

Representing Measurement Uncertainty by Gaussian Mixtures

Luisa Hoffmann¹ ; Mohamed Yamen Alayoubi¹; Ali Darijani^{2,3} ; Jürgen Beyerer^{2,3} ;
Michael Heizmann¹ 

¹ Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT)

² Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB)

³ Karlsruhe Institute of Technology, Vision and Fusion Laboratory (IES)

Abstract

This paper presents a probabilistic framework for uncertainty quantification in nonlinear measurement systems based on Gaussian mixture models. Instead of relying on local linearization as often used according to the Guide to the Expression of Uncertainty in Measurement (GUM), the proposed approach represents the measurement model as a joint probability density function of input and output quantities. This joint distribution is approximated using a Gaussian mixture, the parameters of which are estimated from data via the Expectation–Maximization (EM) algorithm. Posterior distributions of the measurand are obtained by conditioning the Gaussian mixture on observed input values, enabling the characterization of non-Gaussian, asymmetric, and multimodal uncertainty structures. This formulation allows for a flexible and computationally efficient alternative to classical uncertainty propagation methods, particularly in the presence of nonlinear relationships. The approach is demonstrated using fatigue strength estimation of materials, where uncertainties in hardness, residual stress, and surface roughness are propagated through a nonlinear model. Monte Carlo simulation is used as a reference, and the resulting distributions are approximated by Gaussian mixtures. The results show that the proposed method provides a consistent and compact representation of uncertainty while preserving relevant distributional features.

Keywords: Gaussian mixture, Measurement uncertainty, Circular Factory

1 Introduction

1.1 Motivation

In the Circular Factory¹, used products are reintegrated into production and reconditioned for subsequent product generations. Products in such a system are subject to numerous sources of uncertainty, including unknown prior usage conditions, measurement inaccuracies, incomplete material knowledge, and imperfect models. These uncertainties must be rigorously accounted for in decision-making processes. Probability density functions (PDFs) provide a principled mathematical framework for representing uncertainty [1]. In contrast to single-value uncertainty measures, PDFs capture the full stochastic behavior of a quantity of interest. However, their practical application requires representations that are both expressive and computationally efficient. Gaussian mixtures (GMs) constitute such a representation. As a parametric family characterized by weights, means, and variances, they enable compact descriptions of complex datasets. GMs can approximate arbitrary distributions, including multimodal and asymmetric ones, while maintaining computational tractability [2]. Moreover, many key oper-

ations such as marginalization, conditioning, and Bayesian fusion can be realized via closed-form or efficiently computable solutions in the Gaussian mixture domain. This makes them particularly suitable for nonlinear estimation and sensor fusion tasks [1]. A fundamental component of uncertainty quantification is the measurement model, which describes the relationship between an observed value and the underlying true quantity. Such models may be defined analytically (e.g., density derived from mass and volume), specified in datasheets, or obtained empirically from measurement data.

1.2 Related Work

The current state of the art in measurement uncertainty evaluation is the *Guide to the Expression of Uncertainty in Measurement* (GUM) [3], published by the International Organization for Standardization. The classical GUM framework relies on linearization of the measurement model and assumes approximately Gaussian output distributions. These assumptions are violated for strongly nonlinear models or asymmetric distributions.

To overcome these limitations, GUM Supplement 1 introduces a Monte Carlo method for propagating full probability distributions numerically. This approach avoids linearization and Gaussian assumptions, and instead approximates the output distribution through repeated sampling. Harris and Cox

¹The vision of the Collaborative Research Center 1574 "Circular Factory for the Perpetual Product" is to enable integrated linear and circular production on an industrial scale. (www.sfb1574.kit.edu)

[4] describe this method as a numerical implementation of the "propagation of distributions" and discuss practical aspects such as convergence and computational cost. Comparative studies, for example [5] [6], show that GUM and Monte Carlo methods yield consistent results for linear or near-linear models, while Monte Carlo methods become essential for significantly nonlinear systems.

However, while the Monte Carlo approach provides accurate distributional propagation, it typically yields only numerical representations of the resulting distributions. This can limit interpretability and computational efficiency in downstream processing, motivating the need for compact parametric representations of uncertainty.

In this paper, we propose a method to determine the posterior uncertainty of a true quantity given a single observed measurement and an arbitrary measurement model. The resulting uncertainty is represented as a probability density function in the form of a Gaussian mixture. This approach generalizes classical GUM-based uncertainty propagation by accommodating non-Gaussian, nonlinear, and multimodal structures while preserving computational efficiency.

