrice répond bien au premier besoin d'identifier infailliblement les pièces qui ne vérifient pas les exigences de productions.

L'approche conservatrice présente cependant un coût : la fiabilité de la garantie est payée au prix d'une globale surqualité. Mettre l'accent pour filtrer avec une très grande sévérité l'erreur de mesure peut revenir à investir dans l'usinage de pièces d'une qualité inutilement supérieure. Cette approche limite aussi la quantité d'information : en travaillant avec des intervalles, on limite la connaissance à ses bornes. La méthode présente en conséquence une certaine rigidité, puisqu'il y a peu de données pour l'exploitation, la décision et notamment l'ajustement des risques de production.

1.2 La modélisation probabiliste pour intégrer l'incertitude

1.2.1 Les failles et critiques du standard

Le modèle proposé par le GUM ne fait pas l'unanimité. Les divergences ne sont pas surprenantes après avoir présenté les discussions inachevées autour des concepts même qui décrivent les quantités en jeu. L'approche probabiliste évoquée en note dans le VIM est elle aussi évocatrice : la présence de cette alternative suggère que d'autres méthodologies ont aussi leurs places et qu'il n'y a pas aujourd'hui un modèle qui prévaut. On trouvera beaucoup de travaux critiques publiés par Kacker du National Institute of Standards and Technology (USA) [80], au sujet du GUM. Une incohérence est relevée autour de la valeur vraie, dont la définition oscille entre "une valeur unique" ou "un ensemble de valeurs potentielles", les deux définitions ne permettant pas d'appliquer la notion d'un intervalle de recouvrement sans accrocs. Les hésitations autour de certaines définitions se concrétisent par le déploiement de méthodes mélangeant l'approche fréquentiste et celle bayésienne [4].

Bien que l'équation fondamentale (1.1) soit répandue dans la littérature, il existe des alternatives pour modéliser la mesure [3]. Les modélisations en question se distinguent notamment dans la modélisation de l'erreur de mesure, par des modèles physiques plus complexes et ainsi plus explicites sur la nature et l'impact des différentes composantes. On peut notamment lire une critique adressée au GUM par Kacker qui reproche une conception de l'erreur de mesure qui est déconnectée de la mesure [81]. Il est en effet tout à fait concevable que l'erreur de mesure soit fonction de la mesure.

Enfin, en plus de l'identification et de la modélisation, c'est aussi la gestion de l'erreur de mesure qui est discutable. Il se trouve qu'il y a une erreur intrinsèque à l'approche classique, lors de la recherche de la valeur vraie. La valeur vraie est perçue comme étant aléatoirement dispersée autour de la mesure alors c'est la mesure qui est aléatoirement dispersée autour de la valeur vraie. En effet, c'est la mesure qui est le résultat d'une erreur de mesure commise en mesurant la valeur vraie (Figure 1.4). Or l'intervalle contenant la valeur vraie est construit classiquement centré sur la mesure. Par exemple, un courant de 220 volts mesuré à 218, 224 puis 221 volts rappelle que la dispersion de l'erreur de mesure s'opère de la valeur vraie vers la mesure. Cette approche mène à une impasse dans la pratique : il faudrait appliquer l'erreur de mesure à la valeur vraie qui est inconnue (et précisément celle que l'on recherche), pour estimer la valeur d'une mesure qui est déjà connue. Autrement dit, l'intervalle de recouvrement a pour but d'estimer la position de la

valeur vraie, et il faut connaître la valeur vraie pour l'identifier parfaitement. Faute de pouvoir faire mieux, cette impasse explique que l'approche classique soit malgré tout adoptée.

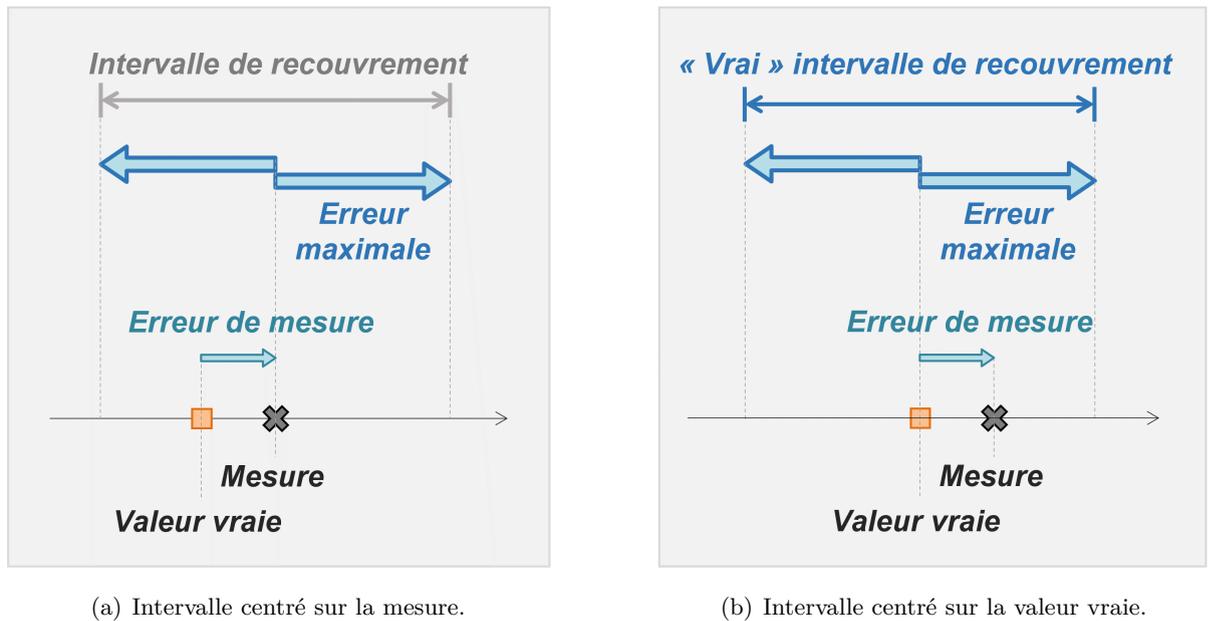


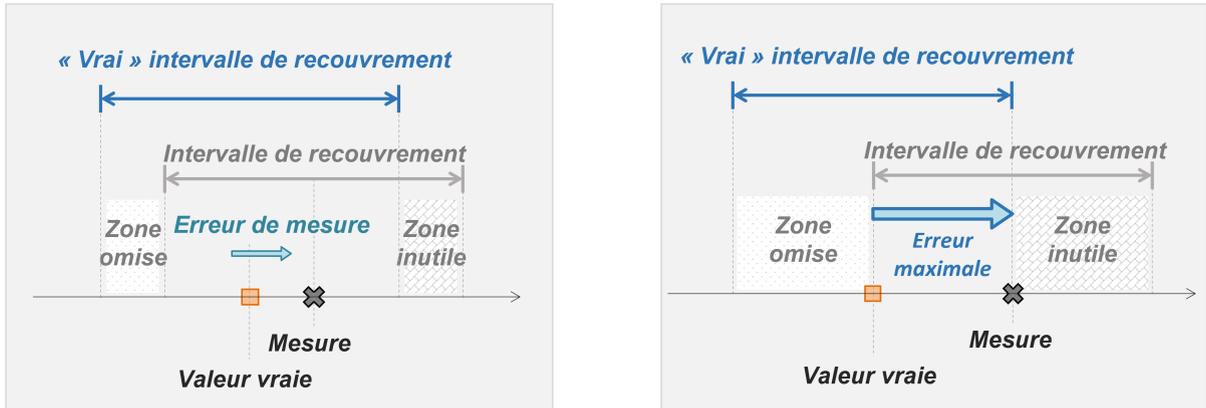
FIGURE 1.4: Illustrations de l'inversion du concept dans la construction de l'intervalle de recouvrement.

La conséquence est illustrée dans la Figure 1.5 pour un cas où $\mu_E = 0$. L'intervalle de recouvrement n'est a priori pas où il devrait être : dans l'unique cas où la mesure est superposée à la valeur vraie, l'intervalle des valeurs suspectées est le bon. Dans les autres cas, il y a une part des valeurs recouvertes qu'il n'était pas pertinent de considérer, et surtout une même part de valeurs qui n'a pas été envisagée. Dans le pire des cas, c'est la moitié de la zone de recouvrement qui est pertinente (et la moitié des valeurs pertinentes à cibler qui est omise).

Même si l'approche classique engendre une erreur lors de la construction de l'intervalle de recouvrement, il ne s'agit de la rejeter du but en blanc. Il faut donc distinguer :

- D'une part l'idée première, d'appliquer des bornes maximales de l'erreur de mesure de part et d'autre de la mesure pour cerner la valeur vraie, qui n'est pas optimale mais correcte.
- D'autre part, l'application de la loi des erreurs sur la mesure pour le calcul de probabilité pour décrire la position de la valeur vraie. Cette démarche présente deux erreurs.
 - Le première erreur concerne l'intervalle de recouvrement, (comme explicité ci-dessus 1.2.1), qui ne sera correctement identifié que dans un seul cas de figure, celui où la mesure obtenue est celle espérée ($m^{(i)} = t^{(i)} + \mu_E$). La probabilité que cette configuration apparaisse est égale à la densité de la moyenne de l'erreur de mesure. Une loi normale de l'erreur de mesure favorise ce coup de chance. Mais une loi uniforme le raréfie, et ce qui d'autant plus fréquent que la loi uniforme est la modélisation standard lorsque seules les bornes de l'erreur de mesure sont connues.
 - La seconde erreur concerne la loi de l'erreur de mesure qui ne peut être appliquée telle quelle à la mesure (la valeur vraie n'est pas une réalisation de la mesure, mais inversement, comme mentionné en Section 1.2.1). Il faut utiliser la fonction de vraisemblance,

dont la définition dépend de la modélisation (la notion est explicitée et illustrée dans le cadre de l'inférence bayésienne en Section 1.3.2).



(a) Exemple d'intervalle de recouvrement erroné.

(b) Exemple d'intervalle de recouvrement plus gravement erroné.

FIGURE 1.5: Illustrations de la mauvaise identification de l'intervalle de recouvrement.

1.2.2 Modèles probabilistes en métrologie

Modélisation des mesures

Dans cette thèse, une unique mesure est associée à chaque cote clef. Il n'est pas rare que plusieurs soient réalisées en pratique, notamment pour se prémunir d'une erreur flagrante, mais cela requiert un investissement qui n'est pas toujours souhaité. Le nombre de mesures par dimensions clef est généralement inférieur à la dizaine, en sachant que 10 est une référence en contrôle dans la procédure d'exécution d'une étude de répétabilité et de reproductibilité (R & R). L'unique mesure retenue pour représenter la mesure peut par exemple représenter la moyenne d'un ensemble de mesure, ou bien son minimum, selon ce qui est pertinent par rapport à l'instrumentation ; ou encore une "mesure indirecte" qui concerne une cote physiquement inaccessible et qui est estimée par la combinaison d'autres mesures (via une formule géométrique par exemple). De plus, il n'y a qu'une seule cote clef par pièce. Bien qu'il puisse y en avoir plusieurs, cela n'en est pas moins une hypothèse réaliste. Au final, à une pièce est associée une unique mesure. Pour un lot de N pièces on s'intéresse aux N mesures : $\{m^{(1)}, m^{(2)}, \dots, m^{(N)}\}$. Ces N mesures du lot sont perçues comme étant N réalisations de la variable aléatoire représentant les mesures : M .

Les méthodes développées dans cette thèse traitent de "lots" de mesures. Il faut noter que la fabrication par lots engendre une spécificité majeure : les caractéristiques statistiques (moyenne et écart-type) des dimensions pour chaque lot évoluent en fonction de l'instant de fabrication [8, 23]. La notion de lot sous-entend que les mesures qui le composent sont toutes de même nature et comparables. D'une part, les mesures doivent être homogènes (il ne peut y avoir dans un même lot la mesure de la masse et celle d'une épaisseur de matière) et se référer à une même cotation (les mesures de la profondeur de deux trous distincts ne doivent pas être mélangées dans un même lot). D'autre part, pour une industrie qui présenterait plusieurs étapes d'usinage, cela restreint les mesures à provenir d'un même endroit de la chaîne de production (les mesures de l'état de surface d'une même pièce usinée puis traitée en surface doivent être placées dans deux lots distincts). La définition de lot reste ensuite à affiner selon le contexte. Sur une chaîne qui produit en continu, deux

lots peuvent être distingués par une opération de maintenance. Si la production est discontinuée, les cycles de production peuvent être plus pertinents que la maintenance.

Enfin, les solutions ont été pensées pour l'industrie manufacturière et les méthodes développées concernent des mesures unidimensionnelles. Pour les quantités multi-dimensionnelles, comme la mesure de l'aire d'une surface, les méthodes peuvent être appliquées soit aux composantes individuellement, soit au résultat final. Cependant, les phénomènes de corrélation ne sont pas abordés. Le supplément 2 du GUM ainsi que d'autres articles de la littérature [97, 30]) traitent de cette problématique.

Modélisation probabiliste de l'erreur de mesure

Les méthodes développées dans cette thèse utilisent une description probabiliste de l'erreur de mesure. L'erreur de mesure est désormais représentée par une variable aléatoire, nommée E , et caractérisée par une densité de probabilité. L'erreur de mesure $e^{(i)}$ associée à la mesure $m^{(i)}$ est perçue comme une réalisation de cette variable aléatoire. Un lot de N mesures présente N erreurs de mesure.

L'usage d'une densité de probabilité permet d'intégrer les différentes composantes qui ont été identifiées (Figure 1.6). Cela permet d'inclure la partie aléatoire de l'erreur de mesure ainsi que sa partie systématique (cette dernière correspond à l'espérance de la grandeur E). Cela permet de manipuler l'erreur de mesure comme un tout, comme évoqué par Kyburg : "le traitement de ce type d'erreurs [aléatoires] suppose de saisir les erreurs collectivement [...]" [91]. Une densité de probabilité est jugée plus bénéfique que l'approche standard pour la manipulation des composantes des deux types (A et B) [142]. Il faut remarquer cependant que les conditions à réunir pour décrire l'erreur de mesure dans son environnement sont contraignantes. La question de l'erreur de mesure ne manque pas d'intérêt mais le manque de moyens et la complexité des méthodes sont souvent un frein. Dans l'usage, par soucis d'efficacité, la liste des composantes de l'erreur de mesure est souvent réduite à celle de l'outil de mesure (la composante la plus flagrante et dont l'incertitude est déjà indiquée par le constructeur).

Une densité de probabilité n'est pas seulement une solution efficace pour modéliser l'erreur de mesure, c'est aussi facile à obtenir. Une fois que les composantes de l'erreur de mesure sont modélisées, la loi peut être obtenue par simulation. La méthode de Monte Carlo par exemple [77, 20]), facile d'accès, reconstruit un très grand nombre d'erreur de mesure possible. L'ensemble des valeurs permet finalement d'obtenir la loi de l'erreur de mesure (avec un lissage à noyau par exemple [75, 113]).

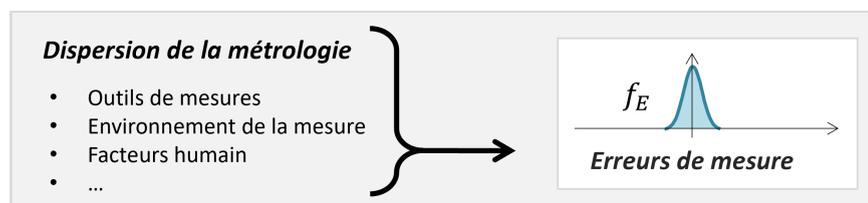


FIGURE 1.6: L'erreur de mesure est décrite par une variable aléatoire.

Modèle probabiliste de la valeur vraie

La valeur vraie est elle aussi considérée comme une variable aléatoire (Figure 1.7). Les procédés de fabrication introduisent une inévitable incertitude sur la géométrie des pièces produites. Il est d'une part impossible qu'une cote clef de la pièce usinée soit exactement égale à la valeur nominale définie en bureau d'étude : on parle d'un défaut de précision. Il est d'autre part également impossible que les valeurs vraies d'une même cote clef de deux pièces distinctes soient égales : on parle d'un défaut de fiabilité du processus. Toutes les valeurs vraies $t^{(i)}$ issues du processus de production sont regardées comme des réalisations de cette variable aléatoire.

A l'instar de l'erreur de mesure, les causes physiques qui provoquent cette dispersion de la valeur vraie sont très variées et proviennent de l'environnement du processus de production. L'étape de repositionnement de la pièce entre deux étapes d'usinages (un exemple de réglage non numérique assuré par les opérateurs sur des machines) peut être une cause de dispersion de la valeur vraie. Le serrage des pièces est aussi un exemple par ses éventuels problèmes de réversibilité (un phénomène d'hystérésis peut apparaître lors l'ajustement d'une position par l'utilisation d'une manivelle à vis). Les vibrations que subit la machine sont une source de dispersion provenant de l'environnement de la machine. La machine d'usinage opère elle-même avec des outils qui ne présentent jamais deux niveaux d'usure identiques.

La dispersion de la valeur vraie pourrait faire l'objet d'une étude en soi : l'identification des causes de la dispersion demande un travail important, assez proche de celui requis pour l'erreur de mesure. La loi de la valeur vraie, c'est-à-dire celle de la production, n'est pas ici considérée comme une donnée d'entrée (ce n'est pas une donnée nécessaire), mais les méthodes proposées seraient tout de même aptes à prendre en compte cette information.



FIGURE 1.7: Les valeurs vraies issues de la production sont décrites par une variable aléatoire.

1.2.3 Modélisation probabiliste des données métrologiques

Une densité de probabilité est un outil idéal pour modéliser la valeur vraie et l'erreur de mesure. Une fonction permet évidemment d'intégrer plus d'informations qu'un unique indicateur statistique. Cet outil permet de résoudre une limite fréquentiste pointée par Gregis [64] : "Les erreurs aléatoires sont considérées comme une population (typiquement gaussienne), à laquelle les outils statistiques classiques peuvent être appliqués. Par exemple, la variance d'une population est généralement évaluée par la variance d'un échantillon. Cette procédure ne peut pas être utilisée pour les incertitudes liées à des biais, car il n'y a pas de population derrière une erreur systématique et, par conséquent, il n'y a pas d'échantillon derrière un biais."

L'utilisation d'une densité de probabilité n'est pas une innovation propre à ces travaux. Le sujet est depuis longtemps traité par le JCGM et le GUM propose déjà, dans le supplément n°1, des méthodes pour l'intégration de lois de probabilité [10]. Il y est présenté l'application de la méthode de simulation de Monte Carlo pour créer la loi de l'erreur de mesure. Une autre méthode,

parue récemment, est celle de la "bande de garde". Une illustration de la méthode est donnée en Figure 1.9. Le fait est que toutes les mesures ne sont pas autant impactées par l'erreur de mesure. Les mesures à risques sont celles qui sont proches des bornes de l'intervalle de tolérance (voir Figure 1.8). L'idée est alors de rajouter une contrainte pour la sélection des mesures. Les mesures qui sont incluses dans l'intervalle de tolérance doivent désormais être à une distance minimale de ses bornes. La distance minimale est la bande de garde. Sa construction requiert de connaître la densité de probabilité de l'erreur de mesure et de fixer un seuil de risque acceptable. L'intervalle de la nouvelle sélection est nommée intervalle d'acceptance. En filtrant des mesures à l'intervalle de tolérance, il ne peut que réduire que le risque client et augmenter le risque fournisseur. Cette thèse vise à obtenir plus de la connaissance loi de l'erreur de mesure.

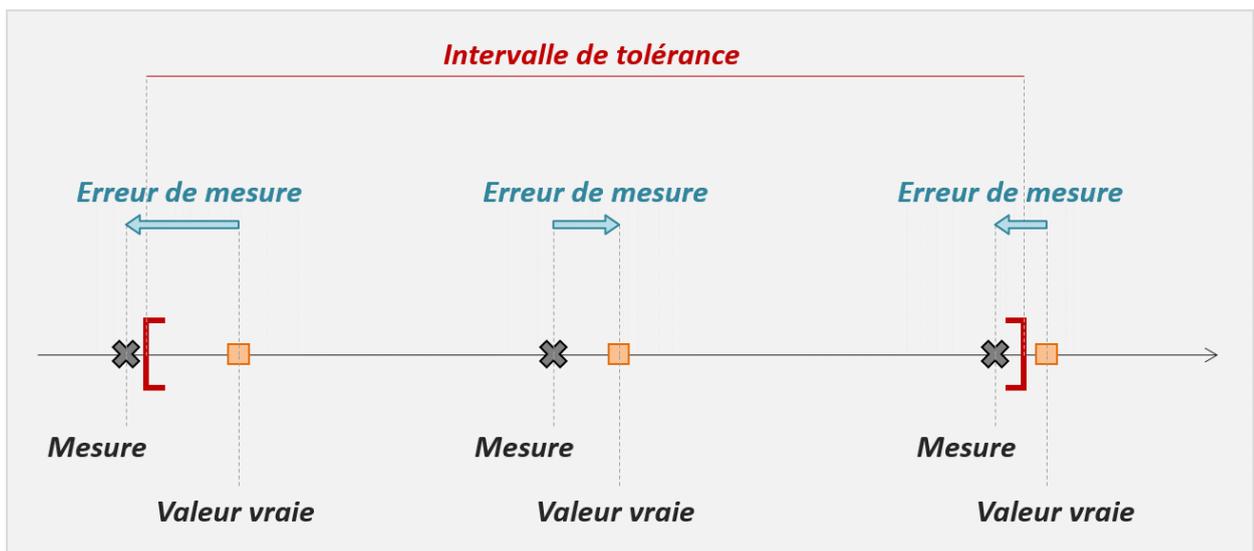


FIGURE 1.8: Schéma des mesures à risques. Les mesures dont la conformité est brouillée par l'erreur de mesure sont celles proches des bornes de l'intervalle de tolérance.

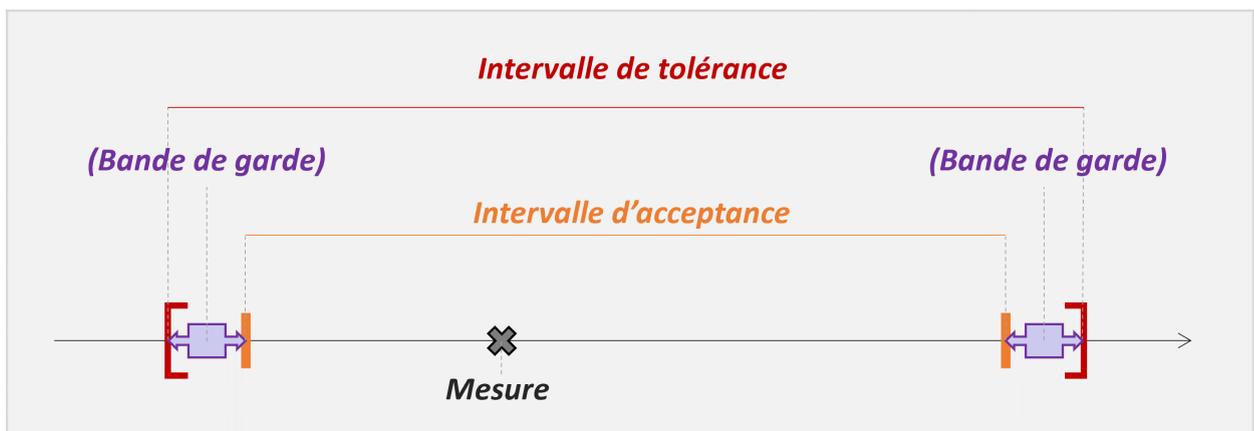


FIGURE 1.9: Schéma du concept de bande de garde.

1.3 Méthodes probabilistes pour améliorer la mesure

1.3.1 Recherche de la loi de la valeur vraie

Le produit de convolution, ou la reconstruction de la mesure

La vision probabiliste implique que l'équation fondamentale (1.1) est désormais une équation reliant trois variables aléatoires :

$$M = T + E \quad (1.2)$$

avec M , E et T les variables aléatoires de la mesure, de l'erreur de mesure et de la valeur vraie. Dans le cadre général, la loi d'une variable aléatoire définit comme étant égale à la somme de deux autres indépendantes est donnée par le produit de convolution [104, 101] :

$$f_M(m) = \int_{-\infty}^{+\infty} f_T(m - e) f_E(e) de \quad (1.3)$$

avec f_T la densité de probabilité de la valeur vraie produites par le processus de fabrication, f_E celle de l'erreur de mesure qui est commise lors de la mesure, et f_M celle des mesures m lues en métrologie. L'hypothèse de l'indépendance entre la valeur vraie et l'erreur de mesure est une donnée précieuse pour l'application de cette méthode. Cette affirmation se retrouve dans la littérature sans rencontrer d'opposition [101]. La fabrication de la pièce et l'étape de métrologie sont en effet des étapes distinctes réalisées avec des outils différents. Il faut cependant relever qu'ici, c'est l'environnement métrologique tout entier qui est considéré dans cette étude pour décrire l'erreur de mesure, et que cet environnement pourrait partiellement coïncider avec celui de la production (hygrométries similaires, mêmes opérateurs).

Le produit de convolution qui définit la loi f_M est une intégrale qui peut être calculée numériquement. La méthode de Monte Carlo est la solution la plus répandue [102, 34]. Elle est très accessible et présente une grande robustesse. A moins de contraintes externes (généralement celle du nombre de calculs possibles), elle s'impose généralement comme la solution de premier choix, et est déjà implémentée dans la majorité des logiciels scientifiques. Mais beaucoup d'autres méthodes se prêtent à cet exercice et notamment les quadratures gaussiennes [5] qui seront abordées au Chapitre 2.

La variable d'intégration e est l'erreur de mesure. Pour une mesure m , il s'agit d'essayer toutes les valeurs possibles de l'erreur de mesure (entre moins et plus l'infini). La probabilité d'occurrence de chaque valeur testée est pesée avec la loi f_E . Chaque valeur de e implique une valeur vraie complémentaire ($t = m - e$) dont la probabilité d'occurrence est évaluée avec f_T . La fréquence d'apparition est calculée pour tous les couples des valeurs (t, e) . L'ensemble des fréquences est ensuite sommé pour aboutir à la probabilité d'obtenir la mesure m en métrologie.

La déconvolution, ou la déconstruction de la mesure

La convolution ne répond pas à la problématique de la recherche de la valeur vraie. Similairement à ce qui a été exposé dans la Section 1.2.1, c'est la quantité f_T qui est recherchée tandis la quantité f_M est déjà observable. La recherche de la valeur vraie demande d'identifier l'un des termes impliqués dans l'intégrale du produit de convolution. Selon Wahab [133] : "la déconvolution ou le problème du filtrage inverse cherche à défaire la convolution [...] et à extraire le signal

réel de celui observé." Cette opération, qui est l'identification d'une loi marginale, est nommée "déconvolution" et présente une réelle difficulté.

Il existe aujourd'hui beaucoup de méthodes de déconvolution. Les cas d'usage les plus répandus concernent les séries temporelles continues, les traitements d'images [17] et les traitements de signaux [99, 92], par exemple ceux tracés par les sismographes pour lesquels la déconvolution "fait généralement référence à l'extraction de pics" [133]. Certaines méthodes de déconvolution font appels aux théories bayésiennes [118]. Les méthodes de déconvolution répondent à un besoin précis : les solutions qu'elles apportent sont fortement liées au domaine d'application, à la modélisation du problème et aux données disponibles. Or peu de méthodes sont dédiées à la métrologie, domaine qui propose un très grand nombre de modélisations pour la mesure et une connaissance partielle de l'erreur de mesure.

Dans cette thèse, la modélisation est celle présentée en Section 1.1.2 : $m = t + e$. Il existe peu de travaux dans la littérature qui utilisent à cette modélisation, notamment pour l'erreur de mesure, souvent plus complexe (ne serait-ce que linéaire) [24, 25, 107, 3]. La densité de probabilité de l'erreur de mesure est une connaissance admise : cette information, qui est une denrée rare [89] impact significativement le choix de la méthode de déconvolution. En métrologie chaque mesure est une donnée à part entière, un phénomène en soi : l'échantillon de mesures ne peut pas être traité par filtrage. Les mesures ne sont pas considérées comme corrélées : l'échantillon n'est pas traité comme une série de données temporelles ordonnées (les mesures peuvent être observées dans un autre ordre que celui de la production). Une mesure est associée à chaque pièce : c'est un problème statistique uni-varié. Les simplifications induites par le modèle uni-varié (par rapport au multi-varié) sont explicités par Weckenmann [139].

Dans la littérature, il est fréquent que la loi qui est égale au produit de convolution des deux autres soit connue par des données observables [41, 101, 13] - tout comme ici ce sont les mesures m issues de la métrologie qui permettent de connaître la loi f_M . Mais beaucoup de méthodes reprochent aux données observées d'être imparfaites (de contenir une part d'erreur) : qu'un bruit de mesure extérieure à la convolution vienne troubler la lecture du produit de convolution [33]. Dans ces travaux, l'erreur de mesure fait partie intégrante de la définition de la mesure m et ne constitue pas un problème en soi. Ceci ferme la porte aux nombreuses méthodes s'adressant aux erreurs sur les variables (plus communément rencontrées sous le nom anglais d'*errors-in-variable*) [101, 136] et cantonne la recherche aux solutions classiques de déconvolution.

Les méthodes de déconvolution classiques sont datées du siècle dernier. L'approche la plus connue consiste à traiter le problème dans le domaine de Fourier [41, 101, 13] où la déconvolution est triviale. Les lois y sont basculées et combinées, puis le résultat est récupéré par transformation inverse de Fourier. Mais la transformation inverse est soumise à des conditions exigeantes [101] (la loi uniforme est un exemple de loi qu'il faut traiter singulièrement) et les alternatives ne permettent pas de généraliser la solution. L'estimation par densité de la solution est une alternative qui butte sur l'identification des paramètres de la solution et perd rapidement en efficacité lorsque la loi de l'erreur de mesure est lisse (*smooth* dans la littérature anglophone), comme la loi normale ou d'autres lois usuelles paramétrées [41, 136, 137, 101, 68, 143]. La régression non-paramétrique quant à elle, offre une solution plus généralement applicable et adaptable à la problématique métrologique (quelque que soit le type de loi de l'erreur de mesure) [101] et est donc retenu.

Il existe donc un besoin de développer des méthodes plus générales et plus accessibles pour l'implémentation de la déconvolution dans le contexte de la métrologie, qui répondent à une modélisation spécifique, notamment caractérisée par la connaissance de la loi de l'erreur de mesure.

1.3.2 Recherche de la valeur vraie d'une mesure

Les statistiques bayésiennes pour l'exploitation de deux informations

La recherche de la valeur vraie nécessite d'intégrer de l'information. La première information concernée est la loi de l'erreur de mesure f_E , dont la connaissance est admise. Les méthodes de déconvolution spécialement développées permettent d'intégrer la loi f_M observée sur tout un échantillon de mesures pour obtenir une estimation de celle des valeurs vraies f_T . Dans un second temps, il est alors possible de confronter une unique mesure d'un lot aux deux quantités qui la définissent : non seulement à la loi de l'erreur de mesure pour ce qui concerne $e^{(i)}$ mais aussi à la loi estimée des valeurs vraies pour ce qui concerne $t^{(i)}$.

L'intégration de deux connaissances, dont l'une décrivant une connaissance "*a priori*", constitue la plus grande motivation et la majeure plus-value de la statistique bayésienne. L'exploitation de l'historique métrologique apparaît dans la liste des sujets prometteur dressée par Reifenberg en 2014 : "L'analyse approfondie de certaines méthodes de qualification, l'exploitation de l'historique métrologique pour quantifier le risque d'apparition de non-conformité, et adapter la stratégie de contrôle, l'analyse de la criticité d'un instrument dans un processus de mesure, les analyses de risques sont autant d'outils qui mériteraient d'être plus largement exploités pour que la discipline gagne en efficacité au niveau national." [117]. L'étude des méthodes présentées dans cette thèse a été menée jusqu'à l'analyse de la conformité et l'analyse des risques.

Les méthodes bayésiennes sont aujourd'hui activement développées et très répandues. Les applications dans le domaine de la métrologie sont nombreuses, à commencer par le GUM lui-même qui a introduit cette approche dernièrement [12]. Selon Göller, "la procédure de mise à jour bayésienne fournit un outil rigoureux et puissant pour résoudre des difficultés" relatives "aux erreurs de mesure" [59]. Un article écrit par Carvajal [19] expose une application des méthodes bayésiennes dans le but de faciliter la calibration des outils de mesures. L'intégration de la connaissance *a priori* a pour conséquence de compléter les observations, et ainsi d'obtenir une calibration de bonne qualité même avec peu de données. Partout où les statistiques classiques (dites fréquentistes) sont déployées, partout l'approche bayésienne se propose en alternative. Le caractère novateur des méthodes bayésiennes permet aussi de croire que l'on pourra y trouver des solutions aux points faibles trouvés aux méthodes fréquentistes (parce que présentes de longue date et donc largement éprouvées), qui peuvent aujourd'hui faire l'objet d'un risque d'abus [145]. La capacité des méthodes à valoriser une connaissance *a priori* (par la combinaison de la connaissance théorique avec les observations), représente une grande opportunité qui s'étend aux domaines des sciences humaines et sociales [53].

Les statistiques bayésiennes permettent d'intégrer une connaissance *a priori* du résultat [29, 6]. Les formules découlent du théorème de Bayes qui définit la probabilité conditionnelle telle que :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1.4)$$

avec A et B deux événements, et B ayant une probabilité non nulle. Cette formule peut facilement s'identifier dans le cas de l'inférence bayésienne [59] :

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{\int_{-\infty}^{+\infty} P(\theta)P(Y|\theta)d\theta} \quad (1.5)$$

avec Y les données observées qui sont combinées avec θ la loi *a priori*. L'autre expression de

l'inférence bayésienne, plus répandue, utilise les fonctions de densité de probabilité ([111, 119]) :

$$p(\theta|Y) = \frac{\pi(\theta)\mathcal{L}(Y|\theta)}{\mathcal{C}} \quad (1.6)$$

avec π les lois de la connaissance *a priori* et \mathcal{L} celle de la vraisemblance, dont le produit est normé par la constante d'intégration \mathcal{C} pour donner la loi *a posteriori* p . L'application de ce théorème permet de mettre à jour la connaissance *a priori* par les observations.

L'innovation bayésienne est donc de donner un crédit à une connaissance *a priori* [119]. Cette connaissance sort du cadre classique de la statistique fréquentiste, qui n'extrait d'informations que des données observées. La loi *a priori* est une donnée externe, imposée arbitrairement et souvent qualifiée d'expertise, nom qui met en avant sa qualité et sa légitimité. Par exemple, un chef de production peut être certain que le diamètre des alésages usinés n'est pas inférieur à une valeur cible (par expérience, ou par un système de contrôle non-destructif par prise d'empreinte de matériau plastique) ; tandis que la combinaison théorique naïve des mesures et de l'erreur de mesure peuvent laisser croire que certaine proportion des diamètres produits sont inférieurs à la valeur cible. Cet exemple se traduit comme une troncature de la loi de la valeur vraie. L'approche bayésienne "contraste fondamentalement du point de vue fréquentiste" [119].

Mais le nom "d'expertise" n'est pas une garantie pour autant. Contrairement à la connaissance des données expérimentales, qui ont pour elle une légitimité qui est intuitivement admise, la connaissance *a priori* est très discutée. Dans certaines applications d'inférences bayésiennes, la loi *a priori* est dite neutre ou non-informatif, et apprend des observations. Sinon, elle comporte une information qui influera le résultat. Or sa définition, étant posée de façon subjective [119], ouvre ainsi la porte à des erreurs (à un apport d'information biaisée ou erronées) ou à des abus dans le pire des cas. La présence d'incertitudes autour de la connaissance *a priori*, et donc à la base même d'une méthode qui est probabiliste, ne manque pas de rajouter une certaine couche de scepticisme [4]. La quantification de son impact est ainsi un sujet de grand enjeu : il fait aujourd'hui toujours l'objet d'études, comme celle proposée par Ghaderinezhad [56], étude contemporaine qui l'évalue via la méthode de Stein.

Alternatives à l'approche bayésienne en métrologie

Il existe bien sur des approches autres que celle bayésienne pour intégrer des informations. L'utilisation de "Signal Relation Diagram" dans la métrologie est proposée par Ruhm pour agréger les informations données par les mesures et celles observées dans leurs environnements [120]. Il est affirmé que l'application de cette méthode apporte des informations autour de la mesure, du processus et de son monitoring, mais aussi à propos de quantités non mesurables.

Un autre domaine plus largement dédié à l'intégration d'information est la "Data assimilation". Parmi les méthodes les plus répandues on retrouve : *Blue*, *Bats* et *Sals* [108]. Un article de Otle [110] présente aussi des méthodes de *data assimilation* fondées sur la statistique bayésienne, tel le Filtre de Kalman, une méthode de filtre à particules ainsi que la "Variational assimilation". Cet article contient aussi une présentation de la data assimilation, définie comme "englob[ant] toutes les techniques qui permettent de combiner de manière optimale les informations mathématiques contenues dans les données, qui permettent de combiner de façon optimale les informations mathématiques contenues dans les équations d'un modèle, et les informations physiques issues des

observations, afin de reconstruire les états d'un système".

Il existe enfin des alternatives pour traiter des problématiques proches dans un cadre fréquentiste. La différence entre l'approche bayésienne et l'approche classique est un sujet extrêmement bien documenté. On pourra retrouver dans les publications récentes de Nievergelt [83] et Wood [51] des exposés très détaillés. Les travaux de Fierens [47] proposent une vision où une mesure seule n'est pas exploitable (elle n'est qu'une composante, seul l'échantillon contient des données) et font appel à des méthodes de robustesse [69]. Walley a publié une méthode à la frontière de ces "deux philosophies reconnues comme incompatibles entre elles", pour intégrer dans le domaine fréquentiste la fonction de vraisemblance et tous les bénéfices de l'information qu'elle apporte. Sadeghi [121] propose une méthodologie pour l'intégration d'une connaissance *a priori* et la création d'une information *a posteriori* dans un cadre fréquentiste. Ces outils n'intègrent pas la densité de probabilité de l'erreur de mesure, et ne requiert donc pas de la connaître (contrairement aux objectifs et méthodes de cette thèse).

Conclusion

Ce premier chapitre a présenté la définition de la mesure et de ses composantes. Cette étape primordiale est toujours activement discutée et se révèle être une première source d'incertitudes. Puis, deux approches ont été présentées pour améliorer la mesure : la première est celle des standards actuels, la seconde est celle probabiliste, qui diffère dans la modélisation et dans les méthodes.

Il a été montré que l'identification de l'erreur de mesure est la connaissance clef pour l'amélioration de la mesure. L'approche standard est limitée dans la description de l'erreur de mesure, et ses méthodes le sont de surcroît. La modélisation probabiliste se distingue particulièrement en décrivant l'erreur de mesure comme étant une variable aléatoire, donc caractérisée par une densité de probabilité. Cette approche peut contenir plus d'informations que les descriptions statistiques et permet de reconstruire la loi de l'erreur de mesure par simulation.

Dans le but d'améliorer la mesure, il existe de nombreuses méthodes de déconvolution mais peu sont adaptées à la métrologie et permettent d'autant moins d'exploiter une connaissance de la densité de probabilité de l'erreur de mesure. Il existe donc un besoin en méthodologie dans cette configuration particulière. Enfin, des méthodes probabilistes pour l'assimilation de plusieurs informations ont été présentées pour estimer la valeur vraie d'une mesure, et l'approche bayésienne a été retenue.

Chapitre 2

Méthodes pour la correction des erreurs de mesure

Points clefs

- Recherche de la loi de la valeur vraie
 - Recherche de la valeur vraie de chaque mesure
 - Étude de la conformité et des risques de production
-

Sommaire

Introduction	50
2.1 Estimation de la loi de la valeur vraie	51
2.1.1 Cas de variables aléatoires gaussiennes	51
2.1.2 Identification de la loi de la valeur vraie	53
2.1.3 Reconstitution par approximations locales	58
2.2 Recherche de la valeur vraie d'une mesure	61
2.2.1 Application de l'inférence bayésienne	61
2.2.2 Le cas particulier des variables aléatoires gaussiennes	65
2.2.3 Quantification de l'amélioration	67
2.3 Étude de la conformité et des risques associés	69
2.3.1 Lecture de la conformité	69
2.3.2 Probabilités et risques de mal-classifier	74
Conclusion	82

Introduction

Ce chapitre présente des méthodes développées pour la métrologie et dont l'objectif est d'exploiter au mieux la connaissance de la densité de probabilité de l'erreur de mesure. Cet objectif se traduit par la recherche de la valeur vraie ou de sa loi et de toute nouvelle information provenant de ces résultats, conformément à l'énoncé du VIM [11] : "Le résultat d'un mesurage est l'ensemble de[s] valeurs attribuées à un mesurande, complété par toute autre information pertinente disponible pour l'amélioration des mesures."

La déconvolution est une étape complexe. Plus la méthode de déconvolution est fiable et robuste, plus elle requiert des hypothèses exigeantes. De même qu'une méthode résout idéalement un problème donné, ainsi plusieurs solutions de déconvolution sont présentées pour plusieurs cas de figure, et dont la qualité des résultats et la rigidité dans la modélisation vont de pair.

L'inférence bayésienne, elle, s'applique quelle que soit la méthode de déconvolution utilisée car ses exigences ne portent que sur la qualité des informations assimilées. La densité de probabilité de l'erreur de mesure est admise comme connue; celle de la valeur vraie est obtenue suite à la déconvolution.

Ces méthodes ont à retrouver la valeur vraie est évaluée par différents indicateurs statistiques et à travers la lecture de la conformité, qui reste l'aboutissement pratique indissociable au savoir-faire métrologique. Cette étude de la conformité est réalisée avec des outils de data science.

La première section décrit les méthodes proposées dans cette thèse pour réaliser la déconvolution dans le contexte de la métrologie. Les méthodes visent à corriger les effets des erreurs de mesure sur la loi de probabilité associée à un échantillon. La deuxième section présente une application de l'inférence bayésienne en métrologie. Dans ces travaux de thèse, la correction s'applique cette fois à chacune des mesures de l'échantillon, pour en rechercher la valeur vraie. La correction s'appuie sur la connaissance de la densité de probabilité de l'erreur de mesure ainsi que sur la loi déconvoluée qui caractérise l'échantillon de mesures étudié. La dernière section présente les outils statistiques utilisés pour évaluer l'amélioration des mesures et les outils de data science pour la gestion de la production.

2.1 Estimation de la loi de la valeur vraie

2.1.1 Cas de variables aléatoires gaussiennes

Que cela réponde à une réalité, à une approximation ou à un choix arbitraire, les différentes raisons qui motivent la modélisation gaussienne dans le cadre de la métrologie ont déjà été synthétiquement explicitées par Gregis : "L'hypothèse gaussienne constitue le cas le plus traditionnel, et celui qu'on retrouve dans le GUM. Depuis les travaux de Laplace et Gauss sur les probabilités, la loi normale s'est imposée comme loi archétypale, soutenue en cela par le théorème central limite [112]. La suprématie de la loi normale en tant que loi « naturelle » a, depuis, été régulièrement contestée et son statut, notamment dans le GUM, est désormais celui d'un simple modèle, utile mais imparfait. La loi normale reste cependant l'une des lois les plus usitées, notamment du fait de sa simplicité mathématique" [64]. La modélisation gaussienne est une hypothèse à la fois forte et répandue.

Les erreurs de mesure et les valeurs vraies sont représentées par les variables aléatoires E et T . Les fonctions de probabilité de densité qui les définissent sont ici supposées gaussiennes. Elles sont entièrement définies par une moyenne μ et un écart-type σ :

$$\begin{cases} E \sim \mathcal{N}(\mu_E, \sigma_E) \\ T \sim \mathcal{N}(\mu_T, \sigma_T) \end{cases} \quad (2.1)$$

Lorsque que les deux variables qui composent la mesure suivent des lois gaussiennes, la déconvolution est résolue avec l'usage des fonctions caractéristiques. Les fonctions caractéristiques des valeurs vraies et des erreurs de mesure sont combinées selon le produit de convolution pour donner la fonction caractéristique des mesures.

Les fonctions caractéristiques sont des fonctions associées à toute variable aléatoire qui suit une loi. Ces fonctions sont généralement utilisées pour définir la loi de la somme de deux variables aléatoires ou réciproquement pour démontrer que deux variables aléatoires suivent une même loi. L'une des définitions est donnée par :

$$\phi_Z(x) = \int_{-\infty}^{+\infty} f_Z(z) \exp(ixz) dz \quad (2.2)$$

où ϕ est la fonction caractéristique de la variable aléatoire Z et i l'unité imaginaire ($i^2 = -1$). On pourra reconnaître dans cette formule la propriété suivante : si la variable aléatoire Z admet une densité de probabilité, alors le calcul de la fonction caractéristique revient à celui de la transformée de Fourier [101] (et c'est généralement dans ce cadre qu'elles sont utilisées en métrologie statistique).

L'expression analytique de la fonction caractéristique est connue pour un spectre de loi (loi normale, binomiale, gamma, logistique, exponentielle, ou encore de Bernoulli). Dans cette étude, les deux variables aléatoires en jeu suivent une loi normale. L'expression de la fonction caractéristique d'une loi normale est :

$$\phi_Z(z) = \exp\left(\mu_Z iz - \sigma_Z^2 \frac{1}{2} z^2\right) \quad (2.3)$$

avec ϕ la fonction caractéristique de la variable aléatoire Z de moyenne μ_Z et d'écart-type σ_Z .

C'est lorsque les deux variables aléatoires de la somme sont indépendantes que l'usage des fonctions caractéristiques est particulièrement intéressant. Dans ce cas particulier, le produit de

convolution des deux fonctions de probabilité de densité revient à multiplier les fonctions caractéristiques [101]. En utilisant les notations des variables aléatoires de la mesure, de l'erreur de mesure et de la valeur vraie on a :

$$\phi_M(x) = \phi_{T+E}(x) = \phi_T(x) \phi_E(x) \quad (2.4)$$

L'hypothèse d'indépendance entre les deux variables aléatoires réduit la convolution à un produit usuel [66, 101], et fait ainsi de la déconvolution une simple division :

$$\begin{aligned} \phi_T(x) &= \frac{\phi_M(x)}{\phi_E(x)} \\ &= \frac{e^{ix(\mu_M) - \frac{1}{2}x^2(\sigma_M^2)}}{e^{ix(\mu_E) - \frac{1}{2}x^2(\sigma_E^2)}} \\ &= e^{ix(\mu_M - \mu_E) - \frac{1}{2}x^2(\sigma_M^2 - \sigma_E^2)} \\ &= e^{ix(\mu_T) - \frac{1}{2}x^2(\sigma_T^2)} \end{aligned} \quad (2.5)$$

La fonction caractéristique obtenue suite à la déconvolution est celle d'une loi gaussienne [127], dont les paramètres peuvent être lus par identification :

$$T \sim \mathcal{N}\left(\mu_M - \mu_E, \sqrt{\sigma_M^2 - \sigma_E^2}\right) \quad (2.6)$$

L'application de la déconvolution à deux fonctions caractéristiques gaussiennes a permis d'identifier une solution, qui est aussi gaussienne. Ce résultat se retrouve aussi avec des fonctions de type Dirac et du χ^2 . Pour le reste, il y a très peu de couples de lois dont la combinaison des fonctions caractéristiques peut être reconnue par identification. La recherche de ce qui peut être vu comme un équivalent des fonctions *a priori* conjuguées (dans le cadre bayésien) pourrait étendre cette solution à d'autres modélisations que celle gaussienne.

Bien que l'on ait affirmé que l'erreur de mesure n'est pas corrélée avec la valeur vraie dans cette étude, elles peuvent être auto-corrélées. La valeur de l'erreur de mesure à un instant donné peut dépendre de celle à l'instant précédent. Par exemple, deux mesures de température réalisées dans un laboratoire de métrologie et espacées d'une milliseconde sont évidemment liées entre elles : il est improbable d'obtenir un écart de $10^\circ C$ dans ce laps de temps. Cet exemple met en exergue un facteur d'échelle : l'observation de la température à la même heure et à un même poste peut sembler aléatoire d'une semaine à l'autre. L'exemple rapporte aussi les grandeurs dans leur contexte : la température requiert d'être particulièrement maîtrisée dans les laboratoires de mesure, ainsi les variations qui y seront observées ne seront pas du même ordre de grandeur que dans l'industrie.

De même, les valeurs vraies peuvent aussi être auto-corrélées. La remarque concernant les valeurs des composantes de l'environnement de la production est tout aussi valable. Dans la Section 1.2.2, l'exemple de l'usure de la machine-outil a été évoqué. L'usure (du foret par exemple) évolue dans le temps : cela n'impactera pas les valeurs vraies de deux pièces successives de la même façon, mais laisse légitimement penser qu'il peut y avoir une corrélation (les perçages de deux pièces rapprochées présenteront probablement une valeur proche). Les effets de l'auto-corrélation sont étudiés dans le Chapitre 3.

2.1.2 Identification de la loi de la valeur vraie

La méthode présentée dans cette section se fonde sur le fait que la valeur vraie suit une loi usuelle et qu'il ne s'agit donc que de l'identifier. Contrairement à la méthode de déconvolution précédente, cette approche permet que la valeur vraie suive une loi autre que gaussienne.

L'approche est la suivante : la fonction recherchée est celle qui permettra de reconstruire au mieux l'échantillon de mesures. La reconstruction est réalisée avec l'équation (1.3) explicitée dans la Section 1.3.1. La formule de la convolution met en jeu trois densités de probabilité. La loi de l'erreur de mesure est connue (hypothèse pour ces travaux), la loi de la mesure aussi (via l'échantillon de mesures) : la loi de la valeur vraie est l'unique inconnue restante. L'approche consiste alors à tester un panel de solutions : on fait une hypothèse sur la nature de la loi de la valeur vraie, puis on propage l'erreur de mesure et on évalue l'adéquation du mélange avec les mesures observées.

Pour ce faire, il faut dresser une liste de lois que l'on appelle "familles candidates". Cette liste peut être extrêmement longue, car toutes les fonctions de probabilité de densité peuvent être testées. Le nombre de fonctions est principalement limité par des contraintes de temps dans la phase de développement (il faut implémenter chaque famille candidate manuellement) et dans la phase d'utilisation (chaque famille candidate augmente le nombre de calculs). Dans le cadre de cette thèse, six fonctions ont été retenues et sont détaillées dans le Tableau C.1.

Six fonctions de densité de probabilité sont implémentées dans cette étude pour tenter d'identifier la loi de la valeur vraie. Elles sont importées depuis le module python *OpenTurns* [7].

Parmi les lois à support fini sur \mathbb{R} (dont les densités de probabilité non nulles en dehors d'un intervalle donné), la loi uniforme modélise une valeur vraie équiprobablement répartie. Par rapport à la loi uniforme, la loi arcsinus indique que la valeur vraie est plus probablement proche de ses bornes. À l'inverse, la loi triangulaire est une loi unimodale dont les densités de probabilité nulles à ses bornes (la valeur vraie la plus probable est décrite par le mode).

Parmi les lois définies sur \mathbb{R} entier, la loi normale est unimodale et symétrique. Elle modélise une dispersion centrée sur une valeur cible (et d'autres avantages ont été énumérées dans la Section 2.1.1). Les lois uniforme, arcsinus, triangulaire et normale figurent parmi les standards du GUM [10]. Les deux dernières lois utilisées dans ces travaux ont été choisies pour modéliser des profils dissymétriques : la loi lognormale pour l'asymétrie positive (avec une moyenne supérieure au mode) et celle de Weibull pour l'asymétrie négative (avec une moyenne étant inférieure au mode). Bien que la loi de Weibull présente d'ordinaire une asymétrie positive, il existe les deux solutions dans le module *OpenTurns* qui sont distinguées par les noms "WeibullMin" et "WeibullMax" (l'asymétrie négative est celle désignée sous le nom "WeibullMax"). Toutes ces distributions sont détaillées en Annexe C.

Recherche des meilleurs paramètres

Avant de pouvoir confronter les familles candidates entre elles, il est nécessaire d'identifier la loi qui représentera chaque famille. Il existe par exemple une infinité de lois arcsinus, selon le nombre infini de paramétrages possibles. Identifier une loi parmi celles possibles, c'est identifier un jeu de paramètres qui permet la meilleure reconstruction de l'échantillon de mesures (Figure 2.2).

Le jeu de paramètres d'une loi est représenté par le vecteur θ . Si la loi est de type uniforme, le vecteur est $\theta = [a, b]$ avec a et b les bornes inférieurs et supérieurs de la loi uniforme. Le meilleur

jeu de paramètres Θ est celui qui offre la reconstruction la plus fidèle de l'échantillon de mesures. Pour chacune des familles candidates, la recherche du meilleur jeu de paramètres est réalisée par optimisation. L'optimisation est un processus itératif qui teste successivement différentes solutions jusqu'à atteindre l'optimale. L'optimisation se fonde d'une part sur le calcul d'un score pour évaluer la qualité de la solution testée, et d'autre part sur une stratégie pour définir le prochain vecteur de paramètres θ à tester en connaissance des derniers scores.

Pour ce qui est du score, le raisonnement débute au vecteur θ qui concatène les paramètres à optimiser de la famille candidate. Lorsque la valeur de θ a été fixée, la densité de probabilité des valeurs vraies est entièrement définie. Il est alors possible de calculer les densités f_M de chacune des mesures $m^{(i)}$ correspondant au produit de convolution de l'Équation (1.3). La méthode utilisée est décrite dans la Section 2.1.2. Ces densités calculées sont ensuite comparées à celles observées dans l'échantillon de mesures. Si les densités calculées sont toutes proches de zéro, alors soit l'échantillon de mesures n'était pas représentatif des résultats obtenus en métrologie, soit la loi des valeurs vraies n'est pas adéquate. La représentativité de l'échantillon de mesures étant une hypothèse requise pour le reste de l'étude, il faut donc que les densités calculées soient les plus élevées possibles. La "vraisemblance" est un score qui évalue ce critère : le score utilisé pour l'optimisation.

La vraisemblance est le produit de toutes ces densités entre elles [93] :

$$L(\theta) = \prod_{i=1}^n (f_{M,\theta}(m^{(i)})) \quad (2.7)$$

Le calcul de la vraisemblance est cependant largement implémenté sous sa forme logarithmique, entre autres pour éviter qu'une unique valeur nulle annule le score tout entier. Le score de référence utilisé dans l'optimisation est la log-vraisemblance l donnée par :

$$l(\theta) = \ln(L(\theta)) = \sum_{i=1}^n \ln(f_{M,\theta}(m^{(i)})) \quad (2.8)$$

Pour ce qui est de la stratégie d'optimisation, plusieurs algorithmes de sélection ont été retenus. Le choix de multiples algorithmes est principalement motivé par le fait que l'optimisation est un processus qui peut échouer, car un algorithme n'est généralement adapté qu'à un seul type de problème. Une optimisation peut ainsi exiger plusieurs tentatives avec différents algorithmes avant de réussir. Il est aussi possible de tous les appliquer indépendamment et de conserver le meilleur résultat. La sélection et l'ajustement des algorithmes justifient un sujet de recherche à part entière, mais non nécessaire par rapport aux objectifs de cette thèse. L'implémentation de l'optimisation s'est fondée sur le module *OpenTurns* qui propose un large choix d'algorithmes préécrits et préconfigurés.