As a practical example, we consider measurement data related to the fatigue strength of materials. This example demonstrates how complex measurement uncertainties can be systematically represented and propagated within the proposed Gaussian mixture framework. We compare the method to the GUM uncertainty propagation method.

2 Method

2.1 Estimating the model

The objective of the proposed method is to represent the entire measurement model as a joint PDF of the output quantity and all relevant input quantities using a multivariate GM. Let $\mathbf{x} = [x_1, x_2, \dots, x_i]^T$ denote the unknown output quantity of interest, the measurand. The input variables, which are measured or observed, are described by $\mathbf{y} = [y_1, y_2, \dots, y_j]^T$. Instead of describing the measurement model as a deterministic function, the stochastic relationship between input and output quantities is modeled through the joint distribution. This joint density is approximated by a Gaussian mixture model using the random variables $\mathbf{y}$, $\mathbf{x}$. This formulation allows modeling complex and nonlinear dependencies between the variables to be represented within a compact parametric framework:

$$p(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}, \mathbf{y} \mid \mu_k, \Sigma_k) \quad (1)$$

where the mixture weights satisfy

$$\pi_k \geq 0, \quad \sum_{k=1}^K \pi_k = 1. \quad (2)$$

The mean vector and covariance matrix of the k -th component are given by

$$\mu_k = \begin{bmatrix} \mu_{k,x} \\ \mu_{k,y} \end{bmatrix}, \quad \Sigma_k = \begin{bmatrix} \Sigma_{k,xx} & \Sigma_{k,xy} \\ \Sigma_{k,yx} & \Sigma_{k,yy} \end{bmatrix}. \quad (3)$$

This representation allows us to model arbitrary nonlinear and multimodal dependencies between variables.

To estimate the parameters of the Gaussian mixture, the Expectation-Maximization (EM) algorithm [7], [8] is employed. The EM algorithm is a standard and efficient approach for maximum likelihood estimation of mixture models and provides closed-form updates for the mixture parameters. The algorithm iteratively alternates between two steps. In the expectation step (E-step), the posterior probability that a data sample belongs to each mixture component is computed. In the maximization step (M-step), the mixture parameters, including component weights, means, and covariance matrices, are updated based on these findings. Thus, the parameters of the mixture model are learned directly from the observed data samples by maximizing the likelihood of the dataset.

To prevent overfitting, the number of Gaussian components is not chosen arbitrarily but determined using model selection criteria. In particular, the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) [8] are evaluated for models with different numbers of components. Both scores penalize the number of components. The model with the minimum information criterion is selected, providing a balance between model accuracy and complexity.

The AIC is defined as

$$\text{AIC} = 2p - 2 \ln(\mathcal{L}), \quad (4)$$

where p denotes the number of free parameters of the model and $\mathcal{L}$ the maximized likelihood of the data.

The BIC is given by

$$\text{BIC} = p \ln(N) - 2 \ln(\mathcal{L}), \quad (5)$$

where N is the number of data samples. The preferred model is the one that minimizes the respective information criterion. Figure 2 shows an example of a nonlinear measurement curve and its joint GM fit. The number of components was set to five, based on the minimum AIC score shown in Figure 1.

2.2 Calculate the posterior distribution

Once the joint probability density $p(\mathbf{x}, \mathbf{y})$ has been estimated, inference can be performed by conditioning on observed input values. Conditioning corresponds to observing specific measurement values and evaluating the resulting posterior distribution of the output quantity. Conceptually, this can be