Paramétrage de l'optimisation

Pour que le processus d'optimisation puisse être lancé, il faut que la valeur initiale $\theta^{(0)}$ soit arbitrairement fixée. Or la qualité de l'optimisation dépend du choix de ce point de départ. Cela est dû au fait qu'une optimisation fait généralement face à plusieurs solutions, et plus le point de départ est loin de la solution optimale, plus il augmente les chances que l'algorithme n'aboutisse pas sur la solution "globale", mais sur une autre dite "locale". Telle une ascension au pied de la

chaîne des puys à la recherche du plus haut sommet dans un brouillard n'offrant aucune visibilité, un algorithme d'optimisation grimpera la pente qui se proposera à lui, jusqu'au sommet, sans aucune garantie que ce soit le plus haut (qui est le Puy de Dôme). Seule la garantie d'un départ proche du plus haut sommet lui permettrait de croire qu'il l'a bel et bien réalisé la plus grande ascension. De la même façon, il faut avoir une certaine idée de la solution optimale pour maximiser les chances de la trouver, c'est-à-dire un bon point de départ pour l'optimisation.

Pour fixer la valeur initiale de $\theta^{(0)}$, il a été choisi de se référer aux moments des variables aléatoires, pour leur propriété d'additivité. Si les deux variables aléatoires sont indépendantes (hypothèse admise dans cette étude pour l'erreur de mesure et la valeur vraie), alors cette propriété s'étend au moment centré d'ordre deux (la variance) et au moment centré d'ordre trois. Par exemple, quelle que soit la loi que suit la valeur vraie (loi de Weibull, normal, arcsinus, *etc.*) si la moyenne des mesures est de 202.5 et la moyenne de l'erreur de mesure -0.01 , alors sa moyenne est assurément de 202.4.

Le point de départ choisi pour l'optimisation est donc un vecteur de paramètres $\theta^{(0)}$ qui vérifient au moins ces deux contraintes d'égalités. Pour la loi normale, l'identification de la moyenne et la variance suffit. Pour identifier les paramètres des autres lois, il est possible d'appliquer la méthode du maximum de vraisemblances.

Calcul des densités f_M

L'optimisation requiert le calcul des densités théoriques f_M . Ce calcul est celui de l'intégrale du produit de convolution (Équation (1.3)). La méthode numérique la plus connue pour estimer la valeur d'une intégrale est celle de Monte-Carlo, une méthode de simulation intuitive et accessible mais qui requiert généralement un trop grand nombre de calculs. Le tirage d'importance est une alternative parmi les plus connues, un peu plus complexe et qui se couple bien avec le besoin de l'optimisation d'avoir une certaine idée de la solution pour réduire le coût de calcul. Mais la solution retenue figure parmi les méthodes de quadratures, plus complexes et qui donnent une meilleure estimation du résultat. Il est intéressant de noter qu'il serait tout de même possible d'appliquer la méthode du tirage d'importance dans le cadre d'une quadrature gaussienne si la question de la réduction du coût de calcul devait se poser [39].

L'approche par quadrature est de remplacer le calcul intégral (réalisé en tout point du domaine de définition) par l'évaluation et la pondération de certains points clefs (appelés "points d'intégration"). Cette substitution s'applique lorsque l'intégrale est celle du produit de deux fonctions dont l'une pouvant être reconnue parmi les fonctions de pondération de la littérature. L'estimation s'écrit :

$$\int_{-\infty}^{+\infty} h(x)\phi(x)dx \simeq \sum_{j=1}^k \omega^{(j)}h(\xi^{(j)}) \quad (2.9)$$

avec h une fonction quelconque (dont l'intégrale existe), $\omega^{(j)}$ les poids associés aux points d'intégration $\xi^{(j)}$. Les valeurs des poids et des points sont données de façon standard selon la nature de la fonction ϕ . Dans le cas d'une quadrature gaussienne, ϕ correspond à la loi normale centrée réduite, aussi dite "standard" (densité de probabilité gaussienne de moyenne nulle et d'écart-type égal à un). Lorsque la loi des erreurs de mesure peut être identifiée comme la fonction de pondération de la quadrature gaussienne, l'estimation du produit de convolution s'écrit dans le cadre de ces travaux comme :

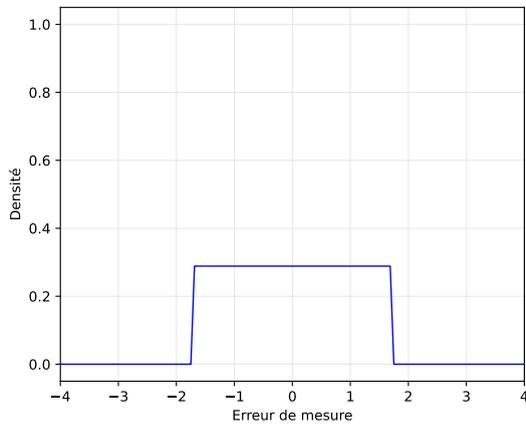
$$f_M(m, \theta) \simeq \sum_{j=1}^k \omega^{(j)} f_T(m - \xi^{(j)}, \theta) \quad (2.10)$$

avec f_M la loi des mesures $m^{(i)}$ obtenues en métrologie, f_T la loi des valeurs vraies dont les paramètres sont donnés par le vecteur θ .

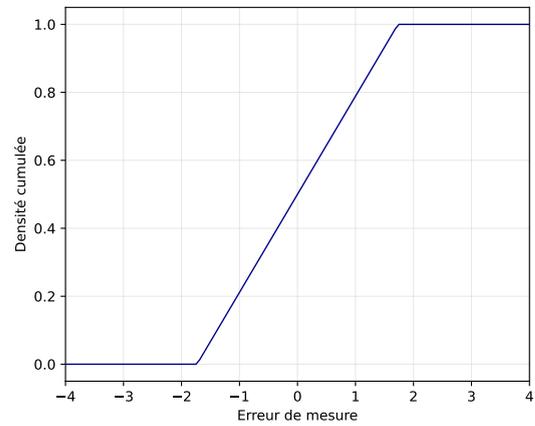
Il est donc possible de calculer les densités f_M avec une quadrature gaussienne, dans le cas où la densité de probabilité des erreurs de mesure suit une loi normale centrée réduite. On pourrait étendre cette méthode de calcul en considérant toutes les autres fonctions de pondération dans la littérature, mais il est bien plus avantageux d'utiliser la transformation de Nataf [95, 105, 96] sur la quadrature gaussienne. A l'instar de transformées de Fourier, cette transformation projette une loi d'un espace à un autre. Toutes les lois que l'on définit et manipule ordinairement sont considérées comme appartenant au premier espace dit "physique". Par cette transformation, toutes les lois physiques peuvent être projetées sous la forme d'une loi normale standard dans le second espace, dit "standard" (et par la transformation inverse, toutes les lois sont ramenées à leurs formes d'origines). La fonction de pondération de la quadrature gaussienne peut ainsi être lue comme étant la transformée dans l'espace standard de la loi réelle des erreurs de mesure et les points comme étant ceux dans l'espace standard. La transformation inverse de Nataf de ces points aboutit alors sur la solution physique du problème :

$$\hat{\xi}^{(j)} = F_E^{-1} \left(\Phi \left(\xi^{(j)} \right) \right) \quad (2.11)$$

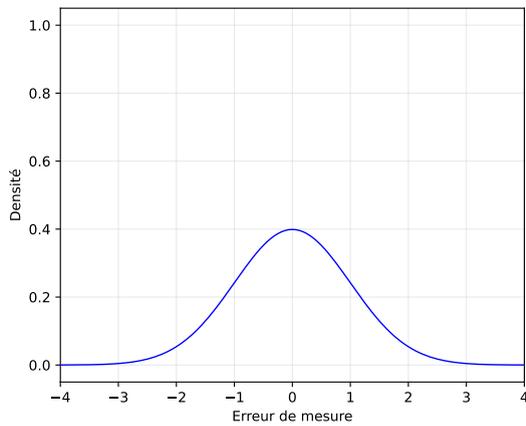
avec F_E et Φ , les fonctions de répartition de l'erreur de mesure et de la loi normale standard. Il est bon de noter qu'il n'y a aucune contrainte concernant la loi des erreurs de mesure avec cette approche, pourvu que l'on puisse avoir accès à sa fonction de répartition F_E (exemples en Figure 2.1). Cette condition est beaucoup moins restrictive que les hypothèses de modélisation gaussiennes de la première méthode. En plus de toutes les lois usuelles (telles que celles sélectionnées pour les candidates), cela permet par exemple d'accéder à des lois plus chaotiques, comme celles multimodales ou directement calquées d'un histogramme (Figure 2.1(f)).



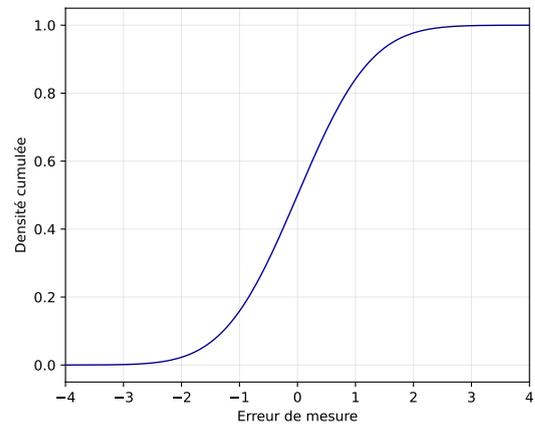
(a) Densité de probabilité (loi uniforme)



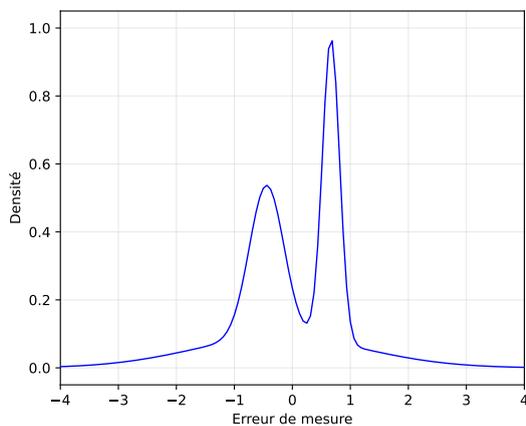
(b) Fonction de répartition (loi uniforme)



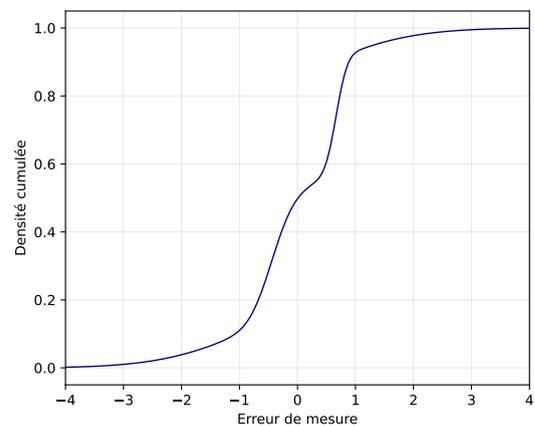
(c) Densité de probabilité (loi normale)



(d) Fonction de répartition (loi normale)



(e) Densité de probabilité (loi quelconque)



(f) Fonction de répartition (loi quelconque)

FIGURE 2.1: Exemples de fonctions de répartition de différentes lois. A gauche, trois exemples de lois de moyenne 0 et d'écart-type 1. A droite, leurs fonctions des répartitions respectives.

Choix de la meilleure famille

Chaque famille candidate propose une solution : une loi dont les paramètres Θ ont été obtenus par optimisation. Il s'agit désormais de comparer ces solutions entre elles, afin de retenir la meilleure identification de la valeur vraie (Figure 2.2). Cette analyse se fonde sur le critère d'information bayésien [124], de son acronyme *BIC*, défini tel que :

$$BIC = -2 \times l(\Theta) + (m - 1) \times \ln(N) \quad (2.12)$$

avec $l(\Theta)$ la vraisemblance calculée pour le vecteur optimal Θ contenant un nombre m de paramètres, et N le nombre de mesures dans l'échantillon. Comme précédemment, identifier la meilleure loi consiste à trouver celle qui permet de reconstruire au mieux l'échantillon de mesures. D'une part, le score *BIC* favorise logiquement les représentants avec les meilleurs scores de vraisemblances : aussi peut-on déduire par la présence du coefficient -2 dans l'Équation (2.12) que le meilleur représentant est celui qui a le plus petit score *BIC*. D'autre part, il pénalise le nombre de paramètres nécessaires pour définir la loi : autrement dit, le score *BIC* privilégie le modèle le plus simple pour un même résultat, conformément au principe de parcimonie (aussi appelé rasoir d'Ockham) [59]. Ce terme de pénalisation est plus important que celui du critère d'information d'Akaike [2] *AIC* dont il est inspiré. Dans l'usage, il y a une tendance à préférer le critère bayésien [1, 86].

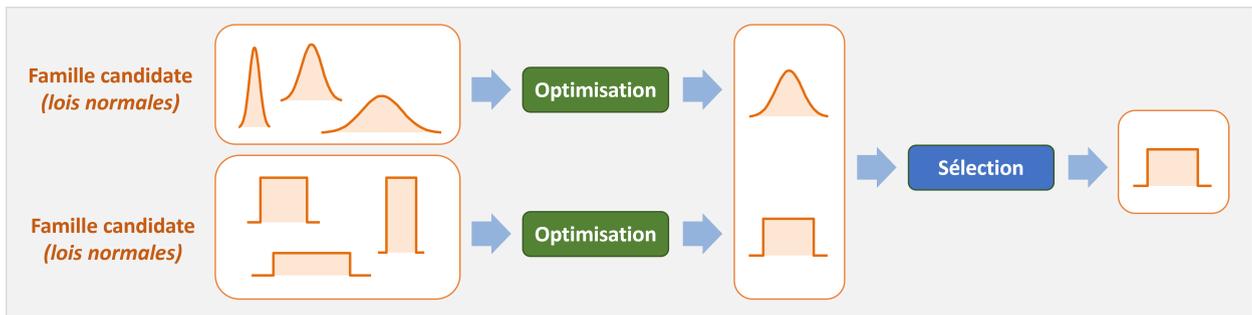


FIGURE 2.2: Schéma de la seconde méthode de déconvolution, pour deux familles candidates.

2.1.3 Reconstitution par approximations locales

La troisième méthode appartient aux méthodes de régressions non-paramétriques. Elle s'adresse aux cas non couverts par les deux premières : les cas où la densité de probabilité des valeurs vraies ne peut pas être modélisée par une loi usuelle (parfois dite non-paramétrique [101]). En fonderie, par exemple, on utilise parfois des moules multi-empreintes : avec un moule à 6 empreintes, les pièces sont produites par six (et il n'est pas forcément possible de garder une trace de l'empreinte d'origine de chaque pièce). Or il est peu probable que toutes les empreintes du moule soient parfaitement identiques et qu'elles présentent les mêmes conditions de production (les empreintes les plus excentrées auront une vitesse de refroidissement vraisemblablement plus rapide que celles au centre du moule). Le lot finalement obtenu est alors un vrac de pièces dont la densité de probabilité est définie par un mélange de 6 lois qui reflètent chacune une empreinte du moule. Et plus il y aura une grande dispersion de production entre chaque empreinte du moule, plus le caractère multimodal de la loi de la production sera visible.

La méthode par approximations locales consiste à reconstruire la loi de la valeur vraie en évaluant sa densité point par point. L'évaluation est fondée sur l'approximation par quadrature

gaussienne du produit de convolution (Équation (2.10)) exposée dans la Section 2.1.2. A x donné, l'idée est :

1. d'estimer la loi des mesures, pour que la quantité $f_M(x)$ soit connue ;
2. de substituer la loi des valeurs vraies f_T par une fonction à n paramètres, pour que le calcul d'une quantité $f_T(x)$ soit réduit à la résolution d'un système à n inconnues .

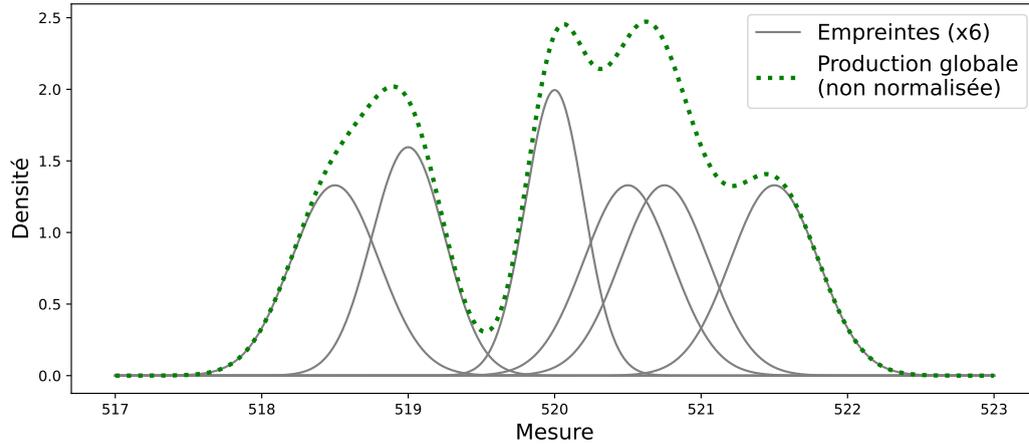


FIGURE 2.3: Illustration d'un mélange de lois pour le cas du moulage multi-empreintes.

Estimation de f_M

L'estimation de la loi des mesures obtenues en métrologies (f_M) est réalisée sur l'échantillon de mesures et selon la méthode de Parzen-Rosenblatt, aussi appelée estimation par noyaux [113, 75] (*kernel smoothing* en anglais). L'expression de l'estimation de la densité est :

$$\begin{cases} f_M(m^{(i)}) \simeq f_M^{KS}(m^{(i)}) \\ f_M^{KS}(m^{(i)}) = \frac{1}{Nh} \sum_{j=1}^N K\left(\frac{m^{(i)} - m^{(j)}}{h}\right) \end{cases} \quad (2.13)$$

avec f_M^{KS} la probabilité d'obtenir la mesure $m^{(i)}$, N la taille de l'échantillon de mesures, K le noyau et h est le paramètre de fenêtre. Un petit échantillon n'est probablement pas représentatif de la loi dont il est tiré. Pour que l'échantillon de mesures soit représentatif de celles qui sont globalement obtenues en métrologie, il faut que N soit suffisamment grand. La loi du noyau est usuellement fixée comme la loi standard. La valeur du paramètre de fenêtre h est fixée selon la règle de Silvermann [126] :

$$\begin{cases} K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \\ h = 1.06\hat{\sigma}N^{-1/5} \end{cases} \quad (2.14)$$

avec $\hat{\sigma}$ l'écart-type de l'échantillon de mesures.

L'application du lissage permet donc au final d'estimer la densité de probabilité des mesures. En couplant cette estimation avec l'approximation du produit de convolution exprimée dans l'Équation (2.10), on obtient :

$$f_M^{KS}(m^{(i)}) \simeq \sum_{j=1}^k \omega^{(j)} f_T(m^{(i)} - \xi^{(j)}) \quad (2.15)$$

avec $(\xi^{(j)}, \omega^{(j)})$ les k points et poids de la quadrature, $f_M^{KS}(m^{(i)})$ la densité lue depuis l'estimation par noyau, et $f_T(m^{(i)} - \xi^{(j)})$ les seules quantités inconnues de cette équation. Cette relation peut être décrite comme une équation avec k inconnues, k correspondant au nombre d'appel de la fonction f_T dans la somme de l'Équation 2.15.

Substitution de f_T

Il serait possible de définir complètement la loi des valeurs vraies puis d'évaluer avec l'Équation (2.15) sa capacité à reconstruire les densités observées dans l'échantillon. Cette approche nous reconduirait dans les traces de la précédente méthode, mais dont les différentes difficultés seraient amplifiées : unicité de la solution, divergences et minima locaux des algorithmes d'optimisation... Elle serait d'autant plus fastidieuse à implémenter que l'objectif est ici de coller point par point à des lois non identifiables : il y aurait un très grand nombre de solutions à prédéfinir et à tester afin d'anticiper tous les cas possibles, ce qui ne serait pas viable.

L'approche utilisée pour l'identification de la loi de la valeur vraie n'est donc pas adaptée. Contrairement à cette approche qui recherche une solution globale, la stratégie est ici de reconstruire la solution par une multitude d'évaluations. Pour ce faire, la loi des valeurs vraies est d'abord modélisée par une fonction dont la valeur des paramètres est recalculée à la résolution de chaque système d'équation. La fonction linéaire, qui est la modélisation la plus simple et qui ne requiert l'identification que d'un seul paramètre, est utilisée ici pour illustrer la démarche : si la loi des valeurs vraies est fixée comme étant la fonction affine $f_T(x) = \lambda x$, $\lambda \in \mathbb{R}, \forall x \in \mathbb{R}$, alors la valeur du paramètre λ est directement donnée en fonction de $m^{(i)}$:

$$\begin{aligned} f_M^{KS}(m^{(i)}) &\simeq \sum_{j=1}^k \omega^{(j)} f_T(m^{(i)} - \xi^{(j)}) \\ \Leftrightarrow f_M^{KS}(m^{(i)}) &\simeq \sum_{j=1}^k \omega^{(j)} \lambda (m^{(i)} - \xi^{(j)}) \\ \Leftrightarrow \lambda &\simeq f_M^{KS}(m^{(i)}) / \left(\sum_{j=1}^k \omega^{(j)} (m^{(i)} - \xi^{(j)}) \right) \end{aligned} \quad (2.16)$$

La modélisation qui a été retenue dans ces travaux est une fonction exponentielle qui présente des avantages mathématiques connus : reconstruction de courbe, définie sur \mathbb{R} tout entier, dérivable, sans racine ni asymptote verticale, et semblable au noyau de l'estimation. La fonction exponentielle a l'avantage d'être positive quelle que soit la valeur de ses paramètres (contrairement à la fonction affine précédente qui pouvait potentiellement aboutir à des densités négatives si λ était négatif). Le paramétrage de la fonction est donné par les coefficients d'un polynôme du deuxième degré pour offrir une plus grande finesse dans l'approximation locale.

$$f_M^{KS}(m^{(i)}) \simeq \sum_{j=1}^k \omega^{(j)} \exp \left(a (\xi^{(j)})^2 - b (\xi^{(j)}) + c \right) \quad (2.17)$$

Contrairement à l'Équation (2.16) de la fonction linéaire, la finesse de cette modélisation requiert cependant d'identifier trois inconnues. Deux équations sont ajoutées au système en considérant des points très proches de part et d'autre de la valeur $m^{(i)}$. Un paramètre de faible valeur h est introduit à cet effet :

$$\begin{cases} f_M^{KS}(m^{(i)} - h) \simeq \sum_{j=1}^k \omega^{(j)} \exp \left[a(\xi^{(j)} - h)^2 - b(\xi^{(j)} - h) + c \right] \\ f_M^{KS}(m^{(i)}) \simeq \sum_{j=1}^k \omega^{(j)} \exp \left[a(\xi^{(j)})^2 - b(\xi^{(j)}) + c \right] \\ f_M^{KS}(m^{(i)} + h) \simeq \sum_{j=1}^k \omega^{(j)} \exp \left[a(\xi^{(j)} + h)^2 - b(\xi^{(j)} + h) + c \right] \end{cases} \quad (2.18)$$

Au final, l'application de cette méthode n'est pas réservée à l'évaluation des densités aux valeurs $m^{(i)}$. L'approximation locale peut ainsi être réalisée en un grand nombre de points d'une plage de valeurs arbitrairement choisies, pour reconstruire avec précision la loi de la valeur vraie. Aussi, si la loi de la valeur vraie est modélisée par une fonction avec n paramètres à identifier, il suffira de réaliser un même nombre d'évaluations locales pour que le système reste solvable [57].

2.2 Recherche de la valeur vraie d'une mesure

2.2.1 Application de l'inférence bayésienne

Les méthodes de déconvolution développées dans la Section 2.1 ont pour but d'obtenir une meilleure lecture de l'échantillon de mesures, et donc une meilleure approximation de la loi qui décrit les valeurs vraies issues de la production. Dans cette section, il s'agit d'améliorer la lecture de chaque mesure prise individuellement, en utilisant les connaissances de la loi de la production et celle de l'erreur de mesure. Si la déconvolution permet d'accéder à une bonne estimation de la loi de la production, cette connaissance peut aussi être le fruit d'une expertise (théorique, expérimentale) et donc renseignée arbitrairement.

L'inférence bayésienne est utilisée ici dans le but d'intégrer deux sources d'information (la loi de la production et celle de l'erreur de mesure) pour avoir une idée plus précise de la position de la valeur vraie d'une mesure donnée. L'équation 1.6 exposée dans le Chapitre 1 découle du théorème de Bayes. L'application de l'inférence bayésienne dans le contexte de cette recherche donne :

$$p(x|m^{(i)}) = \frac{1}{C} \mathcal{L}(m^{(i)}|x) f_T(x) \quad (2.19)$$

avec $m^{(i)}$ une mesure ciblée, p la loi *a posteriori*, π celle *a priori*, \mathcal{L} celle de la vraisemblance et C la constante d'intégration. D'un point de vue pratique, la constante C n'est pas un sujet d'intérêt aussi la formule qui reflète le mieux l'implémentation numérique du problème est :

$$p(x|m^{(i)}) \propto \mathcal{L}(m^{(i)}|x) f_T(x) \quad (2.20)$$

Le terme $p(x|m^{(i)})$ est la densité de probabilité conditionnelle qu'une valeur x soit la valeur vraie du mesurande pour laquelle on a mesuré $m^{(i)}$. Selon Goller : "L'approche bayésienne met à jour la plausibilité relative de chaque modèle dans un ensemble de modèles candidats, qui est quantifiée par la loi de probabilité postérieure" [59]. En effet, l'application de l'inférence bayésienne est la confrontation de la probabilité de deux implications qui découlent de cette proposition : l'une est

d'avoir produit la valeur vraie x , l'autre, en conséquence, est d'avoir commis l'erreur de mesure sachant x .

La loi de la vraisemblance \mathcal{L} évalue la probabilité d'obtenir une mesure pour une valeur vraie donnée. Elle est fortement liée à la loi de l'erreur de mesure, et sa définition exacte dépend de la modélisation. Un schéma de sa construction est proposé en Figure 2.4. Par exemple, avec la modélisation de l'erreur de mesure de cette thèse, si la concentration d'une solution a pour valeur vraie $0.10gL^{-1}$, la probabilité qu'elle soit mesurée à $+0.11gL^{-1}$ est :

$$\mathcal{L}(0.11|0.10) = f_E(0.11 - 0.10) =$$

En généralisant l'expression on a :

$$\mathcal{L}(m^{(i)}|x) = f_E(m^{(i)} - x) \quad (2.21)$$

La loi de la vraisemblance peut être utilisée comme une estimation de la position de la valeur vraie. Conformément à ce qui est dénoncé dans le premier chapitre, il est alors fort probable qu'elle soit mal centrée. Ce n'est justement pas cette information brute qui est utilisée pour calculer des probabilités dans le cadre bayésien, mais un produit enrichi de la connaissance *a priori* : la loi *a posteriori*. C'est cet enrichissement qui distingue la méthode bayésienne de l'approche classique.

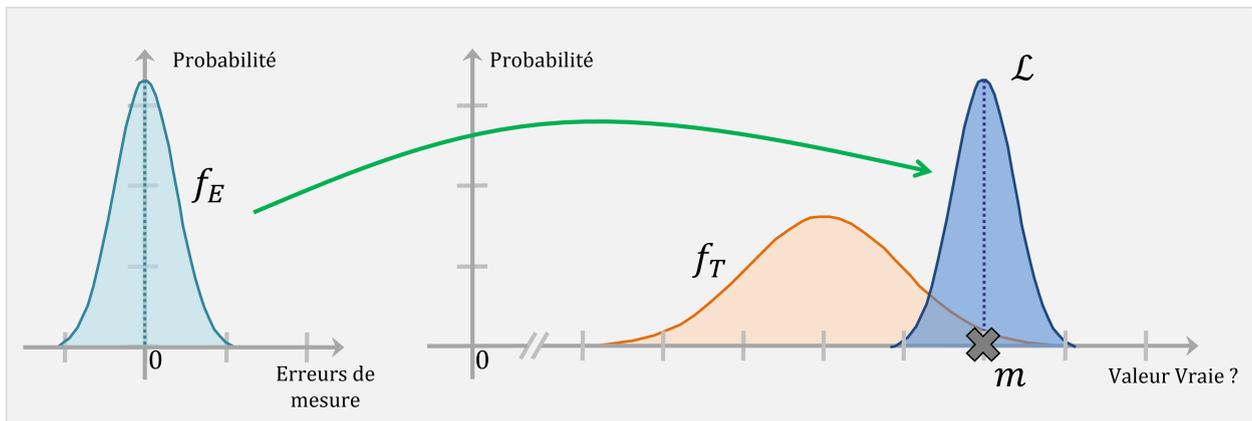


FIGURE 2.4: Schéma de la construction de la loi de vraisemblance.

Boreux [15] prévient que l'application de l'inférence bayésienne à une donnée qui a aussi contribué à construire la loi *a priori* n'est pas conforme aux règles statistiques, que la donnée s'auto-corrigerait en partie. Cela se traduit dans cette étude par le fait que la loi déconvoluée d'un échantillon de mesures ne doit pas être utilisée comme connaissance *a priori* pour l'application de l'inférence bayésienne à chacune de ses mesures.

La mesure $m^{(i)}$ sur laquelle est appliquée l'inférence bayésienne doit donc faire partie d'un lot de mesures "comparable" à celui qui a été utilisé lors de la déconvolution. Deux lots sont comparables si leurs mesures sont issues d'une même loi de l'erreur de mesure et d'une même loi de la valeur vraie. Du point de vue pratique, des lots mesurés sont comparables s'ils sont produits de la même façon et dans un même environnement. La différence entre ces deux lots pourrait presque se réduire à la date de production.

L'inférence bayésienne est appliquée une unique fois et identiquement à toutes les mesures de l'échantillon (les calculs peuvent être exécutés en parallèle). Il n'y a pas d'aléatoire dans cette méthode : deux exécutions parallèles d'un même calcul aboutissent nécessairement au même résultat.

La Figure 2.5 propose des illustrations pour trois exemples d'application de l'inférence bayésienne. D'autres illustrations de cette méthode seront très facilement trouvables dans la littérature [83, 51]. Dans le premier exemple (Figure 2.5(a)), la densité de probabilité des erreurs de mesure est définie par une loi uniforme centrée sur zéro et dont les bornes sont -0.3 et $+0.3$. Cette loi signifie d'une part que la dispersion introduite par les erreurs de mesure est nécessairement comprise dans ces bornes, et d'autre part que toutes les valeurs possibles dans cet intervalle sont équiprobables. La vraisemblance est l'application de la loi de l'erreur de mesure à une mesure donnée. Comme montré en Figure 2.5(d), si la mesure est égale à 101.2 , alors la valeur vraie est contenue dans un intervalle de recouvrement dont les bornes sont connues et valent 100.9 et 101.5 . Cependant, il n'y a pas d'information supplémentaire pour estimer la valeur vraie dans cet intervalle. La seconde information est la loi *a priori*, la loi de la valeur vraie (ou du moins une bonne approximation, qui peut être obtenue suite à la déconvolution). Cette loi est fixée arbitrairement comme suivant une loi normale de moyenne 101 et d'écart-type 0.4 pour les trois exemples. Dans la Figure 2.5(g), les lois de la vraisemblance et de l'*a priori* sont combinées pour donner la loi *a posteriori*. Cette dernière est une troncature de l'*a priori* donnant des informations complémentaires sur la valeur vraie qui, toujours en étant comprise entre 100.9 et 101.5 , est probablement inférieure à 100.2 . Sans surprise, la valeur vraie la plus probable est 101 . On remarquera au passage que si l'une des deux lois suit une loi uniforme entre $-\infty$ et $+\infty$, alors il n'y a aucun gain d'information. Si c'est le cas de l'*a priori* on parle alors d'un *a priori* "non-informatif".

Le second exemple en Figure 2.5(b) met en jeu une loi de l'erreur de mesure triangulaire, symétrique et centrée sur zéro. Cette fois un mode apparaît clairement : l'erreur de mesure est généralement proche de zéro. Cette loi garantit aussi que l'erreur de mesure est strictement contenue dans ses bornes -0.5 et $+0.5$ et que, dans cet intervalle, plus elle est loin de zéro moins elle est probable. L'application de cette loi triangulaire à une mesure (comme montré en Figure 2.5(e)) donne une vraisemblance de la même forme. En conséquence, si la mesure est 100.2 , il est plus probable que la valeur vraie vaille 100.3 plutôt que 100.4 (et symétriquement 100.1 plutôt que 100.0), de la même manière qu'il est plus probable que les erreurs de mesure valent 0.1 plutôt que 0.2 (et -0.1 plutôt que -0.2). La loi *a posteriori* en Figure 2.5(h) ressemble aussi à une loi triangulaire mais n'est plus symétrique : en considérant la loi *a priori*, il est plus probable que la valeur vraie soit inférieure à la mesure, que supérieure. Pour autant, la valeur vraie la plus probable reste très proche de la mesure.

La loi de l'erreur de mesure du troisième exemple (Figure 2.5(c)) a été fixée avec une évidente asymétrie pour que l'effet miroir de la définition de la vraisemblance donnée en Équation (2.21) soit nettement visible en Figure 2.5(f). Il est fort probable que l'erreur de mesure soit égale à 0.5 , ainsi il est tout à fait vraisemblable que la mesure 101.2 provienne de la valeur vraie 101.7 . Cette conclusion n'est cependant pas en accord avec la loi *a priori* qui suggère une valeur vraie plus probablement située autour de 101.1 . La loi *a posteriori* affichée en Figure 2.5(i) affiche un compromis entre ces deux informations pour conclure que la valeur vraie la plus probable est autour de 101.35 .

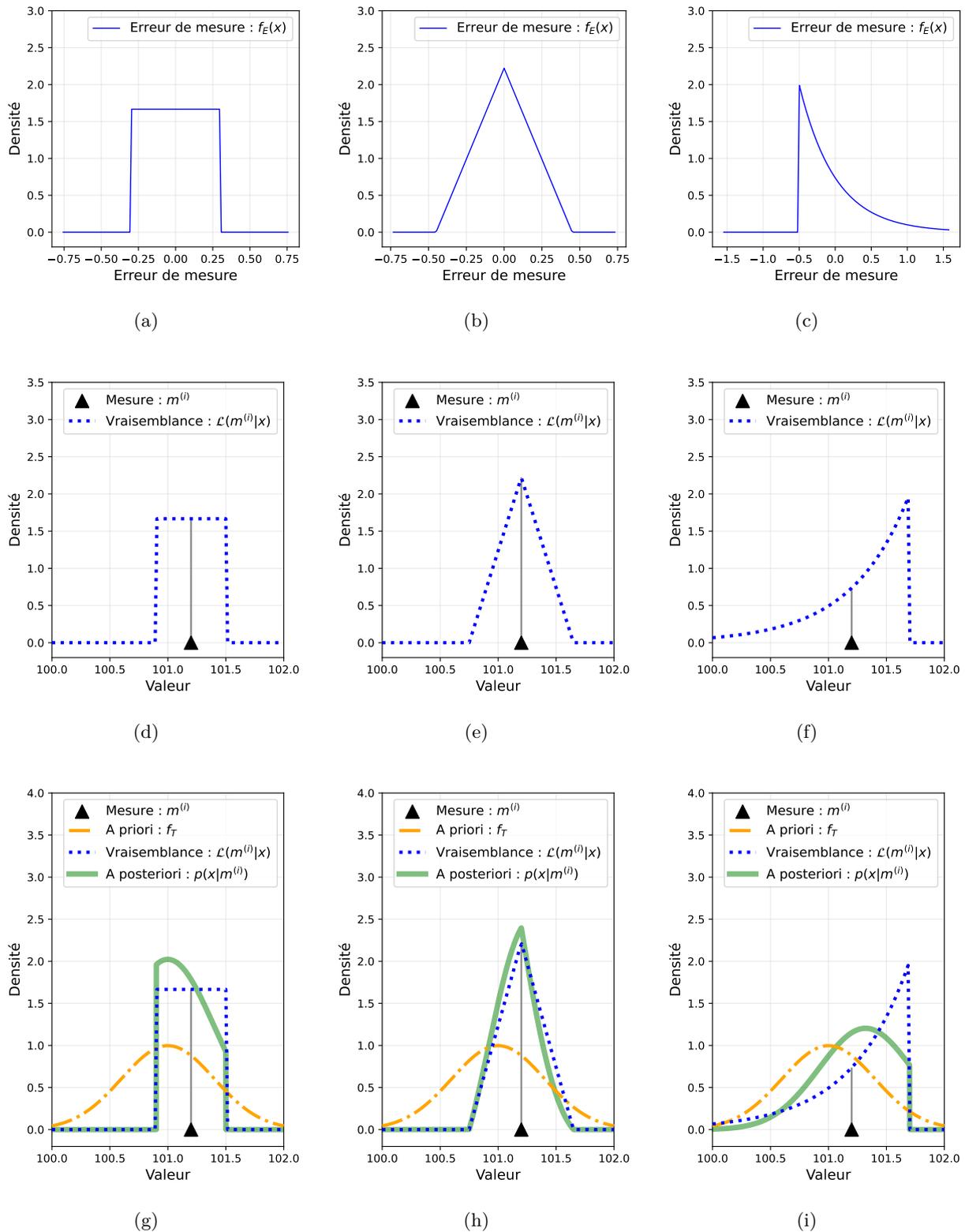


FIGURE 2.5: Trois exemples d'application.

Par l'application de l'inférence bayésienne, "le savoir initial [...] est [mis] à jour en utilisant les données expérimentales" [59]. Ici, la connaissance *a priori* de la production est mise à jour pour chaque mesure afin d'obtenir la densité de probabilité dite *a posteriori*.

Lorsque l'inférence bayésienne est appliquée à plusieurs mesures d'un même lot, seule la position de la loi *a posteriori* varie (mode, écart-type et autres propriétés de la loi sont inchangés), et ce parce que les loi *a priori* et la loi de l'erreur de mesure sont les mêmes quelle que soit la mesure.

Connaissant la loi *a posteriori* il est possible de calculer des probabilités cumulées (ou inversement des quantiles) pour connaître le risque qu'une mesure soit hors tolérance (ou connaître la dernière mesure acceptable selon un seuil d'acceptation donné). Cela permet une application affinée de stratégies comme celle de la bande-de-garde [9].

Il est aussi possible de lire au maximum de la loi *a posteriori* la valeur vraie la plus probable, une quantité résultant de la révision statistique et qui peut se substituer à la mesure initiale. Cette quantité est appelée la valeur révisée, notée $r^{(i)}$. Cette quantité peut être lue comme la "restitution" de "l'observation", pour faire écho aux travaux de Cox sur la même thématique [29]. C'est une approximation de la valeur vraie, le représentant le plus raisonnable, qui est "en moyenne" plus proche de la valeur vraie (il n'y pas de garantie au cas par cas, il est donc tout-à-fait possible qu'une valeur révisée soit pire que la mesure).

2.2.2 Le cas particulier des variables aléatoires gaussiennes

Le cas de variables aléatoires suivant des lois normales a été justifié et étudié dans le cadre de la déconvolution (Section 2.1.1). La simplicité de l'expression de la densité pour les lois normales permet aussi d'aboutir à des résultats simples et analytiques, notamment pour l'*a posteriori* dans le cadre de l'inférence bayésienne, qui requiert d'ordinaire la présence de lois conjuguées [119].

Lorsque les lois de la valeur vraie et de l'erreur de mesure sont normales, la loi *a posteriori* l'est elle aussi, et est donnée par [127] :

$$\left\{ \begin{array}{l} p(x|m^{(i)}) \sim \mathcal{N}(\mu_P, \sigma_P) \\ \mu_P = m^{(i)} \left(\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} \right) + \left(\frac{\mu_T \sigma_E^2 - \mu_E \sigma_T^2}{\sigma_T^2 + \sigma_E^2} \right) \\ \sigma_P = \sqrt{\frac{\sigma_T^2 \sigma_E^2}{\sigma_T^2 + \sigma_E^2}} \end{array} \right. \quad (2.22)$$

avec $p(x|m^{(i)})$ la densité *a posteriori* que x soit la valeur vraie en ayant mesuré $m^{(i)}$, μ_X et σ_X la moyenne et l'écart-type de la variable X qui est soit celle la valeur vraie T soit celle de l'erreur de mesure E . L'écart-type de la loi *a posteriori* vérifie une propriété énoncée dans la section précédente (Section 2.2.1) : sa valeur est constante et donc indépendante de la mesure $m^{(i)}$.

Résultats sur la valeur révisée

La valeur la plus probable d'une loi gaussienne est toujours sa moyenne. La valeur révisée est donc toujours égale à la moyenne de la loi *a posteriori*. Son expression est :

$$r^{(i)} = m^{(i)} \left(\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} \right) + \left(\frac{\mu_T \sigma_E^2 - \mu_E \sigma_T^2}{\sigma_T^2 + \sigma_E^2} \right) \quad (2.23)$$

Il est intéressant de relever que tous ces paramètres sont identiques, quelle que soit la mesure $m^{(i)}$. Il y a ainsi une relation linéaire entre la valeur révisée et la mesure.

Dans le cas (plutôt fréquent) où la moyenne des erreurs de mesure est nulle, l'expression de la valeur révisée est alors celle d'un barycentre :

$$\mu_E = 0 \Rightarrow r^{(i)} = \frac{m^{(i)}\sigma_T^2 + \mu_T\sigma_E^2}{\sigma_T^2 + \sigma_E^2} \quad (2.24)$$

Ce résultat signifie que la valeur révisée sera nécessairement comprise entre la mesure et la moyenne (estimée par déconvolution) de la valeur vraie. Elle sera à équidistance de ces deux grandeurs si la dispersion des erreurs de mesure est égale à celle de la valeur vraie. Comme la dispersion des erreurs de mesure est généralement inférieure à celle de la valeur vraie en réalité (et que l'étude des mesures obtenues en métrologie n'est réellement pertinente que dans cette configuration-là), il est attendu que la valeur révisée soit plus proche de la mesure que de la moyenne de la valeur vraie.

L'application de l'inférence bayésienne permet d'obtenir des valeurs révisées qui sont en moyenne plus proches des valeurs vraies que les mesures. Il est pertinent de les utiliser à la place des mesures pour évaluer la qualité de la production. Il faut cependant anticiper que si la valeur révisée est utilisée afin de juger de la conformité, alors le risque de se tromper peut atteindre 50% dans le pire des cas. Ce risque maximum, illustré en Figure 2.6, advient lorsque la valeur révisée coïncide avec l'une des bornes de l'intervalle de tolérance, et donc que la moitié de la loi *a posteriori* est hors tolérance. Il sera donc préférable de prendre des décisions sur la conformité avec l'information de la loi *a posteriori*, qui permet de connaître et de décider du pourcentage de risque d'erreur.

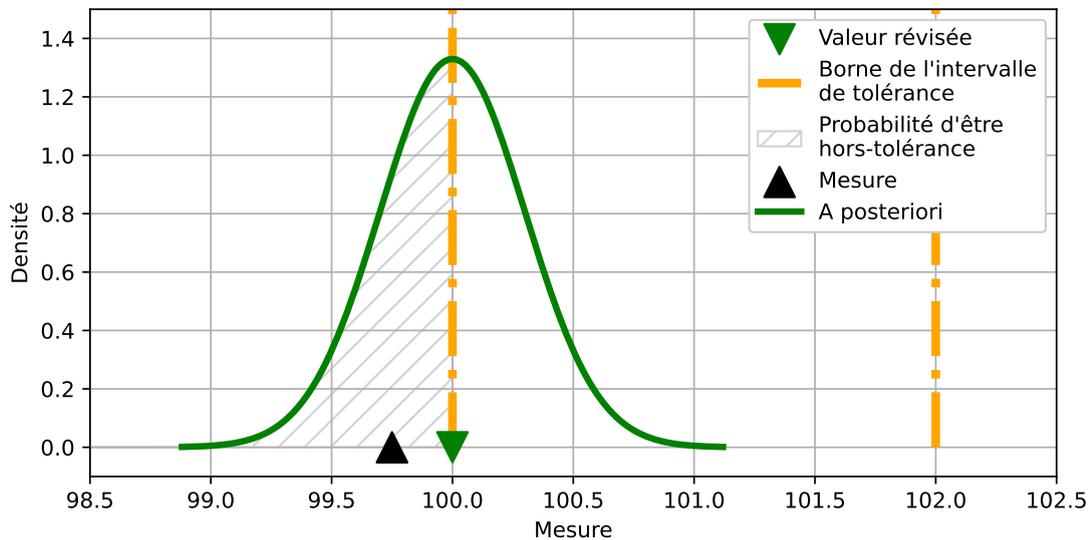


FIGURE 2.6: Le risque maximum (dans le cas gaussien)

Intervalle de tolérance équivalent

Pour juger si une mesure est conforme ou non, il faut la confronter à l'intervalle de tolérance. Et pour avoir un meilleur jugement, on peut transformer la mesure par inférence bayésienne (et obtenir la valeur révisée) avant de la confronter à l'intervalle de tolérance. Or dans le cas gaussien, il se trouve que la fonction de la transformation est un simple polynôme du premier degré. Plutôt que confronter une mesure linéairement transformée à un intervalle, il est aussi possible de confronter

la mesure à un intervalle de tolérance transformé, nommé ici l'intervalle de tolérance "équivalent" :

$$\begin{aligned}
 r^{(i)} \in \text{IT} &\Leftrightarrow r^{(i)} \in [l; u] \\
 &\Leftrightarrow a_0 * m^{(i)} + a_1 \in [l; u] \\
 &\Leftrightarrow m^{(i)} \in [(l - a_1) / a_0; (u - a_1) / a_0] \\
 &\Leftrightarrow m^{(i)} \in [l_{EQ}; u_{EQ}] \\
 &\Leftrightarrow m^{(i)} \in \text{IT}_{EQ}
 \end{aligned} \tag{2.25}$$

avec a_0 et a_1 les deux coefficients du polynôme - facilement identifiables dans l'Équation (2.23), IT l'intervalle de tolérance défini par les bornes inférieure et supérieure l et u , et enfin IT_{EQ} l'intervalle de tolérance équivalent défini par les bornes inférieure et supérieure l_{EQ} et u_{EQ} .

L'utilisation de l'intervalle de tolérance équivalent est une alternative essentiellement pratique par rapport à l'utilisation des valeurs révisées. Cette reformulation permet de simplement communiquer la sélection des mesures conformes plutôt que d'intégrer la notion de valeurs révisées dans tout le reste de la production. Il est important de remarquer du point de vue de la conformité que l'intervalle de tolérance équivalent est fondé sur les valeurs révisées et qu'il n'incorpore donc pas pleinement l'information de la loi *a posteriori*.

2.2.3 Quantification de l'amélioration

L'application de la déconvolution et de l'inférence bayésienne permet d'obtenir une estimation de la valeur vraie (la valeur révisée). Cette section décrit les différents indicateurs déployés pour confronter la qualité des estimations aux mesures initiales : le coefficient de détermination, les résidus, l'erreur quadratique moyenne et le pourcentage de mesures améliorées.

Le coefficient de détermination

Le coefficient de détermination linéaire de Pearson [84] est un score qui permet d'évaluer la qualité des prédictions d'une régression (linéaire ou non). L'idée dans cette application est de considérer que la mesure et la valeur révisée sont des prédictions de la valeur vraie. Le coefficient de détermination permet alors d'évaluer et de comparer la qualité de chacune de ces prédictions :

$$R^2 = 1 - \frac{\sum_{i=1}^n (t^{(i)} - x^{(i)})^2}{\sum_{i=1}^n (t^{(i)} - \mu_T)^2} \tag{2.26}$$

où les $x^{(i)}$ sont soit les mesures, soit les valeurs révisées.

Si l'une de ces deux prédictions est parfaite alors toutes ses valeurs prédites $x^{(i)}$ sont égales aux vraies valeurs $t^{(i)}$ et son score est égal à un. Plus le modèle prédictif sera mauvais et plus le score sera petit. Les incertitudes présentes dans ce modèle font que la recherche de solutions parfaites n'est pas pertinente : les valeurs révisées seront satisfaisantes à partir du moment où leur score sera plus élevé que celui des mesures.

L'expression pour le coefficient de détermination permet que le score R^2 soit nul voire négatif : un tel cas signifie qu'il serait préférable pour l'utilisateur d'abandonner la prédiction (la mesure ou la valeur révisée) et de s'en tenir à la moyenne de la valeur réelle μ_T dont le score R^2 vaut zéro (donc en pratique à l'estimation de μ_T obtenue lors de la déconvolution).

Les écarts types des erreurs de mesure et des valeurs réelles ont un impact de premier ordre sur le coefficient de détermination. On peut en effet voir dans l'Équation (2.26) que μ_T est soustrait dans le calcul, aussi une variation de la moyenne de la valeur vraie n'affecte pas le score R^2 .

On remarque aussi qu'une augmentation de la moyenne de l'erreur de mesure ne détériore que le coefficient de détermination des mesures (Annexe A.7). La sensibilité des méthodes d'amélioration à la moyenne de l'erreur de mesure ne sera pas plus étudiée.

Les résidus des prédictions

Le résidu est l'écart d'une prédiction à la valeur vraie. L'objectif d'amélioration des mesures peut être exprimé comme la réduction de leurs résidus. Cette quantité est choisie pour l'affichage et la comparaison des échantillons des mesures et des valeurs révisées. Les résidus sont obtenus pour chaque échantillon avec la formule :

$$\epsilon = x^{(i)} - t^{(i)} \quad (2.27)$$

ϵ étant le résidu et $x^{(i)}$ correspondant soit aux mesures soit aux valeurs révisées. Dans le cas des mesures, le résidu est l'erreur de mesure.

L'erreur quadratique moyenne

L'erreur quadratique moyenne (ou risque quadratique) est un score qui évalue les écarts de prédiction et qui a l'avantage d'être très largement répandu. Il est utilisé ici en plus du coefficient de détermination. Sa formule est donnée par :

$$MSE_n = \frac{1}{n} \sum_{i=1}^n (x^{(i)} - \mu_T)^2 \quad (2.28)$$

avec les $x^{(i)}$ correspondant soit aux mesures $m^{(i)}$ soit aux valeurs révisées $r^{(i)}$.

Le pourcentage de mesures améliorées

On calcule aussi le pourcentage P_B des mesures pour lesquelles l'inférence bayésienne a été bénéfique, c'est-à-dire dont la valeur révisée est plus proche de la valeur vraie. Ce score apporte une information qui complète celle des deux précédents : en effet, il est possible que l'inférence bayésienne soit globalement favorable, mais que le nombre de cas améliorés soit tout de même plus faible que ceux détériorés (comme il est possible de perdre un match de tennis en ayant inscrit le plus de sets, lorsque ceux perdus sont perdus de peu et que ceux remportés le sont très largement). En utilisant une fonction indicatrice I , l'expression de P_B est :

$$P_B = \frac{\sum_{i=1}^n I^{(i)}}{n}, \quad I^{(i)} = \begin{cases} 1 & \text{si } |r^{(i)} - t^{(i)}| > |m^{(i)} - t^{(i)}| \\ 0 & \text{sinon} \end{cases} \quad (2.29)$$

2.3 Étude de la conformité et des risques associés

2.3.1 Lecture de la conformité

Matrice de confusion

Il est tentant d'affirmer qu'avoir une meilleure idée des valeurs vraies implique un regard plus juste sur la conformité. Or une amélioration globale n'implique pas forcément une meilleure déclaration de conformité (tout comme il y a un besoin de calculer le pourcentage des cas qui ont été amélioré). Il se pourrait, par exemple, que les valeurs révisées les plus profitables ne concernent que les mesures dont la conformité n'était pas discutée. Les précédents scores statistiques, comme le coefficient de détermination ou l'erreur quadratique moyenne, ne permettent pas d'étudier l'évolution de la déclaration de conformité des pièces.

La matrice de confusion est un outil pertinent pour synthétiser et visualiser efficacement la déclaration de conformité. Les matrices de confusion sont un outil très répandu en data science [42, 128, 103, 62, 37]. Elles sont utilisées pour illustrer une "classification", qui tout à fait comparable à une déclaration de conformité. Comme on trierait les pièces d'une production par couleur, la classification est l'affectation de catégories à un jeu de données.

La matrice de conformité est une matrice carrée. Sur chaque axe apparaît la liste des catégories, identiquement ordonnée. Le nombre de catégories d'une matrice de confusion est fixé par l'utilisateur (on pourra distinguer la couleur de chaque pièce ou uniquement le noir du reste) et les catégories sont cumulables (certaines pièces peuvent être multicolores).

Un exemple de matrice de confusion est présenté en Figure 2.7. Les données sont simulées et l'exemple utilisé (fictif) est celui du tri réalisé par une machine imparfaite, d'un lot de pièces de quatre couleurs (blanc, vert, bleu, noir). Une donnée est placée dans la colonne de la catégorie dans laquelle elle est prédite, et dans la ligne de la catégorie dans laquelle elle appartient vraiment. Pour connaître le total par catégorie prédite, on somme les cases de la matrice par colonne ; pour connaître le total par vraie catégorie, on somme les cases de la matrice par ligne. On peut remarquer que la matrice de confusion n'est pas symétrique (si le bleu est parfois confondu avec du noir, ce n'est pas le cas inversement) et qu'elle permet ainsi de déceler un total trompeur (il y a bien 25 linges classés vert au final mais seulement 3/5 à juste titre). La matrice de confusion serait parfaite s'il n'y avait aucune erreur de tri. Dans ce cas, et si les catégories ne sont pas cumulables, il n'y a pas de termes extra-diagonaux (la matrice de confusion est diagonale et sa trace est égale au nombre total de donnée).

Couleur des pièces					Somme (vrai)		Somme (prédiction)					
Vrai	Blanc	25	0	0	0	Σ Vrai	Blanc	25				
	Vert	0	15	10	0		Vert	25				
	Bleu	0	10	10	5		Bleu	25	25	25	20	30
	Noir	0	0	0	25		Noir	25				
		Blanc	Vert	Bleu	Noir							
		Tri de la machine						Σ Tri de la machine				

FIGURE 2.7: Exemple de matrice de confusion de pièces triées par une machine.