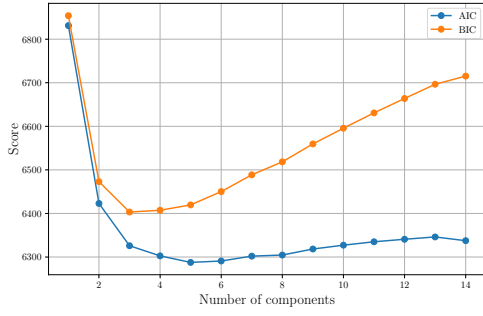


Figure 1: AIC and BIC scores of the GM for different numbers of components

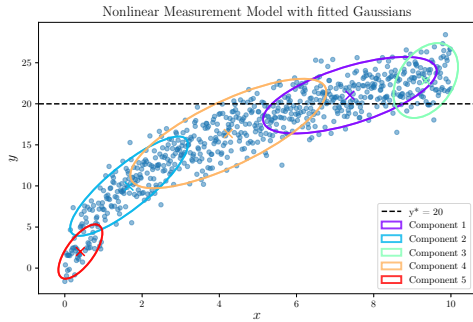


Figure 2: Measurement model with an observed value $y^* = 20$

interpreted as taking a slice through the joint Gaussian mixture along the observed dimensions. Assume that the input variables are observed as $\mathbf{y}^* = [y_1^*, y_2^*, \dots, y_i^*]^T$. The posterior distribution of the output quantity x is then obtained using Bayes' rule

$$p(\mathbf{x} | \mathbf{y}^*) = \frac{p(\mathbf{x}, \mathbf{y}^*)}{p(\mathbf{y}^*)}. \quad (6)$$

The denominator represents the marginal density of the joint distribution evaluated at the observed input values. It is obtained by integrating the joint density over the output variable

$$p(\mathbf{y}) = \int p(\mathbf{x}, \mathbf{y}) d\mathbf{x}. \quad (7)$$

Due to the properties of Gaussian mixtures, this conditioning step can be performed analytically for each mixture component. As a result, the posterior distribution $p(\mathbf{x} | \mathbf{y}^*)$ is again represented as a Gaussian mixture of lower dimension. This enables efficient inference of the output uncertainty for arbitrary observations of the input variables.

The conditional distribution of the output variable for component k is therefore given by (see [2] [1]):

$$\mu_{k,x|y} = \mu_{k,x} + \Sigma_{k,xy} \Sigma_{k,yy}^{-1} (y^* - \mu_{k,y}), \quad (8)$$

$$\Sigma_{k,x|y} = \Sigma_{k,xx} - \Sigma_{k,xy} \Sigma_{k,yy}^{-1} \Sigma_{k,yx}. \quad (9)$$

The mixture weights are updated according to the likelihood of the observation under each component

$$L_k = \mathcal{N}(y^* | \mu_{k,y}, \Sigma_{k,yy}), \quad (10)$$

leading to the normalized conditional weights

$$\tilde{\pi}_k = \frac{\pi_k L_k}{\sum_{j=1}^K \pi_j L_j}. \quad (11)$$

The posterior distribution $p(\mathbf{x} | \mathbf{y}^*)$ is therefore again a Gaussian mixture with updated parameters $\{\tilde{\pi}_k, \mu_{k,x|y}, \Sigma_{k,x|y}\}_{k=1}^K$.

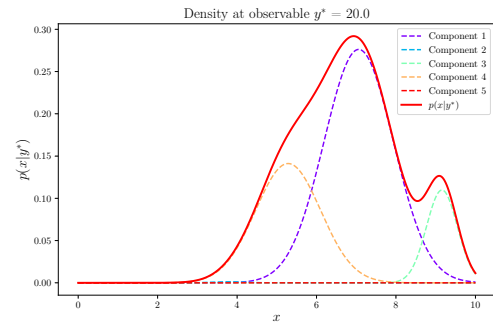


Figure 3: Conditional posterior $p(x | y^*)$ GM for an observed value $y^* = 20$

Figure 3 shows the posterior PDF after an observation for the given exemplary measurement model.

3 Application

3.1 Generation of fatigue strength samples

Determining the local fatigue strength $R_w(x)$ of materials in used components is an important quality indicator for assessing the product's remaining service life and plays an important role in the context of the circular factory in determining whether a material can be reused. Fatigue strength is a quantity that depends on the depth r into the material. However, fatigue strength cannot be measured directly, instead, other parameters must be measured to infer this quantity. Those measurable parameters are surface roughness R_z , hardness HV , and residual stress σ_{ES} .

In the experiment, each of these quantities was measured repeatedly, yielding a mean value μ and a standard deviation σ that describe the measurement uncertainty. Table 1 summarizes the obtained statistics. In this experiment, all quantities are considered at a depth of $r = 0$, corresponding to the material surface. The fatigue strength was calculated using nonlinear analytical relations that describe the influence of the considered surface and material parameters. The formulas to calculate fatigue strength stem from Winderlich [9]:

Table 1: Measured input variables (depth $r=0$).