La matrice de confusion confronte donc des prédictions à la vérité. Elle s'applique ici en considérant que les conformités lues depuis les mesures ou les valeurs révisées sont toutes les deux des prédictions (de la vraie conformité donnée par les valeurs vraies). Il est aussi possible de construire une matrice de confusion avec les valeurs vraies, si elles sont accessibles, pour obtenir la matrice de référence. Comme il est impossible de remonter à la véritable conformité des données réelles, l'évaluation de l'apport des méthodes de déconvolution et d'inférence bayésienne est réalisée sur des données simulées dans le Chapitre 3.

Dans le cadre de cette étude, les deux seuls statuts non cumulables qui nous intéressent sont "conforme" et "non conforme". Cette configuration est très commune et désignée sous le nom de "classificateur binaire". Elle admet communément le vocabulaire suivant :

- Parmi les résultats correctement prédis, on nomme :
 - Vrai positif (TP, true positive) : une pièce conforme classée conforme.
 - Vrai négatif (TN, true negative) : une pièce non conforme classée non conforme.
- Parmi les prédictions défectueuses, on nomme :
 - Faux positif (FP, false positive) : une pièce non conforme classée conforme.
 - Faux négatif (FN, false negative) : une pièce conforme classée non conforme.

Dans le cas des faux positifs, on parle de "risque client" car c'est au client que sont laissées les conséquences d'une pièce certifiée à tort. Dans le cas des faux négatifs on parle de "risque fournisseur" car c'est au niveau de la production que l'on subit les conséquences de pièces erronément rebutées ou retouchées.

A défaut d'informations au sujet de l'erreur de mesure, ce sont les mesures incluses dans l'intervalle de tolérance qui sont dites conformes. Dans la matrice de confusion, c'est le caractère positif ou négatif qui reflète la conformité de la mesure par rapport à l'intervalle de tolérance. Un élargissement de la sélection ou un rétrécissement par rapport à l'intervalle de mesure impacte le nombre de faux positifs et le nombre de faux négatifs.

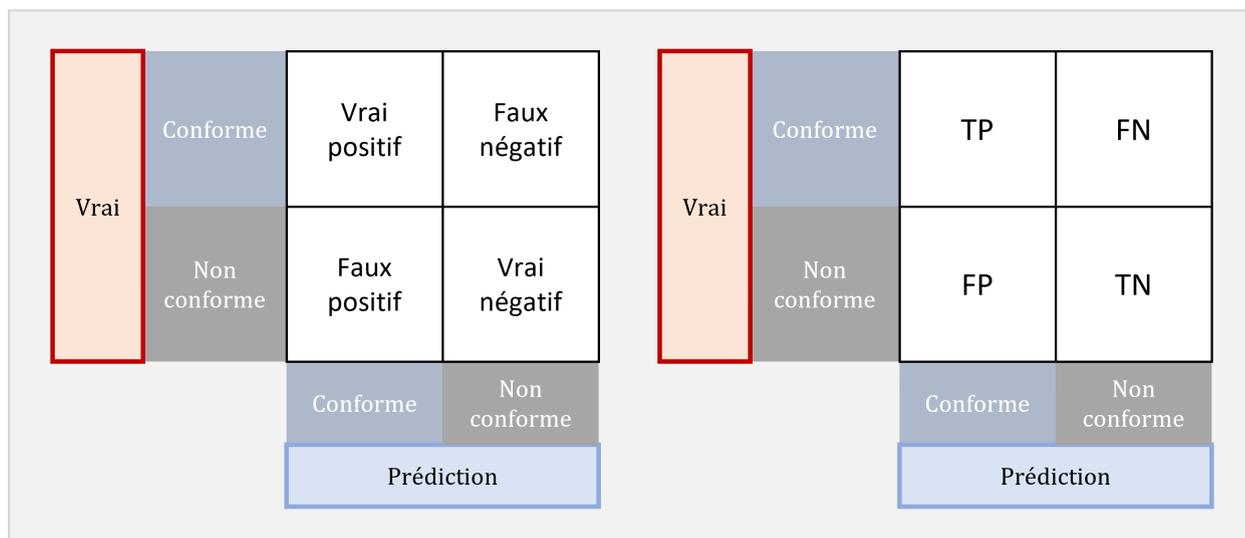


FIGURE 2.8: Disposition des termes de la matrice de confusion (classification binaire).

Une matrice de confusion peut aussi être visualisée dans un repère cartésien [125]. En Figure 2.9 est montré un schéma de ce que l'on peut obtenir (avec toujours la classification prédite sur l'axe des abscisses et celle véritable sur l'axe des ordonnées).

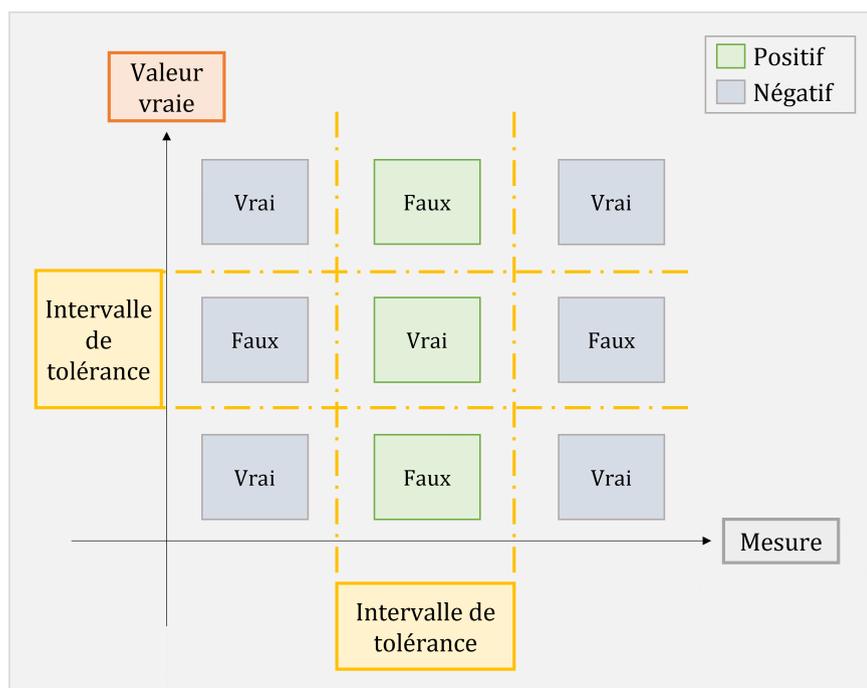


FIGURE 2.9: Schéma de la lecture graphique de la conformité.

Courbe ROC

Il a été affirmé en Section 2.2.2 que le risque maximum assumé en se référant uniquement à la valeur révisée dans le cas gaussien s'élève à 50%; qu'en conséquence il était préférable de juger de la conformité en utilisant la loi *a posteriori* en fixant un seuil acceptable de probabilité d'être hors-tolérance. Suivant que le seuil vaut 0.01, 0.5 ou 3 pourcents, la matrice de confusion change.

La courbe Receiver Operating Characteristic (ROC), traduite "fonction d'efficacité du récepteur", est dessinée par l'ensemble des matrices de confusions pour tous les seuils de décision possibles (entre 0 et 100%). Une courbe ROC est en quelque sorte l'affichage de toutes les performances des prédictions possibles, et permet ainsi de visualiser rapidement le plein potentiel de décision, notamment la matrice de confusion la plus profitable (et donc le seuil associé). De même que la courbe ROC est indissociable des matrices de confusion, elles sont présentées de pair dans les sources bibliographiques mentionnées en Section 2.3.1.

Pour représenter une matrice de confusion en un seul point dans un graphique en deux dimensions, il faut condenser ses quatre informations (TP, FP, FN, TN) en deux. Il existe alors de nombreuses combinaisons possibles. Les informations sont communément regroupées pour calculer des taux, suivant une lecture verticale ou horizontale de la matrice. Et les quantités les plus communément adoptées sont le taux de vrais positifs (TP_R , true positive rate) et le taux de faux positifs (FP_R , false positive rate) qui correspondent à une lecture horizontale de la matrice (Figure 2.10), et dont la valeur est comprise entre zéro et un. Ils sont définis tels que :

$$TP_R = \frac{TP}{TP + FN} \quad (2.30)$$

$$FP_R = \frac{FP}{FP + TN}$$

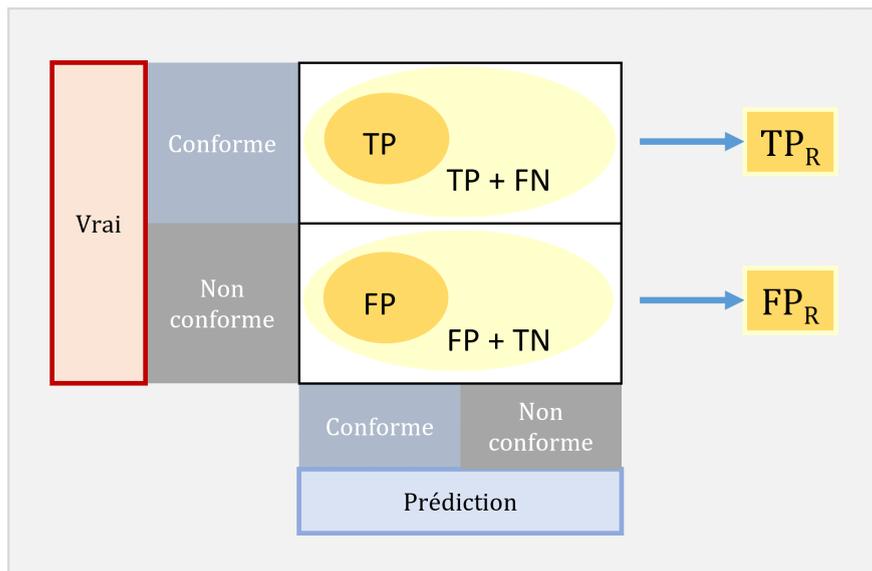


FIGURE 2.10: Illustration des ratios utilisés pour tracer la courbe ROC.

Le taux de vrais positifs est affiché sur l'axe des ordonnées. Il caractérise la proportion de pièces prédites conformes parmi toutes celles qui le sont vraiment. En pratique, à moins d'une sélection très sévère, la valeur de ce taux est très proche de un car le nombre de faux négatifs est généralement négligeable devant le nombre de vrais positifs. La très grande majorité des pièces produites sont en effet conformes (dans l'industrie on compte souvent les pièces défectueuses en *ppm* - pièces par millions) et reconnues comme tel.

Le taux de faux positifs est affiché en abscisse. Il caractérise la proportion de pièces que l'on croit conforme parmi toutes celles non conformes. En pratique, sa valeur varie facilement entre 0

et 1 car le taux est d'autant plus sensible qu'il y a peu de pièces non-conformes.

En utilisant les taux de vrais positifs et de faux positifs pour construire la courbe ROC, l'ajout d'un faux positif est nettement plus visible que l'ajout d'un faux négatif. La courbe ROC oriente donc la lecture vers un coût client plus impactant que le coût fournisseur. En pratique, il est en effet plus dangereux d'omettre une erreur dont on ne maîtrisera pas les conséquences que d'opter pour une sélection plus stricte qui garantit le bon déroulé de la production.

La matrice de confusion de coordonnées $(0, 1)$ (voir Figure 2.11) est donc celle d'une prédiction parfaite. Située en haut à gauche du graphique, elle constitue le pôle idéal vers lequel tout modèle prédictif doit tendre. A l'opposé, dans le coin en bas à droite se situe la pire des matrices de confusion possibles, celle de coordonnées $(1, 0)$ qui prédit avec une grande justesse l'inverse de la vérité.

Dans le cas où l'on accepterait toutes les pièces à l'aveugle, les coordonnées de la matrice de confusion sont $(1, 1)$. Et si elles sont refusées aveuglément, les coordonnées sont $(0, 0)$. Un équilibreur de décisions aléatoire passe par ces deux coins du graphique. L'ensemble des points situés sur cet équilibre est celui des matrices de confusion de classifications utilisant diverses proportions d'acceptation et de rejet, et dont la décision s'applique aléatoirement sur les individus de la population. La performance de ces classifications n'est qu'une question de chance.

Une courbe ROC décrit les résultats de tous les niveaux d'exigences pour la stratégie de sélection, allant du rejet total à l'acceptation totale. Un indicateur fréquemment utilisé pour quantifier l'efficacité du récepteur (ici de notre courbe *a posteriori*) est l'aire sous la courbe (AUC, "Area Under Curve" en anglais). L'aire de la courbe est comprise entre zéro et un :

- Si toutes les pièces sont vraiment conformes alors la courbe ROC est confondue avec l'axe des ordonnées et son aire est nulle.
- Si toutes les pièces sont vraiment non-conformes alors la courbe ROC est confondue avec l'axe des abscisses et son aire est nulle.
- Tant qu'il y a au moins une pièce vraiment conforme et une pièce vraiment non-conforme, chaque courbe ROC commence à la coordonnée $(0, 0)$ et aboutit à $(1, 1)$. L'aire d'un récepteur parfait tend vers 1. L'aire d'un récepteur aléatoire est 0.5.

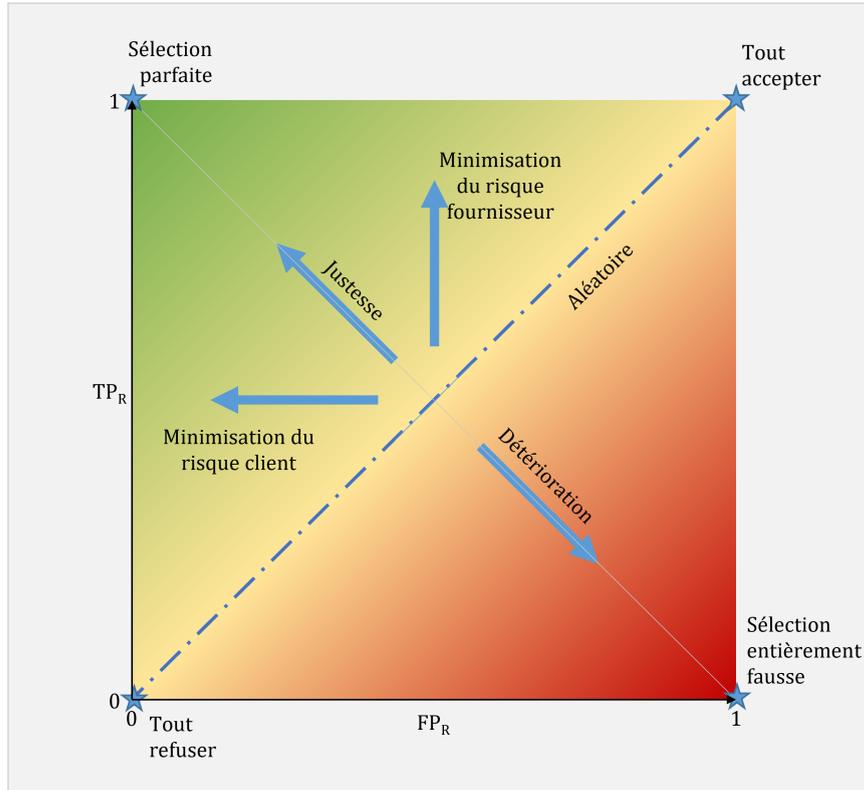


FIGURE 2.11: Différents axes et pôles du graphique de la courbe ROC.

2.3.2 Probabilités et risques de mal-classifier

Reconstruire l'origine de la mesure

Cette section traite de la conformité de chaque mesure $m^{(i)}$, là où l'utilisation des matrices de confusion et de la courbe ROC ont pour but de caractériser la production toute entière, en extrayant l'information d'un échantillon de très grande taille.

Une mesure est la somme d'une valeur vraie et d'une erreur de mesure (Équation (1.1)). Les combinaisons $(t^{(i)}, e^{(i)})$ dont la somme vaut $m^{(i)}$ sont infinies mais non équiprobables : certaines seront fréquentes quand d'autres seront improbables, voire impossibles. La reconstruction des différentes combinaisons est triviale, avec un seul degré de liberté :

$$\begin{cases} t^{(i)} = m^{(i)} - \lambda \\ e^{(i)} = \lambda \end{cases}, \quad \forall \lambda \in \mathbb{R} \quad (2.31)$$

A l'instar de l'inférence bayésienne, il faut connaître les lois de la valeur vraie et de l'erreur de mesure pour calculer la probabilité qu'un couple $(t^{(i)}, e^{(i)})$ apparaisse : il faut connaître la loi de la valeur vraie pour savoir densité de probabilité de la composante $t^{(i)}$, et celle de l'erreur de mesure pour savoir la densité de probabilité de la composante $e^{(i)}$. Dans cette étude la loi de l'erreur de mesure est une donnée d'entrée connue. La loi de la valeur vraie, elle, n'est parfaitement connaissable que dans le cadre de la simulation. Mais la déconvolution présentée au Chapitre 2.1 permet cependant d'en avoir une estimation en réalité.

La question de la classification intervient après l'application d'un intervalle de tolérance. Il faut alors discerner deux informations :

- La conformité de la mesure : la conformité de la mesure est état de fait, non discutable, qui est tranché par la position de la mesure par rapport à l'intervalle de tolérance.
- L'acceptation de la mesure : la décision d'accepter ou rejeter une mesure est une décision libre, qui peut être réalisée indépendamment de la conformité de la mesure et donc discutable.

La conformité de la mesure étant une chose établie, chaque décision au sujet de la conformité ne peut déboucher que sur deux issues. Dans le but d'affiner la maîtrise des risques de production, les deux événements d'intérêt sont :

- La probabilité d'accepter à tort la mesure (de générer un faux positif - FP, d'augmenter le risque client)
- La probabilité de rejeter à tort la mesure (de générer un faux négatif - FN, d'augmenter le risque fournisseur).

Ces problèmes pourraient aussi être étudiés sous l'angle de la "probabilité d'avoir raison" et les résultats seraient simplement intervertis car ces événements sont complémentaires.

La probabilité de prendre une mauvaise décision

Pour chaque combinaison potentiellement à l'origine de la mesure $m^{(i)}$, on peut lire si la valeur vraie est conforme ou non (appartient à l'intervalle ou non). Parce que les combinaisons sont des reconstructions théoriques, cette information est accessible indépendamment du fait que les données soient réelles ou simulées.

Pour connaître la probabilité de prendre une mauvaise décision, il suffit alors de cumuler la probabilité d'occurrence de toutes les combinaisons dont le statut de la valeur vraie n'est pas en accord avec la décision.

Un exemple discret est présenté Figure 2.12. L'erreur de mesure suit une loi uniforme discrète définie entre -0.3 et $+0.3$. La valeur vraie suit une loi uniforme discrète définie entre 100 et 102.2. L'intervalle de tolérance a pour bornes $[100.1, 101.5]$. La mesure étudiée (101.7) marquée d'une croix noire sur l'axe des abscisses est non-conforme. En dessous, toutes les combinaisons possibles qui en sont potentiellement à l'origine sont représentées (elles sont au nombre de 7, les autres ont une probabilité nulle de survenir). Toutes les combinaisons sont équiprobables dans cet exemple.

- Si la pièce est déclarée non-conforme (donc rejetée), la probabilité de prendre une mauvaise décision est de 2 chances sur 7 car il existe 2 combinaisons dont la valeur vraie est conforme.
- Si la pièce est déclarée conforme (donc acceptée), la probabilité de prendre une mauvaise décision est de 5 chances sur 7 car il existe 5 combinaisons dont la valeur vraie est non-conforme.

Cet exemple permet d'observer que les probabilités de prendre une mauvaise décision en acceptant ou en rejetant la mesure sont celles de deux événements complémentaires - dont la somme vaut un.

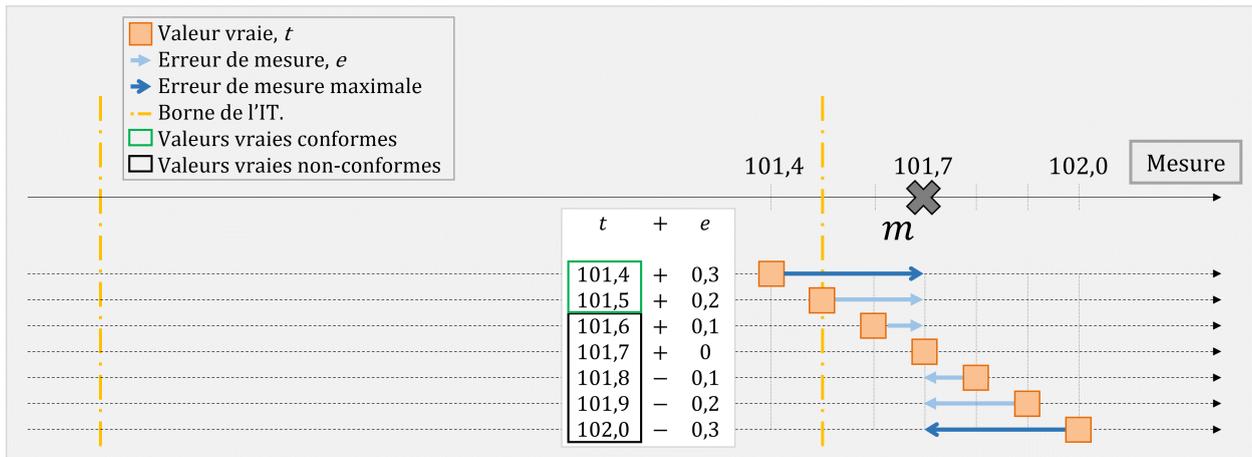


FIGURE 2.12: Schéma des combinaisons possibles.

Les courbes de probabilité de prendre une mauvaise décision

Puisqu'il est possible de calculer les probabilités de prendre une mauvaise décision en acceptant et rejetant une mesure donnée, il est possible de tracer deux courbes en envisageant toutes les mesures possibles. En guise d'exemple, on considère la production définie arbitrairement et décrite dans le Tableau 2.1. Les bornes de l'intervalle de tolérance sont fixées à $[723, 751]$.

TABLEAU 2.1: Paramétrage de l'exemple.

Variable aléatoire	Distribution	Minimum	Maximum
Valeur vraie	Uniforme	728	760
Erreur de mesure	Uniforme	-4	+4

On trace en Figure 2.13 les deux courbes de probabilité de prendre une mauvaise décision. On peut remarquer en particulier le "domaine des mesures", qui est l'intervalle de toutes les mesures possibles et qu'il est aisé d'identifier dans le cas de lois uniformes (ici entre 724 et 764). Pour des fonctions définies sur \mathbb{R} entier, ce domaine peut être approximé en considérant un très faible quantile (de l'ordre de 10^{-5}). On constate que :

- La borne inférieure de l'intervalle de tolérance n'est pas dans le domaine de mesures et qu'elle ne soulève donc aucun doute vis à vis de la conformité.
- La borne supérieure l'est :
 - Si la mesure est confondue avec à la borne supérieure de l'intervalle de tolérance, il est impossible de savoir s'il est plus probable que la valeur vraie soit conforme ou non.
 - De part et d'autre de cette borne supérieure les tendances sont intuitives : une mesure hors tolérance est plus probablement le signe d'une valeur vraie hors tolérance, et vice-versa.
 - Les tracés de ces tendances sont comparables à la fonction de répartition de l'erreur de mesure -ou à son complémentaire (le caractère linéaire correspond bien à l'intégrale d'une fonction uniforme).

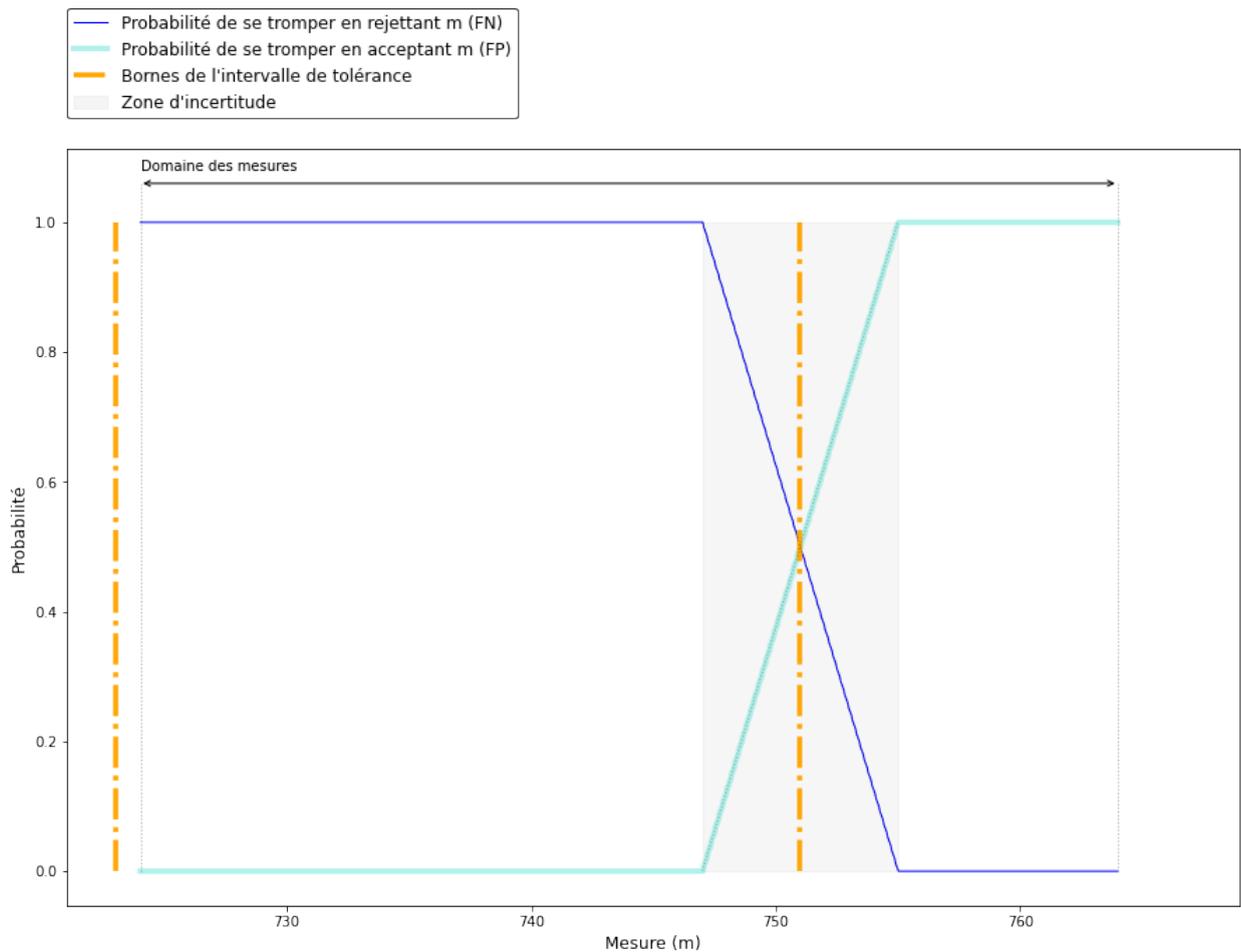


FIGURE 2.13: Exemple de courbes de probabilité de prendre une mauvaise décision (lois uniformes).

Zone non homogène

Selon la configuration de l'étude, il se peut que la liste des combinaisons possibles derrière une mesure $m^{(i)}$ soit inhabituellement tronquée.

Dans la Figure 2.12 précédemment exposée, la liste des combinaisons possible est limitée par la loi de l'erreur de mesure (les combinaisons ne s'étendent en effet pas au-delà des erreurs de mesure maximales). Dans ce cas, on dit que la mesure appartient à la "zone homogène".

Mais il arrive aussi que la loi de la valeur vraie soit encore plus contraignante que celle de l'erreur de mesure. La taille de la liste des combinaisons à étudier est alors inférieure à celle dans la zone homogène et varie selon la valeur de la mesure. Ces mesures font alors partie de ce que l'on nomme ici : la "zone non-homogène".

La Figure 2.14 reprend le schéma des combinaisons possibles précédemment exposé (Figure 2.12). Désormais, la valeur vraie suit une loi uniforme discrète définie entre 100 et 101.7. Les combinaisons dont la valeur vraie était supérieure à 101.7 ne sont plus possibles, ce qui modifie les probabilités de prendre une mauvaise décision. La probabilité d'un faux négatif est accrue au détriment de celle de générer un faux positif. Il y a désormais autant de combinaisons équiprobables ayant des valeurs vraies conformes et non-conformes, donc la même chance de prendre une mauvaise décision quelle

que soit la décision (acceptation ou rejet).

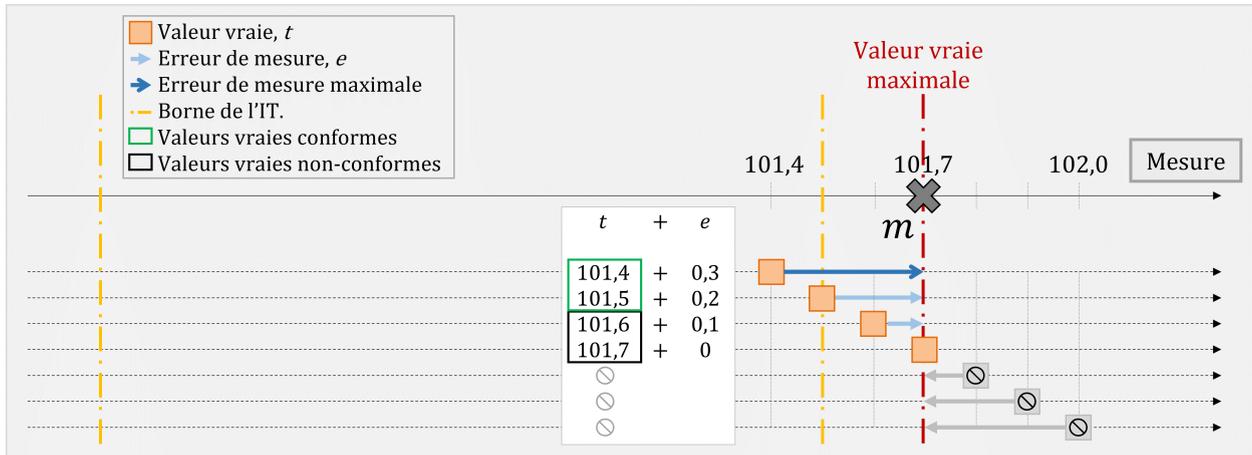


FIGURE 2.14: Exemple d'une mesure dans la zone non homogène.

Les risques d'une mauvaise classification

Il est ensuite possible d'évaluer les risques client et fournisseur de chacune des mesures. Le lecteur trouvera une étude très similaire parmi les travaux de Forbes [49]. Le risque est défini en pondérant les probabilités par le coût réel. Ainsi deux événements équiprobables diffèrent en risque à proportion de leurs conséquences. Le coût réel est propre à chaque cas, et est arbitrairement défini par l'utilisateur. En pratique, les conséquences relatives au risque client sont généralement plus importantes que celles relatives au risque fournisseur (généralement plus maîtrisables) [87].

Suite à la pondération, les courbes de probabilité de prendre une mauvaise décision sont modifiées. Là où le maximum des courbes était de 1 (qui correspond à 100%), il est désormais égal au coût associé. Une conséquence particulièrement intéressante est que les deux courbes ne se croisent plus au même endroit (effet qui peut se cumuler avec celui de la zone non homogène).

La Figure 2.15 montre un exemple de courbes de risques où l'on estime qu'une erreur vis-à-vis du client coûte 10 fois plus qu'une erreur pour le fournisseur. Cet exemple reprend le paramétrage présenté dans le Tableau 2.1 dont les courbes de probabilités d'erreurs sont exposées en Figure 2.13. Les courbes ont désormais pour maximum 1 ou 10 et ne se croisent plus au niveau de la borne supérieure. On pourra remarquer que seul le ratio des coûts importe : il revient au même d'étudier des courbes de risques avec des coûts (1, 10) qu'avec des coûts (50, 500). On peut aussi relever que les deux intersections des courbes FN et FP deviennent plus proches lorsque le coût de l'erreur est plus grand vis-à-vis du client qu'au niveau de la production, et vice-versa.

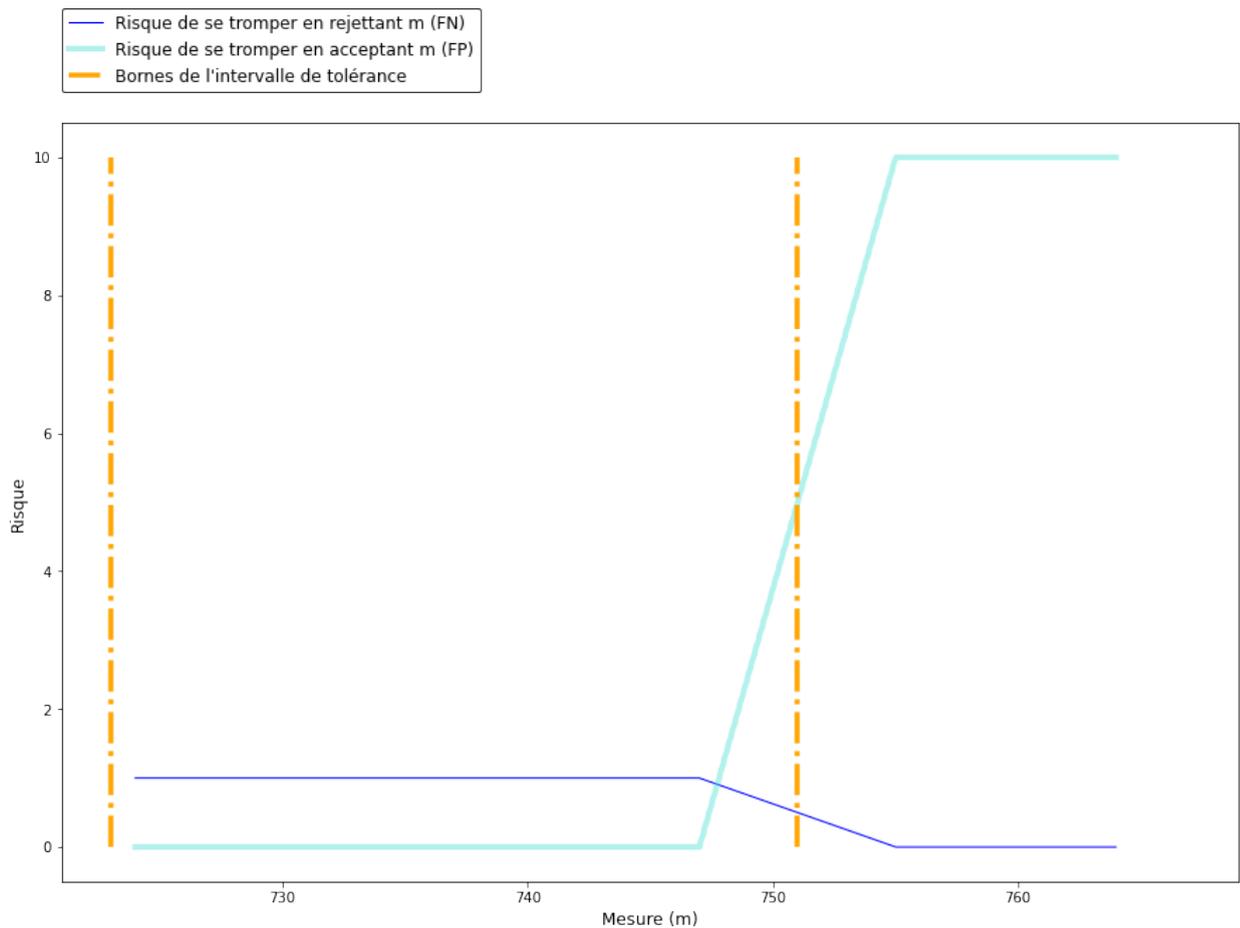


FIGURE 2.15: Exemple de courbes des risques d'erreur (lois uniformes).

La sélection optimale

La "sélection" est la désignation arbitraire de l'ensemble des mesures jugées arbitrairement conformes. Si la connaissance exacte de la conformité de certaines mesures est possible avec des lois uniformes, elle est impossible pour des lois définies sur \mathbb{R} entier : tout choix est risqué. Quelque que soit la connaissance utilisée à l'appui (et même avec une connaissance de la loi *a posteriori*), la sélection se réduit à une plage de mesures, à un intervalle et donc à deux bornes.

En présence d'incertitudes, les mesures sélectionnées ne sont pas celles assurément conformes mais celles dont on peut prendre le risque de les accepter. Lorsque l'on accepte une mesure, c'est que l'on juge que le risque de se tromper plus faible qu'en la refusant. Décider de la conformité d'une mesure c'est choisir la courbe de risques sur laquelle on parie.

La connaissance des courbes de risques permet d'avoir conscience des risques engagés par une sélection (ce qui est un résultat très important en soi, découlant de la déconvolution). Elle ouvre la porte à l'identification de la sélection optimale. On trouvera des travaux similaires publiés par Fernandez dans le but d'équilibrer les risques client et fournisseur [43]. L'étude de la sélection peut se faire dès les courbes de probabilité de prendre une mauvaise décision, mais elles ne seront abordées ici qu'avec les courbes de risques (car il suffit de mettre des coûts égaux à un pour revenir à des courbes de probabilités). La sélection optimale est ici définie comme celle qui minimise le risque total (il est possible de choisir d'autres critères d'optimalité).

En notant α et β les bornes inférieure et supérieure de la sélection arbitrairement effectuée, et r_{FP} et r_{FN} les courbes de risques associées aux probabilités de générer un faux-positif et un faux-négatif, la courbe du risque r_T est :

$$r_T = \begin{cases} r_{FP}(m) & \text{si } m \in [\alpha, \beta] \\ r_{FN}(m) & \text{sinon} \end{cases}, \quad \alpha \in \mathbb{R}, \beta \in [\alpha, +\infty[\quad (2.32)$$

Le risque total R_T est le cumul de tous les risques associés à chaque mesure, qui peut se formaliser sous la forme intégrale :

$$R_T = \int_{-\infty}^{+\infty} r_T(m) dm \quad (2.33)$$

La formule explicite du risque total est décrite dans les travaux de Cox sous le nom de "risque globale" [29].

La Figure 2.16 reprend le précédent exemple dont les courbes de risques étaient données en Figure 2.15. La sélection est celle qui se fonde sur la classification des mesures par rapport à l'intervalle de tolérance $((\alpha, \beta) = (l, u))$. On constate visuellement que le risque total serait minimisé si la borne supérieure de la sélection était placée au niveau de l'intersection des deux courbes. C'est en effet le résultat général : les bornes de la sélection qui minimise le risque total correspondent aux deux intersections des courbes de risques (s'il n'y a pas d'intersection, c'est qu'il n'y a pas de bornes à fixer et donc pas lieu de rejeter les mesures). Le risque total de la sélection optimale pourra être confronté à celui d'une sélection établit selon les bornes de l'intervalle de tolérance.

Il est utile de rappeler que cette information n'est accessible que si l'on connaît la loi des valeurs vraies. L'étape de déconvolution permettant d'en avoir une estimation, il est ensuite possible de construire les courbes des probabilités d'erreur, des risques et finalement la sélection optimale.

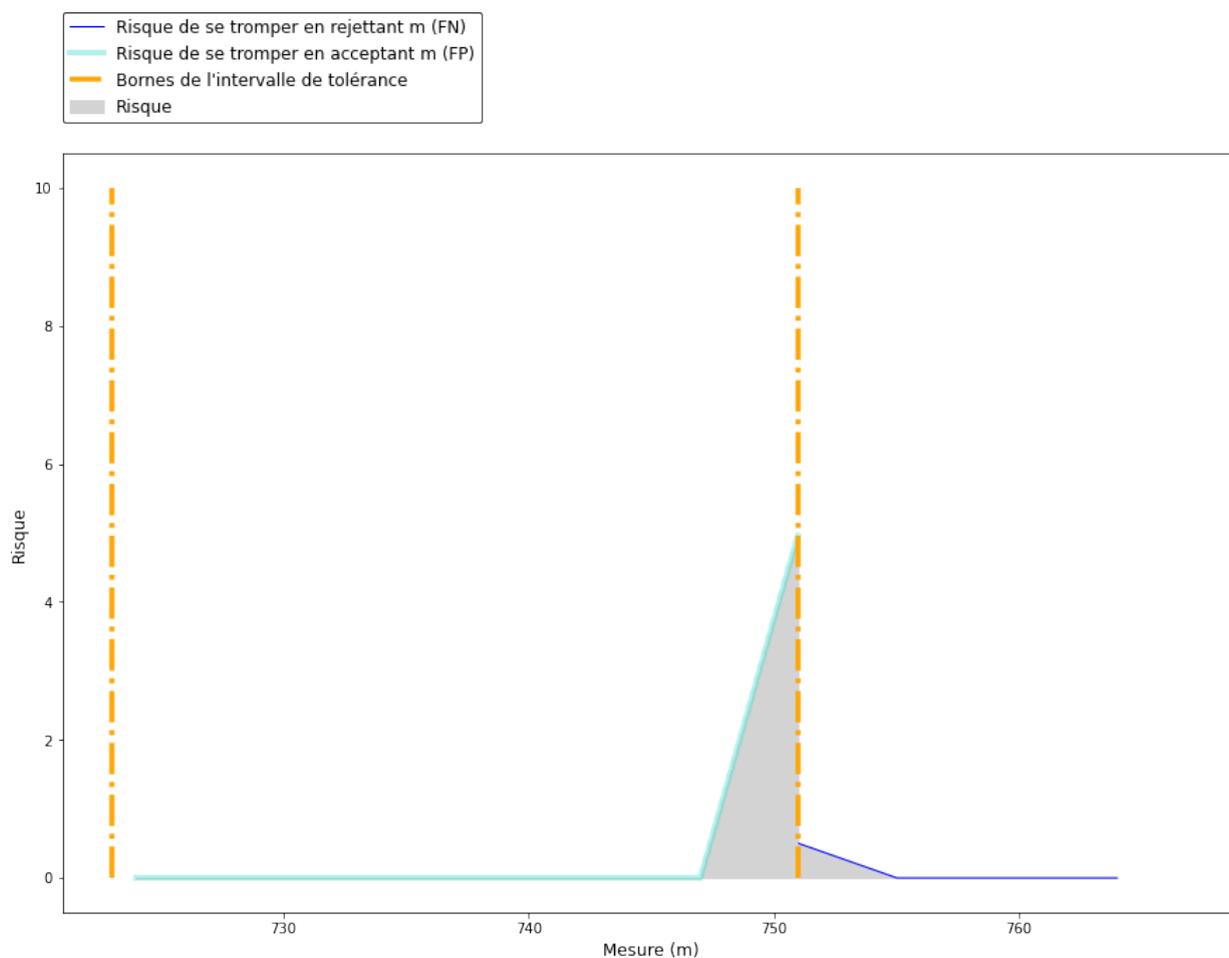


FIGURE 2.16: Illustration du risque total.

Conclusion

Ce chapitre a présenté plusieurs méthodes probabilistes pour l'amélioration de la mesure. Ces méthodes interviennent soit au niveau de la déconvolution soit pour l'estimation de la valeur vraie. Sont détaillées trois méthodes de déconvolution dans la première section, puis l'inférence bayésienne dans la seconde, et enfin les divers outils statistiques pour l'évaluation de ces méthodes dans la dernière.

La première méthode de déconvolution requiert des hypothèses fortes de modélisation : les lois de l'erreur de mesure et de la valeur vraie doivent être gaussiennes. Dans ce cas spécifique, tout le problème peut être résolu de façon analytique et ce jusqu'à l'expression de l'estimation de la valeur vraie ($r^{(i)}$). Cette spécificité permet d'appliquer de façon équivalente la transformation de l'inférence bayésienne à un intervalle de tolérance plutôt qu'aux mesures, pour obtenir la nouvelle sélection (IT_{EQ}). L'implémentation de la méthode et l'exploitation du résultat est facile.

La deuxième méthode de déconvolution ne permet pas une résolution analytique mais est bien plus souple que la première en termes de modélisation : la loi de la valeur doit être identifiable à une fonction paramétrée usuelle et rien n'est requis quant à la loi de l'erreur de mesure. Cette méthode teste un ensemble de lois implémentées, et en recherche le meilleur paramétrage par optimisation du score de vraisemblance des densités f_M calculées avec une quadrature de Hermite-Gauss. La loi ayant le plus grand score BIC parmi celles testées est conservée pour la valeur vraie. Cette méthode permettra par exemple de retrouver que la variable aléatoire de la valeur vraie suit une loi log-normale, dont les paramètres sont eux aussi identifiés. On remarquera que cette méthode d'identification peut inclure des lois particulièrement malléables, telles que les lois *Beta* et *Gamma*, qui permettent à elles seules de couvrir un très grand nombre de cas. D'une façon générale, plus les lois implémentées sont nombreuses et variées, plus il y a de chances d'obtenir une solution correcte. Les limites de cette méthode sont donc principalement liées aux moyens déployés pour l'implémentation des lois à tester.

La dernière méthode de déconvolution par approximation locale s'adresse au cas le plus général : aucune hypothèse de modélisation n'est nécessaire pour les lois de la valeur vraie et de l'erreur de mesure. La solution s'appuie aussi sur la quadrature de Hermite-Gauss ainsi que sur un lissage à noyau de la loi de l'échantillon de mesures. L'identification de la loi de la valeur vraie est discrétisée : point par point, un système d'équations est résolu pour l'approximer localement. Cette dernière méthode a l'avantage de pouvoir identifier une loi de la valeur vraie ayant n'importe quelle forme, comme une loi multimodale par exemple. Mais plus cette dernière est complexe, plus la taille de l'échantillon de mesures $m^{(i)}$ doit être grand pour vraiment en retrouver la forme. La méthode est aussi limitée par la précision du lissage à noyau.

Les méthodes de déconvolution présentées ici constituent un résultat majeur de ces travaux de recherche et prétendent (particulièrement pour les deux dernières) contribuer à la littérature de l'application de méthodes probabilistes en métrologie pour l'amélioration de la mesure. Le Tableau 2.2 synthétise les principales informations concernant les trois méthodes de déconvolution ainsi que l'inférence bayésienne. La colonne intitulée "complexité du code" donne un indice relatif de la difficulté que soulève l'implémentation des différentes solutions.

L'inférence bayésienne permet l'assimilation de deux lois. Dans ce cadre métrologique, les lois de l'erreur de mesure et de la valeur vraie estimée par déconvolution sont agrégées pour estimer

la position de la valeur vraie d'une mesure donnée. L'inférence bayésienne est appliquée à chacune des mesures d'un échantillon et aboutit pour chacune d'elles à une loi dite *a posteriori*. Cette loi est celle de la position de la valeur vraie $t^{(i)}$ et constitue un deuxième résultat majeur pour l'amélioration de la mesure. De cette loi peut être extraite la valeur vraie la plus probablement à l'origine de la mesure, nommée "valeur révisée" $r^{(i)}$.

Plusieurs outils statistiques ont été décrits pour évaluer avec complémentarité les résultats des méthodes probabilistes présentées dans les deux premières sections. Ces outils explicitent le gain apporté par la déconvolution et l'inférence bayésienne pour la maîtrise des risques en production. Le calcul du risque total est probablement une autre grande plus-value de l'information obtenue par la connaissance de la densité de probabilité de l'erreur de mesure.

TABLEAU 2.2: Synthèse des méthodes probabilistes utilisées.

Objectif	Méthodes	Hypothèse et modélisation	Données d'entrées	Données de sortie
Identifier la loi de la valeur vraie	Fonctions caractéristiques	La valeur vraie suit une loi normale L'erreur de mesure suit une loi normale L'échantillon de mesure est suffisamment grand	Échantillon de mesures Distribution de l'erreur de mesure	Connaissance analytique de la loi de la valeur vraie
Identifier la loi de la valeur vraie	Optimisation et quadrature	La valeur vraie suit une loi identifiable L'échantillon de mesure est suffisamment grand Lissage à noyau sur l'échantillon de mesure	Échantillon de mesures Distribution de l'erreur de mesure	Connaissance analytique de la loi de la valeur vraie
Identifier la loi de la valeur vraie	Reconstruction locale	L'échantillon de mesure est suffisamment grand Lissage à noyau sur l'échantillon de mesure	Échantillon de mesures Distribution de l'erreur de mesure Valeur de pas	Connaissance discrète de la loi de la valeur vraie
Identifier la valeur vraie la plus probable d'une mesure donnée	Inférence bayésienne	-	Une mesure Distribution de l'erreur de mesure Distribution de la valeur vraie	Connaissance discrète de la loi <i>a posteriori</i>

Chapitre 3

Applications et évaluation des méthodes probabilistes

Points clefs

F

- Simulation de données
 - Évaluation des différentes méthodes probabilistes
 - Discussions des paramètres influents
-

Sommaire

Paramétrage de la simulation	86
3.1 Application à des variables aléatoires gaussiennes	87
3.1.1 Paramétrage et données d'entrée	87
3.1.2 Application et résultats de la déconvolution	87
3.1.3 Application et résultat de l'inférence bayésienne	91
3.2 Application des méthodes pour une loi de la valeur vraie identifiable .	95
3.2.1 Paramétrage et données d'entrée	95
3.2.2 Application de la déconvolution	95
3.2.3 Application et résultat de l'inférence bayésienne	100
3.3 Application des méthodes pour des données quelconques	103
3.3.1 Paramétrage et données d'entrées	103
3.3.2 Application et résultat de la déconvolution	104
3.3.3 Application et résultat de l'inférence bayésienne	107
Conclusion	112

Paramétrage de la simulation

Ce chapitre est centré sur l'évaluation de l'apport des méthodes probabilistes présentées précédemment (Sections 2.1 et 2.2). L'objectif est d'améliorer les mesures obtenues en métrologie : il n'est pas raisonnable, en présence d'incertitudes, de prétendre à une parfaite identification des valeurs vraies. Un exemple est spécifiquement introduit pour chacune des trois méthodes de déconvolution. L'évaluation des méthodes est réalisée sur des exemples construits avec des données simulées, afin de connaître les solutions vers lesquelles converger.

Une donnée majeure pour l'application de ces méthodes est la densité de probabilité de l'erreur de mesure, qui a été supposée connue pour l'ensemble de ces travaux. On notera que cette hypothèse est impossible à vérifier en réalité mais qu'elle sera tout de même utilisée dans ces exemples pour tester pleinement le potentiel des méthodes développées. Le cas d'une loi mal identifiée sera discuté dans un second temps. Les lois de l'erreur de mesure utilisées sont fixées de sorte à simuler la réalité. Les moyennes de ces lois sont choisies nulles : un biais identifié ne pose pas de problème en réalité, il est simplement soustrait à la mesure. De plus, la correction systématique du biais représente une amélioration de peu d'intérêts. La seconde donnée nécessaire pour l'application des méthodes probabilistes est la loi de la valeur vraie. Les lois utilisées dans les exemples sont fixées de sorte à simuler la réalité. La moyenne de la valeur vraie n'affecte pas l'amélioration des mesures et n'est donc pas particulièrement discutée. L'impact de la valeur de l'écart-type est lui prédominant [139] : des écarts-types égaux remettent en cause la pertinence de la mesure, et mesurer avec un écart-type de l'erreur de mesure supérieur à celui de la valeur vraie est sans intérêt (cela revient à mesurer du bruit). A l'inverse, si l'écart-type de l'erreur de mesure est négligeable devant la valeur vraie alors il n'y a plus de nécessité d'améliorer la mesure. En compromis, l'écart-type de la valeur vraie a été choisi deux fois plus grand que celui de l'erreur de mesure pour les trois exemples.

Les échantillons de mesures utilisées pour les exemples sont construits en additionnant un échantillon de valeurs vraies à un échantillon d'erreurs de mesure. Comme dans le reste de ces travaux, les variables aléatoires de la valeur vraie et de l'erreur de mesure sont supposées indépendantes. Le phénomène d'auto-corrélation des échantillons de la valeur vraie ou de l'erreur de mesure sera discutée dans un second temps. La méthode de tirage pour l'échantillonnage des lois est celle de Monte-Carlo. La taille de l'échantillon de mesures est fixée à 1000 pour les trois exemples. Cette quantité a été choisie pour réaliser un compromis entre les exigences des méthodes statistiques et les contraintes matérielles ; pour assurer une bonne représentation des lois des variables aléatoires en jeu et que la collecte des mesures soit réalisable en pratique. Une étude est réalisée pour montrer la sensibilité des méthodes à la taille de l'échantillon. En ce qui concerne la conformité, le nombre de pièces non-conformes dans l'industrie se comptent généralement en pièces par million (ppm)[36]. Avec des échantillons de 1000 mesures, il serait difficile d'évaluer la qualité de la déclaration de non-conformité avec un tel taux. Les bornes de l'intervalle de tolérance sont ainsi fixées en fonction de chaque exemple de sorte à ce que le taux de non-conformité soit de l'ordre de grandeur de 5%.

3.1 Application à des variables aléatoires gaussiennes

3.1.1 Paramétrage et données d'entrée

Cet exemple répond aux exigences de modélisation pour l'application de la première méthode de déconvolution. Les variables aléatoires de la valeur vraie et de l'erreur de mesure suivent des lois gaussiennes. Les paramètres de ces lois sont synthétisés dans le Tableau 3.1. Les bornes de l'intervalle de tolérance sont fixées égales à 100.2 et 101.8.

TABLEAU 3.1: Paramétrage des lois d'entrées.

Variation aléatoire	Loi	Moyenne	Écart-type
Valeur vraie	Normale	101.0	0.4
Erreur de mesure	Normale	0.0	0.2

3.1.2 Application et résultats de la déconvolution

La méthode de déconvolution décrite dans la Section 2.1.1 est appliquée à un échantillon de 1000 mesures et avec une parfaite connaissance de la densité de probabilité de l'erreur de mesure dans le but d'estimer la loi de la valeur vraie. L'histogramme de l'échantillon de mesures est tracé en arrière-plan d'un gris léger. La loi obtenue est représentée d'un trait épais et vert sur la Figure 3.1 avec celle de la valeur vraie d'un trait rouge et pointillé. Le tracé noir et fin est celui d'une loi gaussienne inférée à l'échantillon afin de mieux visualiser l'apport de la déconvolution. L'axe des ordonnées indique les densités de probabilité, et l'axe des abscisses les valeurs vraies. Les lignes verticales orangées en semi-pointillé représentent les bornes de l'intervalle de tolérance.