Variable	Mean	Standard deviation
R_z	$6.66 \mu\text{m}$	$0.89 \mu\text{m}$
HV	677.5 HV	11.4 HV
σ_{ES}	348.4 MPa	4.15 MPa

The static tensile strength is proportional to the hardness:

$$R_m(r) = 3.3 \cdot HV(r) \quad (12)$$

Subsequently, the surface influence factor is:

$$\omega = 1 - 0.22 \cdot \log(R_z) \left(\frac{\log(R_m(r=0))}{20} - 1 \right). \quad (13)$$

The following quantity named Local stress-free alternating load resistance is computed:

$$R_w^0(r) = \frac{1.98HV(r) - 0.0011 \cdot HV(r)^2}{1 + 20.7/HV(r)} \quad (14)$$

Finally, the fatigue strength is:

$$R_w(r) = \omega \cdot R_w^0 \cdot \log(R_z) \cdot \left(1 - \frac{\sigma_{ES}(r)}{R_m(r)} \right). \quad (15)$$

Because the model equations are nonlinear, analytical uncertainty propagation is not straightforward. Therefore, a Monte Carlo simulation was employed to propagate the input uncertainties through the fatigue strength model. Each input quantity was modeled as a normal distribution $\mathcal{N}(\mu, \sigma)$ defined by its mean and standard deviation, from which samples were drawn for every simulation run. The sampled inputs were then inserted into the fatigue strength equations, resulting in a large set of simulated fatigue strength realizations. This procedure yields a numerical approximation of the joint distribution of the input variables and the resulting fatigue strength. In this case, the output quantity depends on three input variables. Therefore the measurement model is four-dimensional.

3.2 Estimation of the probabilistic measurement model

The generated dataset of input-output samples was used to estimate a probabilistic measurement model of the fatigue strength. For this purpose, the joint distribution of the fatigue strength and the input variables was approximated by a four-dimensional GM. The parameters of the mixture model were estimated using the EM algorithm. To determine the optimal number of components, the AIC and BIC were calculated for different numbers of Gaussian components. The trend of the scores over the number of components is shown in Figure 5. The optimal number of components according to BIC is 30. For this case, the joint PDF is demonstrated in Figure 4. To visualize the four-dimensional distribution

effectively, the marginal distribution was plotted for each input variable.

The resulting Gaussian mixture provides a compact parametric representation of the joint probability density of the fatigue strength and the influencing input variables. In contrast to deterministic models, this probabilistic representation captures nonlinear relationships as well as possible multimodal characteristics of the distribution.

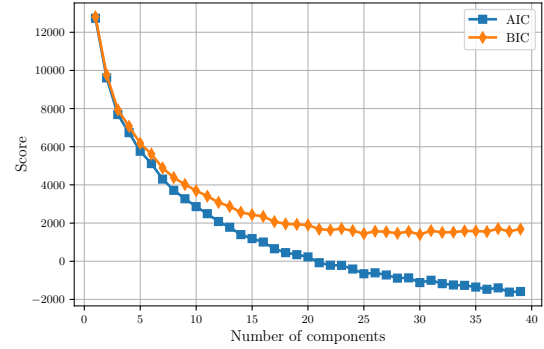


Figure 5: AIC and BIC score for different numbers of components of the 4-dimensional GM

3.3 Posterior inference of fatigue strength

Once the probabilistic measurement model has been estimated, it can be used to infer the probability distribution of the fatigue strength for observed input parameters. Observations of surface roughness, hardness, and residual stress correspond to conditioning the joint Gaussian mixture model on the measured input variables.

By conditioning the joint model on the observed values of the input variables, the posterior distribution of the fatigue strength is obtained. This posterior distribution represents the uncertainty of the fatigue strength given the measured material parameters. Figure 6 shows an example of the posterior probability density of the fatigue strength for a specific set of observed input variables, meaning $HV^* = 670 \text{ MPa}$, $\sigma_{ES}^* = 345 \text{ MPa}$ and $R_z^* = 7 \mu\text{m}$. The resulting distribution reflects both the uncertainty of the measurements and the nonlinear propagation of uncertainty through the fatigue strength model. The results demonstrate that the proposed probabilistic framework allows the fatigue strength distribution to be efficiently inferred for arbitrary combinations of observed material parameters. This provides a flexible measurement model that can directly quantify the uncertainty of the fatigue strength based on experimentally observed input quantities. [9].