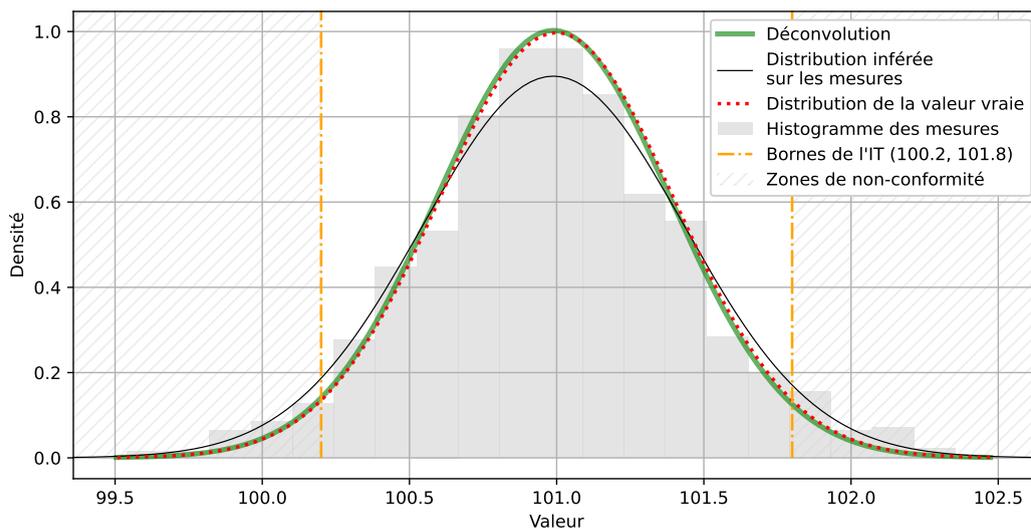


FIGURE 3.1: Déconvolution de la loi des mesures obtenues en métrologie.

La courbe obtenue par déconvolution est très proche de celle de la valeur vraie, ce qui est très satisfaisant. Le Tableau 3.2 montre que les paramètres obtenus via la méthode de déconvolution sont très proches de ceux qu'il fallait identifier. L'écart est de l'ordre d'un pourcent de l'écart-type

de la valeur vraie. L'écart observé avec la loi de la valeur vraie peut statistiquement être associé à la taille de l'échantillon de mesures.

Étant efficacement épurée de l'erreur de mesure, la loi déconvoluée se substitue à la lecture ordinaire du lot de mesures (via un histogramme ou une inférence) pour caractériser la production. Le Tableau 3.3 montre que la probabilité de produire hors tolérance calculée après déconvolution est plus faible (conformément à ce que l'on peut voir en Figure 3.1 au niveau des bornes de l'intervalle de tolérance) et bien plus proche de la valeur de référence que celle déduite de l'échantillon de mesures. Il est constaté que le score obtenu est légèrement en deçà de celui attendu, ce qui rappelle que la déconvolution est une méthode probabiliste dont les résultats ne sont pas nécessairement encadrés par ceux des mesures et de la valeur vraie.

TABLEAU 3.2: Précision de l'identification des paramètres.

Paramètre	Valeur	Erreur ($\%/\sigma_T$)
Moyenne	100.989	2.779
Écart-type	0.398	0.525

TABLEAU 3.3: Comparaison des probabilités de produire hors tolérance (p_{HT}).

Distribution	p_{HT}
Valeur vraie	0.0455
Mesure	0.0725
Déconvolution	0.0444

Une fois qu'une estimation de la valeur vraie est connue, il est possible de tracer les courbes de probabilité de se tromper sur la classification des mesures. La Figure 3.2 présente celles obtenues dans ce cas gaussien. Les courbes obtenues vérifient bien qu'il est très probable de se tromper en acceptant une mesure loin des bornes de l'intervalle de tolérance et qu'il est très probable de se tromper en refusant celles largement incluses. Les transitions, de zéro à un, correspondent bien au tracé de la fonction de répartition d'une loi normale (celle de l'erreur de mesure). Les intersections des deux courbes sont attendues au niveau des bornes de l'intervalle de tolérance mais sont déplacées à l'extérieur par l'effet de la zone non-homogène.

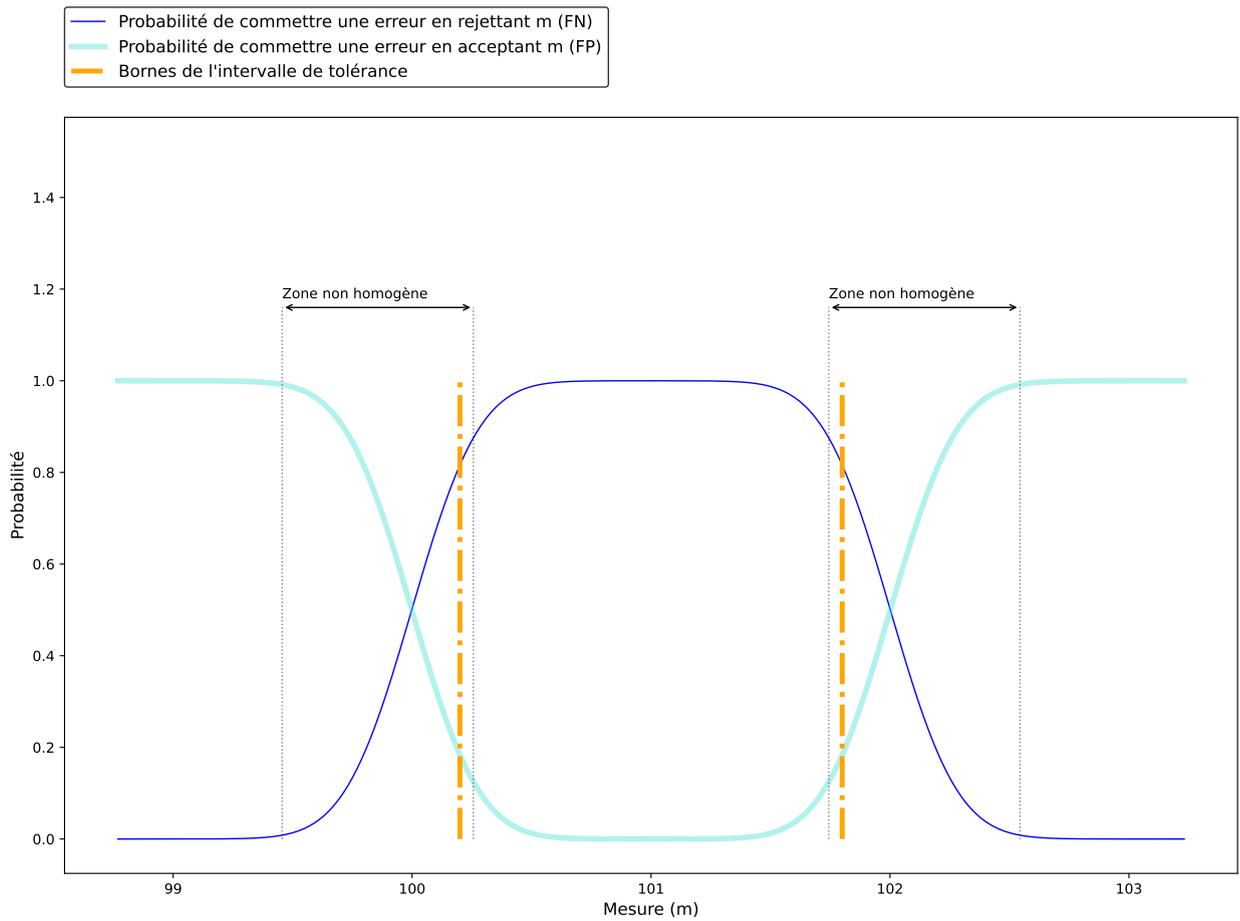


FIGURE 3.2: Probabilités qu'une mesure soit mal classifiée.

Les courbes de risque sont tracées en Figure 3.3 avec un coût client 10 fois plus élevé à celui fournisseur. Lorsque le coût client est plus grand que celui fournisseur, les bornes de la sélection qui minimisent le risque total ont tendance à se resserrer (l'exigence augmente par crainte d'accepter une pièce à tort). Ici, les intersections des courbes de risque se retrouvent logiquement rapprochées et ont même été déplacées à l'intérieur de l'intervalle de tolérance.

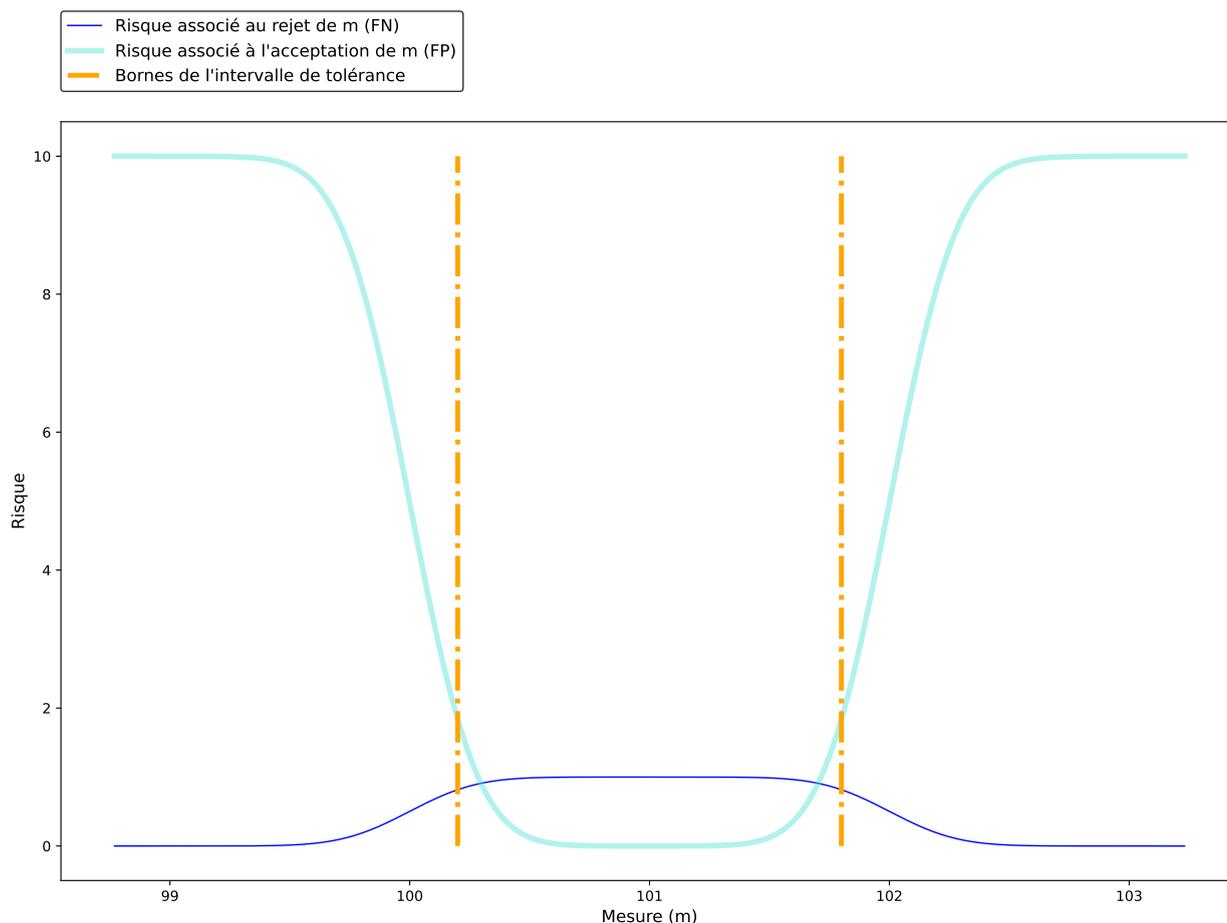


FIGURE 3.3: Risque qu'une mesure soit mal classifiée.

Les courbes de risque permettent d'avoir un regard critique sur la conformité de la production. La sélection classique confronte les mesures à l'intervalle de tolérance. Les bornes de la sélection optimale sont les intersections des courbes de risques, connues grâce à la déconvolution. Les bornes de la sélection qui minimise parfaitement les risques sont les intersections des courbes de risque que l'on construirait en connaissance de la loi de la valeur vraie. Le Tableau 3.4 décrit les valeurs des bornes et le risque total associés à chaque intervalle de sélection.

TABLEAU 3.4: Risque total des différentes sélections.

Base de sélection des mesures	Borne inférieure	Borne supérieure	Risque total
Intervalle de tolérance	100.2	101.8	0.896
Sélection optimale (Déconvolution)	100.301	101.703	0.806
Sélection parfaite (Loi de la valeur vraie)	100.301	101.699	0.806

Le risque total est largement amélioré suite à l'application de la déconvolution. La sélection optimale post-déconvolution est exacte à 10^{-3} près. Les bornes inférieures sont identiques (à 10^{-3} près) et les bornes supérieures ne divergent que pour quatre millièmes. L'apport de la déconvolution pour la gestion des risques de production est très net dans le cas gaussien.

3.1.3 Application et résultat de l'inférence bayésienne

La méthode d'inférence bayésienne décrite dans la Section 2.2.1 est appliquée pour obtenir la loi *a posteriori* et la valeur révisée de chacune des mesures de l'échantillon. La Figure 3.4 présente les nuages de points des résidus des échantillons des mesures et des valeurs révisées. La ligne bleue verticale en pointillé est celle de la moyenne de la valeur vraie. La ligne horizontale rouge est celle où la prédiction est égale à la valeur vraie (le meilleur des cas).

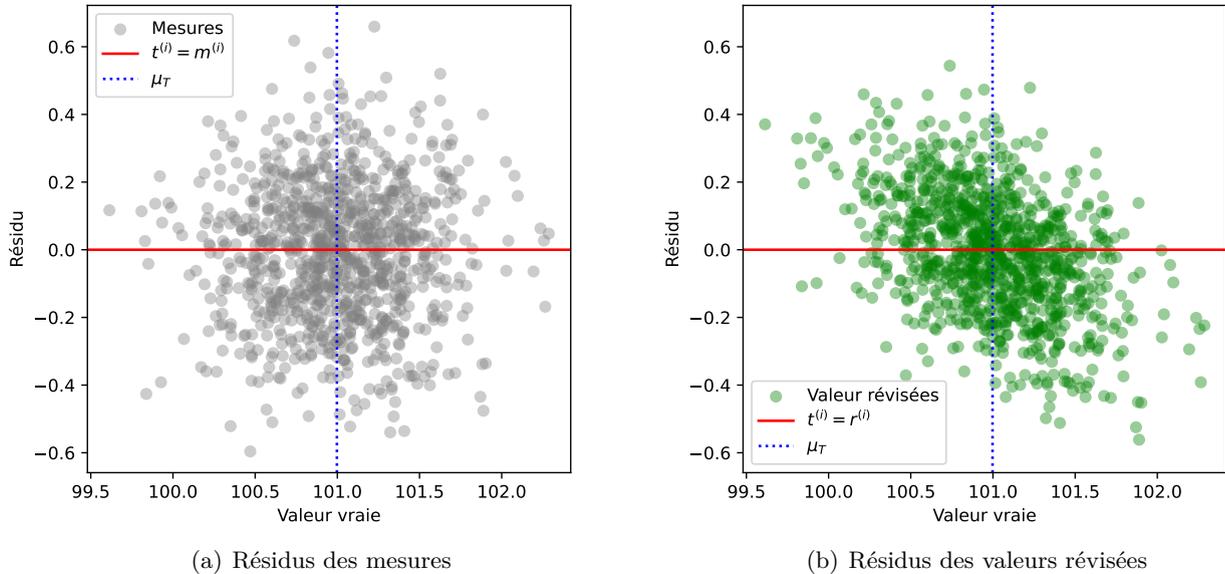


FIGURE 3.4: Comparaison des résidus des mesures et des valeurs révisées.

Le nuage de points des résidus des mesures (Figure 3.4(a)) révèle les origines gaussiennes des tirages des valeurs vraies et des erreurs de mesure. La densité du nuage est concentrique et présente une symétrie ponctuelle autour de la moyenne des valeurs vraie et de celle de l'erreur de mesure - qui, étant nulle, est confondue avec la ligne d'équation $t^{(i)} = m^{(i)}$.

Suite à l'application de l'inférence bayésienne, le nuage de points des valeurs révisées (Figure 3.4(b)) montre une corrélation négative qui correspond à la transformation linéaire explicitée au Chapitre 2.23. Les mesures les plus améliorées se situent dans les quarts bas-gauche et haut-droit du graphique, ainsi qu'autour de la droite verticale d'équation $T = \mu_T$. Dans le reste du graphique et particulièrement pour les valeurs extrêmes de T , l'inférence bayésienne a tendance à aggraver les résidus des valeurs vraies. La population des mesures améliorées par la révision est présentée dans l'Annexe B.2. La méthode présente une balance globalement en faveur des valeurs révisées.

Le coefficient de détermination (R^2) et l'erreur moyenne quadratique sont utilisés pour quantifier la qualité des prédictions. Le Tableau 3.5 présente les scores des prédictions de la valeur vraie sur la base des mesures et sur la base des valeurs révisées. Les deux scores sont en faveur des valeurs révisées (l'écart moyen est plus faible et le coefficient est plus proche de un). Le Tableau 3.5 expose aussi le pourcentage des mesures améliorées qui est supérieur à 50%. Il est donc possible de

conclure que l'application des méthodes probabilistes a permis une amélioration globale concernant la majorité des mesures.

TABLEAU 3.5: Indices de qualité des mesures et des valeurs révisées.

Score	Mesures	Valeurs révisées
Erreur quadratique moyenne	0.041	0.032
R^2	0.749	0.802
Pourcentage des mesures améliorées	-	58.6

Dans le cas de variables aléatoires gaussiennes, il est possible de définir l'intervalle de tolérance équivalent. Les bornes des différents intervalles ainsi que le coefficient et la constante du polynôme utilisé pour la transformation sont donnés dans le Tableau 3.6. L'intervalle équivalent obtenu est centré sur l'intervalle de tolérance et élargit de 0.2 à chacune des bornes. La sélection des mesures sur la base des valeurs révisées revient donc à accepter plus de mesures, résultat qui avait déjà été établi lors de l'étude de conformité via des méthodes statistiques [94].

TABLEAU 3.6: Bornes de l'intervalle de tolérance équivalent.

	Borne inférieure	Borne supérieure	a_0	a_1
Intervalle de tolérance	100.2	101.8	-	-
IT équivalent	100	102	0.8	20.2

Les différentes déclarations de conformité peuvent être visualisées dans la Figure 3.6. La première des matrices de confusion en Figure 3.6(a) est celle de référence (prédiction des valeurs vraies). Cette véritable classification peut être visualisée dans un repère cartésien dans la Figure 3.5. La grande majorité des pièces est conforme (95.8%). La deuxième matrice montre la déclaration de conformité réalisée en confrontant les mesures à l'intervalle de tolérance (Figure 3.6(b)) : 936 pièces sont correctement classées, 55 prédites conformes à tort et 9 prédites non-conformes à tort. La troisième matrice représente la confrontation des valeurs révisées à l'intervalle de tolérance (Figure 3.6(c)). Un plus grand nombre de pièces est correctement classé (967) grâce aux méthodes proposées. Cette amélioration est la somme d'un nombre de faux négatifs largement diminué et d'un nombre de faux positifs augmenté. La tendance d'accepter plus de mesures (déjà déduite avec l'intervalle équivalent) améliore la déclaration de conformité mais s'accompagne de plus de pièces acceptées à tort. Ces résultats sont favorables à l'application des méthodes probabilistes, et donc à la substitution des mesures par les valeurs révisées si le coût du risque client est égal à celui du risque fournisseur.

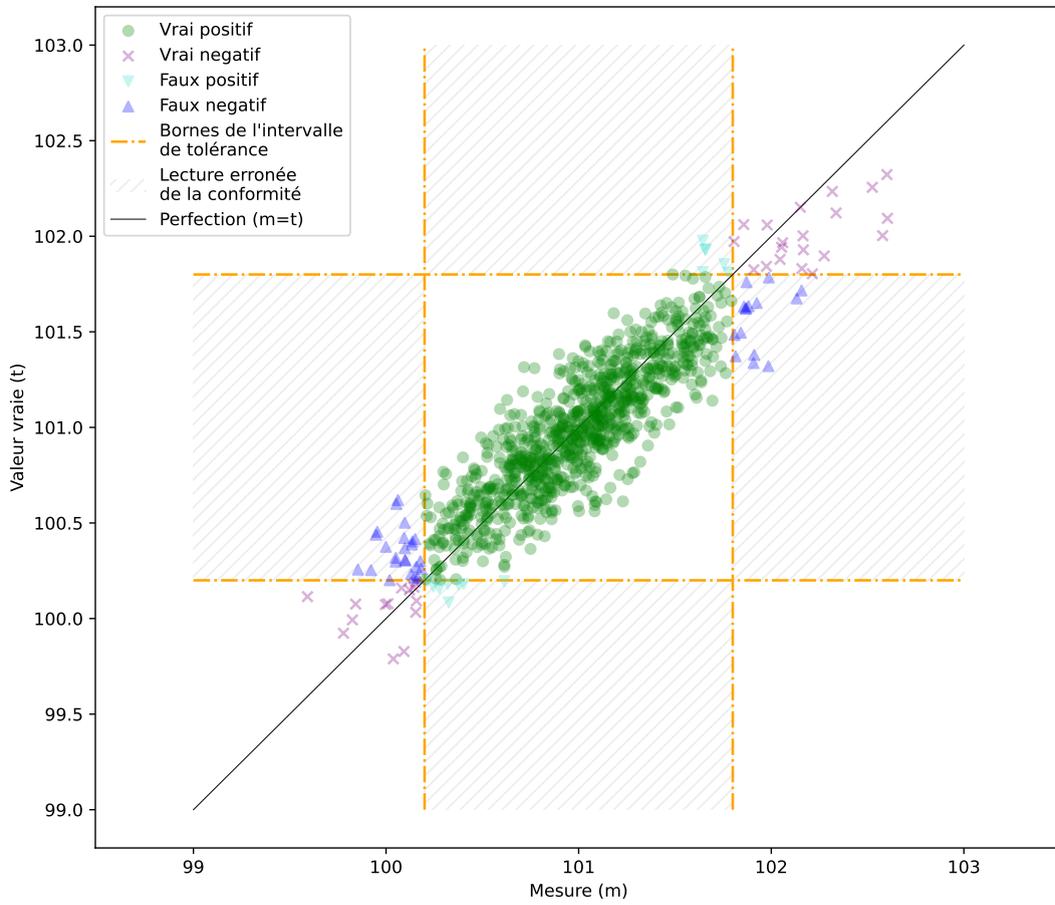


FIGURE 3.5: Classification de l'échantillon.

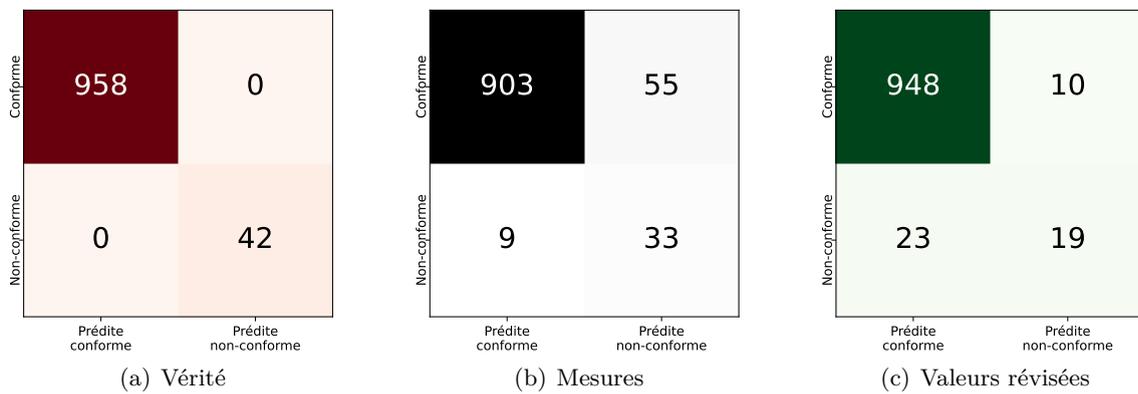


FIGURE 3.6: Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas gaussien).

Une classification fondée sur les valeurs révisées implique que le seuil du risque maximum accep-

table est égal à 50%. La courbe ROC présentée en Figure 3.7 permet de visualiser la performance des lois *a posteriori* pour toutes les valeurs de seuils entre 0 et 100%.

Le tracé de la courbe ROC est fondé sur la connaissance des lois de l'erreur de mesure et de la valeur vraie. La qualité du tracé dépend de la précision de ces deux informations. La connaissance du comportement de l'erreur de mesure étant admise, la qualité dépend principalement de la bonne estimation de la loi de la valeur vraie obtenue par déconvolution. Dans le but d'évaluer le potentiel de ces méthodes et dans le cadre de la simulation, la Figure 3.7 est tracée avec une connaissance parfaite de la loi de la valeur vraie.

Les trois matrices de confusion présentées précédemment (Figure 3.6) sont aussi incluses dans le graphique. La matrice des valeurs vraies a logiquement les coordonnées (0, 1). Par rapport à celle des mesures, les valeurs révisées présentent plus de vrais positifs et plus de faux négatifs, conformément à ce que les matrices montraient.

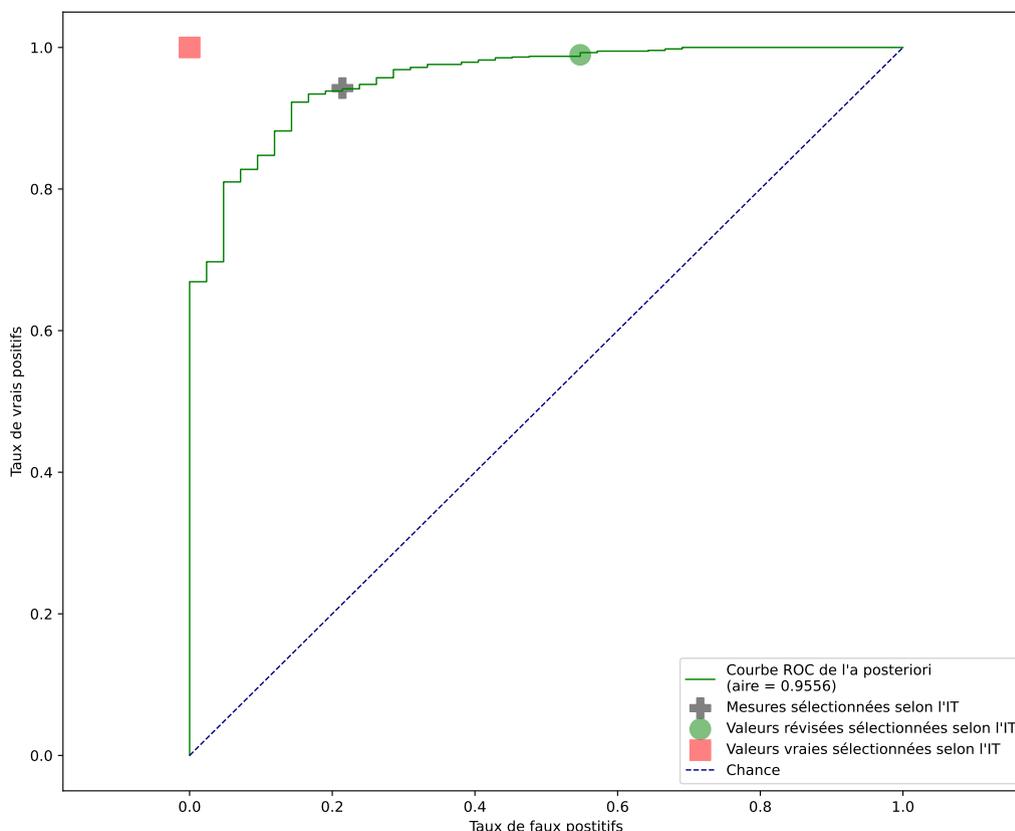


FIGURE 3.7: Tracé de la courbe ROC (cas gaussien).

Les lois *a posteriori* donnent des meilleurs résultats que les mesures lorsque le seuil de risque maximum est entre 18% et 33%. Le point de la courbe ROC le plus proche de la sélection parfaite (de coordonnée (0, 1)) est obtenu pour un seuil de 29%. Le taux de faux positifs est nul tant que

le risque maximum acceptable est inférieur ou égal à 12%. Le taux de vrais positifs est égal à un dès que le seuil est supérieur ou égal à 64%. En pratique, le caractère optimal de la sélection reste relatif à des critères arbitraires (par exemple, $TP_R > 0.99$ ou $FP_R < 0.0001$).

3.2 Application des méthodes pour une loi de la valeur vraie identifiable

3.2.1 Paramétrage et données d'entrée

Les méthodes probabilistes ont été testées dans la section précédente avec des variables aléatoires gaussiennes. Dans cet exemple, la loi de la valeur vraie est désormais log-normale, et présente une asymétrie (*skewness*) positive. La densité est donnée par $\ln(X - \gamma) = \mathcal{N}(\lambda, \zeta)$. La définition des paramètres λ et ζ est :

$$\begin{cases} \zeta = \sqrt{\ln\left(1 + \frac{\sigma_T^2}{(\mu_T - \gamma)^2}\right)} \\ \lambda = \ln(\mu_T - \gamma) - \frac{\zeta^2}{2} \end{cases} \quad (3.1)$$

avec μ_T et σ_T la moyenne et l'écart-type de la valeur vraie et γ est le paramètre de position. La loi de l'erreur de mesure est maintenue gaussienne, et son écart-type a été modifié à 0.3 pour qu'il soit toujours inférieur de moitié à celui de la valeur vraie. Le paramétrage des deux lois est synthétisé dans le Tableau 3.7. Enfin les bornes de l'intervalle de tolérance sont changées et fixées égales à [99.5, 102] dans le but de conserver une probabilité de défaillance comparable au précédent exemple.

TABLEAU 3.7: Paramétrage des lois d'entrées (cas log-normal).

Variable aléatoire	Type	Paramètres			Moyenne	Écart-type
Valeur vraie	Log-normale	$\mu_{log}=0.01$	$\sigma_{log}=0.5$	$\gamma=99.5$	100.645	0.610
Erreur de mesure	Normale	-	-	-	0	0.3

μ_{log}, σ_{log} : la moyenne et l'écart-type du logarithme de T
 γ , le paramètre de position.

3.2.2 Application de la déconvolution

La méthode de déconvolution et la liste des lois candidates utilisées dans cet exemple sont celle décrites dans la Section 2.1.2. Le nombre de points d'intégration pour calculer la quadrature est fixé à 25.

Le meilleur score BIC de chacune des différentes familles de lois testées est présenté en Figure 3.8. Les valeurs vraies étant définies sur un domaine semi-infini, les lois bornées (uniforme, triangulaire et arcsinus) obtiennent logiquement les plus mauvais scores. La loi normale est définie sur \mathbb{R} mais n'est pas asymétrique, et dépasse à peine la triangulaire. Parmi les deux lois restantes, le meilleur score est celui de la famille des lois log-normale et vaut 1882.51, ce qui montre que la loi originale a bien été identifiée.

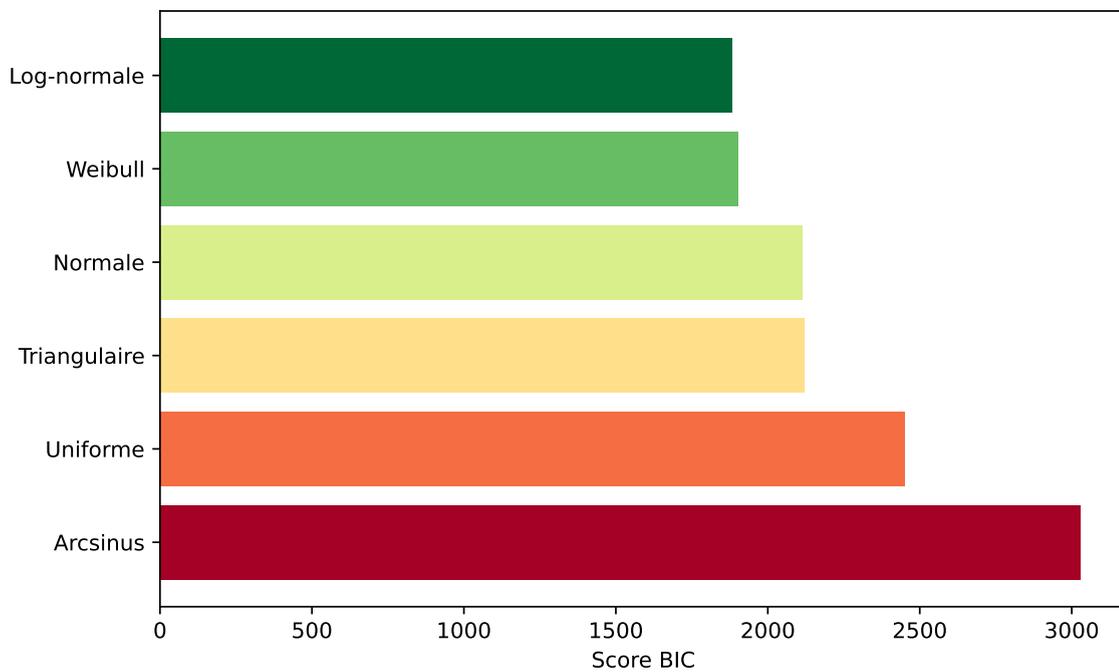


FIGURE 3.8: Score BIC du représentant de chaque famille candidate (cas log-normal).

L'histogramme de l'échantillon de mesures est tracé en arrière-plan de la Figure 3.9 d'un gris léger. La loi déconvoluée est représentée d'un trait épais et vert avec celle de la valeur vraie d'un trait rouge et pointillé. Le tracé noir et fin est celui d'une loi log-normale inférée à l'échantillon afin de mieux visualiser l'apport de la déconvolution. L'axe des ordonnées indique les densités de probabilité, et l'axe des abscisses les valeurs vraies. Les lignes verticales orangées en semi-pointillé représentent les bornes de l'intervalle de tolérance.

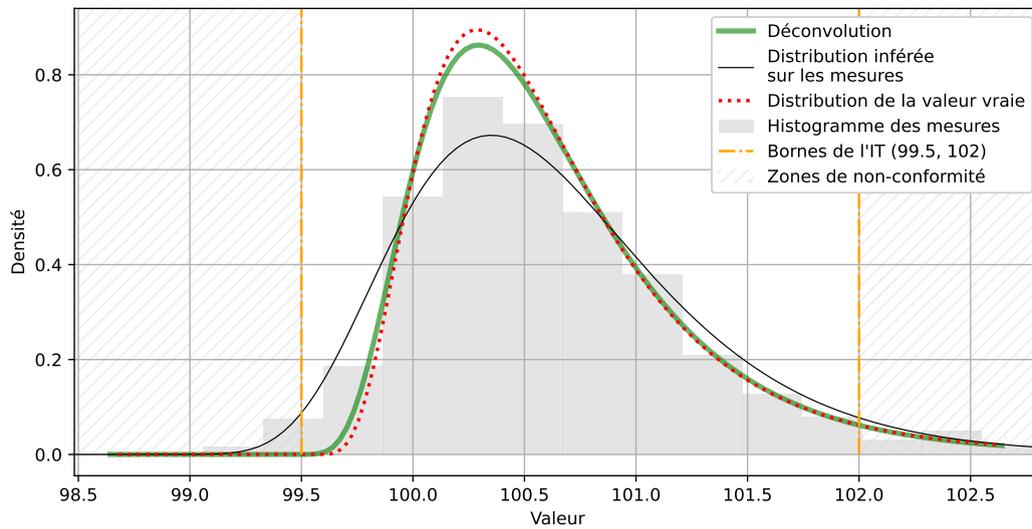


FIGURE 3.9: Déconvolution de la loi des mesures obtenues en métrologie (cas log-normal).

La courbe obtenue par déconvolution est très proche de celle de la valeur vraie, ce qui est très satisfaisant. Les paramètres de cette loi optimale sont décrits dans le Tableau 3.8. Là encore, l'écart de ces deux premiers moments statistiques est de l'ordre d'un pourcent de l'écart-type de la valeur vraie. L'écart observé entre les courbes peut statistiquement être associé à la taille de l'échantillon de mesures.

La loi déconvoluée est efficacement corrigée de la dispersion induite par l'erreur de mesure. Elle peut donc se substituer à la lecture ordinaire du lot de mesures (via un histogramme ou une inférence) pour caractériser la production. Le Tableau 3.9 montre que la probabilité de produire hors tolérance calculée après déconvolution est plus faible (conformément à ce que l'on peut voir en Figure 3.9 au niveau des bornes de l'intervalle de tolérance) et bien plus réaliste que celle déduite de l'échantillon de mesures.

TABLEAU 3.8: Évaluation de la précision de l'identification (cas log-normal).

Paramètres	Valeur	Erreur ($\%/\sigma_T$)
Moyenne	100.632	2.033
Écart-type	0.599	1.857

TABLEAU 3.9: Comparaison des probabilités de produire hors tolérance (p_{HT}) (cas log-normal).

Distribution	p_{HT}
Valeur vraie	0.0349
Mesure	0.0650
Déconvolution	0.0320

Lorsqu'une estimation de la loi de la valeur vraie est connue, il est possible de tracer les courbes de probabilité de prendre une mauvaise décision. La Figure 3.10 présente celles obtenues dans ce cas log-normal. Les courbes obtenues vérifient bien qu'il est très probable de se tromper en acceptant une mesure loin des bornes de l'intervalle de tolérance et qu'il est très probable de se tromper en refusant celles largement incluses. Les transitions, de zéro à un, correspondent bien au tracé de la fonction de répartition d'une loi normale (celle de l'erreur de mesure).

Dans cet exemple, les intersections des deux courbes ne s'opèrent pas au niveau des bornes de l'intervalle de tolérance : cela s'explique du fait que la moyenne et la médiane d'une loi log-normale sont distinctes. A la borne droite de l'intervalle de tolérance, il n'y a pas d'effet de zone non-homogène. A la borne gauche, l'intersection est déportée à l'infini - signifiant ainsi qu'il n'y a pas de rejet à faire de ce côté.

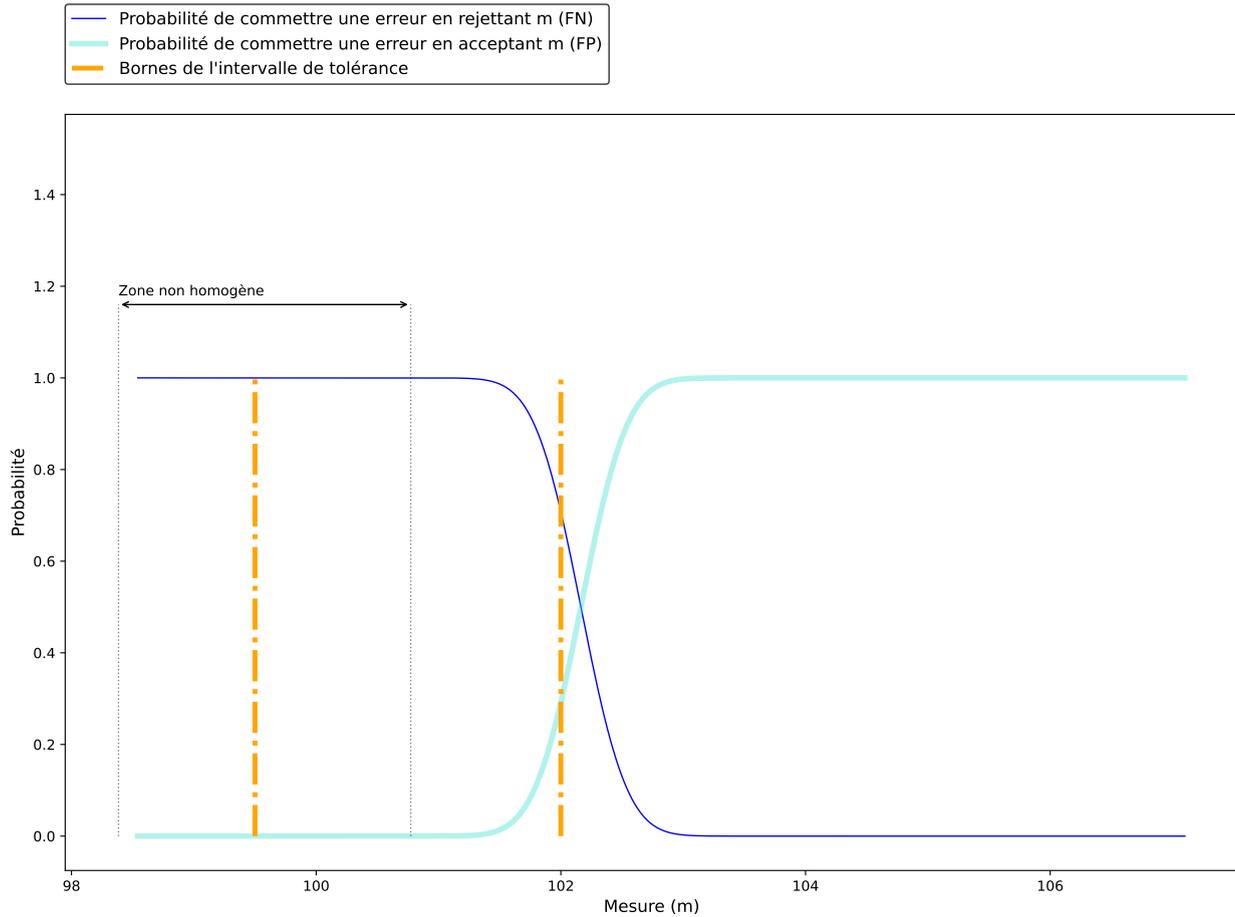


FIGURE 3.10: Probabilités qu'une mesure soit mal classifiée (cas log-normal).

Les courbes de risque sont tracées en Figure 3.11 avec un coût client dix fois plus élevé que le coût fournisseur. Comme au précédent exemple, les intersections des courbes de risque se retrouvent logiquement rapprochées et sont aussi déplacées à l'intérieur de l'intervalle de tolérance.

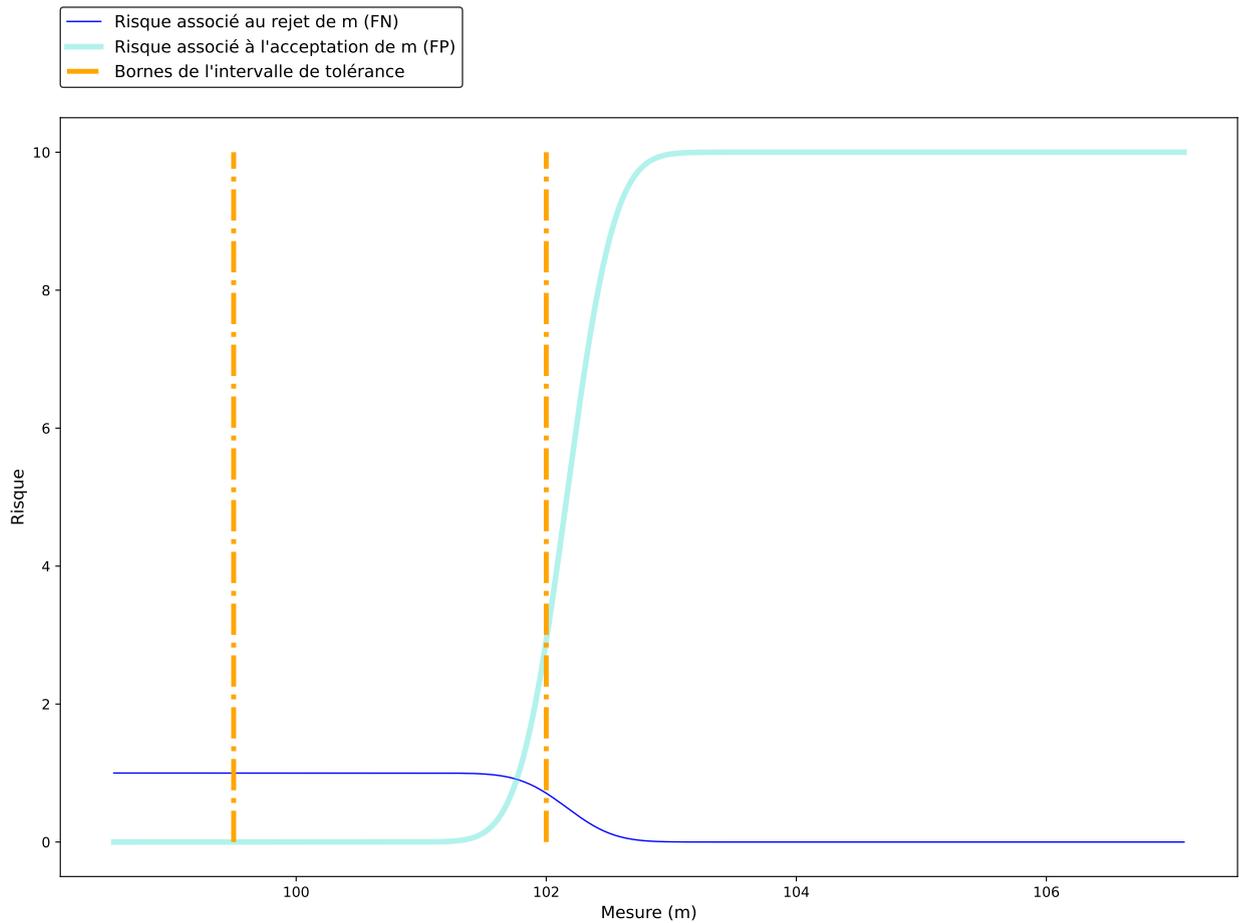


FIGURE 3.11: Risque qu'une mesure soit mal classifiée (cas log-normal).

Les courbes de risque permettent d'avoir un regard critique sur la conformité de la production. La sélection classique confronte les mesures à l'intervalle de tolérance. Les bornes de la sélection optimale sont les intersections des courbes de risques, connues grâce à la déconvolution. Les bornes de la sélection qui minimise parfaitement les risques sont les intersections des courbes de risque que l'on construirait en connaissance de la loi de la valeur vraie. Le Tableau 3.10 décrit les valeurs des bornes et le risque total associés à chaque intervalle de sélection.

TABLEAU 3.10: Risque total des différentes sélections (cas log-normal).

Base de sélection des mesures	Borne inférieure	Borne supérieure	Risque total
Intervalle de tolérance	99.5	102	1.725
Sélection optimale (Déconvolution)	98.486	101.769	0.540
Sélection parfaite (Loi de la valeur vraie)	98.542	101.763	0.539

L'estimation du risque total est très nettement améliorée suite à l'application de la déconvolution. La sélection optimale post-déconvolution est très proche de celle parfaite. Les bornes inférieures acceptent toutes les mesures et les bornes supérieures ne diffèrent que de 6 millièmes. L'apport de la déconvolution pour la gestion des risques de production est très net aussi dans le cas log-normal, où il s'agit d'identifier la loi de la valeur vraie.

3.2.3 Application et résultat de l'inférence bayésienne

La méthode d'inférence bayésienne décrite dans la Section 2.2.1 est appliquée pour obtenir la loi *a posteriori* et la valeur révisée de chacune des mesures de l'échantillon. La Figure 3.12 présente les nuages de points des résidus des échantillons des mesures et des valeurs révisées. La ligne bleue verticale en pointillé est celle de la moyenne de la valeur vraie. La ligne horizontale rouge est celle où la prédiction est égale à la valeur vraie (le meilleur des cas).

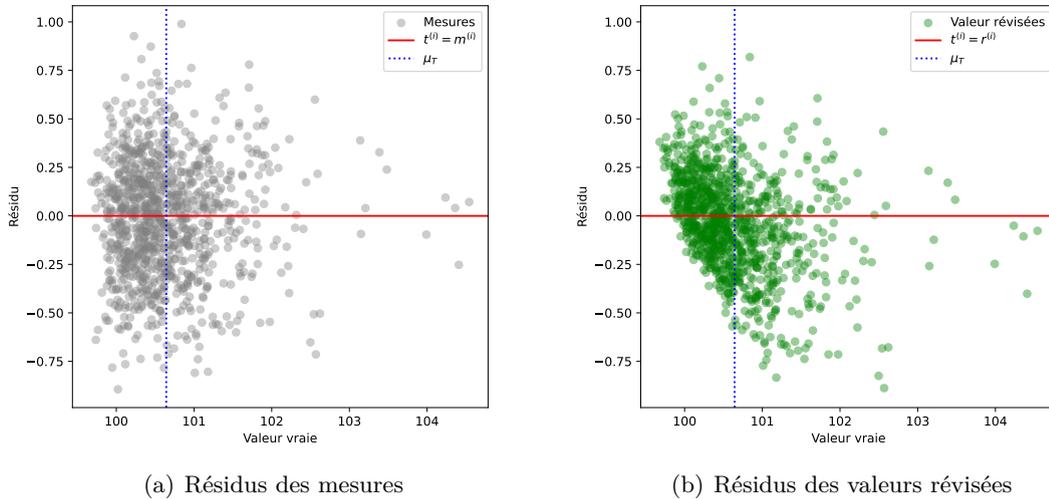


FIGURE 3.12: Comparaison des résidus des mesures et des valeurs révisées (cas log-normal).

Le nuage de points des résidus des mesures (Figure 3.12(a)) permet d'observer en abscisse l'asymétrie de la loi log-normale dans le tirage des valeurs vraies (par rapport à la dispersion normale des erreurs de mesure). La zone la plus dense (le mode) apparaît nettement, à gauche de la moyenne des valeurs vraies.

Suite à l'application de l'inférence bayésienne, le nuage de points des valeurs révisées (Figure 3.12(b)) montre la même corrélation négative que précédemment (Section 3.1.3). Les mesures les plus améliorées se situent dans le quart haut-droit et très nettement dans le quart bas-gauche du graphique, ainsi qu'autour de la droite verticale d'équation $T = \mu_T$. Dans le reste du graphique et particulièrement pour les valeurs extrêmes de T , l'inférence bayésienne a tendance à aggraver les résidus des valeurs vraies. La méthode présente une balance globalement en faveur des valeurs révisées.

Le coefficient de détermination (R^2) et l'erreur moyenne quadratique sont utilisés pour quantifier la qualité des prédictions. Le Tableau 3.11 présente les scores des prédictions de la valeur vraie sur la base des mesures et sur la base des valeurs révisées. Les deux scores sont en faveur des valeurs révisées (l'écart moyen est plus faible et le coefficient est plus proche de un). Le Tableau 3.11 montre aussi le pourcentage des mesures améliorées qui atteint 60%. Il est donc possible de conclure que l'application des méthodes probabilistes a permis une amélioration globale concernant la majorité des mesures.

TABLEAU 3.11: Indices de qualité des mesures et des valeurs révisées (cas non gaussien).

Score	Mesures	Valeurs révisées
Erreur quadratique moyenne	0.092	0.066
R^2	0.763	0.828
% mesures améliorées	-	60.0

Les différentes déclarations de conformité peuvent être visualisées dans la Figure 3.13. La première des matrices de confusion en Figure 3.13(a) est celle de référence (prédiction des valeurs vraies). Elle montre que la grande majorité des pièces est conforme (97.0%). Cette véritable classification peut être visualisée dans un repère cartésien dans la Figure 3.14. La deuxième matrice montre la déclaration de conformité réalisée en confrontant les mesures à l'intervalle de tolérance (Figure 3.13(b)) : 959 pièces sont correctement classées, 7 prédites conformes à tort et 34 prédites non-conformes à tort. La troisième matrice représente la confrontation des valeurs révisées à l'intervalle de tolérance (Figure 3.13(c)). Un plus grand nombre de pièces est correctement classé (980) grâce aux méthodes proposées. Cette amélioration est la somme d'un nombre de faux négatifs largement diminué et d'un nombre de faux positifs légèrement augmenté. Ces résultats sont favorables à l'application des méthodes probabilistes, et donc à la substitution des mesures par les valeurs révisées (lorsque le coût du risque client est égal à celui du risque fournisseur).

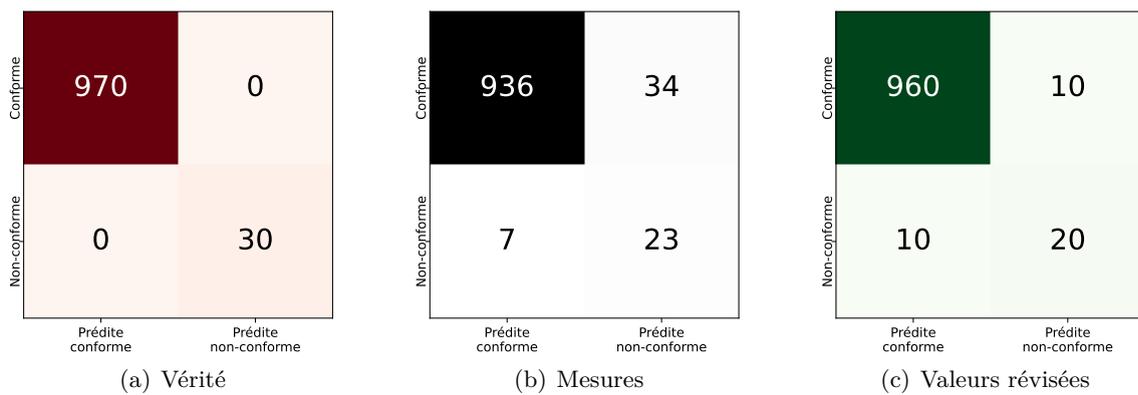


FIGURE 3.13: Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas log-normal).

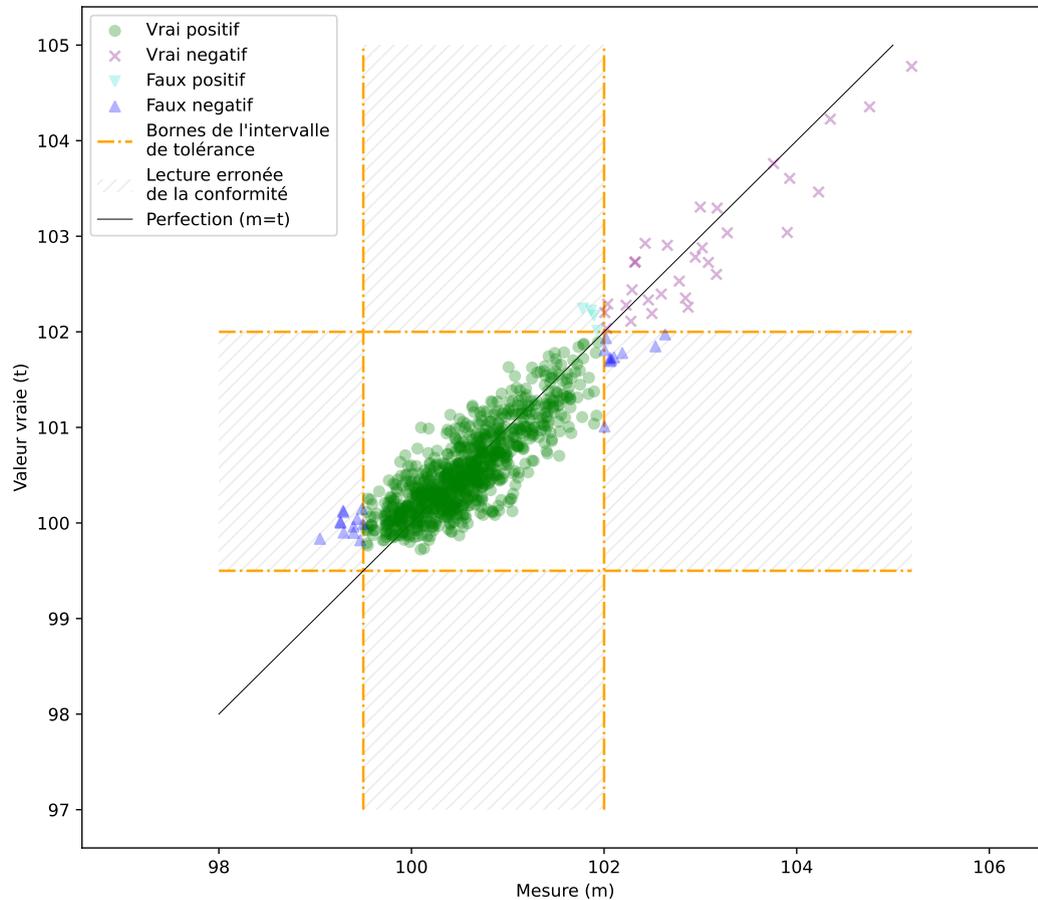


FIGURE 3.14: Classification de l'échantillon (cas log-normal).

Une classification fondée sur les valeurs révisées implique que le seuil du risque maximum acceptable est égal à 50%. La courbe ROC présentée en Figure 3.15 permet de visualiser la performance des lois *a posteriori* pour toutes les valeurs de seuils entre 0 et 100%.

Les trois matrices de confusion présentées précédemment (Figure 3.13) sont aussi incluses dans le graphique. La matrice des valeurs vraies a logiquement les coordonnées (0,1). Par rapport à celle des mesures, les valeurs révisées présentent plus de vrais positifs et plus de faux négatifs, conformément à ce que les matrices montraient.

Les lois *a posteriori* donnent des meilleurs résultats que les mesures pour tous les seuils de risque maximum inférieurs ou égaux à 32%. Le point de la courbe ROC le plus proche de la sélection parfaite (de coordonnée (0,1)) est obtenu pour un seuil de 12%, et le taux de faux positifs est nul tant que le risque maximum acceptable est inférieur ou égal à ce même taux. Le taux de vrais positifs est égal à un dès que le seuil est supérieur ou égal à 93%. La courbe ROC étant toujours au-dessus du point de la matrice des mesures, il peut être conclu que l'utilisation des méthodes probabilistes

sera toujours avantageuse dans cette application, quelle que soit la politique de sélection.

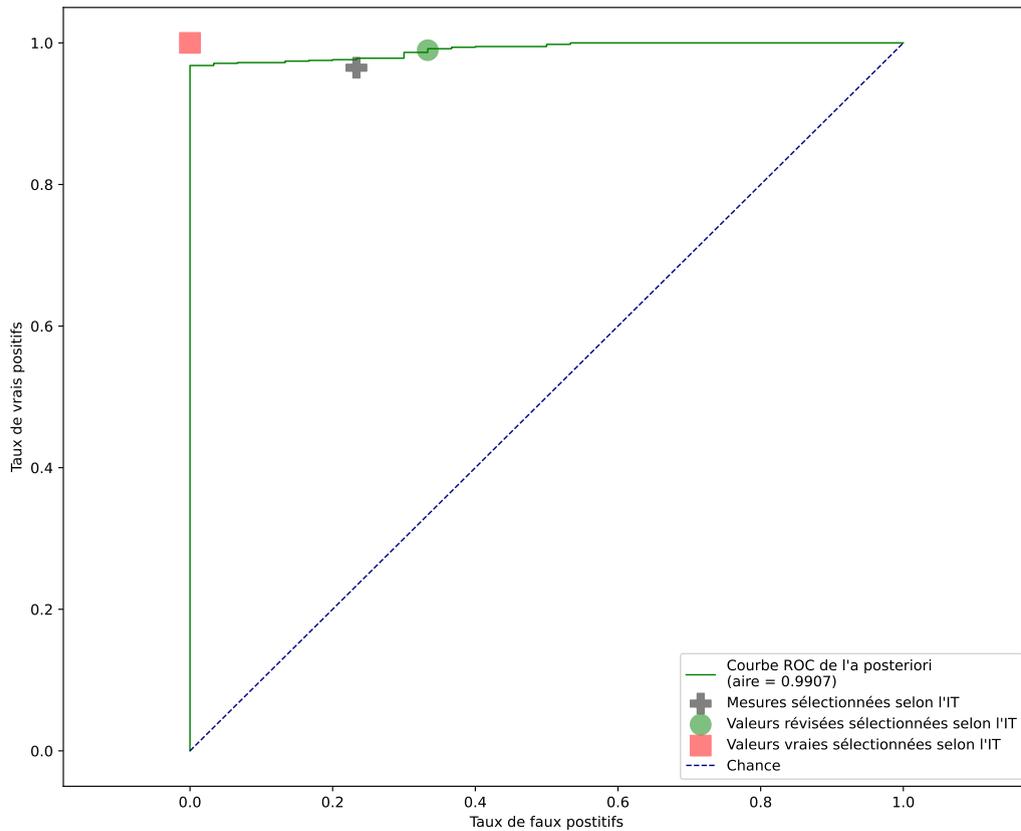


FIGURE 3.15: Tracé de la courbe ROC (cas log-normal).

3.3 Application des méthodes pour des données quelconques

3.3.1 Paramétrage et données d'entrées

Cette section traite d'une variable aléatoire de la valeur vraie (c'est-à-dire de la production) suivant un mélange de lois. Cet exemple permet de confronter la solution probabiliste face à une loi multimodale et fait écho à l'exemple mentionné dans le chapitre précédent (Section 2.1.3). Le mélange, composé de trois lois normales $M1$, $M2$ et $M3$, est décrit dans le Tableau 3.12. Dans le but de maximiser la ressemblance et donc la comparabilité des différents exemples de ce chapitre, la loi de l'erreur de mesure est maintenue gaussienne (mais il n'y a aucune contrainte de modélisation à son sujet). Le ratio des écarts-types a dû être réduit à 0.25 pour que la loi des mesures soit visiblement marquée par le caractère multimodal (sinon la somme converge rapidement vers un profil gaussien).

TABLEAU 3.12: Définitions des lois (cas multimodal).

Composante	Type	Moyenne	Écart-type
M1	Normal	99	0.3
M2	Normal	100	0.2
M3	Normal	101	0.3
Loi des mélanges	Trimodale	100	0.860
E	Normal	0	0.2

L'intervalle de tolérance est changé, légèrement excentré par rapport la loi de la valeur vraie et les bornes fixées égales à $[98.6, 101.8]$ dans le but de conserver une probabilité de défaillance comparable aux précédents exemples.

3.3.2 Application et résultat de la déconvolution

La méthode de déconvolution utilisée est celle décrite dans la Section 2.1.3. La valeur du pas h est fixée à 0.01. L'histogramme de l'échantillon de mesures est tracé en arrière-plan de la Figure 3.16 d'un gris léger. La loi déconvoluée est représentée d'un trait épais et vert avec celle de la valeur vraie d'un trait rouge et pointillé. Le tracé noir et fin est celui d'un lissage à noyau ("semblable à la solution") appliqué à l'échantillon afin de mieux visualiser l'apport de la déconvolution. Le graphique indique la densité de probabilité (axe des ordonnées) d'une valeur vraie donnée (axe des abscisses). Les lignes verticales orangées en semi-pointillé représentent les bornes de l'intervalle de tolérance.

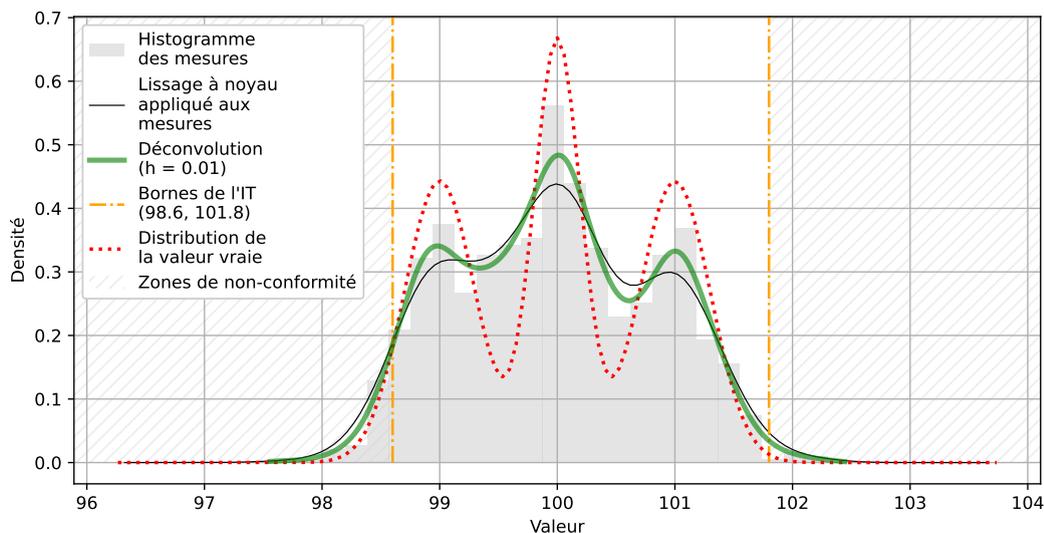


FIGURE 3.16: Déconvolution de la loi des mesures obtenues en métrologie (cas multimodal).

Dans cet exemple, la loi déconvoluée est plus proche de la loi inférée aux mesures que de celle de la valeur vraie. Le résultat, moins significatif, s'explique par la complexité de la valeur vraie recherchée (multimodale). La déconvolution améliore malgré tout la lecture des données, ce qui est toujours satisfaisant. Les deux premiers moments de la loi déconvoluée sont décrits dans le

Tableau 3.13. Là encore, l'écart de ces deux premiers moments statistiques est de l'ordre d'un pourcent de l'écart-type de la valeur vraie, et l'écart observé entre les courbes peut statistiquement être associé à la taille de l'échantillon de mesures.

Le Tableau 3.14 montre qu'en lisant le lot de mesures via la loi déconvoluée, l'erreur sur la probabilité de produire hors tolérance est réduite de moitié (réduction conforme à ce que l'on peut voir en Figure 3.16 au niveau des bornes de l'intervalle de tolérance).

TABLEAU 3.13: Évaluation de la précision de la reconstruction (cas multimodal).

Paramètre	Valeur	Erreur ($\%/\sigma_T$)
Moyenne	100.027	0.031
Écart-type	0.896	0.041

TABLEAU 3.14: Comparaison des probabilités de produire hors tolérance (p_{HT}) (cas multimodal).

Distribution	p_{HT}
Valeur vraie	0.0317
Mesure	0.0627
Déconvolution	0.0475

La Figure 3.17 présente les courbes de probabilité de prendre une mauvaise décision. Les courbes obtenues vérifient bien qu'il est très probable de se tromper en acceptant une mesure loin des bornes de l'intervalle de tolérance et qu'il est très probable de se tromper en refusant celles largement incluses. Les transitions, de zéro à un, correspondent logiquement au tracé de la fonction de répartition d'une loi normale (celle de l'erreur de mesure). Les intersections des deux courbes sont attendues au niveau des bornes de l'intervalle de tolérance mais sont déplacées à l'extérieur par l'effet des zones non-homogènes. La légère excentration de l'intervalle de tolérance offre un tracé différent à chacune des bornes.

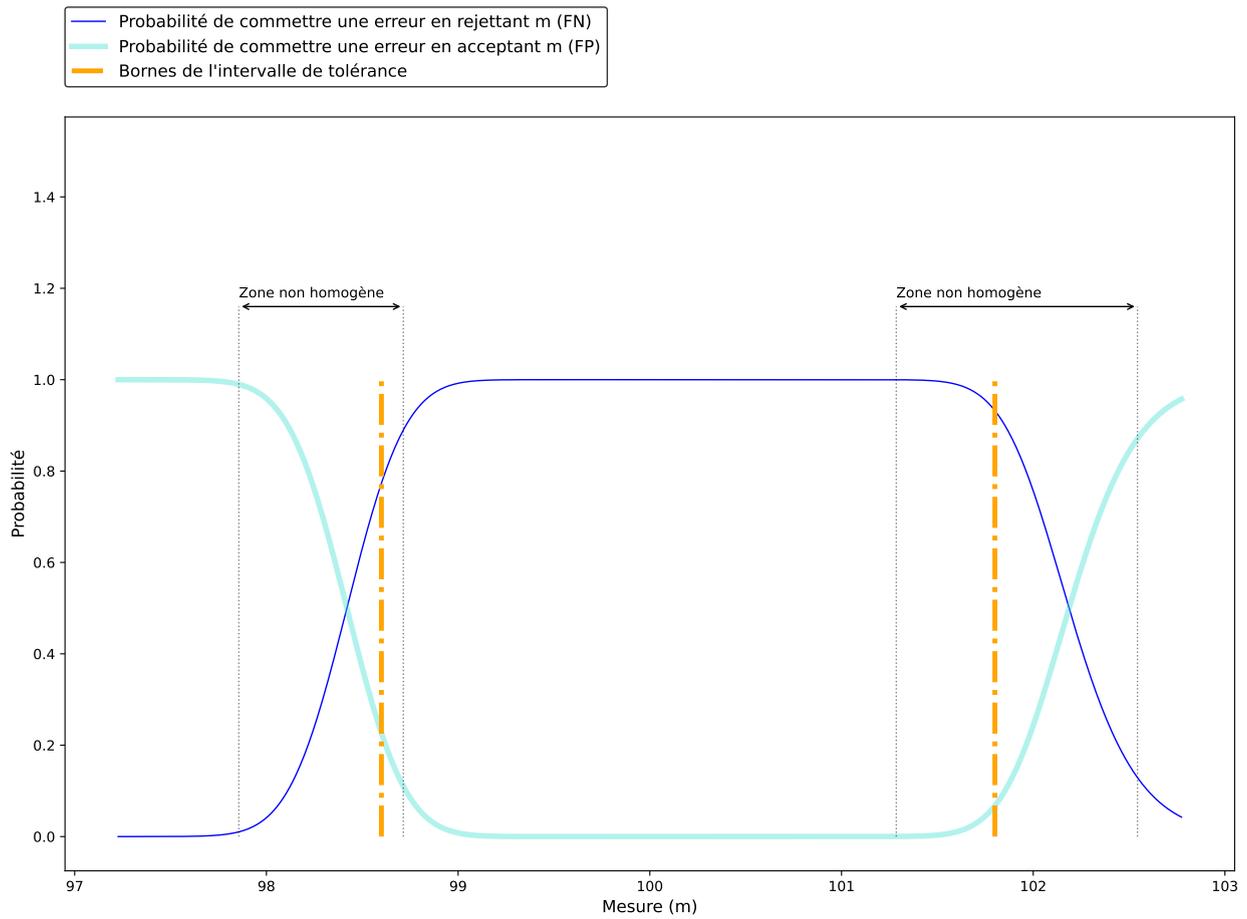


FIGURE 3.17: Probabilités qu'une mesure soit mal classifiée (cas multimodal).

Les courbes de risque sont tracées en Figure 3.18. Là encore, l'asymétrie des coûts en faveur du client tend à resserrer la sélection. Les intersections des courbes de risque se retrouvent logiquement rapprochées et la borne gauche est déplacée à l'intérieur de l'intervalle de tolérance.

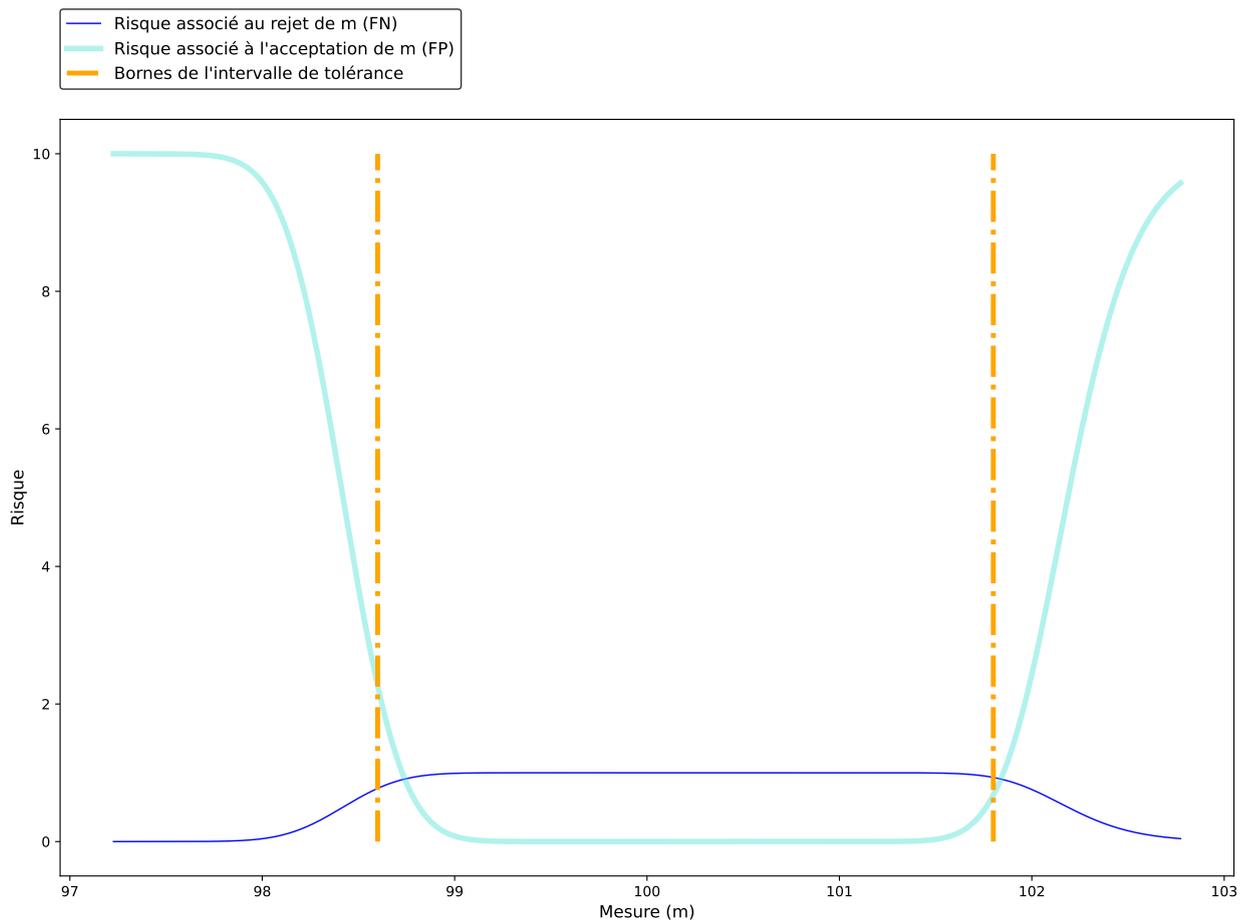


FIGURE 3.18: Risque qu'une mesure soit mal classifiée (cas multimodal).

Le Tableau 3.15 montre que le risque total est grandement modifié suite à l'application de la déconvolution. Non seulement l'idée qu'on se fait de la sélection optimale est très différente de la sélection classique (le risque devient sous-estimé) mais l'écart au risque total réel est aggravé (de quelques points). Le manque de précision dans l'identification de la loi de la valeur vraie ne permet pas d'améliorer la connaissance des risques de production.

TABLEAU 3.15: Risque total des différentes sélections (cas multimodal).

Base de sélection des mesures	Borne inférieure	Borne supérieure	Risque total
Intervalle de tolérance	98.6	101.8	1.009
Sélection optimale (Déconvolution)	98.779	101.774	0.782
Sélection parfaite (Loi de la valeur vraie)	98.742	101.835	0.910

3.3.3 Application et résultat de l'inférence bayésienne

La méthode d'inférence bayésienne décrite dans la Section 2.2.1 est appliquée pour obtenir la loi *a posteriori* et la valeur révisée de chacune des mesures de l'échantillon. La Figure 3.19 présente les nuages de points des résidus des échantillons des mesures et des valeurs révisées. La ligne bleue verticale en pointillé est celle de la moyenne de la valeur vraie. La ligne horizontale rouge est celle où la prédiction est égale à la valeur vraie (le meilleur des cas).

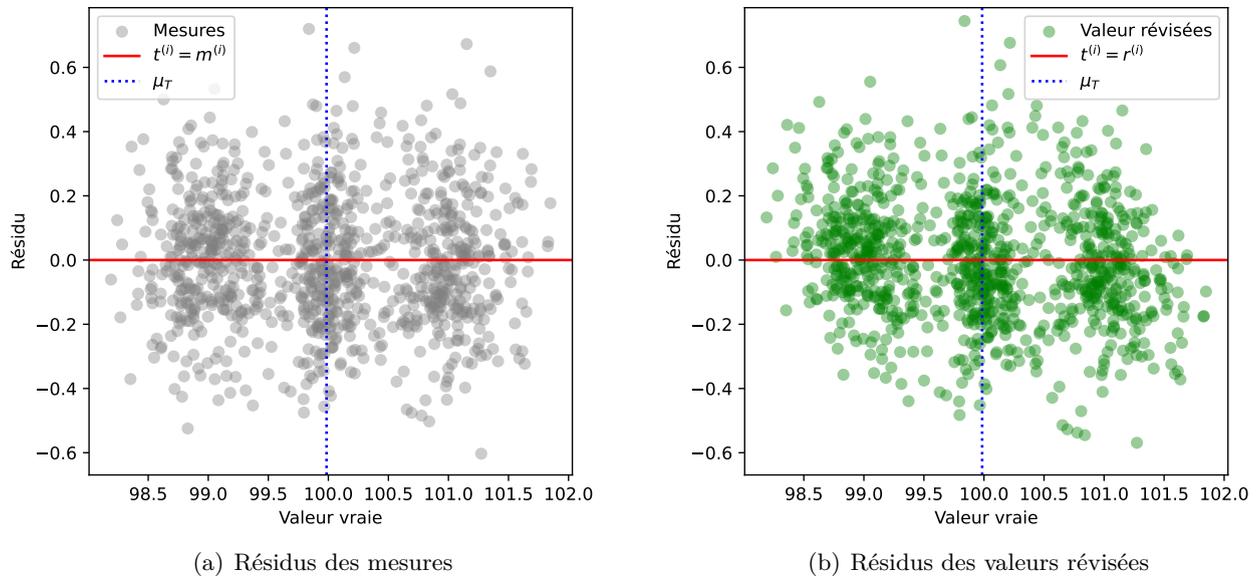


FIGURE 3.19: Comparaison des résidus des mesures et des valeurs révisées (cas multimodal).

Le nuage de points des résidus des mesures (Figure 3.19(a)) permet d'observer nettement les trois marginales de la loi des valeurs vraies qui ont des propriétés similaires.

Le nuage de points des valeurs révisées (Figure 3.19(b)) montre une corrélation négative semblable aux précédents cas (Sections 3.1.3 et 3.2.3). Les mesures les plus améliorées se situent dans les quarts haut-droit et bas-gauche du graphique. L'impact de l'inférence bayésienne n'est pas facilement lisible dans le reste du graphique. La méthode présente une balance légèrement en faveur des valeurs révisées.

Le coefficient de détermination (R^2) et l'erreur moyenne quadratique sont utilisés pour quantifier la qualité des prédictions. Le Tableau 3.16 présente les scores des prédictions de la valeur vraie sur la base des mesures et sur la base des valeurs révisées. Les deux scores sont en faveur des valeurs révisées (l'écart moyen est plus faible et le coefficient est plus proche de un). Le Tableau 3.16 montre aussi le pourcentage des mesures améliorées qui est le plus élevé des trois méthodes : 61%. Il est donc possible de conclure que l'application des méthodes probabilistes a permis une amélioration globale concernant la majorité des mesures.

TABLEAU 3.16: Indices de qualité des mesures et des valeurs révisées (cas multimodal).

Score	Mesures	Valeurs révisées
Erreur quadratique moyenne	0.040	0.035
R^2	0.948	0.954
% mesures améliorées	-	61.0

Les différentes déclarations de conformité peuvent être visualisées dans la Figure 3.21. La première des matrices de confusion en Figure 3.21(a) est celle de référence (prédiction des valeurs vraies). Elle montre que la grande majorité des pièces est conforme (97.0%). Cette véritable clas-

sification peut être visualisée dans un repère cartésien dans la Figure 3.20. La deuxième matrice montre la déclaration de conformité réalisée en confrontant les mesures à l'intervalle de tolérance (Figure 3.21(b)) : 965 pièces sont correctement classées, 8 prédites conformes à tort et 27 prédites non-conformes à tort. La troisième matrice représente la confrontation des valeurs révisées à l'intervalle de tolérance (Figure 3.21(c)). Un plus grand nombre de pièces est correctement classé (978) grâce aux méthodes proposées. Cette amélioration est la somme d'un nombre de faux négatifs largement diminué et d'un nombre de faux positifs légèrement augmenté. Là encore, les résultats sont favorables à l'application des méthodes probabilistes, et donc à la substitution des mesures par les valeurs révisées lorsque le coût du risque client est égal à celui du risque fournisseur.

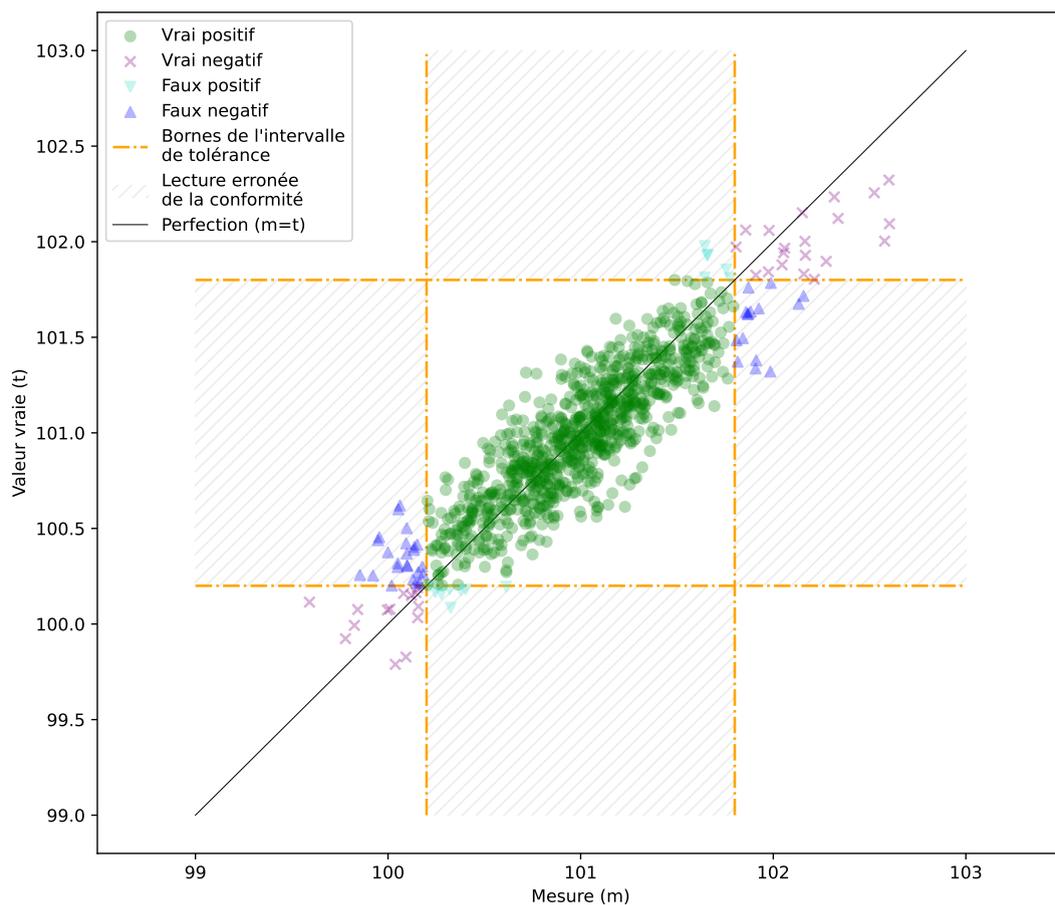


FIGURE 3.20: Classification de l'échantillon (cas multimodal).

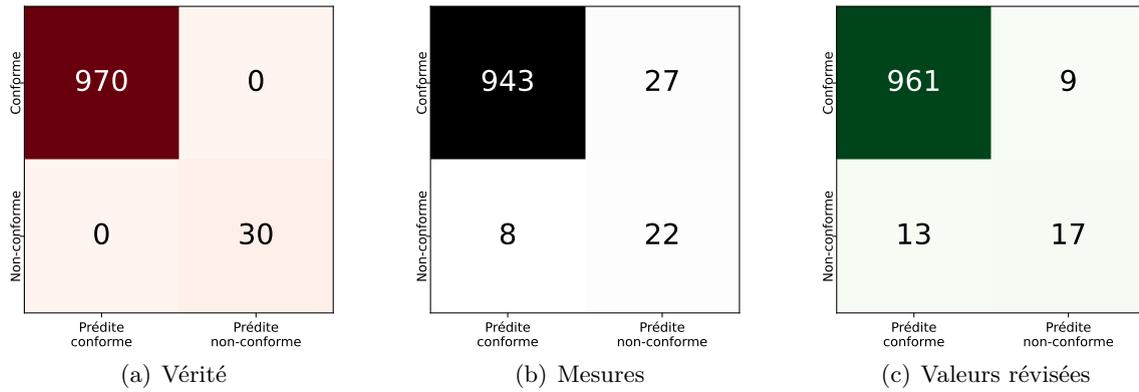


FIGURE 3.21: Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (cas multimodal).

Une classification fondée sur les valeurs révisées implique que le seuil du risque maximum acceptable est égal à 50%. La courbe ROC présentée en Figure 3.22 permet de visualiser la performance des lois *a posteriori* pour toutes les valeurs de seuils entre 0 et 100%. Les trois matrices de confusion présentées précédemment (Figure 3.21) sont aussi incluses dans le graphique. La matrice des valeurs vraies a logiquement les coordonnées (0, 1). Par rapport à celle des mesures, les valeurs révisées présentent plus de vrais positifs et plus de faux négatifs, conformément à ce que les matrices montraient.

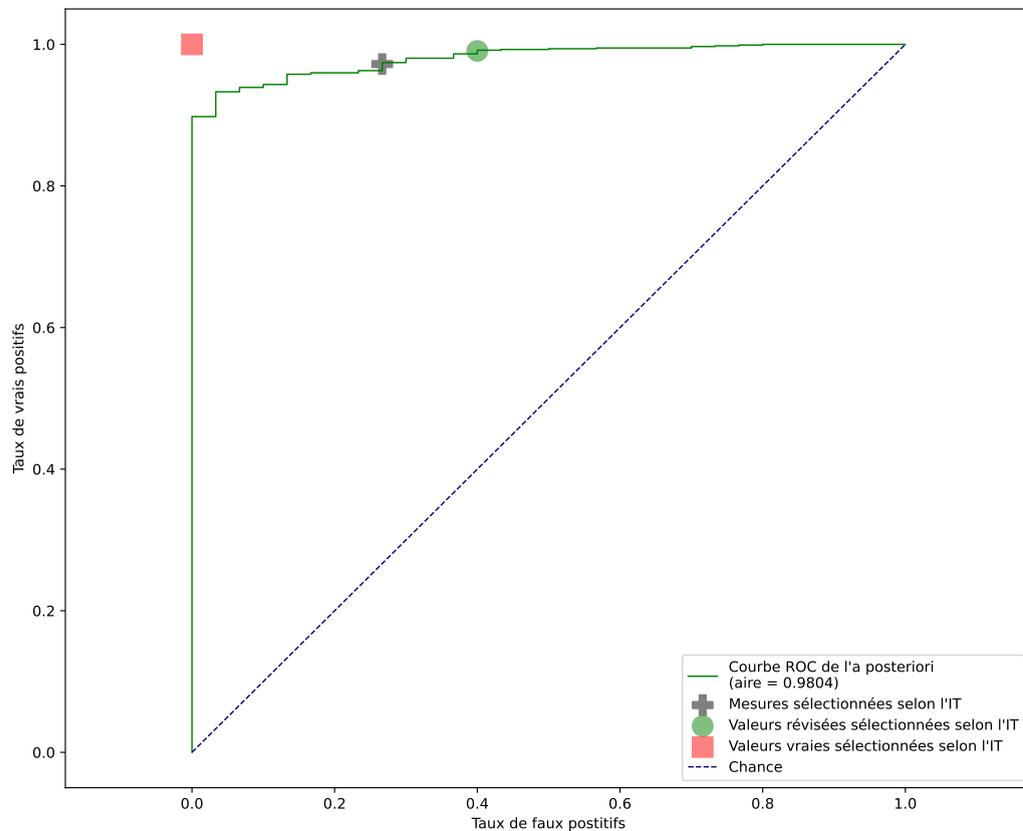


FIGURE 3.22: Tracé de la courbe ROC (cas multimodal).

Les lois *a posteriori* donnent des meilleurs résultats que les mesures pour tous les seuils de risque maximum inférieurs ou égaux à 19%. Le point de la courbe ROC le plus proche de la sélection parfaite (de coordonnée (0,1)) est obtenu pour un seuil de 13%. Le taux de faux positifs est nul tant que le risque maximum acceptable est inférieur ou égal à 6%. Le taux de vrais positifs est égal à un dès que le seuil est supérieur ou égal à 99%.

Conclusion

Ce troisième chapitre a exposé puis discuté les résultats de l'application des méthodes probabilistes présentées dans ces travaux. Trois exemples correspondant aux hypothèses des trois méthodes de déconvolution ont été présentés. Toutes les données utilisées ont été simulées, permettant ainsi de connaître la solution vers laquelle les méthodes devaient tendre. Les différents indicateurs statistiques ont été calculés dans le but de quantifier l'amélioration et surtout de savoir si les résultats obtenus étaient meilleurs que ceux donnés par les mesures.

Les meilleurs résultats ont été obtenus par les méthodes destinées aux données gaussiennes et dans le cas d'une loi de la valeur vraie identifiable. Cette conclusion dépend principalement de la taille des échantillons de mesures qui a été fixée à 1000. Avec des plus petits échantillons, tous les résultats ne seraient pas fiables et seraient de surcroît affectés par une potentielle autocorrélation de l'erreur de mesure. À l'inverse, toutes les méthodes sont très profitables à partir de 10000 échantillons.

La déconvolution et l'inférence bayésienne permettent d'affiner la connaissance de la production et de ses risques, et de les ajuster en fonction de leurs coûts. La déconvolution redéfinit la loi de référence, qui se substitue à la lecture ordinaire de l'échantillon de mesures obtenue en métrologie. Il est alors possible de lire et prédire plus précisément la probabilité de produire hors-tolérance. L'inférence bayésienne construit une loi *a posteriori* pour chaque mesure de l'échantillon. La loi *a posteriori* permet de connaître la valeur vraie la plus probable de chaque mesure, ici nommée "la valeur révisée". Ces estimations des valeurs vraies offrent une nouvelle lecture de la déclaration de conformité, et permet d'identifier correctement un plus grand nombre de pièces. Il est intéressant, en vue d'une application pratique, de noter que cette amélioration s'opère en réduisant les faux négatifs (risque fournisseur) et en augmentant légèrement les faux positifs (risque client), dont les coûts sont généralement différents. La loi *a posteriori* donne aussi accès à la probabilité que la valeur révisée soit hors-tolérance : cette information permet de connaître et d'ajuster le risque maximum acceptable pour la production - dont les performances, en termes de classification, sont synthétisés dans la courbe ROC.

Chapitre 4

Sensibilité des résultats aux paramètres de l'étude

Points clefs

- Variation des paramètres de simulation
 - Altération des données d'entrées
 - Application des méthodes probabilistes
-

Sommaire

Introduction	114
4.1 Impact d'une loi de la valeur vraie non identifiable	115
4.2 Impact d'une loi de l'erreur de mesure mal identifiée	123
4.3 Impact de mesures auto-corrélées	129
4.4 Étude de l'impact de la taille de l'échantillon de mesures	132
4.4.1 Cas de variables aléatoires gaussiennes	132
4.4.2 Déconvolution de données quelconques	135
4.4.3 Discrétisation de la courbe ROC	137
4.5 Impact du ratio des écarts-types de la valeur vraie et de l'erreur de mesure	138
Conclusion	140

Introduction

Le Chapitre 3 a présenté plusieurs applications en accords avec les hypothèses de modélisation des méthodes de déconvolution et d'inférence bayésienne. Les résultats obtenus répondaient aux objectifs d'amélioration de la mesure et de la gestion de la production. L'objectif de ce chapitre est de donner un aperçu de la sensibilité des méthodes en réaction à des données d'entrées imparfaites et à des variations des paramètres de simulation.

Les méthodes de déconvolution et d'inférence bayésienne se fondent toutes les deux sur la connaissance de la loi de l'erreur de mesure. Ce chapitre propose un cas d'application avec une loi de l'erreur de mesure dite "mal-identifiée" : sa moyenne et sa variance sont correctes, mais le tracé de la loi est erroné. L'inférence bayésienne requiert aussi la connaissance de la loi de la valeur vraie. Une deuxième application est construite avec une loi non-identifiable, de sorte à ce que l'étape de déconvolution soit imprécise. L'objectif est d'observer la propagation de cette imprécision à l'inférence bayésienne et donc à la prédiction des valeurs vraies. Une autre hypothèse que les méthodes requièrent est l'indépendance des mesures entre elles. Ce chapitre présente un cas d'application avec un échantillon constitué de mesures dont la valeur est définie en fonction de la précédente, afin d'étudier l'impact d'une éventuelle auto-corrélation.

Enfin, l'impact des principaux paramètres de simulation sont étudiés. Les plus influents sont la taille de l'échantillon et les valeurs des écarts-types de la valeur vraie et de l'erreur de mesure. Les échantillons de grandes tailles sont plus difficiles à obtenir mais sont plus fiables statistiquement. Une application est présentée pour avoir une idée du gain qu'ils peuvent apporter. Les écarts-types modélisent quant à eux différents contextes d'application. Leur étude a pour but d'extraire des tendances plus générales sur l'apport des méthodes, selon que la dispersion de l'erreur de mesure impacte largement ou non la lecture et la gestion de la production.

4.1 Impact d'une loi de la valeur vraie non identifiable

Paramétrage et données d'entrées

Cette analyse s'adresse uniquement à la méthode de déconvolution qui procède par identification. Le paramétrage se fonde sur celui présenté en Section 3.2. La valeur des bornes de l'intervalle de tolérance et la définition de l'erreur de mesure sont conservées, ainsi que le paramétrage utilisé pour l'application de la déconvolution. La principale modification faite au paramétrage initial est l'ajout d'un bruit à la valeur vraie, de sorte à ce que la loi finale ne corresponde pas à une loi usuelle (la solution ne doit pas faire partie de la liste des familles candidates).

L'expression de la valeur vraie est : $T = T_1 + T_2$ avec T_2 la composante du bruit qui suit une loi uniforme et T_1 la composante qui suit une loi très similaire à l'exemple original (sa position est légèrement changée, de sorte à conserver une même moyenne). Le Tableau 4.1 synthétise les paramètres des différentes lois de cet exemple.

TABLEAU 4.1: Paramètres des lois des données d'entrées.

Variable aléatoire	Type	Paramètres			Moyenne	Écart-type
Valeur vraie (T_1)	Log-normale	$\mu_{log}=0.01$	$\sigma_{log}=0.5$	$\gamma=98$	99.145	0.610
Bruit (T_2)	Uniforme	$a=1$	$b=2$	-	1.5	0.287
Erreur de mesure	Normal	-	-	-	0	0.3

a : la borne inférieure, b : la borne supérieure.

μ_{log} , σ_{log} : la moyenne et l'écart-type du logarithme de T

γ : le paramètre de position.

Application et résultat de la déconvolution

Malgré l'altération de la loi log-normale qui définit la valeur vraie, l'application de la déconvolution aboutit à l'identification d'une loi d'une même nature. Le score BIC optimal (2143.25) est supérieur à celui obtenu dans l'exemple d'origine (1882.51, Figure 3.8). Cette détérioration du score est logique puisqu'il n'est plus possible d'identifier la loi de la valeur vraie, mais seulement de l'approximer. L'ensemble des scores BIC est exposé en Figure 4.1. Les autres familles suivent les mêmes tendances que précédemment : les lois bornées sont les moins adéquates et les asymétriques les plus proches de la solution. La ressemblance de chaque tracé à la solution discerne finalement le classement dans chacune de ces catégories : la solution uniforme est meilleure que l'arcsinus, et la solution normale meilleure que celle triangulaire.

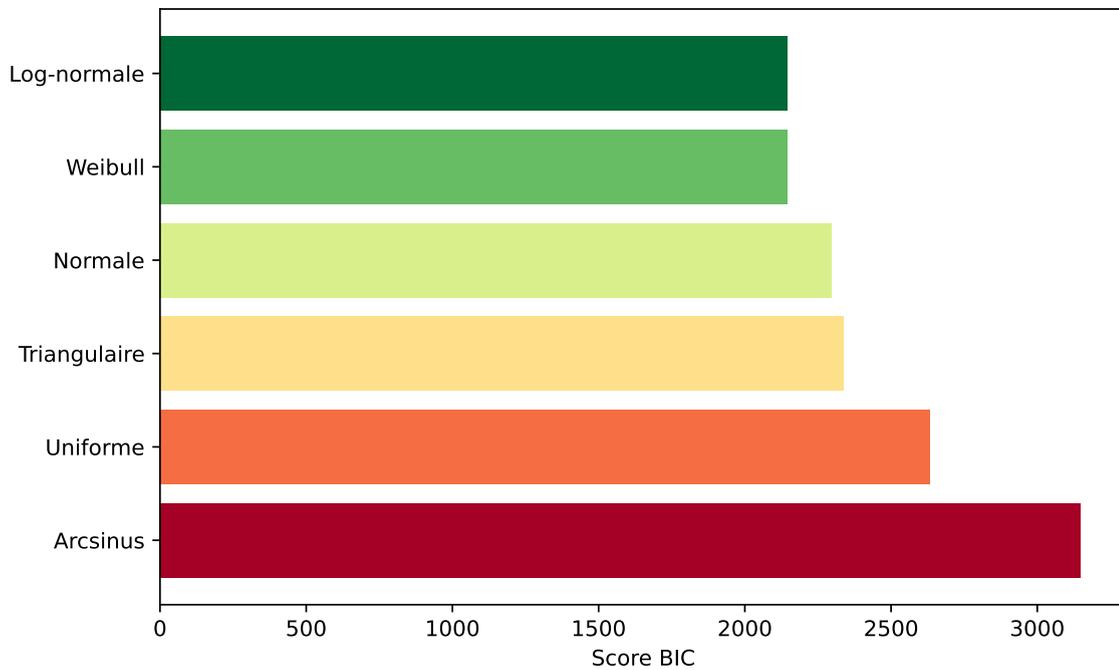


FIGURE 4.1: Score BIC du représentant de chaque famille candidate (cas test).

Sur la Figure 4.2, le gain apporté par la loi déconvoluée est moins évident dans ce cas test. A gauche, les densités dans la queue de loi sont sous-estimées (puis surestimées en se rapprochant du mode). A droite, les densités dans la queue de loi sont nettement identifiées. Le mode est visiblement moins bon que celui donné par l'échantillon de mesures.

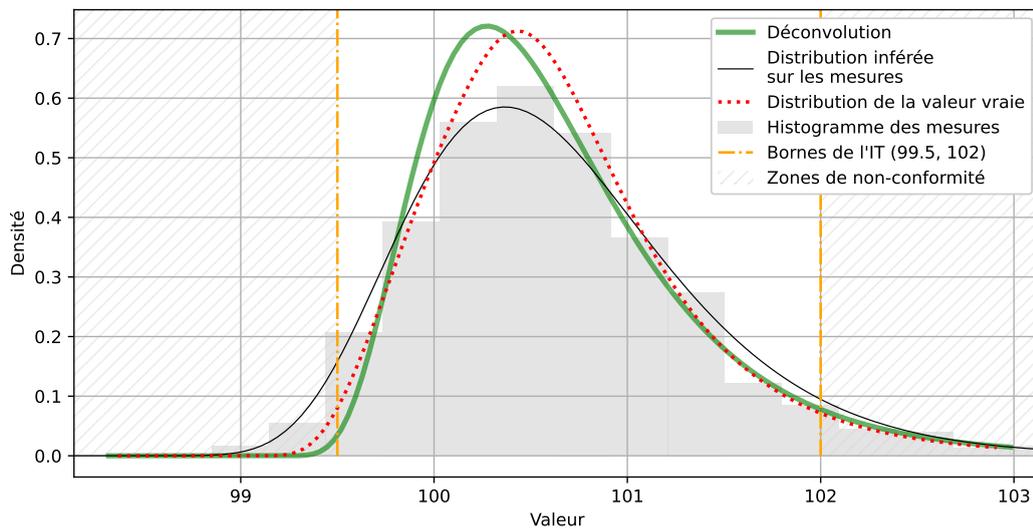


FIGURE 4.2: Déconvolution de la loi des mesures obtenues en métrologie (f_T non identifiable).

Les deux premiers moments statistiques de la loi déconvoluée et leurs pourcentages d'erreur par rapport à l'écart-type de la valeur vraie sont décrits dans le Tableau 4.2. La moyenne identifiée est à mi-chemin entre celle de la valeur vraie et celle de l'échantillon de mesures ($\mu_D = 100.644$, $\mu_T = 100.645$ et $\mu_M = 100.643$). L'écart-type est mieux identifié et s'accompagne d'une tendance à sous-estimer la dispersion ($\sigma_D = 0.760702$, $\sigma_T = 0.693878$ et $\sigma_M = 0.674831$).

Le Tableau 4.3 montre que la probabilité de produire hors tolérance calculée après déconvolution est très proche de la réalité, bien plus que celle qui pourrait être déduite de l'échantillon de mesures avec une loi similaire. La loi déconvoluée est satisfaisante puisqu'elle est plus proche de la vérité que celle des mesures.

TABLEAU 4.2: Évaluation de la précision de l'identification (f_T non identifiable).

Paramètres	Valeur	Erreur (%/ σ_T)
Moyenne	100.642	0.323
Écart-type	0.693	2.684
Mode	100.293	0.011

TABLEAU 4.3: Comparaison des probabilités de produire hors tolérance (p_{HT}) (f_T non identifiable).

Distribution	p_{HT}
Valeur vraie	0.0474
Mesure	0.0879
Déconvolution	0.0476

Les courbes de probabilité de prendre une mauvaise décision sont tracées en Figure 4.3. Le changement de statut correspond bien à la fonction de répartition d'une loi normale (celle de l'erreur de mesure). A la borne gauche de l'intervalle de tolérance, la zone non-homogène réduit nettement les pentes des deux courbes. A la borne droite, il n'y a pas de zone homogène : il apparaît seulement le décalage dû à la loi log-normale, pour laquelle la médiane est distincte de la moyenne.

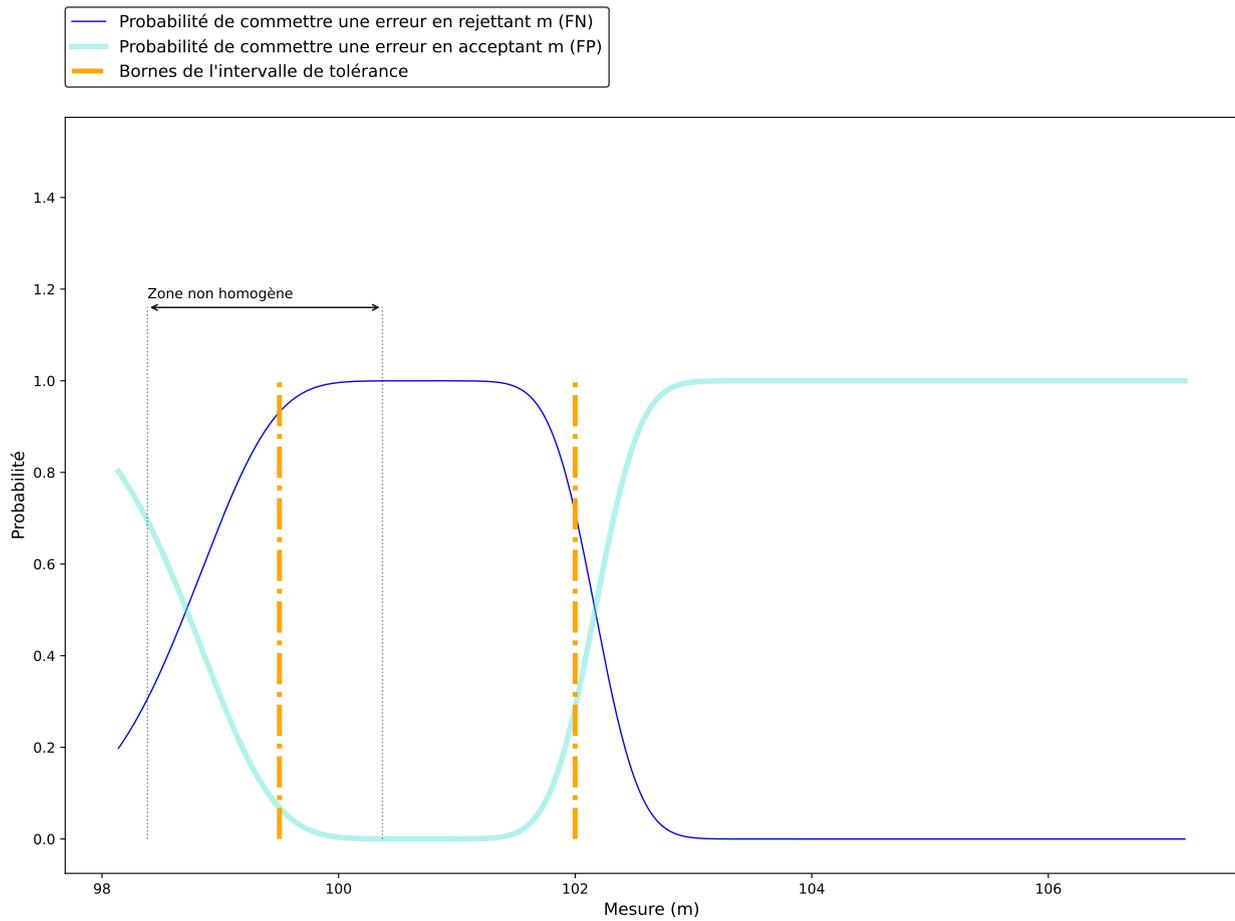
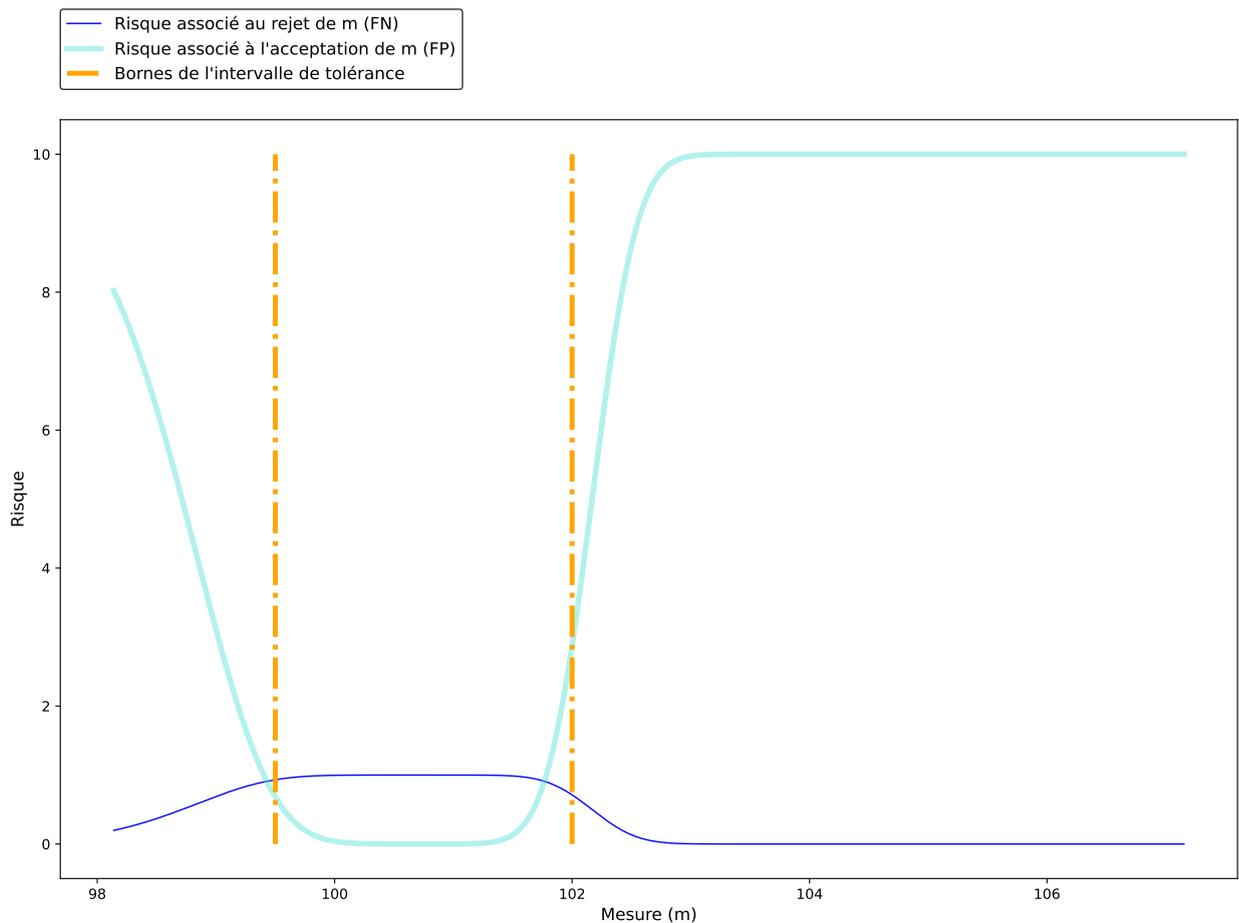


FIGURE 4.3: Probabilités qu'une mesure soit mal classifiée (f_T non identifiable).

Les courbes de risque sont tracées en Figure 4.4. Le ratio de 1 pour 10 qui a été fixé pour tous les exemples fait craindre le risque client et tend toujours à resserrer la sélection. L'intersection à gauche s'est rapprochée de la borne de l'intervalle de tolérance et celle de droite a été déplacée à l'intérieur de l'intervalle de tolérance.

FIGURE 4.4: Risque qu'une mesure soit mal classifiée (f_T non identifiable).

La sélection optimale déduite suite à la déconvolution est comparée à celle qui minimise parfaitement les risques (qui est accessible en connaissance de la loi de la valeur vraie) et à la sélection qui se tient à confronter les mesures brutes à l'intervalle de tolérance. Le Tableau 4.4 décrit les différents intervalles de chacune de ces sélections ainsi que le risque total.

TABLEAU 4.4: Risque total des différentes sélections (f_T non identifiable).