4 Evaluation

In this example, a direct comparison between the proposed method and the GUM framework is not

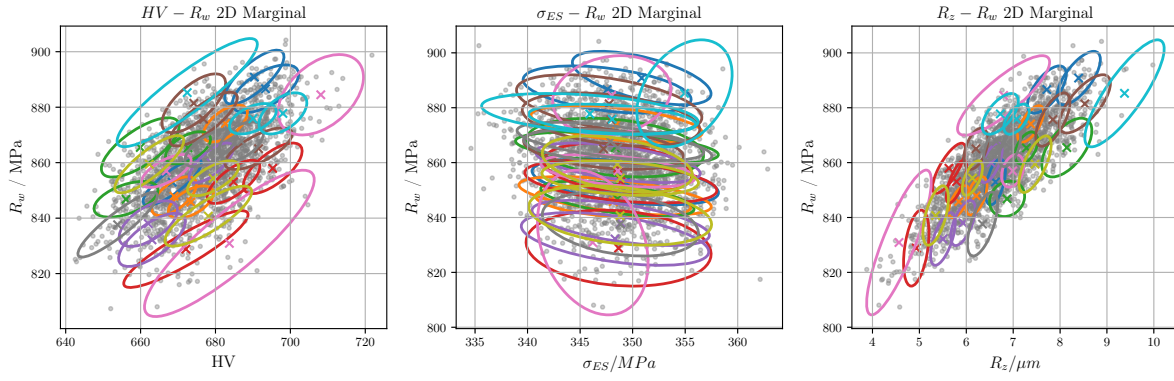


Figure 4: Joint PDF of the Fatigue strength measurement model with 30 GM components

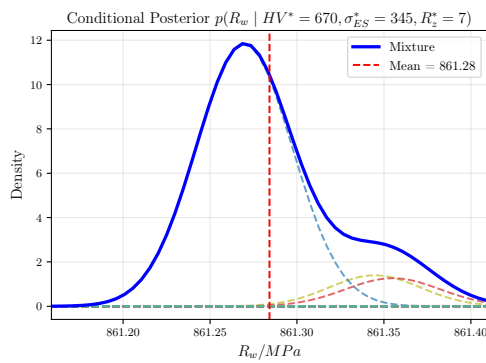


Figure 6: Conditional Posterior with 30 GM components

straightforward, as the two approaches address fundamentally different problem formulations. The classical GUM approach focuses on forward uncertainty propagation, where uncertainties in input quantities are propagated through a measurement model to obtain the uncertainty of an output quantity. In contrast, the method proposed in this work addresses an inverse problem. Starting from an observed measurement, it aims to determine the posterior uncertainty of the underlying true quantity. To this end, both the measurement values and their associated uncertainties are incorporated directly into the probabilistic model, resulting in a multi-dimensional Gaussian mixture representation. From this model, conditional posterior distributions can be derived for given observations.

Applying the GUM framework in this context would require predefined input quantities along with their associated uncertainties. In this example, the standard deviations reported in Table 1 can be interpreted as measurement uncertainties and used within a GUM-based propagation. However, such an approach does not account for the original measurement in a probabilistic inference sense and therefore does not yield a posterior distribution of the true quantity.

Consequently, while both approaches operate on similar data, they differ fundamentally in their objective: the proposed method performs probabilistic inference of the true quantity given an observation, whereas the GUM framework propagates uncertainty from inputs to outputs.

Figure 7 shows the PDF of the observed values including uncertainty. The Monte Carlo-based kernel density estimate (KDE) is approximately Gaussian in this case. Compared to the posterior distribution obtained with the proposed method (Figure 6), differences in the estimated uncertainty become apparent. In particular, the standard deviation obtained using the GUM-based approach is 14.8, while the Monte Carlo method yields a slightly larger value of 15.3, while the standard deviation of the Conditional Posterior in Figure 6 is 0.04.

These differences can be attributed to the fundamentally different treatment of uncertainty. While the GUM and Monte Carlo approaches describe the uncertainty of the observed quantity via forward propagation, the proposed method estimates the posterior distribution of the underlying true quantity. As a result, the proposed approach can lead to a different uncertainty characterization, particularly in the presence of nonlinearities or non-Gaussian effects.