Mesure & Base de sélection des mesures	Borne inférieure	Borne supérieure	Risque total
Intervalle de tolérance	99.5	102	1.665
Sélection optimale (Déconvolution)	99.026	101.748	1.397
Sélection parfaite (Loi de la valeur vraie)	99.43	101.758	1.439

Conformément aux défauts qui ont été observés en Figure 4.2, la borne gauche de la sélection est de mauvaise qualité. A droite, la borne de la sélection est très proche de celle optimale. Le risque total est nettement diminué et même un peu en dessous de celui attendu : malgré une identification imparfaite de la loi de la valeur vraie, les résultats restent en faveur de l'application de cette méthode de déconvolution.

Application et résultat de l'inférence bayésienne

Comme pour les précédents exemples, la loi déconvoluée est utilisée comme *a priori* pour l'inférence bayésienne. Les nuages de points des résidus de l'échantillon de mesures et des valeurs révisées obtenues sont tracés en Figure 4.5. Le nuage propre aux mesures en Figure 4.5(a) permet d'observer que l'asymétrie de la loi log-normale précédemment exposée en Figure 3.12(a) a été réduite dans son mélange avec la composante T_2 . Les résidus des valeurs révisées en Figure 4.5(b) montrent une corrélation négative et sont beaucoup plus compacts.

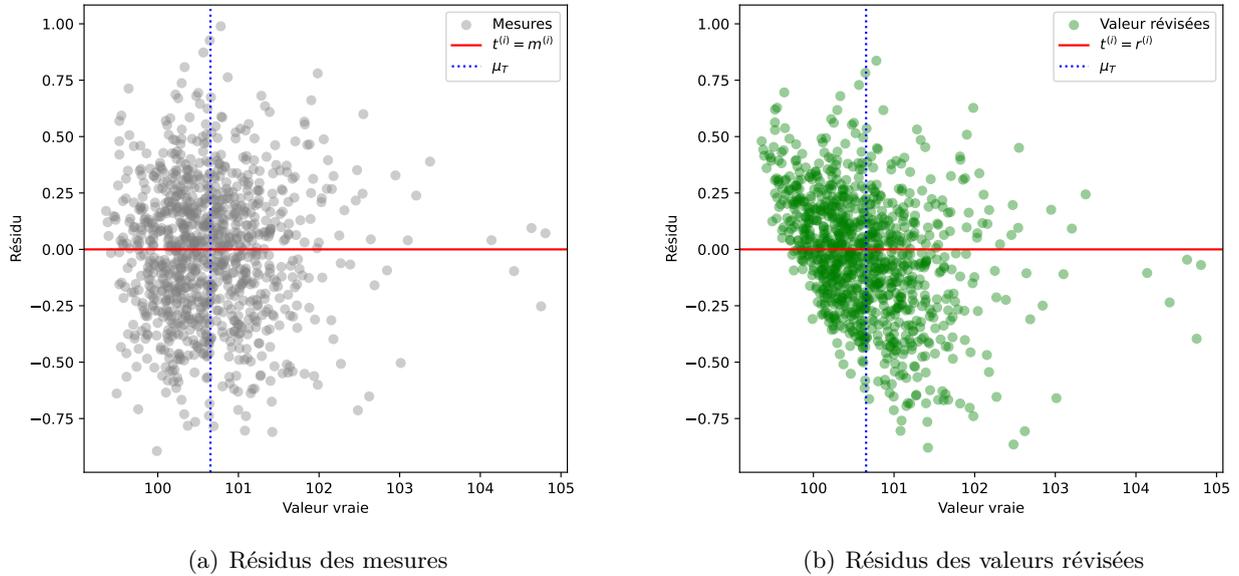


FIGURE 4.5: Comparaison des résidus des mesures et des valeurs révisées (f_T non identifiable).

Le coefficient de détermination (R^2) et l'erreur moyenne quadratique sont utilisés pour quantifier la qualité des prédictions. Le Tableau 4.5 présente les scores des prédictions de la valeur vraie sur la base des mesures et sur la base des valeurs révisées. Les deux scores sont nettement en faveur des valeurs révisées (l'écart moyen est plus faible et le coefficient est plus proche de un). Le Tableau 4.5 montre aussi que le pourcentage des mesures améliorées est égal à 59%, un score proche du meilleur obtenu jusqu'ici (Section 3.3.3). Il est donc possible de conclure que l'application des méthodes probabilistes a permis une amélioration globale concernant la majorité des mesures obtenues en métrologie.

TABLEAU 4.5: Indices de qualité des mesures et des valeurs révisées (f_T non identifiable).

Score	Mesures	Valeurs révisées
Erreur quadratique moyenne	0.092	0.075
R^2	0.806	0.841
% mesures améliorées	-	59.0

Les différentes déclarations de conformité peuvent être visualisées dans la Figure 4.6. La première des matrices de confusion en Figure 4.6(a) est celle de référence (prédiction des valeurs vraies). Elle montre que la grande majorité des pièces est conforme (95.8%). Cette véritable classification peut être visualisée dans un repère cartésien dans la Figure 4.7. La deuxième matrice

montre la déclaration de conformité réalisée en confrontant les mesures à l'intervalle de tolérance (Figure 4.6(b)) : 954 pièces sont correctement classées, 10 prédites conformes à tort et 36 prédites non-conformes à tort. La troisième matrice représente la confrontation des valeurs révisées à l'intervalle de tolérance (Figure 4.6(c)). Un plus grand nombre de pièces est correctement classé (978). Cette amélioration est la somme d'un nombre de faux négatifs très nettement diminué et d'un nombre de faux positifs légèrement augmenté. Ces résultats restent donc favorables à l'application des méthodes probabilistes, et donc à la substitution des mesures par les valeurs révisées si le coût du risque client est égal à celui du risque fournisseur.

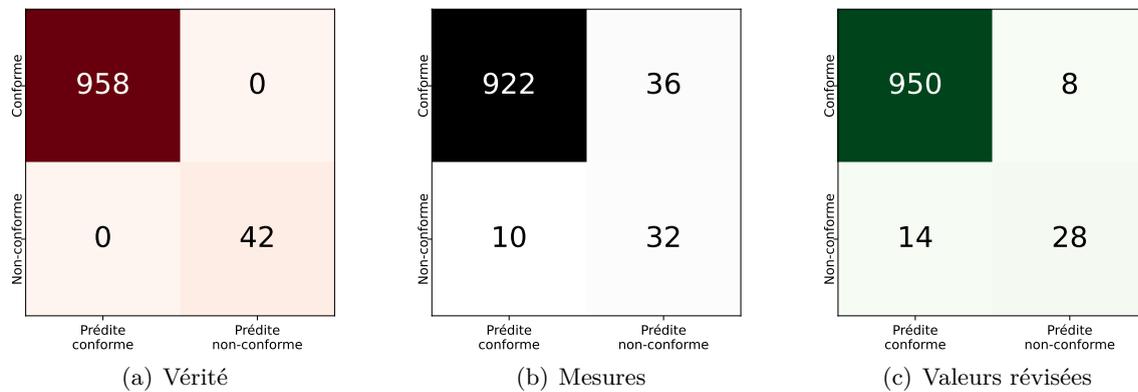


FIGURE 4.6: Comparaison des matrices de confusion des prédictions fondées sur la valeur vraie (a), les mesures (b) et les valeurs révisées (c) (f_T non identifiable).

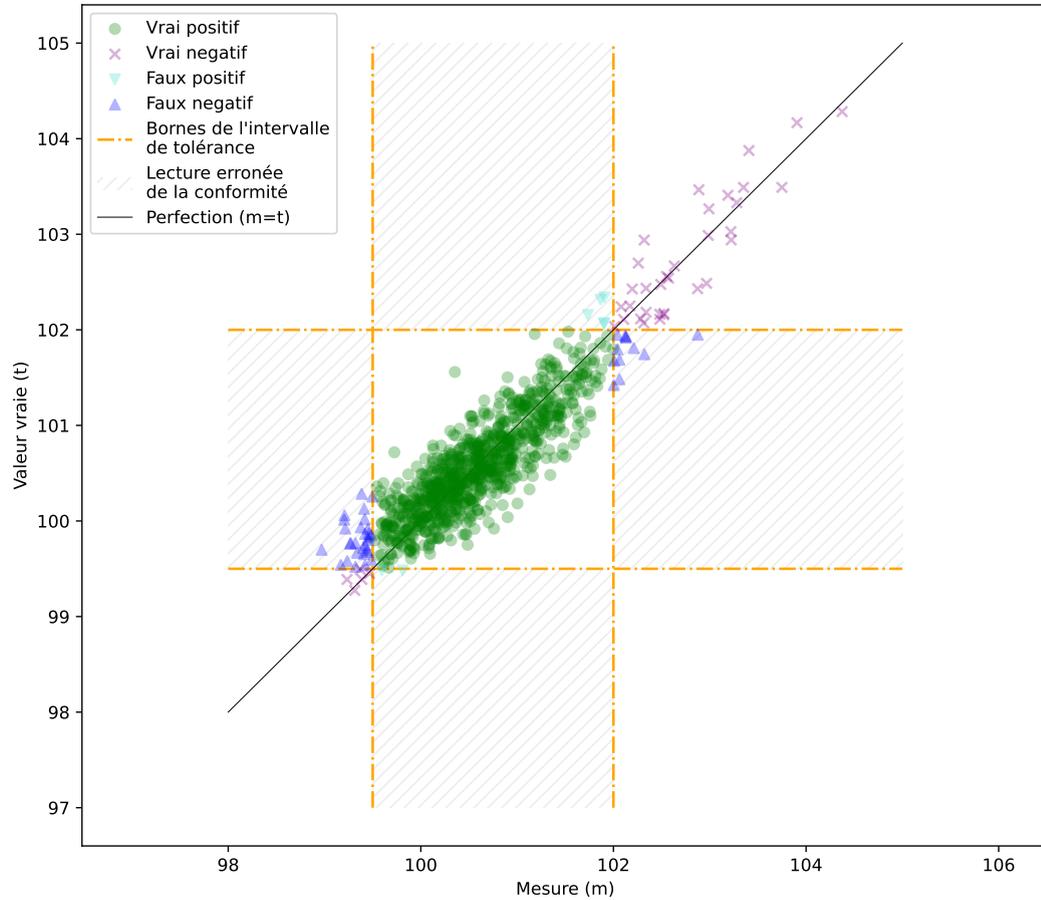


FIGURE 4.7: Classification de l'échantillon (f_T non identifiable).

La courbe ROC présentée en Figure 4.8 permet de visualiser la performance des lois *a posteriori* pour toutes les valeurs de seuils entre 0% et 100%. Les trois matrices de confusion présentées précédemment (Figure 4.7) sont aussi incluses dans le graphique. La matrice des valeurs vraies a logiquement les coordonnées (0,1). La matrice des valeurs révisées présente des taux de faux positifs et de vrais positifs légèrement plus faibles que celle des mesures.

Les lois *a posteriori* donnent des meilleurs résultats que les mesures pour tous les seuils de risque maximum inférieurs ou égaux à 15%. Le point de la courbe ROC le plus proche de la sélection parfaite (de coordonnée (0,1)) est obtenu pour un seuil de 2%. Le taux de faux positifs est nul tant que le risque maximum acceptable est inférieur ou égal à 1%. Le taux de vrais positifs est égal à un dès que le seuil est supérieur ou égal à 98%. La courbe ROC étant toujours au-dessus du point de la matrice des mesures, il peut être conclu que l'utilisation des méthodes probabilistes sera toujours avantageuse dans cette application, quelle que soit la politique de sélection.

Bien que la loi de la valeur vraie n'ait pas de distribution analytique, l'application des méthodes de déconvolution et de révision a tout de même permis d'améliorer les mesures. Les méthodes proposées ont donc montré une robustesse face à l'ajout d'un bruit qui ne permet pas de retrouver la loi de la valeur vraie parmi une liste de lois usuelles, ce qui est encourageant pour le déploiement et l'application de ces méthodes à des données réelles.

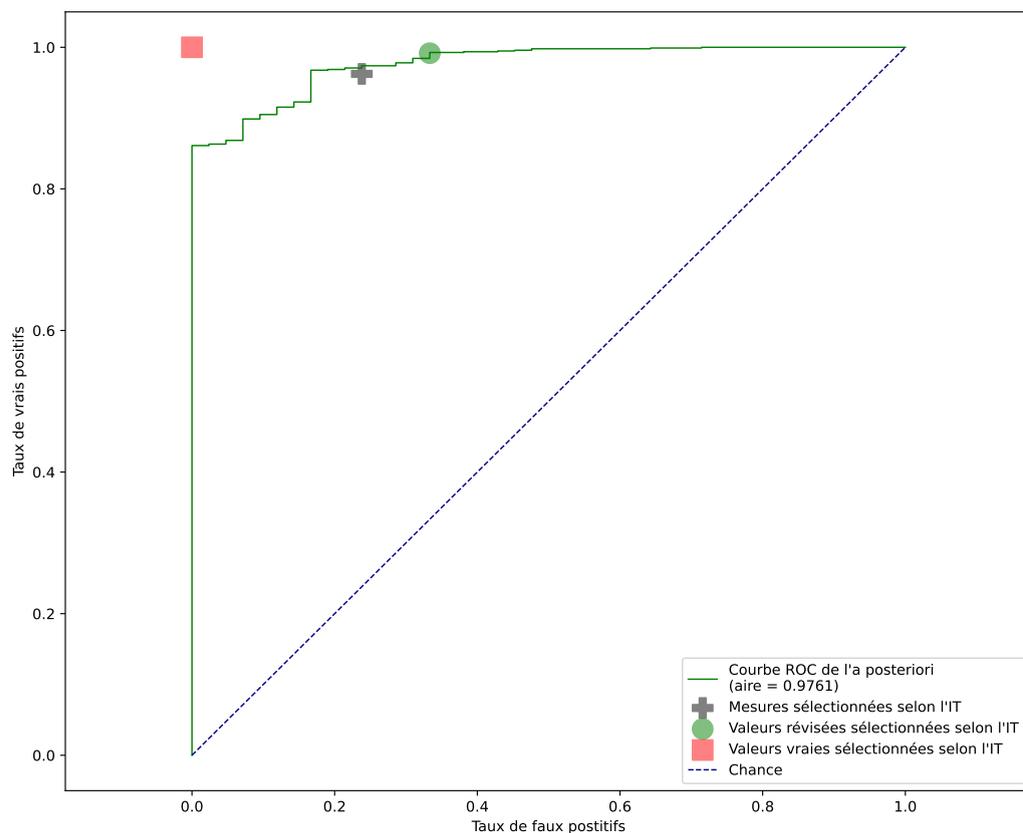


FIGURE 4.8: Tracé de la courbe ROC (f_T non identifiable).

4.2 Impact d'une loi de l'erreur de mesure mal identifiée

Paramétrage et données d'entrées

La densité de probabilité de l'erreur de mesure est une donnée supposée parfaitement connue dans tous ces travaux de thèse. Malgré tous les soins apportés, il est impossible de la connaître parfaitement en réalité. Cette section a pour but d'explicitier la sensibilité des résultats de l'alternative probabiliste à une information imparfaite (Tableau 3.1).

Les données tests sont calquées sur celles de l'exemple gaussien. La définition de la loi de l'erreur de mesure est elle aussi conservée et sera uniquement utilisée pour construire les échantillons de

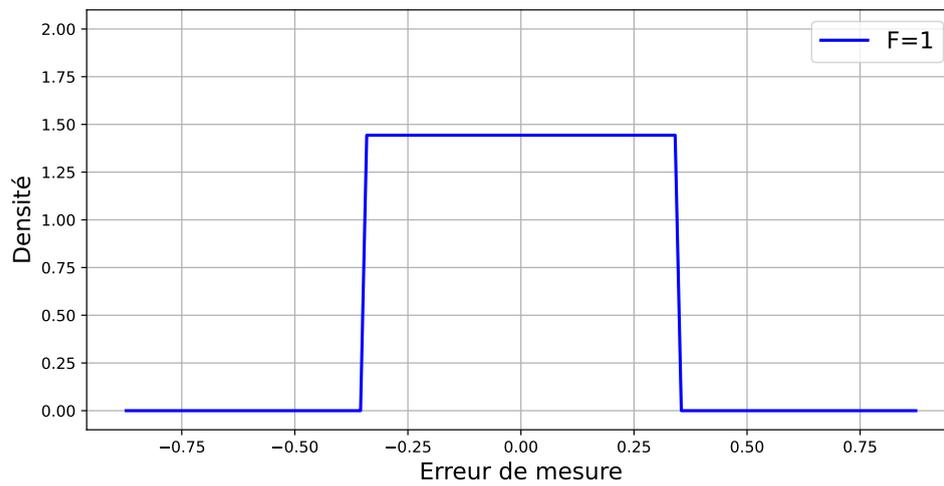
mesures. La loi erronée de l'erreur de mesure est utilisée au moment de l'application des méthodes probabilistes (déconvolution et inférence bayésienne).

Les différentes lois erronées sont tracées en Figure 4.9. Plusieurs degrés d'erreur sont testés. La loi Beta est une loi idéale pour modifier graduellement la géométrie de la courbe erronée. Sa densité de probabilité est définie par :

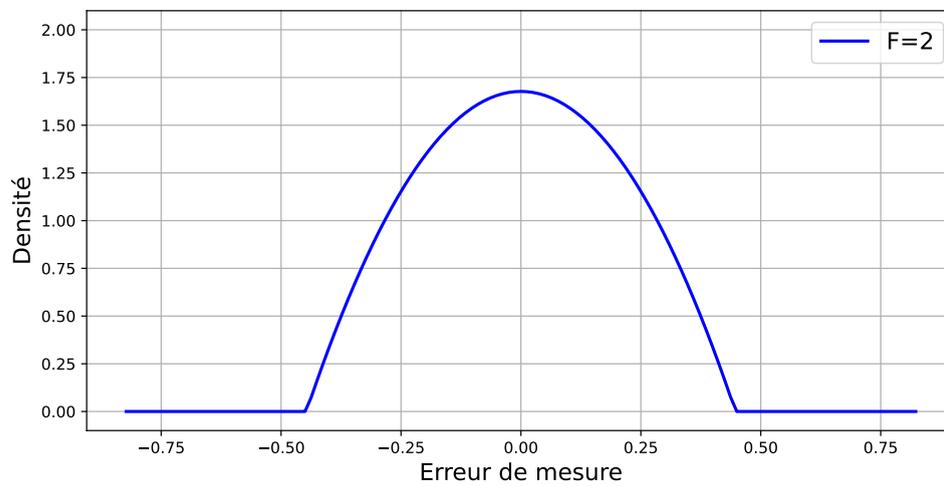
$$f(X) = \frac{(x-a)^{\alpha-1} (b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-1} B(\alpha, \beta)}, \quad x \in [a, b], \quad \alpha > 0, \quad \beta > 0 \text{ et } a < b \quad (4.1)$$

avec B la loi bêta d'Euler. Les valeurs de ses deux paramètres α, β sont fixées identiques et égales à une variable de forme F . Lorsque $F = 1$, la loi Beta correspond à une loi uniforme (Figure 4.9(a)). Lorsque $F = 2$, la loi est toujours bornée mais son allure se rapproche de celle d'une loi normale (Figure 4.9(b)). Lorsque $F = 4$, la loi Beta est similaire à une loi normale (Figure 4.9(c)). Les trois lois sont toutes définies de sorte à avoir la même moyenne et le même écart-type que la véritable loi : seul le tracé diffère.

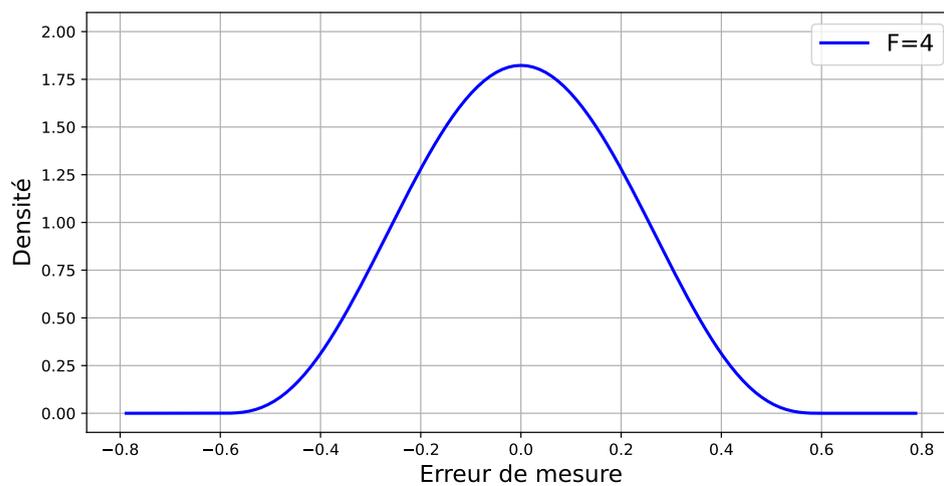
L'application des méthodes probabilistes est réalisée sur 50 échantillons de mesures différents, jusqu'à obtenir les valeurs révisées.



(a)



(b)

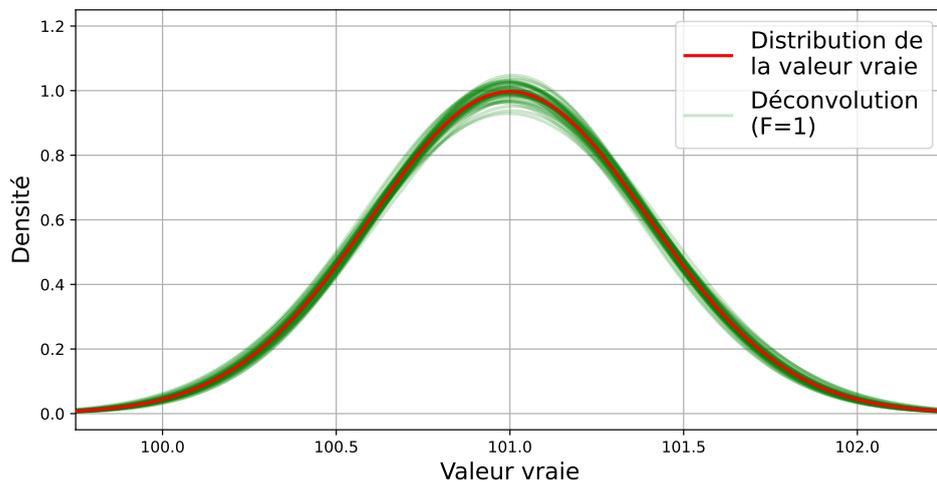


(c)

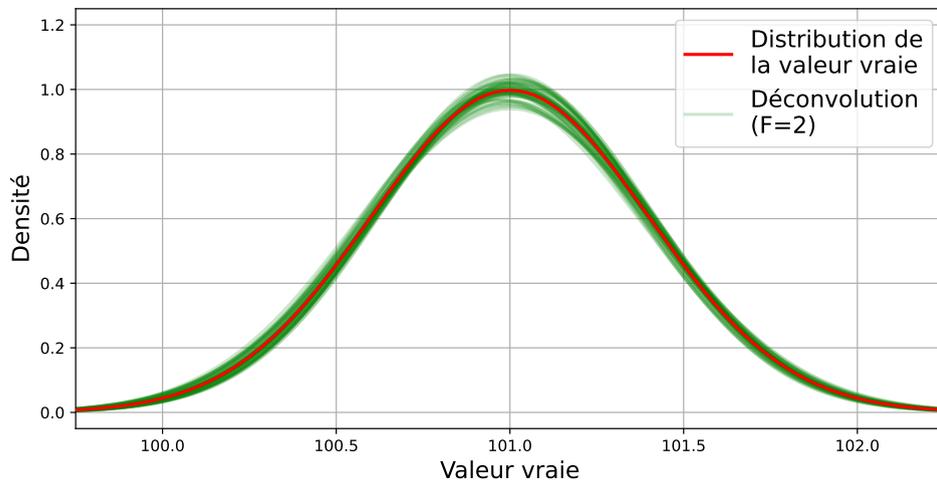
FIGURE 4.9: Tracé des différentes lois erronées de l'erreur de mesure.

Application des méthodes

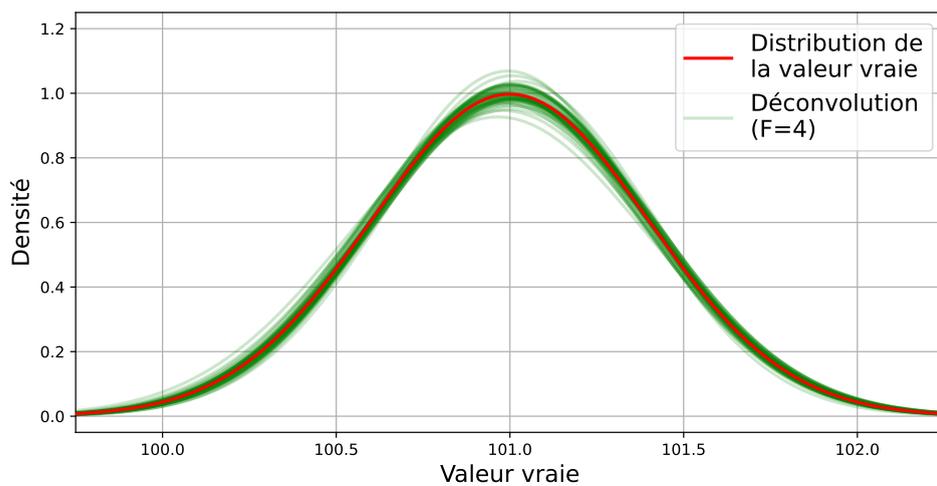
Les lois déconvoluées obtenues pour chaque valeur de F sont tracées en Figure 4.10. Les trois graphiques sont très similaires, quelle que soit la valeur de la variable de forme : avec une modélisation gaussienne la déconvolution ne fait appel qu'aux moyennes et écarts-types. A ce stade, les dispersions ne relèvent toujours que de la taille de l'échantillon.



(a) Déconvolutions réalisées avec une loi de type uniforme.



(b) Déconvolutions réalisées avec une loi de type normale tronquée.



(c) Déconvolutions réalisées avec une loi de type normale.

FIGURE 4.10: Impact d'une mauvaise identification de la loi de l'erreur de mesure lors de la déconvolution de données gaussiennes.

La Figure 4.11 montre les résidus d'un échantillon de valeurs révisées parmi les 50 calculés pour chaque valeur de F . Les nuages de points sont corrélés négativement et se distinguent par une répartition de plus en plus concentrique lorsque la valeur de F augmente.

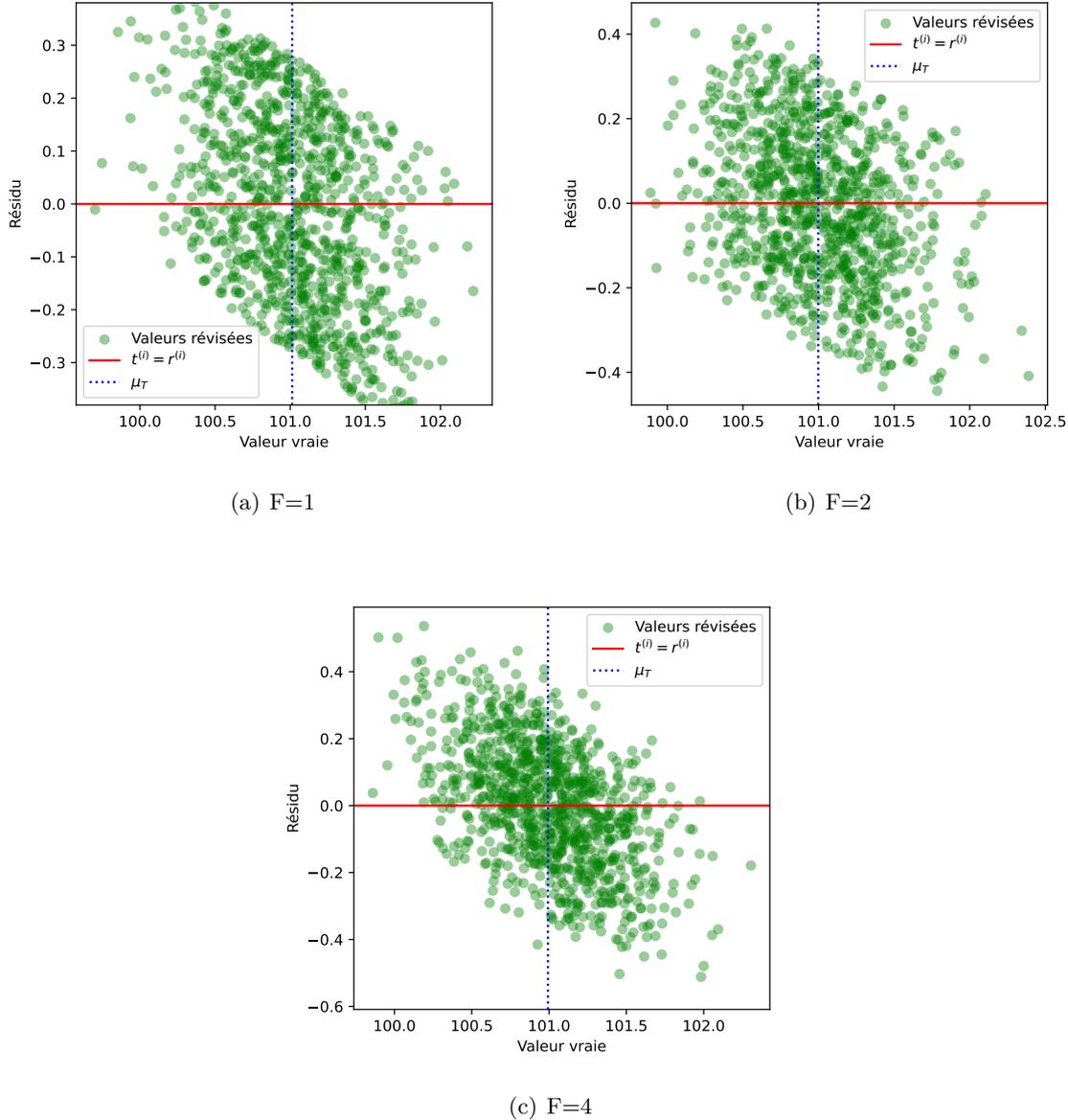


FIGURE 4.11: Nuage de points des résidus de valeurs révisées prises parmi les 50 lots générés.

Le Tableau 4.6 présente la moyenne et l'écart-type des 50 scores R^2 obtenus pour chacune des valeurs de F . Les scores obtenus sont très proches entre eux et ce, quelle que soit la valeur de F (les écarts sont de l'ordre de 10^{-4}). La très faible différence entre ces scores ne permet pas de tirer de conclusion sur une tendance : dans ce cas, elle est principalement due à la taille de l'échantillon des mesures. L'imprécision du tracé de la courbe de l'erreur de mesure influence ainsi les résultats de l'amélioration des mesures dans des proportions négligeables dans le cas d'une modélisation gaussienne.

TABLEAU 4.6: Synthèse des scores R^2 pour les différentes valeurs de F .

Variable de forme	Moyenne (R^2)	Écart-type (R^2)
$F = 1$	0.799	0.010
$F = 2$	0.780	0.010
$F = 4$	0.780	0.010
Vraie loi	0.801	0.009

4.3 Impact de mesures auto-corrélées

Paramétrage et données d'entrées

Deux pièces d'un même lot sont d'autant plus comparables qu'elles se succèdent lors des différentes étapes d'usinage. De nombreux phénomènes physiques dans l'environnement de la mesure évoluent en effet de façon continue : tels la température, le niveau d'usure ou l'accumulation d'un dérèglement. C'est cette idée que recouvre la notion d'auto-corrélation, notion qui peut s'adresser aux trois notions clés de cette étude : la mesure, la valeur vraie ou l'erreur de mesure. Cette section étudie spécifiquement l'auto-corrélation des erreurs de mesure pour évaluer la robustesse des méthodes face à l'hypothèse d'une loi de l'erreur de mesure parfaitement connue.

La valeur de l'erreur de mesure dépend désormais de la précédente. Dans le cadre de ces exemples simulés, la relation qui définit l'auto-corrélation est aussi fixée arbitrairement :

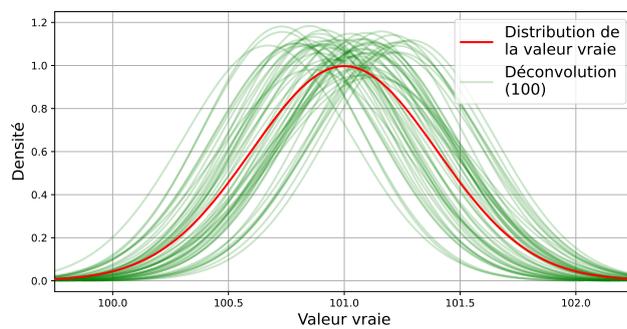
$$e^{(i+1)} = \frac{e^{(i)} + z^{(i)}}{\sqrt{\sigma_E^2 + \sigma_Z^2}} \quad (4.2)$$

avec $e^{(i+1)}$ et $e^{(i)}$ deux erreurs de mesure successives et différenciées par la quantité $z^{(i)}$, la i -ème réalisation de la variable aléatoire Z . La variable aléatoire Z est un artifice de calcul utilisé pour générer l'auto-corrélation tout en maîtrisant les valeurs de l'erreur de mesure. Son rôle est d'ajouter une dispersion aléatoire entre chaque erreur de mesure : elle est fixée indépendante de l'erreur de mesure et suit une loi normale (bruit gaussien) de moyenne nulle. La valeur de son écart-type est fixée inférieure (par quatre fois) à celle de l'écart-type de l'erreur de mesure : si elle est grande devant celle de l'erreur de mesure alors elle annule l'auto-corrélation ; si elle est négligeable alors l'erreur de mesure devient constante (égale à $e^{(0)}$). Le terme $\sqrt{\sigma_E^2 + \sigma_Z^2}$ de l'Équation (4.2) est une normalisation visant à conserver l'écart-type originale de l'erreur de mesure.

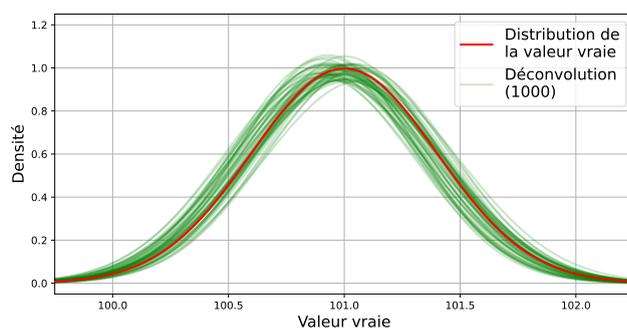
L'impact de l'auto-corrélation est lu sur quatre tailles d'échantillons : 10^2 , 10^3 , 10^4 et 10^5 . En dessous de 100 mesures, les résultats ne sont plus statistiquement pertinents (l'échantillon n'est pas représentatif). Les échantillons de mesures de tailles 10000 et 100000 suffisent à donner la tendance pour les très grands échantillons (et il devient très rare de pouvoir en accéder à plus en pratique). Le reste de l'étude est construit de façon similaire à l'étude de l'impact d'une loi de l'erreur de mesure mal identifiée (Section 4.2). Un nombre de 50 simulations sont réalisées pour chacune de ces tailles, pour lesquelles la déconvolution et l'inférence bayésienne sont appliquées pour l'amélioration des mesures.

Application des méthodes

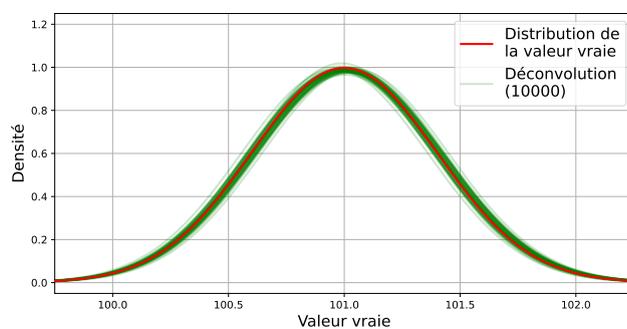
Les résultats de l'étape de déconvolution sont présentés en Figure 4.12. Ils peuvent directement être confrontés à ceux obtenus lors de l'étude de l'impact de la taille de l'échantillon de mesures. On observe que plus la taille de l'échantillon de mesures diminue, plus les tracés des lois déconvoluées sont dispersés. Les résultats obtenus pour une taille 100 sont médiocres : les lois déconvoluées sont largement biaisées et sous-estiment en majorité la dispersion des valeurs vraies. Pour une taille 1000, les biais sont nettement moins marqués et les variances plus fidèles. Avec des échantillons de tailles 10000, l'influence de l'auto-corrélation s'amenuise. Le phénomène d'auto-corrélation aggrave donc significativement les résultats obtenus par déconvolution, tout en étant de second ordre par rapport à l'impact de la taille de l'échantillon.



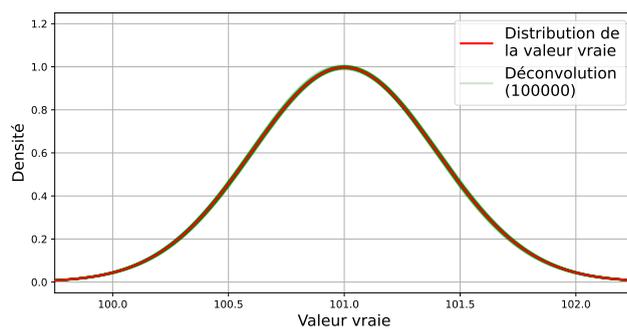
(a) Échantillon taille 100



(b) Échantillon taille 1000



(c) Échantillon taille 10000



(d) Échantillon taille 100000

FIGURE 4.12: Impact de l'auto-corrélation de l'échantillon de mesures sur la déconvolution de lois normales.

Pour chacun des 50 tests, L'inférence bayésienne est appliquée suite à la déconvolution pour obtenir l'échantillon des valeurs révisées. Le score R^2 est calculé pour chacun de ces échantillons. Le Tableau 4.7 présente la moyenne et l'écart-type des scores R^2 pour chaque taille d'échantillon.

TABLEAU 4.7: Synthèse des scores R^2 obtenus avec les échantillons auto-corrélés.

Taille de l'échantillon	Moyenne des R^2	Écart-type des R^2
100	0.714	0.173
1000	0.809	0.034
10000	0.804	0.011
100000	0.801	0.004

En comparaison avec les résultats des échantillons non auto-corrélés (Tableau 4.8), l'auto-corrélation a nettement dégradé la moyenne des scores R^2 obtenue pour des échantillons de taille 100. Pour les tailles supérieures, les résultats sont similaires, d'une qualité légèrement supérieure et en faveur de la taille d'échantillon 1000. Les écarts-types obtenus décroissent quand la taille augmente, et sont du même ordre de grandeur. L'auto-corrélation n'affecte donc pas significativement l'amélioration des mesures pour la taille 100.

4.4 Étude de l'impact de la taille de l'échantillon de mesures

4.4.1 Cas de variables aléatoires gaussiennes

Il a été observé dans les précédents exemples (Sections 3.1.2, 3.2.2 et 3.3.2) que la qualité de la déconvolution (et par conséquent celle de l'inférence bayésienne) dépend de la représentativité de l'échantillon de mesures. Cette condition ne peut être remplie que par un échantillon de taille infinie ou par un improbable hasard. La taille de l'échantillon aura donc toujours un impact négatif sur la qualité de l'amélioration des mesures.

Cette section a pour but d'explicitier l'impact de la taille de l'échantillon sur les méthodes d'amélioration. Les données sont simulées à partir de variables aléatoires gaussiennes, car les résultats des méthodes associées sont connus analytiquement et devraient systématiquement converger vers la solution (si l'échantillon de mesures était représentatif). Quatre tailles sont choisies pour couvrir différents ordres de grandeurs : 10^2 , 10^3 , 10^4 et 10^5 . Les méthodes d'amélioration sont appliquées à 50 simulations de chacune de ces tailles.

Les résultats obtenus à l'étape de la déconvolution sont exposés ci-dessous (Figure 4.13). Les résultats convergent vers la solution à mesure que la taille de l'échantillon augmente.

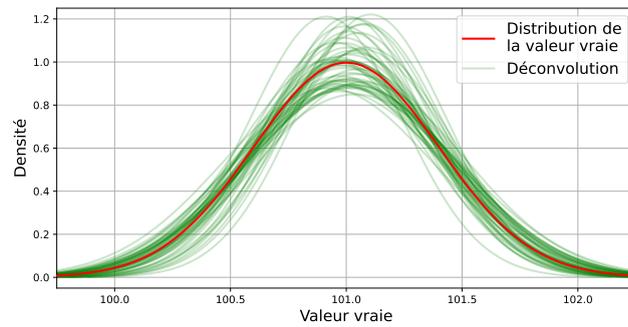
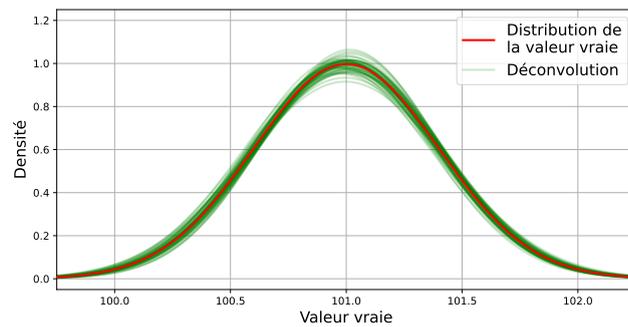
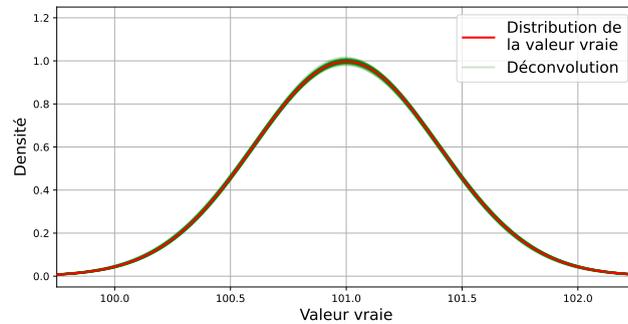
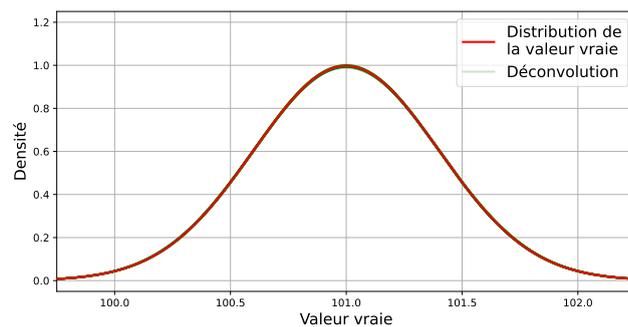
(a) Multiples déconvolutions d'échantillons de 10^2 mesures.(b) Multiples déconvolutions d'échantillons de 10^3 mesures.(c) Multiples déconvolutions d'échantillons de 10^4 mesures.(d) Multiples déconvolutions d'échantillons de 10^5 mesures.

FIGURE 4.13: Impact de la taille de l'échantillon de mesures sur la déconvolution de lois normales.

Le Tableau 4.8 présente une synthèse des scores R^2 des échantillons de valeurs révisées obtenus suite à l'inférence bayésienne. L'écart-type des scores obtenus diminue lorsque la taille de l'échantillon augmente. La moyenne augmente significativement lorsque la taille passe de 10^2 à 10^3 , puis diminue très légèrement. La tendance de cette moyenne est particulièrement visible dans le diagramme en violon présenté en Figure 4.14.

En plus des informations de la classique boîte à moustaches (notamment la médiane et l'écart interquartile), le diagramme en violon permet de visualiser la loi des scores R^2 . Ce graphique comporte deux lois par violon (ce type de graphique est plus généralement présenté avec une seule loi symétriquement dédoublée). La taille des échantillons est en abscisses et les loi des scores (des mesures et des valeurs révisées) en ordonnée (respectivement grises et vertes).

Une autre information apparaît dans le graphique en violon : à partir d'une taille de 1000 mesures, il est improbable que le score R^2 des mesures soit supérieur à celui des valeurs révisées (ce qui confirme le résultat présenté précédemment en Section 4.5). Les quatre cas parmi les deux-cents tests où l'inférence bayésienne a dégradé le score R^2 ont tous eu lieu pour des échantillons de taille 10^2 .

TABLEAU 4.8: Moyenne et écart-type des R^2 en fonction de la taille d'échantillon.

Taille de l'échantillon	Moyenne des R^2	Écart-type des R^2
10^2	0.797	0.035
10^3	0.804	0.010
10^4	0.800	0.003
10^5	0.800	0.001

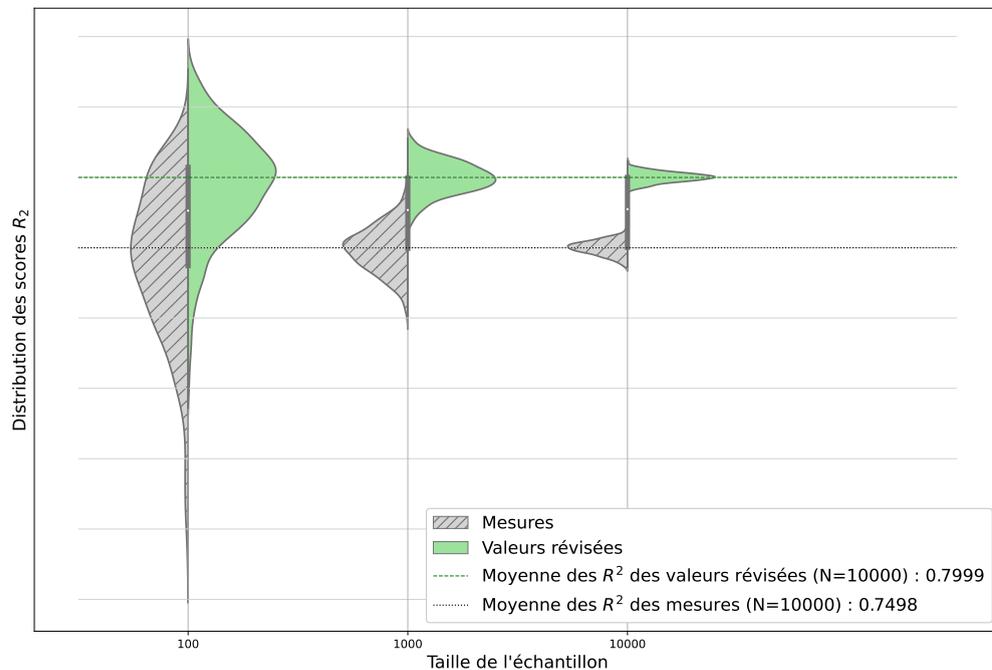
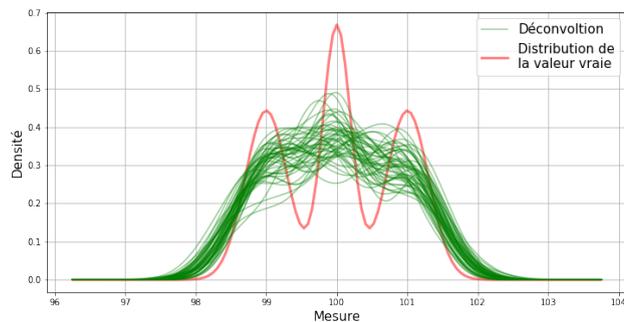


FIGURE 4.14: Distribution des scores R^2 des mesures et valeurs révisées en fonction de la taille d'échantillon.

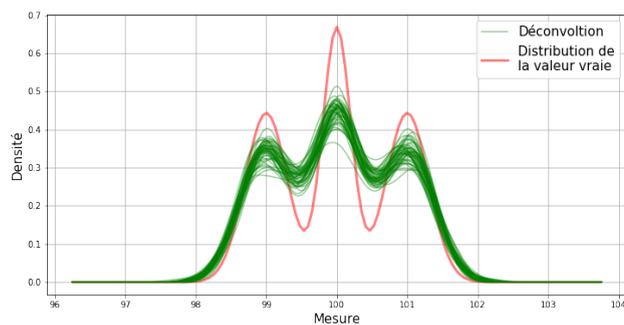
4.4.2 Déconvolution de données quelconques

Parmi les méthodes de déconvolution illustrées dans les exemples précédents, la méthode par approximations locales est celle qui a le moins bien identifié la loi de la valeur vraie (Figure 3.16). La qualité des résultats de cette approche dépend en effet grandement de la taille de l'échantillon de mesures.

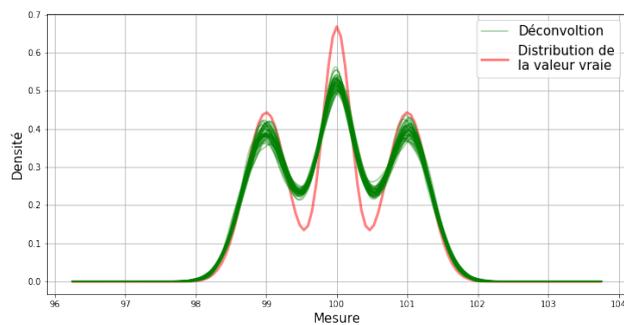
Cette étude reprend le paramétrage de l'exemple d'application des données quelconques (Tableau 3.12). La valeur de la taille de l'échantillon est successivement fixée à 100, 1000, 10000 puis 100000. Pour chacune de ces tailles, 50 échantillons de mesures sont générés par simulation et déconvolués : les 50 courbes obtenues sont superposées et confrontées à la loi de la valeur vraie. Les résultats sont montrés en Figure 4.15.



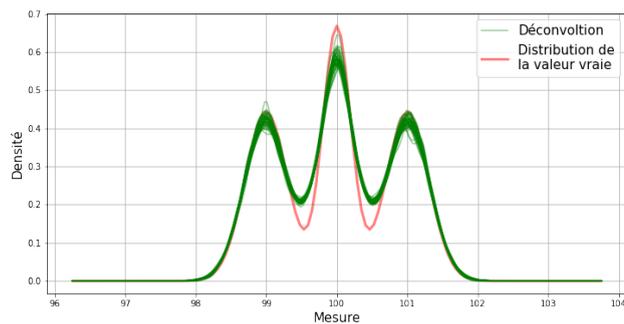
(a) Échantillon de taille 100.



(b) Échantillon de taille 1000.



(c) Échantillon de taille 10000.



(d) Échantillon de taille 100000.

FIGURE 4.15: Impact de la taille de l'échantillon de mesures sur la déconvolution de lois normales.

Pour un échantillon de taille 100 (Figure 4.15(a)), les résultats sont chaotiques et retranscrivent à peine le caractère multimodal de la loi de la valeur vraie. Lorsque la taille de l'échantillon de mesures passe à 1000 (Figure 4.15(b)), les trois modes sont correctement identifiés. Les résultats présente toujours une dispersion mais la qualité moyenne des tracés correspond à celle observée dans l'exemple présenté en Figure 3.16. Quand la taille de l'échantillon augmente encore, avec les tailles 10000 et 100000, les tracés s'améliorent encore, tant dans la régularité que dans la convergence vers la loi de la valeur vraie. Un écart persiste cependant, qui provient en partie des approximations intrinsèques à la méthode du lissage à noyau.

4.4.3 Discrétisation de la courbe ROC

La courbe ROC trace l'évolution des taux de vrais positifs et de faux positifs en fonction d'un seuil de sélection donné. Le dénombrement des mesures selon leurs conformités implique nécessairement que la courbe ROC est discontinue. Plus l'échantillon de mesures est petit, plus les taux varient brusquement et moins la courbe (et notamment son indice AUC) caractérise avec précision la performance des lois *a posteriori*.

Cette étude reprend le paramétrage de l'exemple d'application pour des variables aléatoires gaussiennes présentées précédemment (Tableau 3.1). Un échantillon de taille 10^4 est généré. Les méthodes d'amélioration sont appliquées jusqu'à obtenir la courbe ROC, tracée en Figure 4.16. La courbe obtenue est épurée des fortes discontinuités pour une lecture et donc une gestion plus fine de la production. Ce lissage est à l'avantage du score AUC (l'aire sous la courbe) qui augmente légèrement.

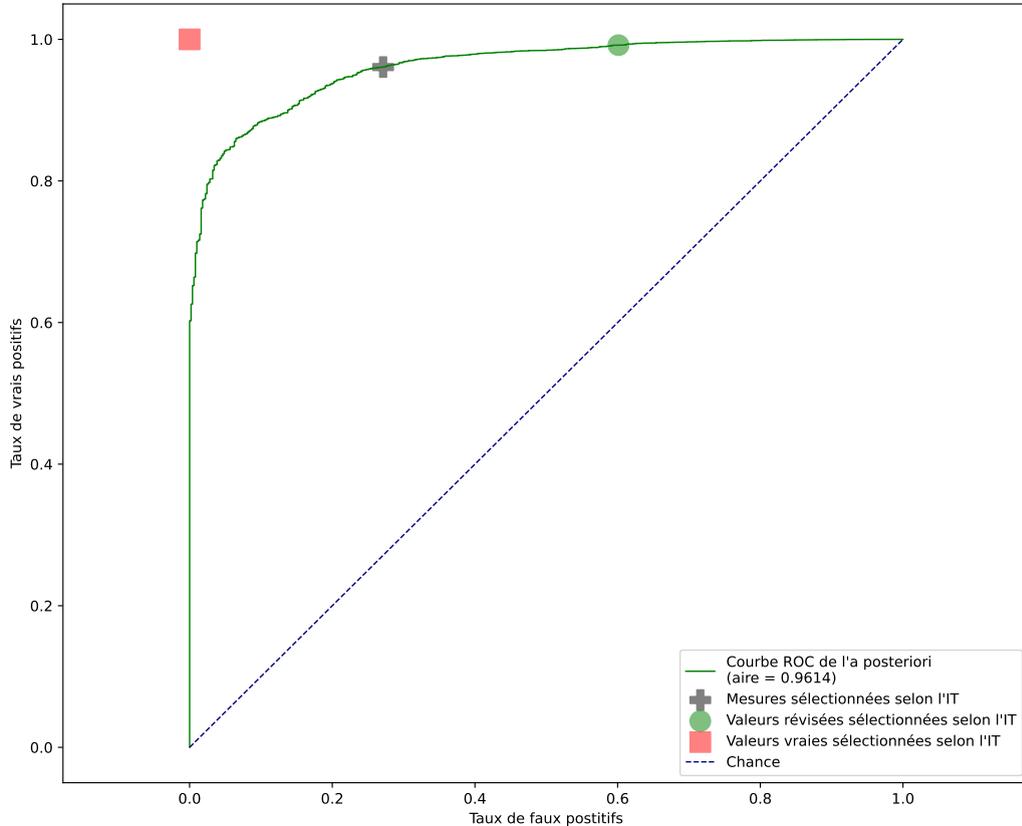


FIGURE 4.16: Courbe ROC pour un échantillon de taille 10^4 (cas gaussien).

4.5 Impact du ratio des écarts-types de la valeur vraie et de l'erreur de mesure

Cette analyse est fondée sur le cas gaussien. Comme le montre l'expression analytique de la valeur révisée dans ce cas-là (Équation 2.23), l'écart-type de la valeur vraie σ_T et l'écart-type de l'erreur de mesure σ_E sont deux paramètres prépondérants. Plus précisément, ce ne sont pas les valeurs individuelles de chacun qui importent mais leurs valeurs relatives. Cette étude explicite l'impact du rapport entre ces deux paramètres sur l'amélioration des mesures.

L'exemple est construit suivant le paramétrage de l'exemple gaussien présenté en Section 3.1. Les méthodes probabilistes relatives à une modélisation gaussienne sont appliquées pour obtenir les valeurs révisées. L'opération est répétée pour différentes valeurs de σ_E , allant de deux fois celui de σ_T à une valeur négligeable.

Les courbes des scores R^2 des mesures et des valeurs révisées sont tracées en Figure 4.17. L'axe des abscisses correspond au ratio $\frac{\sigma_E}{\sigma_T}$. La valeur remarquable $\sigma_E = \sigma_T$ est tracée en bleu pointillé. La ligne horizontale correspondant à un score R^2 nul est tracée en rouge.

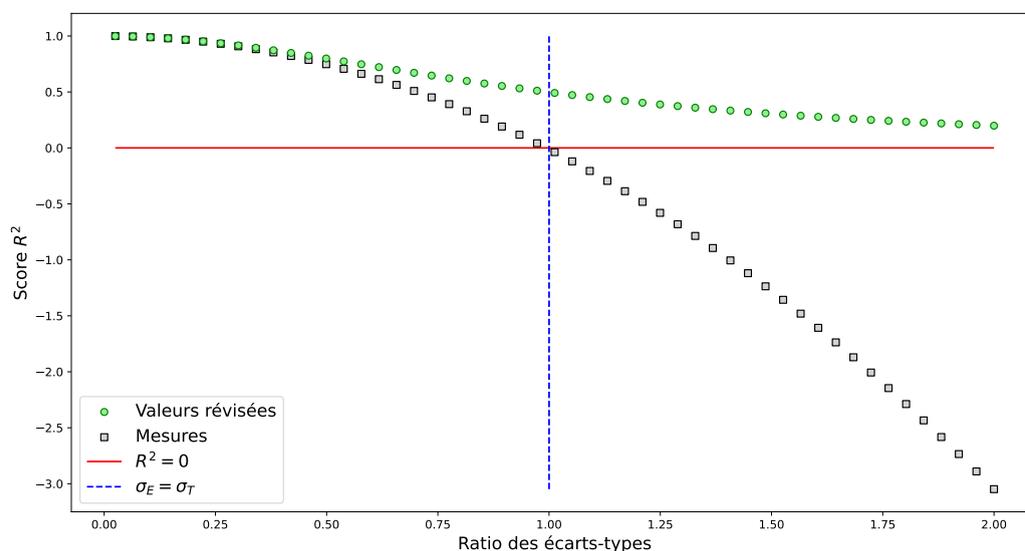


FIGURE 4.17: Évolution du score R^2 en fonction du rapport des écarts-types (application au cas gaussien).

Le graphique montre que :

- Les deux courbes de score R^2 obtenues sont monotones et décroissantes. Cela traduit que la révision ne corrige qu'une proportion de l'erreur de mesure.
- Les scores convergent asymptotiquement vers 1 à mesure que σ_E devient négligeable devant σ_T (atteindre 1 signifierait une correction parfaite, ce qui est impossible pour ces méthodes). Donc plus la mesure est de bonne qualité, moins le gain de la révision est significatif.
- Le score des valeurs révisées est toujours supérieur à celui des mesures. La révision a donc toujours un effet bénéfique pour les mesures, quelles que soient les valeurs des écarts-types en jeu.
- La courbe propre aux valeurs révisées présente un point d'inflexion : elle converge asymptotiquement vers 1 lorsque le ratio tend vers 0, et converge vers 0 lorsque le ratio tend vers $+\infty$. L'asymptote en 0 est la conséquence d'une valeur révisée qui est comprise entre la mesure et la moyenne de la valeur vraie. Or le score d'une prédiction où toutes les valeurs révisées sont égales à la moyenne de la valeur vraie est 0 (Section 2.2.3). Ainsi, plus l'écart-type de l'erreur de mesure est grand devant celui de la valeur vraie, plus l'apport de la révision est significatif.
- La courbe propre aux mesures coupe l'axe des abscisses lorsque $\sigma_E = \sigma_T$. Lorsque la dispersion induite par l'erreur de mesure est comparable à celle de la valeur vraie, alors l'étape de métrologie n'apporte rien à la qualité des mesures ; que cela revient au même d'ignorer la dispersion de la valeur vraie et de ne s'en tenir qu'à sa moyenne (dont le score R^2 est zéro).

Conclusion

Le Chapitre 4 a présenté plusieurs applications pour donner un aperçu de la sensibilité des méthodes de déconvolution et d'inférence bayésienne.

Les trois exemples des Sections 4.1, 4.2 et 4.3 concernent des données d'entrées imparfaites : dans le premier, la loi de la valeur vraie n'est pas identifiable. La loi déconvoluée obtenue est certes moins précise, mais la lecture de la production sont tout de même améliorés. La méthode de déconvolution utilisée dans le cas gaussien est d'ailleurs insensible au tracé de la loi de l'erreur de mesure, tant que sa moyenne et son écart-type sont correctement identifiés. La propagation de l'imprécision de la loi déconvoluée dégrade donc la déclaration de conformité, mais les valeurs révisées restent plus fiables que les mesures. Les deux autres exemples présentent des résultats similaires : l'introduction d'une loi de l'erreur de mesure mal-identifiée ou de mesures corrélées dégrade les performances obtenues dans le cadre idéal, mais restent en faveur de l'application des méthodes d'amélioration dans la majorité des cas. Les valeurs révisées sont meilleures, ou bien d'une qualité similaire aux mesures dans le pire des cas.

Les Sections 4.4 et 4.5 traitent de la variation de paramètres de simulation. Les paramètres les plus influents sont la taille de l'échantillon et les valeurs des écarts-types de la valeur vraie et de l'erreur de mesure. La taille de l'échantillon impact directement la qualité des résultats : les résultats obtenus sont meilleurs et les outils de gestion de la production d'autant plus précis que la taille est grande. L'étude de différentes valeurs du ratio entre les écarts-types a permis d'explicitier des tendances, notamment que la prédiction des valeurs révisées est toujours meilleure que celle des mesures.

Conclusions & Perspectives

Conclusions

Ces travaux ont pour objectif d'exploiter la connaissance de la loi de l'erreur de mesure pour améliorer la qualité des mesures obtenues en métrologie. L'amélioration des mesures consiste à réduire la part d'erreur dans la mesure. Elle doit mener à une meilleure lecture de la conformité de chacune des mesures, et par extension de toute la production.

Ces travaux s'inscrivent dans le projet "We Are Metrology", qui réunit un consortium composé d'industriels et de laboratoires de recherche autour de cette problématique. En étant la première étape scientifique du projet, ils initient la preuve de concept. Les solutions proposées ont pour but d'être implémentées sur une plate-forme connectée, ce qui demande un caractère général, robuste et autonome. Le contexte envisagé est celui de l'industrie manufacturière où les pièces sont produites en grand nombre. La conformité d'une pièce est associée à une unique mesure, et les mesures traitées sont univariées.

Le besoin métrologique soulevé par ce projet est traité en trois temps dans ces travaux de thèse :

- ◇ l'application de l'approche probabiliste en métrologie ;
- ◇ la présentation des méthodes pour la correction de l'erreur de mesure ;
- ◇ l'application et l'évaluation des méthodes probabilistes.

Le premier chapitre expose les définitions de la mesure, de la valeur vraie et de l'erreur de mesure, en faisant appel aux organismes qui font le standard en métrologie. Les trois termes correspondent aux trois composantes de l'équation donnée dans le GUM pour modéliser la mesure, utilisée telle quelle dans ces travaux. La mesure - qui désigne le résultat de mesure dans ces travaux, est plus qu'une simple valeur numérique mais comprend toutes les informations relatives à son environnement (dont les frontières sont encore floues). Bien qu'intuitive, la valeur vraie désigne une quantité qu'il est compliqué de cerner. L'erreur de mesure désigne simplement l'écart entre la mesure et la valeur vraie, mais sa définition est insuffisamment précise pour guider l'identification et la modélisation des interactions de toutes ses composantes. Les imprécisions qui entourent les définitions de la mesure, de la valeur vraie et de l'erreur de mesure posent une réelle difficulté pour leur identification.

Par la multitude des facteurs qui la composent, l'erreur de mesure a un caractère aléatoire. Parce qu'elle ne sera jamais identifiable avec exactitude, elle introduit une incertitude sur la mesure. En réponse au besoin de maîtriser la qualité de la production, l'approche standard consiste à remplacer la mesure par un "intervalle de recouvrement". La construction de cet intervalle est réalisée dans le but d'encadrer la position de la valeur vraie. Elle est centrée sur la mesure et s'appuie

sur des indicateurs statistiques de l'erreur de mesure (généralement les *extrema* ou l'écart-type). Or, la construction devrait être centrée sur la valeur vraie : une impossibilité pratique qui force à utiliser la mesure. La conséquence est que l'intervalle de recouvrement s'étend sur des données qu'il est inutile de considérer, et qu'il omet des données qu'il devrait recouvrir. Cette erreur, ainsi que les incertitudes autour des données déterministes utilisées, nourrissent un scepticisme qui pousse à la construction d'un intervalle de recouvrement "étendu". Cette extension est réalisée via l'application de coefficients fixés arbitrairement. L'approche standard présente une erreur dans sa méthode, et n'est viable pour la gestion de la production que dans le cadre d'une politique rigoureusement conservatrice qui peut devenir coûteuse. L'approche probabiliste est proposée comme solution alternative pour la gestion des incertitudes. L'utilisation de variables aléatoires permet d'intégrer le caractère aléatoire de la mesure, de la valeur vraie et de l'erreur de mesure. Les densités de probabilité permettent de modéliser chacune des composantes plus finement que les indicateurs statistiques. Le cadre probabiliste est idéal pour l'exploitation de la connaissance de la loi de l'erreur de mesure (connue dans ces travaux). Enfin, il ouvre la porte à d'autres méthodes pour gérer la production.

Ce premier chapitre se termine par la présentation des solutions probabilistes face à l'incertitude introduite par l'erreur de mesure. Une première partie cible l'amélioration de la lecture statistique de l'échantillon de mesures. Parmi les différentes approches possibles, la déconvolution est la plus à même d'utiliser la connaissance de la loi de l'erreur de mesure. De très nombreuses méthodes existent dans la littérature. Elles sont fortement liées à la nature des données et donc des domaines d'application, qui sont principalement les traitements de signaux, d'images ou de série temporelles. Elles dépendent toutes fortement de la modélisation (et rares sont celles qui ressemblent à celle du GUM), et ce n'est pas un problème d'erreur dans les variables (*error-in-variable*). Enfin, les solutions potentielles doivent encore être généralisables, pour répondre à une implémentation autonome. La méthode qui a été jugée la plus intéressante est celle de l'estimateur de densité à noyau de déconvolution. Pour le reste, il y a un besoin de développer des méthodes dédiées à la métrologie qui répondent à cette problématique précise. La seconde partie se focalise sur l'amélioration de chaque mesure prise individuellement. Il existe peu de solutions proches de la métrologie permettant d'intégrer des connaissances à la mesure. L'inférence bayésienne a été retenue. C'est une méthode dont l'usage se répand très rapidement aujourd'hui ; bien qu'appréciée et multi-disciplinaire, elle est néanmoins critiquée parce qu'elle ouvre la porte au subjectivisme (donc à de potentiels abus) et sort du cadre fréquentiste. Cette discussion est largement documentée par ailleurs dans la littérature. L'application de l'inférence bayésienne permettra dans ces travaux d'intégrer la connaissance de la loi de l'erreur ainsi que la loi de la production à chaque mesure pour estimer la valeur vraie.

Le deuxième chapitre s'articule en trois parties. La première partie propose des méthodes de déconvolution, qui sont au nombre de trois. Leur objectif commun est d'exploiter la loi de l'erreur de mesure pour améliorer la lecture statistique de l'échantillon de mesures. Toutes requièrent que la taille de l'échantillon soit suffisamment grande pour refléter la loi des mesures obtenues en métrologie. Elles reposent aussi toutes sur l'hypothèse d'indépendance entre la valeur vraie et l'erreur de mesure, qui est raisonnable dans le contexte de la métrologie. Leurs différences relèvent de la modélisation des composantes de la mesure.

- ◇ Dans la première méthode, les variables aléatoires de la mesure et de l'erreur de mesure suivent des lois normales. La motivation de cette modélisation est appuyée par la littérature,

et répond notamment à une loi de l'erreur de mesure définie comme un agrégat d'une multitude de causes. La résolution met en jeu les fonctions caractéristiques de ces deux variables et permet d'identifier la loi de la valeur vraie : sa densité de probabilité est gaussienne et ses paramètres (moyenne et écart-types) sont connus. L'approche de cette méthode est analytique et donc parfaitement robuste.

- ◇ Dans la deuxième méthode, il est supposé que la loi de la valeur vraie appartient à une famille de lois usuelles et paramétrées. Les lois qui constituent les familles mises en œuvre dans ces travaux sont les lois normale, log-normale, de Weibull, triangulaire, arcsinus et uniforme. Le nombre de famille n'est pas limité, et pourrait notamment inclure les lois Beta et Gamma qui présentent des géométries très variées. La contrainte de modélisation qui repose ici uniquement sur la valeur vraie est donc en réalité assez souple. La résolution de la déconvolution s'appuie sur l'équation de la convolution, parce que le calcul des densités des mesures est plus simple que celles des valeurs vraies. L'objectif est de reconstruire la loi des mesures, une fois la loi de la valeur vraie arbitrairement définie. Les densités de référence sont lues sur l'échantillon de mesures. L'une après l'autre, les familles sont envisagées et optimisées pour avoir un jeu de paramètres qui reconstruis au mieux les densités de probabilité des mesures observées. La qualité de la reconstruction est évaluée par le score de vraisemblance. Le score BIC est finalement utilisé pour discerner la meilleure solution offerte par les familles candidates. La qualité de la solution dépend en partie de la qualité de l'optimisation, qui est sensible à la paramétrisation.
- ◇ Dans la troisième méthode, aucune hypothèse de modélisation n'est requise concernant les variables en jeu. Cette méthode permet la reconstruction de lois complexes, telles que les lois multimodales. La déconvolution consiste à reconstruire la loi de la valeur vraie, point par point. L'enjeu est donc d'identifier la valeur de la densité de probabilité localement. La solution se fonde aussi sur le calcul de l'équation de la convolution. Les densités des mesures sont lues sur l'échantillon de mesures. La loi de la valeur vraie est substituée par une loi arbitrairement définie et dont le nombre de paramètres fixe le nombre d'inconnues de l'équation de la convolution. L'approximation locale de la loi de la valeur vraie se réduit ainsi à la résolution d'un système, qui nécessite autant d'évaluations locales de la loi des mesures que de paramètres dans la loi de substitution. La qualité des résultats dépend particulièrement de la qualité du lissage à noyau, qui peut affecter la stabilité du tracé de la loi déconvoluée.

En pratique, si les lois des mesures et de l'erreur de mesure suivent des lois normales, alors la première méthode de déconvolution est privilégiée. Sinon, la troisième méthode de déconvolution qui est la plus générale est appliquée. Si la fonction déconvoluée s'apparente à une famille de loi paramétrée usuelle, alors la deuxième méthode pourra s'appliquer ensuite pour raffiner les résultats obtenus. Ces méthodes de déconvolution transforment la loi des mesures en une loi partiellement épurée de l'erreur de mesure. Cette nouvelle loi, dite "déconvoluée", constitue une bien meilleure estimation de la loi de la valeur vraie. Cette amélioration est mesurable via la probabilité de produire hors-tolérance.

La deuxième partie de ce chapitre concerne l'application de l'inférence bayésienne dans la cadre de la métrologie. L'objectif est d'améliorer chaque mesure de l'échantillon. D'une part, l'inférence bayésienne intègre la loi de l'erreur de mesure. Cette connaissance est requise pour construire la loi de vraisemblance. Les lois de vraisemblances de deux mesures distinctes ne varient que par leurs positions (leurs moyennes). D'autre part, elle intègre la loi qui caractérise la production comme *a priori*. Cette connaissance peut être donnée arbitrairement, si elle est le fruit d'une expertise. Autrement - et comme appliqué dans ces travaux, la loi déconvoluée y répond avec pertinence.

La connaissance *a priori* est la même quelle que soit la mesure traitée. Pour une mesure donnée, l'inférence bayésienne synthétise les lois de la vraisemblance et *a priori* pour construire la loi *a posteriori*. Cette loi caractérise la position de la valeur vraie et remplit pleinement les objectifs de l'intervalle de recouvrement. La densité la plus élevée de l'*a posteriori* est celle de la valeur qui a le plus probablement été mesurée. Elle est l'estimation de la valeur vraie, et est nommée la "valeur révisée". Lorsque la mesure et l'erreur de mesure suivent des lois normales, l'*a posteriori* suit une loi normale aussi ; l'expression de la valeur révisée est connue analytiquement et permet de montrer quelques propriétés ; enfin, il est possible de construire un intervalle de tolérance équivalent. Le gain apporté par la valeur révisée est mesuré avec le coefficient de détermination, les résidus des prédictions, l'erreur quadratique moyenne et le pourcentage de mesures améliorées.

La dernière partie de ce chapitre est dédiée à la présentation d'outils pour la gestion de production. Les outils présentés, existants ou développés, ont pour but de tirer bénéfice des connaissances obtenues par l'amélioration des mesures. Puisqu'ils relèvent de la conformité, ils nécessitent l'application d'un intervalle de tolérance.

- ◇ Suite à l'application de la déconvolution, il est possible de tracer les courbes de probabilité de prendre une mauvaise décision. Pour chaque mesure, il est possible de connaître la probabilité de la mal-classifier. Inversement, il est possible de fixer la probabilité maximale de se tromper, et de lire l'ensemble des mesures acceptables. Les connaissances des lois de la valeur vraie et de l'erreur de mesure permettent donc de connaître la sélection des mesures qui vérifient la probabilité maximale souhaitée. Les courbes de probabilité de prendre une mauvaise décision peuvent être pondérées par les coûts client et fournisseur, et l'on parle alors de risques. Ces travaux s'intéressent particulièrement à la sélection "optimale", celle qui minimise le risque total (cumulé pour toutes les mesures).
- ◇ Suite à l'application de l'inférence bayésienne, les matrices de confusion permettent de comparer les qualités des déclarations de conformités s'appuyant sur les mesures ou sur les valeurs révisées. L'usage des matrices de confusion se borne cependant aux cas simulés, ceux pour lesquels la valeur vraie est connue. La valeur révisée n'est qu'une possibilité pour exploiter la loi *a posteriori*. La sélection d'une mesure peut aussi se fonder sur sa probabilité d'être hors-tolérance. La fonction d'efficacité du récepteur (communément appelée courbe ROC) expose les performances quant à la déclaration de conformité de tous les seuils de probabilité possibles (de zéro à cent pour cent).

Ces travaux se terminent sur la présentation de cas d'applications des méthodes d'amélioration. La déconvolution et l'inférence bayésienne sont appliquées à des cas paramétrés pour imiter la réalité industrielle. Les données utilisées sont simulées pour connaître les valeurs vraies. Les trois premiers cas sont définis pour tester chaque méthode de déconvolution. Tous les indicateurs statistiques ainsi que les méthodes de gestion de production exposés dans le deuxième chapitre sont en faveur de l'application de cette méthode d'amélioration. Pour tous les cas (avec les critères explicités entre parenthèses) :

- ◇ la loi déconvoluée est une meilleure estimation de la loi de la valeur vraie (probabilité de produire hors-tolérance, valeurs de la moyenne et de l'écart type, tracé de la courbe) ;
- ◇ la loi déconvoluée améliore la gestion de la production (risque total de mal-classifier) ;
- ◇ les valeurs révisées sont une meilleure estimation des valeurs vraies (erreur quadratique moyenne, coefficient de détermination, pourcentage de mesures améliorées) ;

- ◇ les valeurs révisées améliorent la lecture de la conformité (matrice de confusion) est plus proche de la vérité ;
- ◇ les valeurs révisées améliorent la gestion de la production (risque total).

Les derniers cas présentés testent la robustesse des méthodes : les résultats, évidemment ternis lorsque les données d'entrée sont inexactes, restent favorables. Au final, dans tous les cas testés, l'application des méthodes d'amélioration ont toujours eu un effet bénéfique. Il est possible de conclure que l'application des méthodes de déconvolution développées et l'application de l'inférence bayésienne prenant pour *a priori* la loi déconvoluée améliorent les mesures obtenues en métrologie et augmentent la maîtrise de la production. Le point majeur soulevé par ces simulations est que les méthodes bénéficient d'échantillons de mesures de grandes tailles. Avec une taille standard fixée à mille dans ces travaux, il faut ainsi envisager en pratique le déploiement de méthodes sur des lignes de production où la prise d'information serait automatisée (par exemple avec des capteurs connectés).

Perspectives

Ces travaux avaient pour objectif de fournir une preuve de concept, pour l'amélioration des mesures par la connaissance de la loi de l'erreur de mesure. Les résultats obtenus sont nettement positifs et permettent d'envisager de plus amples recherches sur cette problématique.

- ◇ Les solutions développées ont été testées et sont prometteuses. Ces travaux n'ont cependant pas la prétention d'avoir pleinement validé leur potentiel. Elles bénéficieraient d'être plus fortement validées du point de vue théorique, avec des démonstrations exhaustives des gains apportés et de toutes les propriétés observées ; mais aussi du point de vue expérimental, par la mise au place d'un plan d'expérience rigoureux qui couvrirait les hypothèses de modélisation et l'étude de la sensibilité à la qualité des données et de la modélisation, et à tous les paramètres de la simulation. Enfin, l'application à un cas réel, et donc à des données réelles sera aussi importante pour valider ces méthodes et leur donner du crédit aux acteurs de l'industrie.
- ◇ Les méthodes présentées dans ces travaux sont appliquées à des mesures uni-variées. Le déploiement de ces méthodes dans l'industrie soulèvera inévitablement la question de données multi-variées : il est fréquent en métrologie que la valeur recherchée soit la combinaison de plusieurs mesures, comme une surface ou un volume. Plus largement, il sera intéressant d'étudier comment cette approche, pensée pour la cotation dimensionnelle, pourrait s'étendre au tolérancement géométrique.
- ◇ Il a été montré que les matrices de confusion pouvaient être représentées dans un graphique cartésien classique (Section 2.3.1). Dans ce graphique, les frontières de la non-conformité sont linéaires : elles se prêtent facilement à l'application de méthodes fiabilistes de simulation pour calculer des probabilités de défaillance - ici, d'être hors-tolérance. La méthode de Monte-Carlo est probablement la méthode la connue d'entre toutes. De plus, les probabilités de produire hors-tolérance dans l'industrie sont très faibles (les pièces non-conformes se comptent en pièces-par-million), et il existe de nombreuses méthodes fiabilistes dédiées au calcul de très faibles probabilités. La méthode FORM-système semble particulièrement adéquate pour exploiter la linéarité de plusieurs frontières. On remarquera que les méthodes

fiabilistes peuvent traiter des problèmes de grandes dimensions, ce qui apparaîtra sûrement avec l'introduction de mesures multi-variées.

- ◇ Bien que favorable pour la majorité, l'application de l'inférence de bayésienne détériore certaines mesures. Il sera intéressant de chercher si, du point de vue statistique, il serait correct de n'appliquer la révision qu'à certaines mesures. Le résultat de la révision serait alors un échantillon hybride, composé à la fois de mesures et de valeurs révisées, qui maximiserait donc la qualité de l'estimation des valeurs vraies. La dernière étude présentée en Annexe B.2 montre qu'il est facile de cerner ces populations dans le cas gaussien. Il s'agirait alors d'identifier les différentes frontières, éventuellement d'élargir les conclusions aux cas non-gaussiens, pour cibler les mesures qu'il faut réviser ou non.
- ◇ Les méthodes proposées dans ces travaux améliorent la qualité des mesures par la connaissance de la loi de l'erreur de mesure. Cette connaissance est ici admise. Elle requière d'identifier toutes les composantes de l'erreur de mesure et de modéliser physiquement leurs impacts. La loi est enfin construite par simulation. Par régression, décomposition de variance ou encore avec des polynômes du chaos, il existe de nombreux outils pour réaliser une analyse de la sensibilité. Cette analyse permettra de cibler les causes d'erreur de mesure qui ont les plus grandes conséquences sur les résultats et donc d'orienter avec pertinence les efforts et les investissements à faire en métrologie. L'analyse de sensibilité contribuera ainsi à améliorer la gestion et à terme la qualité de la production.
- ◇ Il existe aujourd'hui d'autres de façon de traiter les incertitudes. La logique floue, de son nom anglais "fuzzy logic" est un domaine de recherche très actif [98]. Des méthodes d'arithmétiques floues permettent aussi de manipuler des incertitudes, de les combiner et de dresser des prédictions. Ce domaine est tout à fait pertinent pour la métrologie [21], et l'on peut trouver une application pour le traitement d'image dans les travaux de Caufield [22]. La logique floue est principalement fondée sur la notion d'intervalle, une notion certes familière aux métrologues mais inappropriée pour ces travaux qui visent à exploiter toute l'information d'une densité de probabilité (que l'on ne souhaite pas réduire à deux bornes). On rappelle d'ailleurs que la connaissance de la loi de l'erreur de mesure est d'ailleurs une exigence forte : d'autres travaux permettent d'exploiter une information partielle sur l'erreur de mesure [46, 60, 85].

Bilan de publications et communications scientifiques

Publications internationales

- M. Gille, P. Beaurepaire, A. Dumas, N.Gayton and T. Yalamas. Statistical approaches for the reduction of measurement errors in metrology. Article accepté par le journal : "ASME - Journal of Applied Mechanics".

Conférences internationales

- M. Gille, P. Beaurepaire, A. Dumas, N.Gayton and T. Yalamas. A probabilistic approach for the consideration of measurement errors in metrology. 31^{ème} European Safety and Reliability Conference, 19-23 septembre 2021, Angers, France.*
- M. Gille, P. Beaurepaire, F. Taghon, A. Dumas, N.Gayton and T. Yalamas. Integration of the analysis of the error of geometric dimensions modeled with a probabilistic approach. 18^{ème} IPW Conference, 12-14 mai, 2021, Guimaraes, Portugal.*
- M. Gille, P. Beaurepaire, A. Dumas, F. Taghon, N.Gayton and T. Yalamas. On the identification of the probability density function of a dimension from data with measurement errors. 18^{ème} 17th CIRP Conference on Computer Aided Tolerancing, 15-17 Juin, 2022, Metz, France.*

Conférence nationales

- M. Gille, P. Beaurepaire, F. Taghon, A. Dumas, N.Gayton and T. Yalamas. Integration of the analysis of the error of geometric dimensions modeled with a probabilistic approach. 22^{ème} Congrès LambdaMu 2020, octobre 2020 - 19 janvier 2021, Le Havre, France.*

* : *Participation à distance*. Les publications, conférences scientifiques nationales et internationales relatives à ces travaux de recherche n'ont majoritairement pas pu se tenir ordinairement, mais "à distance", suite aux restrictions gouvernementales pour cause de Covid19.

Annexe A

Etudes analytiques

A.1 Expression analytique de la distribution de la valeur vraie dans le cas gaussien

Dans le cas où les variables aléatoires de la valeur vraie et de l'erreur de mesure sont indépendantes et suivent des lois gaussiennes, la fonction caractéristique ϕ_M des valeurs mesurées est définie telle que :

$$\begin{aligned}\phi_M(t) &= \phi_{V+E}(t) \\ &= \phi_V(t) \phi_E(t)\end{aligned}\tag{A.1}$$

Par conséquent :

$$\begin{aligned}\phi_T(t) &= \frac{\phi_M(t)}{\phi_E(t)} \\ &= \frac{e^{it\mu_M - \frac{1}{2}t^2\sigma_M^2}}{e^{it\mu_E - \frac{1}{2}t^2\sigma_E^2}} \\ &= e^{it(\mu_M - \mu_E) - \frac{1}{2}t^2(\sigma_M^2 - \sigma_E^2)}\end{aligned}\tag{A.2}$$

La fonction caractéristique pour la valeur vraie est celle d'une loi normale dont les paramètres peuvent être lu par identification :

$$\begin{cases} T \sim \mathcal{N}(\mu_T, \sigma_T) \\ \mu_T = \mu_M - \mu_E \\ \sigma_T = \sqrt{\sigma_M^2 - \sigma_E^2} \end{cases}\tag{A.3}$$

A.2 Expression de la valeur révisée dans le cas gaussien

La loi de l'*a posteriori* est donnée par la relation :

$$p(x|m) \propto \pi(x)\mathcal{L}(m|x)\tag{A.4}$$

La densité du produit de deux variables aléatoires dont les lois sont normales est décrit par une loi proportionnelle à une loi normale. Normalisée, l'expression de la densité du produit est :

$$p(x | m) \sim \mathcal{N}\left(\frac{\mu_\pi\sigma_{\mathcal{L}}^2 + \mu_{\mathcal{L}}\sigma_\pi^2}{\sigma_\pi^2 + \sigma_{\mathcal{L}}^2}, \sqrt{\frac{\sigma_\pi^2\sigma_{\mathcal{L}}^2}{\sigma_\pi^2 + \sigma_{\mathcal{L}}^2}}\right)\tag{A.5}$$

Dans le cas idéal, la loi de l'*a priori* (approximée lors de la déconvolution) est égale celle de la valeur vraie :

$$\pi(x) = f_T(x)\tag{A.6}$$

La loi de la vraisemblance est définie par :

$$\mathcal{L}(m|x) = f_E(m - x)\tag{A.7}$$

Dans le cas précis où les lois de la valeur vraie et de l'erreur de mesure sont gaussiennes on a :

$$\begin{aligned} T &\sim \mathcal{N}(\mu_T, \sigma_T) \\ E &\sim \mathcal{N}(\mu_E, \sigma_E) \end{aligned} \quad (\text{A.8})$$

Par conséquent :

$$\begin{aligned} \mathcal{L} &\sim \mathcal{N}(\mu_{\mathcal{L}}, \sigma_{\mathcal{L}}) \\ \left\{ \begin{array}{l} \mu_{\mathcal{L}} = m - \mu_E \\ \sigma_{\mathcal{L}} = \sigma_E \end{array} \right. \end{aligned} \quad (\text{A.9})$$

On a donc :

$$p(x | m) \sim \mathcal{N}\left(\frac{\mu_T \sigma_E^2 + (m - \mu_E) \sigma_T^2}{\sigma_T^2 + \sigma_E^2}, \sqrt{\frac{\sigma_T^2 \sigma_E^2}{\sigma_T^2 + \sigma_E^2}}\right) \quad (\text{A.10})$$

En faisant apparaître la relation linéaire entre la moyenne de l'*a posteriori* et la mesure on a :

$$p(x|m) \sim \mathcal{N}\left(m \left(\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}\right) + \left(\frac{\mu_T \sigma_E^2 - \mu_E \sigma_T^2}{\sigma_T^2 + \sigma_E^2}\right), \sqrt{\frac{\sigma_T^2 \sigma_E^2}{\sigma_T^2 + \sigma_E^2}}\right) \quad (\text{A.11})$$

A.3 Extension symétrique de l'intervalle de sélection et corrélation négative entre FN et FP

Expression du problème

Soit IT l'intervalle de tolérance initialement fixé pour la sélection des mesures. Soit IT^δ l'intervalle de tolérance arbitrairement fixé pour la sélection des mesures, avec δ un nombre réel. L'application de l'intervalle IT^δ n'affecte que la prédiction de la mesure mais ne change en rien la réalité de sa conformité. Il est défini tel que :

$$IT^\delta = [x_l - \delta; x_u + \delta] \quad (\text{A.12})$$

avec x_l et x_u les bornes inférieure et supérieure de l'intervalle de tolérance, et δ le paramètre d'ajustement de la sélection. La sélection est agrandie lorsque $\delta > 0$ et réduite lorsque $\delta < 0$.

Implication pour FN

L'expression de FN avec l'usage de l'intervalle IT^δ est :

$$FN^\delta = \int_{x_l}^{x_u} f_T(t) \left(\int_{-\infty}^{x_l - \delta - t} f_E(e) de + \int_{x_u + \delta - t}^{+\infty} f_E(e) de \right) dt \quad (\text{A.13})$$

$$\begin{aligned} FN^\delta = \int_{x_l}^{x_u} f_T(t) &\left(\int_{-\infty}^{x_l - t} f_E(e) de - \int_{x_l - \delta - t}^{x_l - t} f_E(e) de \right. \\ &\left. + \int_{x_u - t}^{+\infty} f_E(e) de - \int_{x_u - t}^{x_u + \delta - t} f_E(e) de \right) dt \end{aligned} \quad (\text{A.14})$$

$$FN^\delta = FN - \int_{x_l}^{x_u} f_T \left(\int_{x_l - \delta - t}^{x_l - t} f_E(e) de + \int_{x_u - t}^{x_u + \delta - t} f_E(e) de \right) dt \quad (\text{A.15})$$

$$FN^\delta = FN + C_{FN}(\delta), \quad \left\{ \begin{array}{l} C_{FN}(\delta) < 0 \text{ si } \delta > 0 \\ C_{FN}(\delta) > 0 \text{ si } \delta < 0 \end{array} \right. \quad (\text{A.16})$$

Implication analogue pour FP

$$FP = \int_{-\infty}^{x_l} f_T(t) \left(\int_{x_l-t}^{x_u-t} f_E(e) de \right) dt + \int_{x_u}^{+\infty} f_T(t) \left(\int_{x_l-t}^{x_u-t} f_E(e) de \right) dt \quad (\text{A.17})$$

$$FP^\delta = \int_{-\infty}^{x_l} f_T(t) \left(\int_{x_l-\delta-t}^{x_u+\delta-t} f_E(e) de \right) dt + \int_{x_u}^{+\infty} f_T(t) \left(\int_{x_l-\delta-t}^{x_u+\delta-t} f_E(e) de \right) dt \quad (\text{A.18})$$

$$\begin{aligned} FP^\delta = FP &+ \int_{-\infty}^{x_l} f_T(t) \left(\int_{x_l-\delta-t}^{x_l-t} f_E(e) de + \int_{x_u-t}^{x_u+\delta-t} f_E(e) de \right) dt \\ &+ \int_{x_u}^{+\infty} f_T(t) \left(\int_{x_l-t}^{x_l-\delta-t} f_E(e) de + \int_{x_u-t}^{x_u+\delta-t} f_E(e) de \right) dt \end{aligned} \quad (\text{A.19})$$

$$FP^\delta = FP + C_{FP}(\delta), \quad \begin{cases} C_{FP}(\delta) < 0 \text{ si } \delta < 0 \\ C_{FP}(\delta) > 0 \text{ si } \delta > 0 \end{cases} \quad (\text{A.20})$$

Synthèse des résultats

On obtient au final que les conditions nécessaires pour réduire un coût (FN ou FP) augmente l'autre (respectivement) :

$$\begin{cases} \text{if } \delta > 0, & FN^\delta > FN \text{ si } FP^\delta < FP \\ \text{if } \delta < 0, & FN^\delta < FN \text{ si } FP^\delta > FP \end{cases} \quad (\text{A.21})$$

Il n'est donc pas possible de réduire FN et FP par une extension symétrique de l'intervalle de sélection.

A.4 L'inférence bayésienne dans le cas gaussien est une extension de l'intervalle de tolérance

L'inférence bayésienne transforme linéairement les mesures en valeurs révisées dans le cas gaussien. En appliquant l'inverse de cette transformation à l'intervalle de tolérance on obtient l'intervalle de tolérance équivalent qui peut être directement confronté aux mesures. On a :

$$\begin{aligned}
& r \in \text{IT} \\
\Leftrightarrow & r \in [\text{IT}_{inf}, \text{IT}_{sup}] \\
\Leftrightarrow & m \times a_1 + a_0 \in [\text{IT}_{inf}, \text{IT}_{sup}] \\
\Leftrightarrow & m \in \left[\frac{\text{IT}_{inf} - a_0}{a_1}, \frac{\text{IT}_{max} - a_0}{a_1} \right] \\
\Leftrightarrow & m \in [\text{IT}_{EQ_{inf}}, \text{IT}_{EQ_{max}}] \tag{A.22} \\
\Leftrightarrow & m \in \text{IT}_{EQ} \\
& \begin{cases} a_1 = \left(\frac{\sigma_T^2}{\sigma_T^2 + \sigma_\varepsilon^2} \right) \\ a_0 = \left(\frac{\mu_T \sigma_\varepsilon^2 - \mu_\varepsilon \sigma_T^2}{\sigma_T^2 + \sigma_\varepsilon^2} \right) \end{cases}
\end{aligned}$$

L'élargissement provoqué par l'inférence bayésienne apparaît en confrontant l'étendue de chaque intervalle :

$$\begin{aligned}
\text{IT}_{EQ_{max}} - \text{IT}_{EQ_{inf}} &= \frac{\text{IT}_{max} - a_0}{a_1} - \frac{\text{IT}_{inf} - a_0}{a_1} \\
&= \frac{1}{a_1} (\text{IT}_{max} - \text{IT}_{min}) \\
\begin{cases} \sigma_T^2 > 0 \\ \sigma_\varepsilon^2 > 0 \end{cases} &\Rightarrow a_1 < 1 \\
&\Rightarrow \frac{1}{a_1} > 1 \\
&\Rightarrow \frac{1}{a_1} (\text{IT}_{max} - \text{IT}_{min}) > \text{IT}_{max} - \text{IT}_{min} \\
&\Rightarrow \text{IT}_{EQ_{max}} - \text{IT}_{EQ_{inf}} > \text{IT}_{max} - \text{IT}_{min} \\
&\Rightarrow \text{IT}_{EQ} > \text{IT}
\end{aligned} \tag{A.23}$$

A.5 La valeur vraie est toujours comprise entre la moyenne de la valeur vraie et la mesure quand la moyenne de l'erreur est nulle

Affirmer que la valeur vraie est toujours comprise entre la moyenne de la valeur vraie et la mesure signifie que :

$$r^{(i)} \in [\mu_T, m^{(i)}], \forall i \in [1, N] \quad (\text{A.24})$$

Cette propriété peut aussi être traduite par :

$$\begin{aligned} r^{(i)} \in [\mu_T, m^{(i)}] &\Leftrightarrow \begin{cases} m^{(i)} - r^{(i)} > 0 &\Rightarrow \mu_T - r^{(i)} < 0 \\ m^{(i)} - r^{(i)} < 0 &\Rightarrow \mu_T - r^{(i)} > 0 \end{cases} \\ &\Leftrightarrow m^{(i)} - r^{(i)} \text{ et } \mu_T - r^{(i)} \text{ sont de signes opposés} \end{aligned} \quad (\text{A.25})$$

$$\mu_E = 0 \Rightarrow r^{(i)} = m^{(i)} \left(\frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \right) + \mu_T \left(\frac{\sigma_E^2}{\sigma_E^2 + \sigma_T^2} \right) \quad (\text{A.26})$$

$$\begin{aligned} \mu_E = 0 &\Rightarrow m^{(i)} - r^{(i)} = m^{(i)} - m^{(i)} \left(\frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \right) - \mu_T \left(\frac{\sigma_E^2}{\sigma_E^2 + \sigma_T^2} \right) \\ &\Rightarrow m^{(i)} - r^{(i)} = (m^{(i)} - \mu_T) \left(\frac{\sigma_E^2}{\sigma_E^2 + \sigma_T^2} \right) \\ &\Rightarrow m^{(i)} - r^{(i)} = (m^{(i)} - \mu_T) C_m, \quad C_m > 0 \end{aligned}$$

$$\begin{aligned} \mu_E = 0 &\Rightarrow m^{(i)} - r^{(i)} = -m^{(i)} \left(\frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \right) + \mu_T - \mu_T \left(\frac{\sigma_E^2}{\sigma_E^2 + \sigma_T^2} \right) \\ &\Rightarrow \mu_T - r^{(i)} = (\mu_T - m^{(i)}) \left(\frac{\sigma_T^2}{\sigma_E^2 + \sigma_T^2} \right) \\ &\Rightarrow \mu_T - r^{(i)} = (\mu_T - m^{(i)}) C_{\mu_T}, \quad C_{\mu_T} > 0 \end{aligned} \quad (\text{A.27})$$

$$\mu_E = 0 \Rightarrow \frac{m^{(i)} - r^{(i)}}{\mu_T - r^{(i)}} = -1 \left(\frac{C_m}{C_{\mu_T}} \right), \quad \left(\frac{C_m}{C_{\mu_T}} \right) > 0$$

Ce résultat implique en particulier :

$$\begin{cases} m^{(i)} = \mu_T \\ \mu_E = 0 \end{cases} \Rightarrow r^{(i)} = m^{(i)} \quad (\text{A.28})$$

L'Equation (A.26) permet de vérifier le résultat :

$$\sigma_T = \sigma_{\mathcal{L}} \Rightarrow \frac{(\mu_T + \mu_{\mathcal{L}})}{2} \quad (\text{A.29})$$

Aussi :

$$\mu_T = \mu_{\mathcal{L}} \Rightarrow r = m \quad (\text{A.30})$$

A.6 Détails concernant la transformation de Nataf et l'identification des points d'intégration et des poids de la quadrature de Gauss-Hermite

La quadrature de Gauss-Hermite est uniquement applicable aux intégrales impliquant une loi normale standard. Leur expression générique est de la forme :

$$\int_{-\infty}^{\infty} y(x)\varphi(x) dx \tag{A.31}$$

où φ désigne la densité de probabilité de la loi normale standard et y peut être n'importe quelle fonction intégrable.

Cependant, de nombreux problèmes d'ingénierie impliquent des variables aléatoires qui ne suivent pas la loi normale standard. L'intégrale associée à de tels problèmes est de la forme :

$$\int_{-\infty}^{\infty} y(x)f_X(x) dx \tag{A.32}$$

où f_X est la densité de probabilité de la variable aléatoire. Dans ce cas, la transformation de Nataf est nécessaire pour utiliser la quadrature de Gauss-Hermite. La transformation consiste à effectuer le changement de variables suivant : $F_X(x) = \Phi(u)$. Φ et F_X désignent les fonctions de répartition de la loi normale standard et de la variable aléatoire impliquées dans l'Equation A.32. Ce changement de variable est équivalent à $f_X(x)dx = \Phi(u)du$. On obtient alors l'intégrale suivante :

$$\int_{-\infty}^{\infty} y\left(F_X^{-1}(\Phi(u))\right)\varphi(u) du \simeq \sum_{j=1}^k \omega^{(j)}y\left(F_X^{-1}\left(\Phi\left(u^{(j)}\right)\right)\right) \tag{A.33}$$

où $\omega^{(j)}$ et $u^{(j)}$ sont les poids et les points d'intégration de la quadrature de Gauss-Hermite. Les termes $F_X^{-1}\left(\Phi\left(u^{(j)}\right)\right)$ peuvent être identifiés comme les points d'intégration d'une variable aléatoire qui ne suit pas une loi normale standard (la procédure ne modifie pas les poids). La transformation de Nataf est ainsi habituellement utilisée pour l'application de la quadrature de Gauss-Hermite pour des lois autres que normales [144].

A.7 L'augmentation de la moyenne de l'erreur de mesure ne dégrade que le coefficient de détermination des mesures

On peut montrer d'une part que l'augmentation de la moyenne de l'erreur de mesure ne dégrade que le coefficient de détermination des mesures, parce que cette même moyenne de l'erreur de mesure est quasi soustraite à la construction des valeurs révisées. Ce comportement est particulièrement visible dans le cas gaussien (que l'on peut décrire analytiquement) si l'on considère que l'écart-type de l'erreur de mesure est négligeable devant celui des valeurs vraies :

$$\begin{aligned} \sigma_T \gg \sigma_E \quad \Leftrightarrow \quad r^{(i)} &\simeq m^{(i)} \left(\frac{\sigma_T^2}{\sigma_T^2} \right) + \left(\frac{-\mu_E \sigma_T^2}{\sigma_T^2} \right) \\ &\Leftrightarrow r^{(i)} \simeq m^{(i)} - \mu_E \end{aligned} \tag{A.34}$$

Ce résultat se retrouve par exemple si l'erreur de mesure ne provoque qu'un décalage systématique lors de la lecture de la mesure (si elle suit une loi de Dirac par exemple). En réutilisant ce résultat et en décomposant l'erreur de mesure $e^{(i)}$ avec la composante moyenne μ_E et celle aléatoire $s_E^{(i)}$, la différence d'impact sur le coefficient de détermination peut être démontrée :

$$e^{(i)} = \mu_E + s_E^{(i)} \quad (\text{A.35})$$

En incorporant la réexpression de l'erreur de mesure dans l'expression du score R^2 des mesures on a :

$$\begin{aligned} t^{(i)} - m^{(i)} &= t^{(i)} - (t^{(i)} + e^{(i)}) \\ &= t^{(i)} - (t^{(i)} + \mu_E + s_E^{(i)}) \\ &= -\mu_E - s_E^{(i)} \end{aligned} \quad (\text{A.36})$$

$$R_{m^{(i)}}^2 = 1 - \frac{\sum_{i=1}^n (\mu_E + s_E^{(i)})^2}{\sum_{i=1}^n (t^{(i)} - \mu_T)^2}$$

En incorporant la réexpression de l'erreur de mesure dans l'expression du score R^2 des valeurs vraies on a :

$$\begin{aligned} t^{(i)} - r^{(i)} &= t^{(i)} - (m^{(i)} - \mu_E) \\ &= t^{(i)} - (t^{(i)} + e^{(i)} - \mu_E) \\ &= t^{(i)} - (t^{(i)} + \mu_E + s_E^{(i)} - \mu_E) \\ &= s_E^{(i)} \end{aligned} \quad (\text{A.37})$$

$$R_{r^{(i)}}^2 = 1 - \frac{\sum_{i=1}^n (s_E^{(i)})^2}{\sum_{i=1}^n (t^{(i)} - \mu_T)^2}$$

La différence entre les deux scores apparaît nettement lorsque le biais devient prépondérant devant la composante aléatoire :

$$s_E^{(i)} \ll \mu_E \Rightarrow (s_E^{(i)})^2 < (\mu_E + s_E^{(i)})^2 \Rightarrow R_2^{r^{(i)}} > R_2^{m^{(i)}} \quad (\text{A.38})$$

Cependant, ce cas avantageux n'est pas le plus fréquent en pratique parce qu'une moyenne de l'erreur de mesure élevée est souvent flagrante et facilement corrigeable une fois identifiée (en la soustrayant [106]).

Annexe B

Etudes graphiques

B.1 Comportement en zone non-homogène

Les probabilités de prendre une mauvaise décision en acceptant et rejetant une mesure donnée peuvent être différentes pour une même mesure si elle appartient à la zone non-homogène. Les Figures B.1, B.2 et B.3 illustre l'effet de cette zone non-homogène à travers 8 configurations. Le paramétrage du problème est le même que celui présenté dans la Tableau 2.1. Seule la position de la borne supérieure u varie.

La Figure B.1(a) présente le cas évident où le domaine de mesure (c'est à dire toutes les mesures qu'il est possible d'obtenir en métrologie) est entièrement inclus dans l'intervalle de tolérance. Cela correspondrait à une production calibrée selon l'approche ordinaire, de sorte à ce qu'aucun risque ne soit pris (ce qui présente évidemment un surcoût).

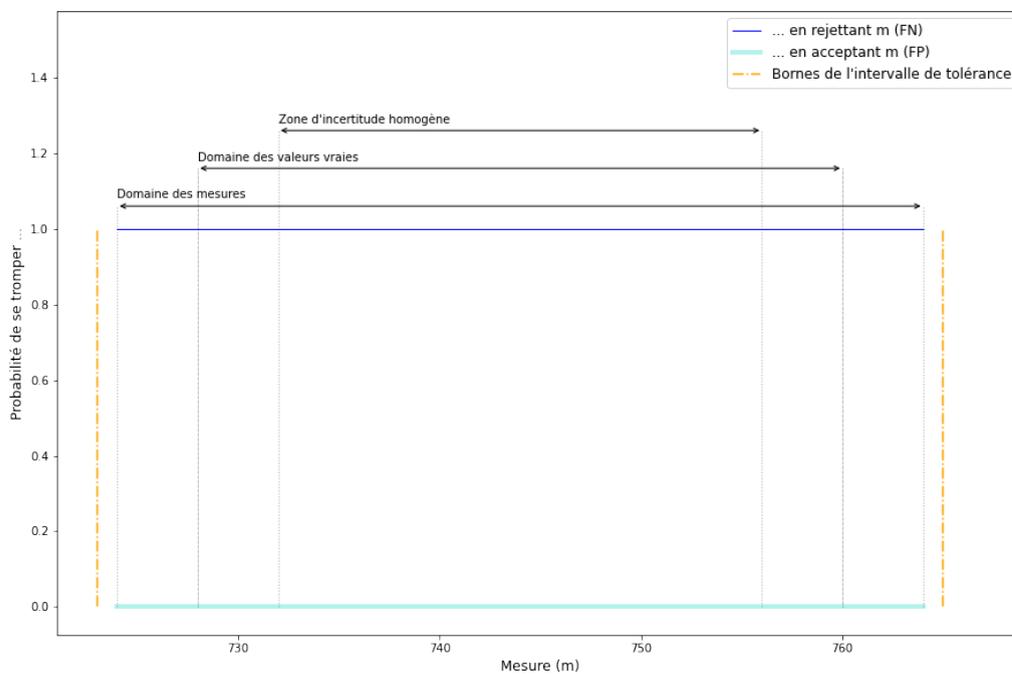
En Figure B.1(b), la borne supérieure de l'intervalle de tolérance est à l'intérieur du domaine des mesures mais le domaine des valeurs vraies (c'est à dire toutes les valeurs vraies qu'il est possible d'obtenir en production) est encore entièrement inclus dans l'intervalle de tolérance. A ce stade, les valeurs vraies sont encore toutes conformes et il faut donc accepter toutes les pièces. S'il y a quelques mesures hors-tolérance, c'est nécessairement à cause de l'erreur de mesure. Cela correspondrait à une production idéalement calibrée (qui ne présente ni incertitudes ni surcoût). La lecture de cette zone n'est faisable que si l'on est en connaissance des lois de la valeur vraie et de l'erreur de mesure (cette information est donc une plus-value de la déconvolution ; autrement, à ce stade la conformité serait déjà discutée).

A partir de la Figure B.2(a) la borne supérieure de l'intervalle de tolérance est incluse dans le domaine de valeurs vraies. Il est désormais possible qu'une valeur vraie soit hors-tolérance :

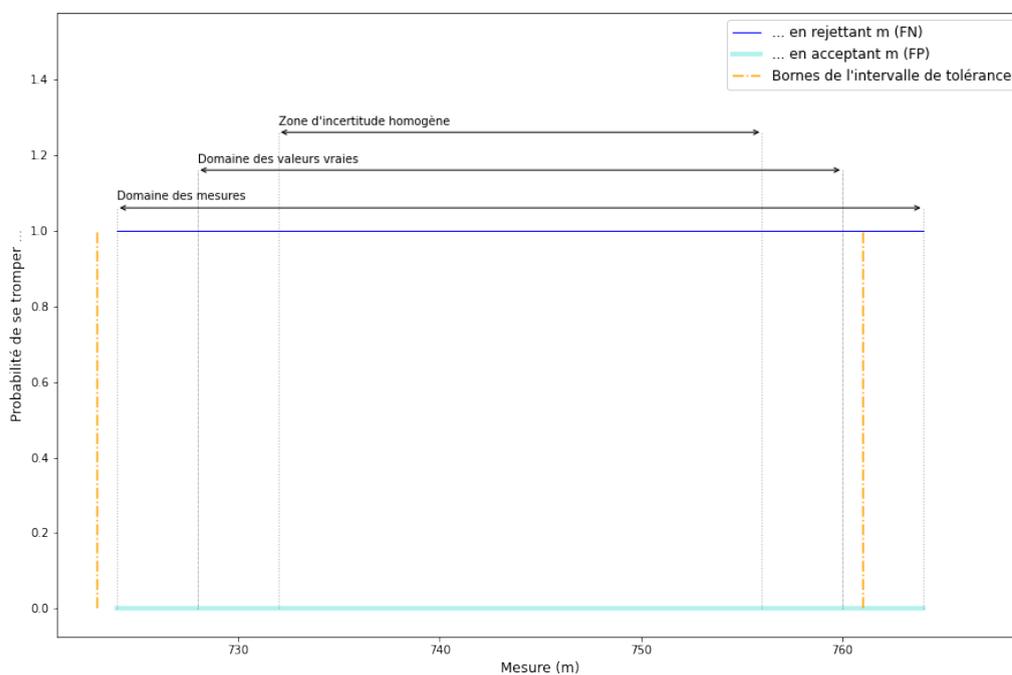
- Toutes les mesures soupçonnées d'avoir une valeur vraie hors-tolérance sont grisées et constituent la *zone d'incertitudes* dont les bornes sont : $[u - e_{max}, u - e_{min}]$ avec e_{min} et e_{max} le minimum et le maximum de l'erreur de mesure (ici $[755, 763]$). Dans cette zone d'incertitudes, les courbes des probabilités de prendre une mauvaise décision devraient être linéaires, comme dans la Figure 2.13 (conformément à la densité de probabilité d'une loi uniforme). Or une irrégularité est visible à partir de la valeur 756, après laquelle toutes les probabilités de faux positifs sont réduites et celles de faux négatifs augmentées. Cette irrégularité correspond à la sortie de la zone homogène dont la borne supérieure est donnée par $t_{max} + e_{min}$ (ici $760 + (-4)$).
- Une conséquence intéressante de la zone non-homogène est que les deux courbes ne se croisent plus au même endroit.
- En dehors de la zone d'incertitude, la conformité de la mesure est en phase avec celle de la valeur vraie.

Parmi les configurations affectées par la zone non-homogène, on remarquera que la Figure B.2(b) correspond à la situation schématisée en Figure 2.14 présentant 50% de probabilités de prendre une mauvaise décision.

Enfin dans la dernière Figure B.4(b), l'effet de la zone non-homogène n'est plus visible. Toutes les mesures que l'on peut suspecter, qui sont de part et d'autre de la borne supérieure de l'intervalle de tolérance, sont dans la zone homogène.