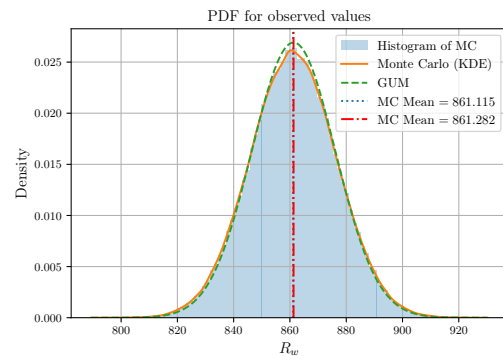


Figure 7: PDF for the GUM and Monte-Carlo case

5 Conclusion

This study presents a robust probabilistic framework for quantifying measurement uncertainty in complex, nonlinear systems, specifically tailored for the challenges of the Circular Factory. The proposed method overcomes significant limitations of conventional approaches by representing the entire measurement model, including the measurand and influencing input quantities, as a joint probability density function approximated by a GM.

In contrast to the standard GUM linearization approach, which relies on local approximations and may become inaccurate for strongly nonlinear or non-Gaussian systems, the proposed method enables a more flexible representation of uncertainty. By leveraging the EM algorithm, the mixture parameters can be estimated directly from data, while model selection criteria provide guidance on the appropriate model complexity. Furthermore, the analytical conditioning of Gaussian mixtures allows the efficient computation of posterior distributions for given observations, offering a computationally attractive alternative to repeated Monte Carlo simulations in scenarios requiring multiple evaluations. The practical application to fatigue strength estimation demonstrates the framework's versatility. By integrating surface roughness, hardness, and residual stress into a unified probabilistic model, we successfully propagated uncertainties through highly nonlinear analytical relations. The results confirm that the method not only yields accurate point estimates but also provides a comprehensive description of the uncertainty distribution, including skewness and multimodality, which are critical for decision-making in reconditioned components. Ultimately, this approach establishes a foundational tool for data-driven decision-making in circular production systems, ensuring that uncertainties are rigorously accounted for when determining the remaining service life of reused components.

Acknowledgement

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the collaborative research center (CRC) 1574 "Circular Factory for the Perpetual Product" with the project ID 471687386.

References

- [1] Ali Darijani et al. *Comprehensive Description of Uncertainty in Measurement for Representation and Propagation with Scalable Precision*. 2026. arXiv: 2603.20365 [stat.ML]. URL: <https://arxiv.org/abs/2603.20365>.
- [2] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 0387310738.
- [3] Joint Committee for Guides in Metrology (JCGM/WG 1). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. 1st edition, corrected version 2010. JCGM 100:2008 (GUM 1995 with minor corrections). Sèvres, France: Bureau International des Poids et Mesures (BIPM), 2008. URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf.
- [4] P M Harris and M G Cox. "On a Monte Carlo method for measurement uncertainty evaluation and its implementation". In: *Metrologia* 51.4 (2014), S176. DOI: 10.1088/0026-1394/51/4/S176. URL: <https://doi.org/10.1088/0026-1394/51/4/S176>.
- [5] Radoslav Marinov. "Comparison of GUM and Monte Carlo Methods for Evaluation the Measurements Uncertainty of DC Voltage, DC Magnetic Flux Density and Dc Resistance". In: *2025 XXXV International Scientific Symposium Metrology and Metrology Assurance (MMA) (2025)*, pp. 1–4. URL: <https://api.semanticscholar.org/CorpusID:284358630>.
- [6] Antonio Piratelli-Filho and Benedito Giacomo. "Uncertainty evaluation in small angle calibration using ISO GUM approach and Monte Carlo method". In: (Jan. 2003).
- [7] Jürgen Beyerer, Raphael Hagmanns, and Daniel Stadler. *Pattern Recognition*. Berlin, Boston: De Gruyter Oldenbourg, 2024. ISBN: 9783111339207. DOI: 10.1515/9783111339207. URL: <https://doi.org/10.1515/9783111339207>.
- [8] Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022. URL: <http://probml.github.io/book1>.
- [9] B. Winderlich. "Das Konzept der lokalen Dauerfestigkeit und seine Anwendung auf martensitische Randschichten, insbesondere Laserhärtungsschichten". In: *Materialwissenschaft und Werkstofftechnik* 21.10 (1990), pp. 378–389. DOI: <https://doi.org/10.1002/mawe.19900211006>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mawe.19900211006>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mawe.19900211006>.