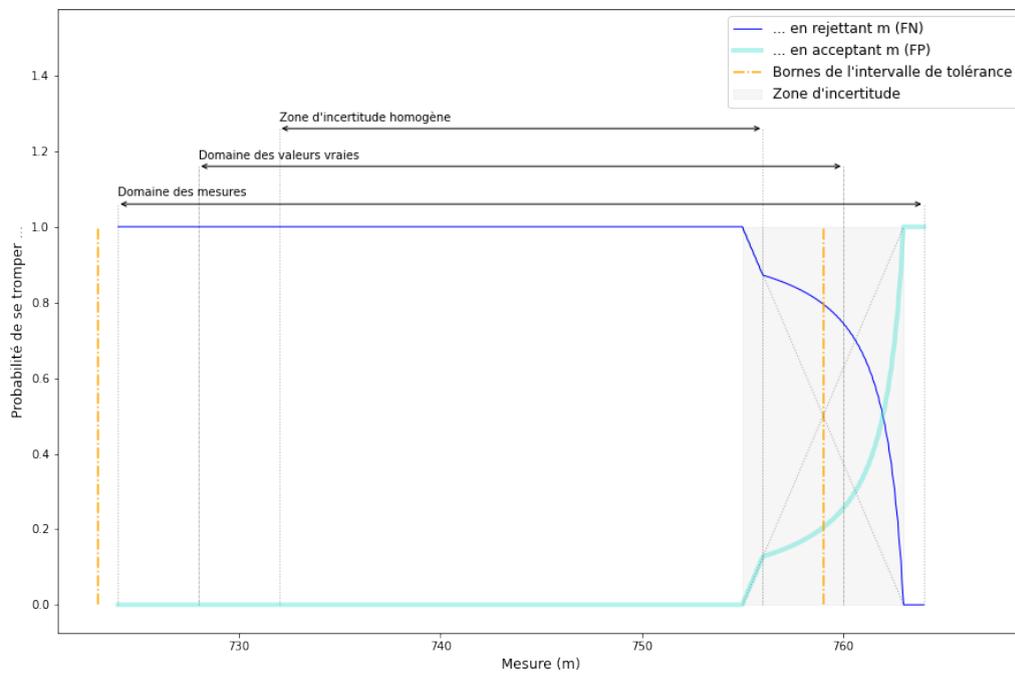


(a)

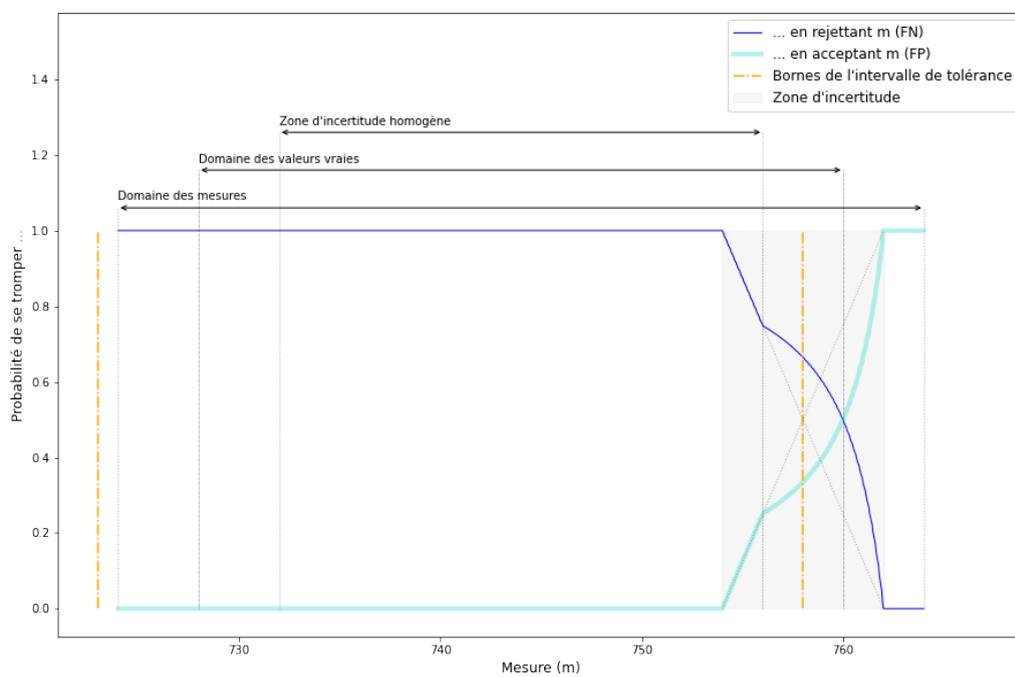


(b)

FIGURE B.1: Illustration des effets de la zone non homogène (1/4).

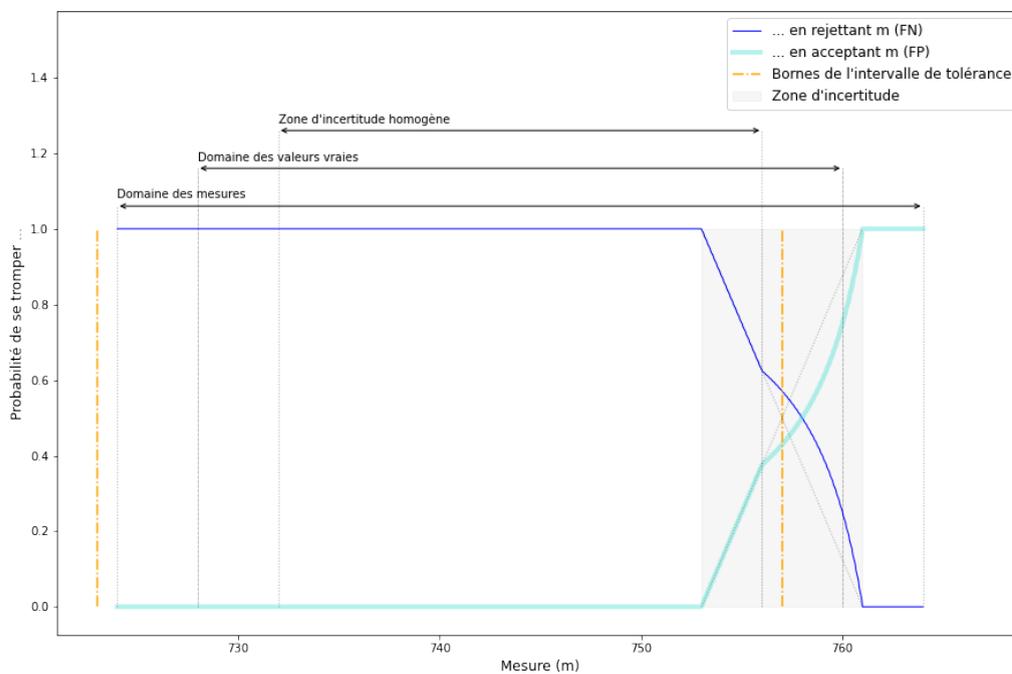


(a)

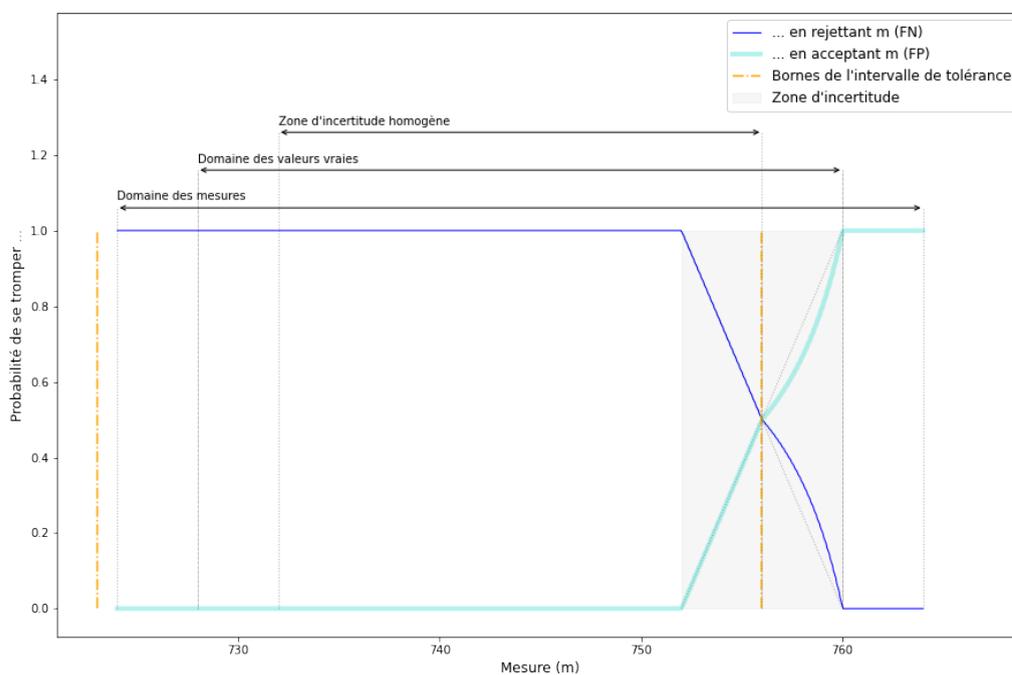


(b)

FIGURE B.2: Illustration des effets de la zone non homogène (2/4).

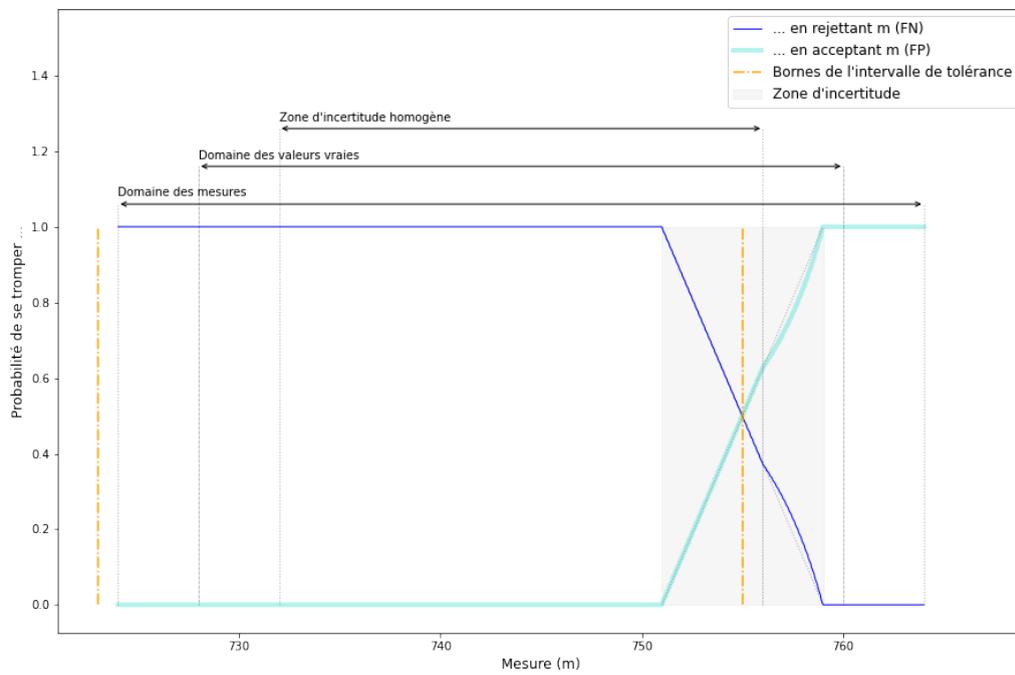


(a)

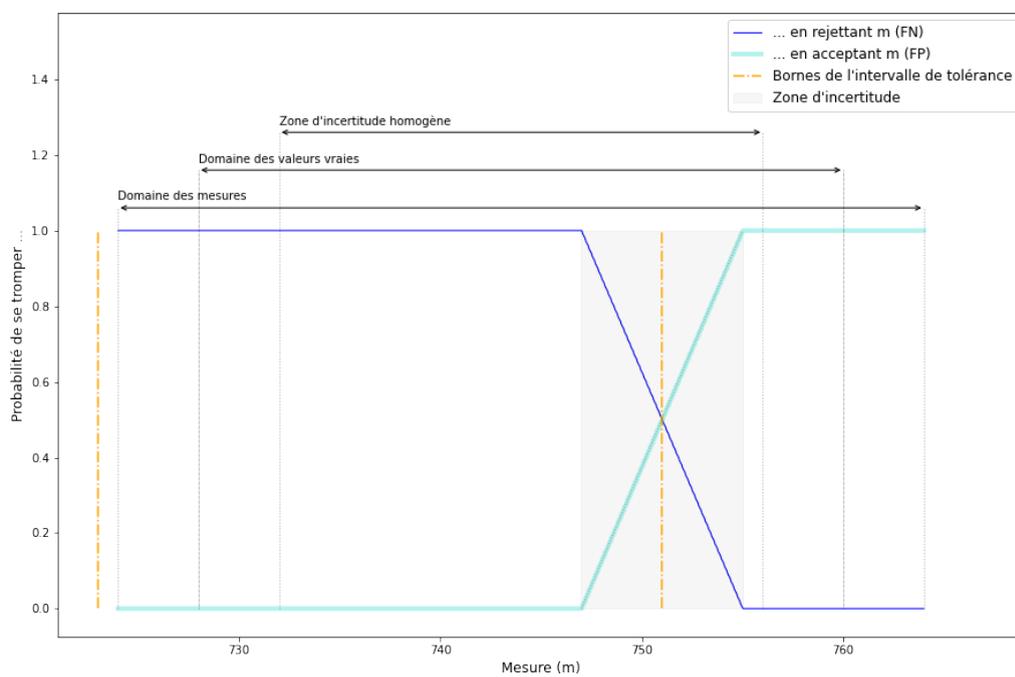


(b)

FIGURE B.3: Illustration des effets de la zone non homogène (3/4).



(a)



(b)

FIGURE B.4: Illustration des effets de la zone non homogène (4/4).

B.2 Population avantagée par la révision

L'application de l'inférence bayésienne permet d'obtenir une valeur révisée pour chaque mesure de l'échantillon (Section 2.2). Le Chapitre 3 a montré que les valeurs révisées obtenues constituaient généralement une prédiction plus précise des valeurs vraies que les mesures.

La Figure B.5 montre un nuage de points, avec en abscisse les mesures et en ordonnée les résidus. La génération de cet échantillon est réalisée selon un paramétrage identique à celui utilisé pour l'exemple gaussien (Section 3.1).

La Figure B.5 montre que le nuage est découpé en quatre populations distinctes :

- les deux populations marquées par des flèches **vertes** : les mesures en bas à gauche (ronds verts) et en haut à droite du graphique (étoiles kaki) bénéficient de la révision : les valeurs révisées obtenues sont plus proches des valeurs vraies ;
- les deux populations marquées par des flèches **rouges** : les mesures en haut à gauche (croix bleues) et en bas à droite du graphique (avec des triangles oranges) perdent à être révisées : les valeurs révisées obtenues sont pires que les mesures initiales.

La Figure B.5 montre aussi la tendance de la révision : une flèche vers le haut signifie que la valeurs révisées obtenues après révision seront plus grandes que les mesures ($r^{(i)} > m^{(i)}$). Une flèches vers le bas indique la tendance opposée.

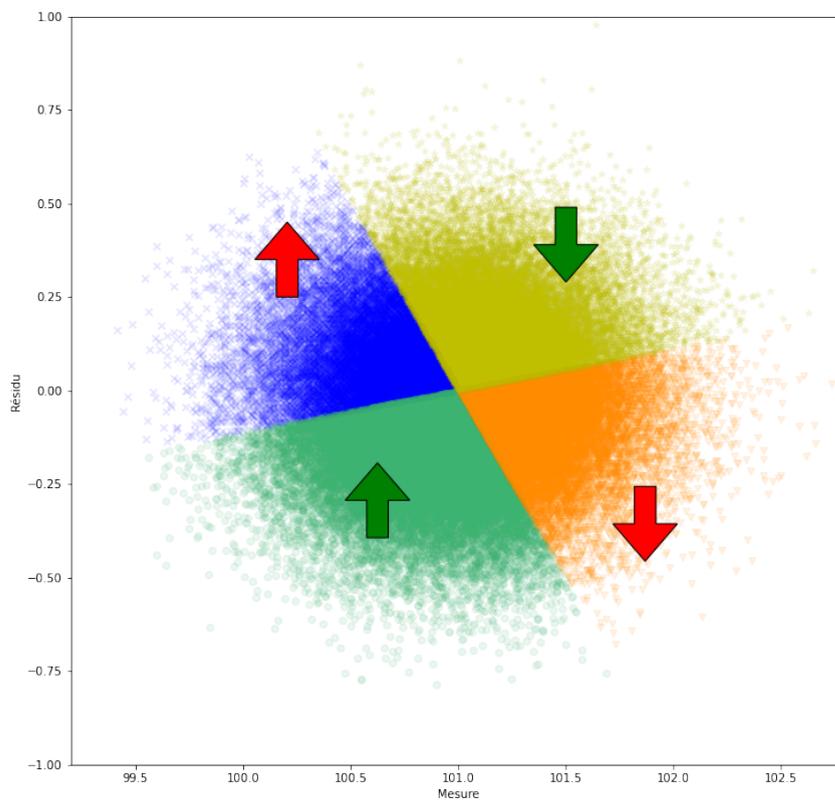
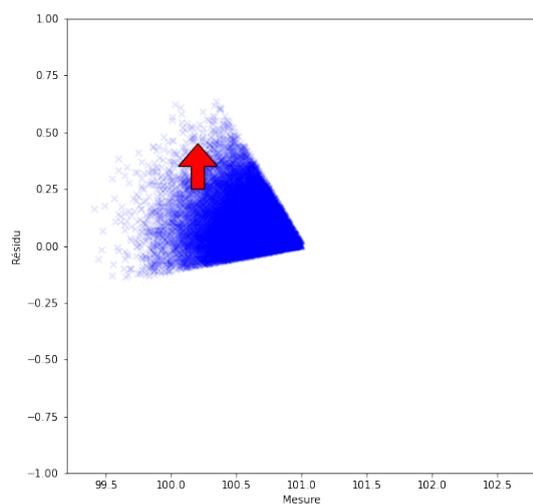


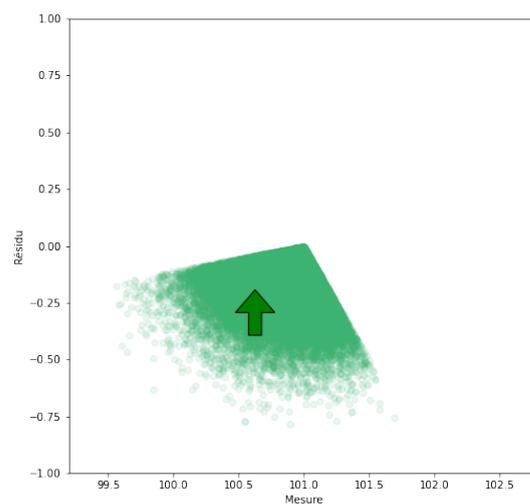
FIGURE B.5: Illustration de l'influence de l'inférence bayésienne (cas gaussien).

L'ensemble des mesures qui bénéficient de la révision est constitué de deux populations opposées et forme un genre de sablier. Il semble être défini par deux droites : l'une correspond clairement à la séparation entre les populations ayant une flèche vers le haut et celles en ayant une vers le bas. L'autre, la plus horizontale, présente une discontinuité au niveau de l'intersection avec la première.

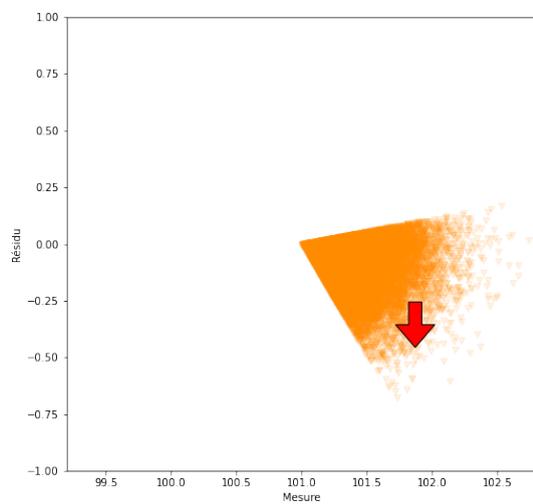
Les quatre populations formées par la combinaison du sens des flèches (haut, bas) et de leurs couleurs (rouge, vert) sont affichées isolément en Figure B.6. Il est intéressant de noter que les quatre populations appartiennent à quatre régions du graphique qui ne se chevauchent pas.



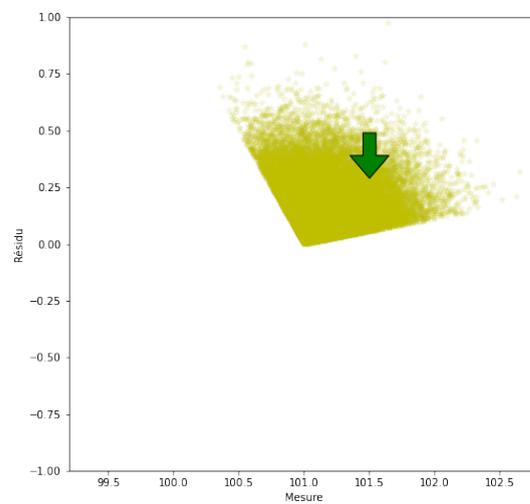
(a)



(b)



(c)



(d)

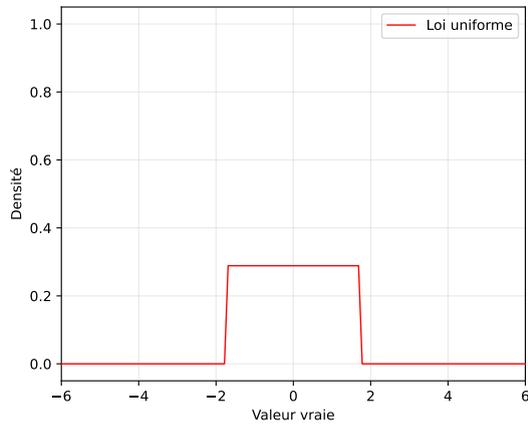
FIGURE B.6: Isolement des différentes catégories de données.

Annexe C

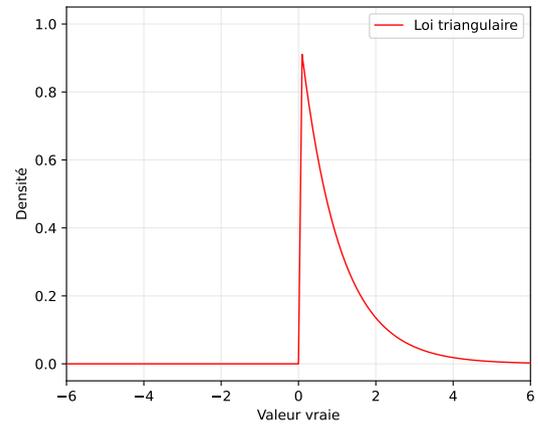
Descriptions des lois

TABLEAU C.1: Définition des lois choisies. Les densités de probabilité prennent une valeur nulle en dehors du support de la variable.

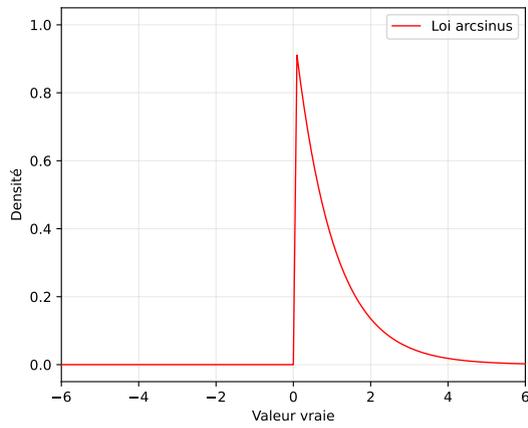
Loi	$f_X(x)$	Définition et support	Paramètres
Uniforme	$\frac{1}{(b-a)}$	$x \in [a, b]$ $a < b$	a Borne inférieure b Borne supérieure
Triangulaire	$\begin{cases} \frac{2(x-a)}{(m-a)(b-a)} & \text{si } a \leq x \leq m \\ \frac{2(b-x)}{(b-m)(b-a)} & \text{si } m < x \leq b \end{cases}$	$x \in [a, b]$ $a \leq m \leq b$ $a < b$	a Borne inférieure m Mode b Borne supérieure
Arcsinus	$\frac{1}{\pi\sqrt{(x-a)(b-x)}}$	$x \in [a, b]$ $a < b$	a Borne inférieure b Borne supérieure
Normale	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$	$x \in \mathbb{R}$ $\sigma > 0$	μ Moyenne σ Écart type
Lognormale	$\frac{1}{\sigma_{\log}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(x-\gamma) - \mu_{\log}}{\sigma_{\log}}\right)^2\right)$	$x \in [\gamma, +\infty[$ $\sigma_{\log} > 0$ $\gamma < \mu_{\log}$	μ_{\log} Espérance (de $\log X$) σ_{\log} Écart type (de $\log X$) γ Position
WeibullMax	$\frac{\alpha}{\beta} \left(-\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(-\frac{x-\gamma}{\beta}\right)^\alpha\right)$	$x \in]-\infty, \gamma]$ $\alpha > 0$ $\beta > 0$	α Facteur de forme β Facteur d'échelle γ Paramètre de position
WeibullMin	$\frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)$	$x \in [\gamma, +\infty[$ $\alpha > 0$ $\beta > 0$	α Facteur de forme β Facteur d'échelle γ Paramètre de position



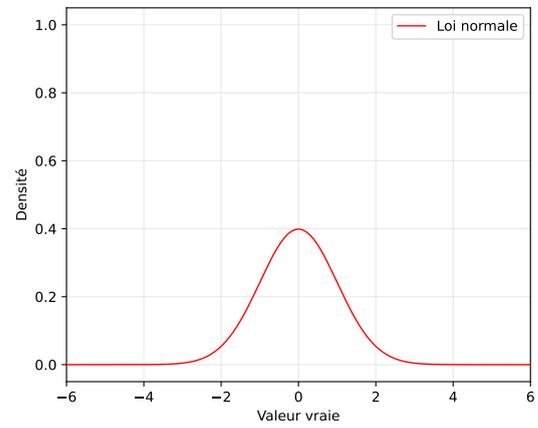
(a) Loi uniforme



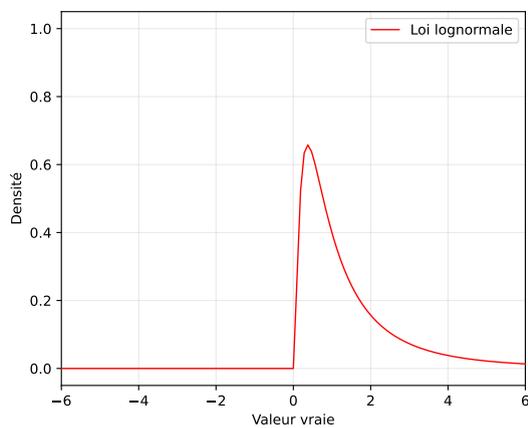
(b) Loi triangulaire



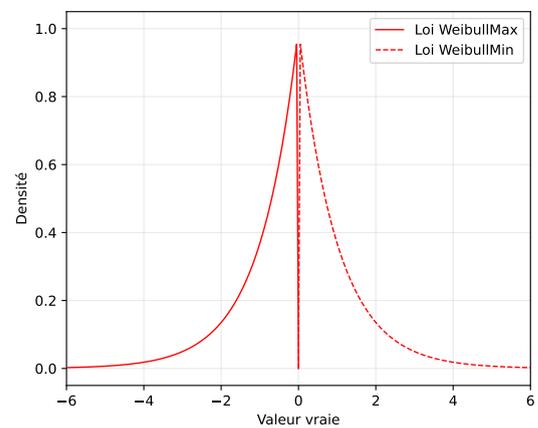
(c) Loi arcsinus



(d) Loi normale



(e) Loi lognormale



(f) Loi de Weibull

FIGURE C.1: Exemple d'une loi de moyenne 0 et d'écart-type 1 pour chacune des fonctions candidates.

Bibliographie

- [1] H. d.-G. Acquah. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of development and agricultural economics*, 2 :001–006, 2010.
- [2] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6) :716–723, 1974.
- [3] F. Alkam, I. Pereira, and T. Lahmer. Qualitatively-improved identified parameters of prestressed concrete catenary poles using sensitivity-based Bayesian approach. *Results in Engineering*, 6 :100104, 2020.
- [4] F. Attivissimo, N. Giaquinto, and M. Savino. A Bayesian paradox and its impact on the GUM approach to uncertainty. *Measurement*, 45(9) :2194–2202, 2012.
- [5] M. L. Bailon and D. J. Horntrop. On the calculation of convolutions with Gaussian Kernels. *Applied Mathematics and Computation*, 176(1) :383–387, 2006.
- [6] D. Banks, A. Cron, and A. Raskind. Bayesian metrology in metabolomics. *Chemometrics and Intelligent Laboratory Systems*, 208 :104208, 2021.
- [7] M. Baudin, A. Dutfoy, B. Iooss, and A.-L. Popelin. OpenTURNS : An Industrial Software for Uncertainty Quantification in Simulation. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of Uncertainty Quantification*, pages 1–38. Springer International Publishing, Cham, 2016.
- [8] P. Beaucaire. *Application des méthodes fiabilistes à l'analyse et à la synthèse des tolérances*. Theses, Université Blaise Pascal - Clermont-Ferrand II, Nov. 2012.
- [9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology, JCGM 100, 2008.
- [10] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data – Supplement 1 to the Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method. Joint Committee for Guides in Metrology, JCGM 101, 2008.
- [11] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Joint Committee for Guides in Metrology (3rd edition), JCGM 200, 2012.

- [12] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Guide to the expression of uncertainty in measurement – Part 6 : Developing and using measurement models. Joint Committee for Guides in Metrology, JCGM GUM-6 :2020, 2020.
- [13] Y. Biraud. Les méthodes de déconvolution et leurs limitations fondamentales. *Revue de Physique Appliquée*, 11(2) :203–214, 1976.
- [14] P. D. Bièvre. The 2007 International Vocabulary of Metrology (VIM), JCGM 200 :2008 [ISO/IEC Guide 99] : Meeting the need for intercontinentally understood concepts and their associated intercontinentally agreed terms. *Clinical Biochemistry*, 42(4) :246–248, 2009.
- [15] J.-J. Boreux, E. Parent, and J. Bernier. *Pratique du calcul bayésien*. Collection statistique et probabilités appliquées. Springer, Paris, Berlin, Heidelberg, 2010.
- [16] P. Brutti, F. D. Santis, and S. Gubbiotti. Predictive measures of the conflict between frequentist and Bayesian estimators. *Journal of Statistical Planning and Inference*, 148 :111–122, 2014.
- [17] C. Caballero-Gaudes, S. Moia, P. Panwar, P. A. Bandettini, and J. Gonzalez-Castillo. A deconvolution algorithm for multi-echo functional MRI : Multi-echo Sparse Paradigm Free Mapping. *NeuroImage*, 202 :116081, 2019.
- [18] S. Carmignato, L. De Chiffre, H. Bosse, R. Leach, A. Balsamo, and W. Estler. Dimensional artefacts to achieve metrological traceability in advanced manufacturing. *CIRP Annals*, 69(2) :693–716, 2020.
- [19] S. A. Carvajal, A. F. Medina, A. J. Bohórquez, and C. A. Sánchez. Estimation of calibration intervals using Bayesian inference. *Measurement*, 187 :110316, 2022.
- [20] H. F. F. Castro. Validation of the GUM using the Monte Carlo method when applied in the calculation of the measurement uncertainty of a compact prover calibration. *Flow Measurement and Instrumentation*, 77 :101877, 2021.
- [21] H. J. Caulfield. Image processing : the nature and value of fuzzy metrology. In M. A. Kahan, editor, *An Optical Believe It or Not : Key Lessons Learned*, volume 7071, page 70710H. International Society for Optics and Photonics, SPIE, 2008.
- [22] H. J. Caulfield, J. Ludman, and J. Shamir. 26 - Fuzzy Metrology. In C. T. Leondes, editor, *Fuzzy Theory Systems*, pages 747–758. Academic Press, San Diego, 1999.
- [23] V. Chabridon, N. Gayton, J.-M. Bourinet, M. Balesdent, and J. Morio. Some Bayesian insights for statistical tolerance analysis. In A. F. de Mécanique, editor, *CFM 2017 - 23ème Congrès Français de Mécanique*, Lille, France, 2017. AFM, Maison de la Mécanique, 39/41 rue Louis Blanc - 92400 Courbevoie.
- [24] X. Chen, H. Hong, and D. Nekipelov. *Measurement Error Models*, pages 1–9. 2008.
- [25] X. Chen, H. Hong, and D. Nekipelov. Nonlinear models of measurement errors. *Journal of Economic Literature*, 49(4) :901–37, December 2011.
- [26] A. Cœur. Accréditation Cofrac : écarts relatifs à l’informatique. *Revue Francophone des Laboratoires*, 2017(497) :24–27, 2017.

- [27] Collège Français de Métrologie (CFM). *Processus de mesure : évaluer les incertitudes 20 exemples*. AFNOR Editions, 2018.
- [28] J.-P. Coton and B. André. Mythologies et réalités de l'accréditation COFRAC en diagnostic histopathologique. *Annales de Pathologie*, 31(5, Supplement) :S157, 2011.
- [29] M. Cox, G. Rossi, P. Harris, and A. Forbes. A probabilistic approach to the analysis of measurement processes. *Metrologia*, 45 :493, 08 2008.
- [30] M. G. Cox, P. M. Harris, and I. M. Smith. Software for GUM Supplement 1 : User Manual. In *NPL Report MS 1*, 2011.
- [31] DeltaMu. Optimu (Version 6.06.001) [Logiciel] <https://www.deltamu.com/fr/logiciel-optimu/>, 2021.
- [32] Deltamu, Phimeca, Perfect Memory, CleRMa, Limos. We Are Metrology (WAM) – La communauté du bien-mesurer. <https://wearemetrology.com/>.
- [33] T. Dockhorn, J. A. Ritchie, Y. Yu, and I. Murray. Density deconvolution with normalizing flows, 2020.
- [34] Y. Dodge. *La méthode de Monte Carlo et ses applications*, pages 111–133. Springer Paris, Paris, 2008.
- [35] M. Dudek-Burlikowska. Quality estimation of process with usage control charts type X-R and quality capability of process Cp, Cpk. *Journal of Materials Processing Technology*, 162-163 :736–743, 2005.
- [36] A. Dumas. *Développement de méthodes probabilistes pour l'analyse des tolérances des systèmes mécaniques sur-contraints*. PhD thesis, Ecole nationale supérieure d'arts et métiers - ENSAM, Dec. 2014.
- [37] I. Düntsch and G. Gediga. Confusion Matrices and Rough Set Data Analysis. *Journal of Physics : Conference Series*, 1229(1) :012055, May 2019.
- [38] S. Eichstädt, A. Keidel, and J. Tesch. Metrology for the digital age. *Measurement : Sensors*, 18 :100232, 2021.
- [39] V. Elvira, P. Closas, and L. Martino. Importance Gaussian Quadrature. *IEEE Transactions on Signal Processing*, PP :1–1, Dec. 2020.
- [40] G. Fabien. Can we dispense with the notion of "true value" in metrology? In *Standardization in measurement : Philosophical, historical and sociological issues*, pages 81–93. Pickering & Chatto, 2015.
- [41] J. Fan and Y. K. Truong. Nonparametric Regression with Errors in Variables. *The Annals of Statistics*, 21(4) :1900 – 1925, 1993.
- [42] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8) :861–874, 2006.
- [43] A. J. Fernández, C. D. Correa-álvarez, and L. R. Pericchi. Balancing producer and consumer risks in optimal attribute testing : A unified Bayesian/Frequentist design. *European Journal of Operational Research*, 286(2) :576–587, 2020.

- [44] A. Ferreira, A. Roussy, and L. Condé. Modèles de métrologie virtuelle pour prédire les mesures physiques dans la fabrication de semi-conducteurs. In *QUALITA 2009. 8ème Congrès international pluridisciplinaire en Qualité et Sécurité de Fonctionnement*, Besançon, France, Mar. 2009.
- [45] A. Ferrero and S. Salicone. An Innovative Method for the Comparison of Measurement Results. In *Transverse Disciplines in Metrology*, pages 741–754. John Wiley & Sons, Ltd, 2009.
- [46] S. Ferson, V. Kreinovich, L. Ginzburg, D. Myers, and K. Sentz. Constructing Probability Boxes and Dempster-Shafer Structures. *Sandia National Laboratories*, Apr. 2003.
- [47] P. I. Fierens, L. C. Rêgo, and T. L. Fine. A frequentist understanding of sets of measures. *Journal of Statistical Planning and Inference*, 139(6) :1879–1892, 2009.
- [48] L. J. Fish. Why Multivariate Methods are Usually Vital. *Measurement and Evaluation in Counseling and Development*, 21(3) :130–137, 1988.
- [49] A. B. Forbes. Measurement uncertainty and optimized conformance assessment. *Measurement*, 39(9) :808–814, 2006.
- [50] A. B. Forbes and J. A. Sousa. The GUM, Bayesian inference and the observation and measurement equations. *Measurement*, 44(8) :1422–1435, 2011.
- [51] I. Fornacon-Wood, H. Mistry, C. Johnson-Hart, C. Faivre-Finn, J. P. B. O’Connor, and G. J. Price. Understanding the Differences Between Bayesian and Frequentist Statistics. *International Journal of Radiation Oncology*Biophysics*Physics*, 112(5) :1076–1082, 2022.
- [52] D. A. S. Fraser and N. Reid. Strong matching of frequentist and Bayesian parametric inference. *Journal of Statistical Planning and Inference*, 103(1) :263–285, 2002.
- [53] J.-M. Galharret. *Statistique bayésienne appliquée en sciences humaines et sociales*. Theses, Université de Nantes Faculté des sciences et des techniques, Dec. 2019.
- [54] D. Gauder, J. Gözl, M. Biehler, M. Diener, and G. Lanza. Balancing the trade-off between measurement uncertainty and measurement time in optical metrology using design of experiments, meta-modelling and convex programming. *CIRP Journal of Manufacturing Science and Technology*, 35 :209–216, 2021.
- [55] N. Gayton. *Méthodes probabilistes pour la conception mécanique*. PhD thesis, Institut Français de Mécanique Avancée, 2012.
- [56] F. Ghaderinezhad and C. Ley. Quantification of the impact of priors in Bayesian statistics via Stein’s Method. *Statistics & Probability Letters*, 146 :206–212, 2019.
- [57] M. Gille, P. Beaurepaire, A. Dumas, F. Taghon, T. Yalamas, and N. Gayton. On the identification of the probability density function of a dimension from data with measurement errors. *Procedia CIRP*, 114 :60–66, 2022. 17th CIRP Conference on Computer Aided Tolerancing (CAT2022).
- [58] E. Goka. *Analyse des tolérances des systèmes complexes - Modélisation des imperfections de fabrication pour une analyse réaliste et robuste du comportement des systèmes*. Theses, Ecole nationale supérieure d’arts et métiers - ENSAM, June 2019.

- [59] B. Goller. *Stochastic Model Validation of Structural Systems*. PhD thesis, Leopold-Franzens – Universität Innsbruck, 2011.
- [60] L. Granvilliers, V. Kreinovich, and N. Müller. Novel Approaches to Numerical Software with Result Verification. In R. Alt, A. Frommer, R. B. Kearfott, and W. Luther, editors, *Numerical Software with Result Verification*, pages 274–305, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [61] N. Greif, H. Schrepf, and D. Richter. Software validation in metrology : A case study for a GUM-supporting software. *Measurement*, 39(9) :849–855, 2006.
- [62] S. Grimnes and Ø. G. Martinsen. Chapter 9 - Data and Models. In S. Grimnes and Ø. G. Martinsen, editors, *Bioimpedance and Bioelectricity Basics (Third Edition)*, pages 329–404. Academic Press, Oxford, third edition edition, 2015.
- [63] A. Grous. Fundamentals of Error Analysis and their Uncertainties in Dimensional Metrology Applied to Science and Technology. In *Applied Metrology for Manufacturing Engineering*, pages 1–84. John Wiley & Sons, Ltd, London : Hoboken, NJ, 2011. Section : 1.
- [64] F. Grégis and N. de Courtenay. Incertitude de mesure et probabilités : la confrontation des approches fréquentiste et bayésienne en métrologie. In I. Drouet, editor, *Le bayésianisme aujourd'hui. Fondements et pratiques*, Collection : Sciences & philosophie. Editions Matériologiques, 2016.
- [65] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585(7825) :357–362, Sept. 2020.
- [66] L. Heinrich. Factorization of the characteristic function of a sum of dependent random variables. *Lithuanian Mathematical Journal - LITH MATH J*, 22 :92–100, Jan. 1982.
- [67] D. B. Hibbert. Evaluation of Measurement Data : The Role of Measurement Uncertainty in Conformity Assessment. *Chemistry International - Newsmagazine for IUPAC*, 35(2) :22–23, 2013.
- [68] X. Huang and H. Zhou. Conditional density estimation with covariate measurement error, 2020.
- [69] P. Huber. Robust Statistics. *International Encyclopedia of Statistical Science*, pages 1248–1251, Jan. 1981.
- [70] D. Idriss, P. Beaurepaire, L. Homri, and N. Gayton. Tolerance Analysis - Key Characteristics Identification by Sensitivity Methods. *Procedia CIRP*, 75 :33–38, 2018.
- [71] ISO 14253-1. Geometrical Product Specifications (GPS) - Inspection by measurement of workpieces and measuring equipments - Part 1 : Decision rules for proving conformance or non-conformance with specifications, 1998.
- [72] ISO 21748. Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation. International Organization for Standardization, Geneva, CH, 2016.

- [73] ISO 5725-1. Accuracy (trueness and precision) of measurement methods and results – Part 1 : General principles and definitions. International Organization for Standardization, Geneva, CH, 1994.
- [74] ISO/IEC Guide 99. Vocabulaire international de métrologie - Concepts fondamentaux et généraux et termes associés (VIM), 2007.
- [75] P. Janssen, J. Swanepoel, and N. Veraverbeke. A note on the behaviour of a kernel-smoothed kernel density estimator. *Statistics & Probability Letters*, 158 :108663, 2020.
- [76] H. V. Jetti, A. Ferrero, and S. Salicone. A modified Bayes' theorem for reliable conformity assessment in industrial metrology. *Measurement*, 184 :109967, 2021.
- [77] A. M. Johansen. Monte Carlo Methods. In P. Peterson, E. Baker, and B. McGaw, editors, *International Encyclopedia of Education (Third Edition)*, pages 296–303. Elsevier, Oxford, third edition edition, 2010.
- [78] J.-M. Judic, J.-M. Pou, L. Leblond, and N. Gayton. Objectif Smart Tolerancing – Des méthodes de tolérancement innovantes et efficaces pour l'usine du futur. In *Livre Blanc, Manufacturing.fr*, page 24, 02 2018.
- [79] R. N. Kacker. Measurement uncertainty and its connection with true value in the GUM versus JCGM documents. *Measurement*, 127 :525–532, 2018.
- [80] R. N. Kacker. Towards correction of the JCGM international vocabulary of metrology. *Measurement : Sensors*, 18 :100063, 2021.
- [81] R. N. Kacker, R. Kessel, and J. Lawrence. Removing divergence of JCGM documents from the GUM (1993) and repairing other defects. *Measurement*, 88, Mar. 2016.
- [82] V. E. Kane. Process Capability Indices. *Journal of Quality Technology*, 18(1) :41–52, Jan. 1986. Publisher : Taylor & Francis.
- [83] F. Korner-Nievergelt, T. Roth, S. von Felten, J. Guélat, B. Almasi, and P. Korner-Nievergelt. Chapter 3 - The Bayesian and the Frequentist Ways of Analyzing Data. In F. Korner-Nievergelt, T. Roth, S. von Felten, J. Guélat, B. Almasi, and P. Korner-Nievergelt, editors, *Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and STAN*, pages 19–31. Academic Press, Boston, Jan. 2015.
- [84] V. Kotu and B. Deshpande. Chapter 4 - classification. In V. Kotu and B. Deshpande, editors, *Data Science (Second Edition)*, pages 65–163. Morgan Kaufmann, second edition edition, 2019.
- [85] V. Kreinovich. Probabilities, Intervals, What Next ? Optimization Problems Related to Extension of Interval Computations to Situations with Partial Information about Probabilities. *Journal of Global Optimization*, 29(3) :265–280, July 2004.
- [86] J. Kuha. AIC and BIC : Comparisons of Assumptions and Performance. *Sociological Methods & Research*, 33 :188–229, Nov. 2004.
- [87] H. Kunzmann, T. Pfeifer, R. Schmitt, H. Schwenke, and A. Weckenmann. Productive metrology - adding value to manufacture. *CIRP Annals*, 54(2) :155–168, 2005.

- [88] E. Kureková. Measurement Process Capability - Trends and Approaches. *Measurement Science Review*, 1(1), 2001.
- [89] D. Kurisu and T. Otsu. On linearization of nonparametric deconvolution estimators for repeated measurements model. *Journal of Multivariate Analysis*, 189 :104921, 2022.
- [90] D. Kurz, J. Pilz, A. Schirru, S. Pampuri, and C. De Luca. A Sampling Decision System for Semiconductor Manufacturing : Relying on Virtual Metrology and Actual Measurements. In *Proceedings of the 2014 Winter Simulation Conference, WSC '14*, pages 2649–2660. IEEE Press, 2014.
- [91] H. E. Kyburg. *Theory and Measurement*. Cambridge, England : Cambridge University Press, 1984.
- [92] D. G. Lainiotis, S. K. Katsikas, and S. D. Likothanassis. Optimal seismic deconvolution. *Signal Processing*, 15(4) :375–404, 1988.
- [93] F. Laporte. *Méthode d'inférence utilisant la vraisemblance empirique basée sur l'entropie pour les modèles de diffusion avec sauts*. Theses, Université Laval, 2019.
- [94] Leblond, Laurent and Pillet, Maurice. Conformity and statistical tolerancing. *Int. J. Metrol. Qual. Eng.*, 9 :1, 2018.
- [95] R. Lebrun. *Contributions à la modélisation de la dépendance stochastique*. Theses, Université Paris-Diderot - Paris VII, May 2013. Issue : 20908922.
- [96] P.-L. Liu and A. D. Kiureghian. Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1(2) :105–112, 1986.
- [97] A. Los and J. R. R. Mayer. Application of GUM Supplement 2 to uncertainty estimation of five-axis CNC machines geometric error parameters identification using the SAMBA method. In *Laser Metrology and Machine Performance X*, 2013.
- [98] L. Lu, F. Goerlandt, O. A. V. Banda, and P. Kujala. Developing fuzzy logic strength of evidence index and application in Bayesian networks for system risk management. *Expert Systems with Applications*, 192 :116374, 2022.
- [99] K. Mahata and M. M. Hyder. Sparse deconvolution via off-grid T.V minimization. *Signal Processing*, 170 :107406, 2020.
- [100] G. Mallet, A. Dugay, M. Leroy, C. Gray, J.-P. Basuyau, and P. Brunelle. Démarche d'accreditation Cofrac dans un laboratoire hospitalier polyvalent. Mise en place et gestion au quotidien. *Immuno-analyse & Biologie Spécialisée*, 20(4) :239–244, 2005.
- [101] A. Meister. *Deconvolution Problems in Nonparametric Statistics*. Lecture Notes in Statistics - 193. Springer Berlin, Heidelberg, 2009.
- [102] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247) :335–341, 1949. PMID : 18139350.
- [103] A. Meyer-Baese and V. Schmid. Chapter 6 - Statistical and Syntactic Pattern Recognition. In *Pattern Recognition and Signal Analysis in Medical Imaging (Second Edition)*, pages 151–196. Academic Press, Oxford, second edition edition, 2014.

- [104] S. Nadarajah and T. Pogány. On the convolution of normal and t random variables. *Statistics*, 47, June 2012.
- [105] A. Nataf. Détermination des distributions de probabilités dont les marges sont données. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris*, 255 :42–43, 1962.
- [106] National Institute of Standards and Technology (NIST) – U.S. Department of Commerce.
- [107] L. Nghiem. *Advances in Measurement Error Modeling*. PhD thesis, Southern Methodist University, 2019.
- [108] C. V. Nguyen and L. Souhac. Data assimilation methods for urban air quality at the local scale. *Atmospheric Environment*, 253 :118366, 2021.
- [109] A. Oppermann, S. Eickelberg, J. Exner, T. Bock, M. Bernien, R. Niepraschk, W. Heeren, O. Baer, and C. Brown. Digital Transformation in Metrology : Building a Metrological Service Ecosystem. *Procedia Computer Science*, 200 :308–317, 2022.
- [110] C. Ottlé and J.-F. Mahfouf. 8 - Data Assimilation of Satellite Observations. In N. Baghdadi and M. Zribi, editors, *Microwave Remote Sensing of Land Surface*, pages 357–382. Elsevier, 2016.
- [111] M. Papananias, T. E. McLeay, M. Mahfouf, and V. Kadirkamanathan. A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing. *Computers in Industry*, 105 :35–47, 2019.
- [112] A. Papoulis and S. U. Pillai. *Probability, random variables and stochastic processes*. McGraw-Hill Electrical and electronic engineering series. McGraw-Hill, Boston Burr Ridge (IL) Du-buque (IA), 4th edition edition, 2002. Publication Title : Probability, random variables and stochastic processes.
- [113] E. Parzen. On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics*, 33 :1065–1076, 1962.
- [114] A. Possolo. Statistics and Metrology. In *Wiley StatsRef : Statistics Reference Online*, pages 1–8. John Wiley & Sons, Ltd, 2018.
- [115] J.-M. Pou and L. Leblond. *La Smart Metrology - De la métrologie des instruments à la métrologie des décisions*. AFNOR, 2022.
- [116] M. Priel. Vocabulaire de la métrologie. *Techniques de l'ingénieur – Organisation et vocabulaire de la métrologie*, base documentaire : TIB421DUO.(ref. article : r115), 2010.
- [117] J.-M. Reifenberg, E. Riout, A. Leroy, and S. Begue. Métrologie à l'EFS : de la production à la raison. *Transfusion Clinique et Biologique*, 21(3) :132–138, 2014.
- [118] E. Rohée, R. Coulon, F. Carrel, T. Dautremer, E. Barat, T. Montagu, S. Normand, and C. Jammes. Benchmark of the non-parametric Bayesian deconvolution method implemented in the SINBAD code for X/γ rays spectra processing. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 836 :91–97, 2016.

- [119] D. Rossat. *Bayesian techniques for inverse uncertainty quantification for multi-physics models of large containment structures*. PhD thesis, Université Grenoble Alpes, Oct. 2022.
- [120] K. H. Ruhm. Measurement plus observation - A new structure in metrology. *Measurement*, 126 :421–432, 2018.
- [121] J. Sadeghi, M. d. Angelis, and E. Patelli. Frequentist history matching with Interval Predictor Models. *Applied Mathematical Modelling*, 61 :29–48, 2018.
- [122] R. Sanchez-Marquez and J. J. Vivas. Building a Cpk control chart - A novel and practical method for practitioners. *Computers & Industrial Engineering*, 158 :107428, 2021.
- [123] K. Sapozhnikova, A. Chunovkina, and R. Taymanov. "Measurement" and related concepts. Their interpretation in the VIM. *Measurement*, 50 :390–396, 2014.
- [124] G. Schwartz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2) :461–464, Mar. 1978.
- [125] L. Separovic, A. M. Saviano, and F. R. Lourenço. Using measurement uncertainty to assess the fitness for purpose of an hplc analytical method in the pharmaceutical industry. *Measurement*, 119 :41–45, 2018.
- [126] B. W. Silverman. *Density Estimation for Statistics and Data Analysis (1st ed.)*. Routledge, New York, 1998.
- [127] J. O. Smith. *Spectral Audio Signal Processing*. W3K Publishing, 2011.
- [128] S. Staelens and I. Buvat. Chapter 5 - Monte Carlo Simulations in Nuclear Medicine Imaging. In P. Verdonck, editor, *Advances in Biomedical Engineering*, pages 177–209. Elsevier, Amsterdam, 2009.
- [129] A. G. Steele and R. J. Douglas. Simplicity with advanced mathematical tools for metrology and testing. *Measurement*, 39(9) :795–807, 2006.
- [130] L. A. Stefanski and R. J. Carroll. Deconvolving kernel density estimators. *Statistics*, 21(2) :169–184, 1990.
- [131] A. Stolfi, M. K. Thompson, L. Carli, and L. De Chiffre. Quantifying the contribution of post-processing in computed tomography measurement uncertainty. *Procedia CIRP*, 43 :297–302, 2016. 14th CIRP CAT 2016 - CIRP Conference on Computer Aided Tolerancing.
- [132] M. Timmermann, T. Werner, and A. Weckenmann. An approach for efficient implementation of uncertainty analysis in production-related measurement processes, 2011.
- [133] M. F. Wahab, F. Gritti, and T. C. O'Haver. Discrete Fourier transform techniques for noise reduction and digital enhancement of analytical signals. *TrAC Trends in Analytical Chemistry*, 143 :116354, 2021.
- [134] P. Walley. Reconciling frequentist properties with the likelihood principle. *Journal of Statistical Planning and Inference*, 105(1) :35–65, 2002.
- [135] C.-H. Wang, M. H. Ali, K.-S. Chen, Y. T. Negash, M.-L. Tseng, and R. R. Tan. Data driven supplier selection as a circular economy enabler : A Taguchi capability index for manufactured products with asymmetric tolerances. *Advanced Engineering Informatics*, 47 :101249, 2021.

- [136] X.-F. Wang and B. Wang. Deconvolution estimation in measurement error models : The r package *decon*. *Journal of statistical software*, 39, 03 2011.
- [137] X.-F. Wang and D. Ye. Conditional density estimation in measurement error problems. *Journal of Multivariate Analysis*, 133 :38–50, 2015.
- [138] A. Weckenmann and M. Knauer. Causes and consequences of measurement uncertainty in production metrology, 1999.
- [139] A. Weckenmann, P. Krämer, and G. Akkasoglu. Metrology - Base for Scientific Cognition and Technical Production. *Advanced Materials Research*, 498 :169 – 176, 2012.
- [140] R. Willink. Principles of probability and statistics for metrology. *Metrologia*, 43(4) :S211–S219, Aug. 2006. Publisher : IOP Publishing.
- [141] R. Willink. An improved procedure for combining type a and type b components of measurement uncertainty. *Int. J. Metrol. Qual. Eng.*, 4(1) :55–62, 2013.
- [142] R. Willink. An accurate and consistent procedure for the evaluation of measurement uncertainty. *Measurement : Sensors*, 18 :100334, 2021.
- [143] R. Yang, D. Apley, J. Staum, and D. Ruppert. Density deconvolution with additive measurement errors using quadratic programming, 12 2018.
- [144] X. Zhang and M. D. Pandey. Structural reliability analysis based on the concepts of entropy, fractional moment and dimensional reduction method. *Structural Safety*, 43 :28–40, 2013.
- [145] C. J. J. v. Zyl. Frequentist and Bayesian inference : A conceptual primer. *New Ideas in Psychology*, 51 :44–49, 2018.