TIM-23-04989R1

# Addressing uncertainty on machine learning models for long-period fiber grating signal conditioning using Monte Carlo method

Felipe Oliveira Barino and Alexandre Bessa dos Santos

***Abstract*—The massive adoption of machine learning (ML) and artificial intelligence models in the field of instrumentation and measurement has raised several doubts concerning the validity of their response and the methodology for estimating their errors. In this study, we revisit ML models that were used to interrogate long-period fiber grating sensors. We used these models to present a comprehensive analysis of the uncertainty propagation through the ML-based optical fiber sensor signal conditioning. The uncertainty propagation was studied using the Monte Carlo method. The results showed the proposed models were capable to damp some optoelectronic noises, don't induce systematic errors under noise, and that the noise-damping effect of the ML models doesn't impact the interrogator's resolution. Moreover, we hope that this work serves as a methodological framework for the evaluation of uncertainty of ML-based optical sensor interrogators.**

***Index Terms* — fiber optic sensors, error statistics, prediction theory, instruments**

## I. INTRODUCTION

In the recent years we have seen consistent progress in machine learning (ML) and data-driven techniques. Consequently, one could note reports of ML applications in several fields, from medicine to stock market [1], [2]. Consequently, this trend was no different for sensors, instrumentation, and measurement. Some applications of ML in the field of instrumentation and measurement are sensor data processing [3], [4], sensor fusion [5]–[7], and sensor calibration [8]–[10]. Consequently, this trend was also adopted for fiber-optic sensors (FOS), from discrete sensors [11]–[16], to distributed optical fiber sensors [17]–[20].

In the FOS field, the signal conditioning is often called interrogation, and involves measuring one or more light characteristics that are then used to calibrate the sensor [21]. For optical fiber sensors that rely on the light's wavelength to estimate the measurand (wavelength-encoded transducers), this task is complicated and demand the use of complex and/or expensive optoelectronics. A notable example is the fiber Bragg gratings (FBGs), in which the reflected light's spectrum is a narrow peak centered at the Bragg wavelength. This Bragg wavelength is a function of temperature and strain, allowing the use of FBGs as transducers. Therefore, one needs the transducer's spectrum to detect the Bragg wavelength and calculate the measurand. Indeed, most of the FBG interrogators use spectrum measurement and peak tracking algorithms to identify the Bragg wavelength [22].

Another type of FOS that requires spectrum measurement is the long-period fiber grating (LPFG). While in FBGs, the Bragg wavelength is a well-defined peak center, in LPFGs the optical spectrum is much more complex. Instead of a single peak, the LPFG spectrum has several rejection bands, each one due to a resonant wavelength, $\lambda_{res}$. In this case, the measurand is calculated by one (or more) resonant wavelengths.

Nonetheless, several works focus on the development of simpler forms of FBG and LPFG interrogation. Note that, the main parameter often needed for sensor calibration is a single wavelength, which represents the FBG's peak or one LPFG's dip. Thus, the use of complex optoelectronics to acquire the whole sensor spectrum might be inefficient in certain cases, due to energy consumption, equipment size, price, and general in-field compatibility. In this scenario, the development of simpler interrogators might improve the massive adoption of FOS, especially for industrial and even consumer-grade applications.

Recent approaches found in the literature focus on single (or sparse) acquisition of optical power, i.e., the use of one (or more) optical filter and the detection of the filtered optical power. One method is based on the use of an arrayed waveguide grating (AWG) to filter the sensor spectrum and estimate the desired parameter. Due to the AWG's dense channel spacing, the FBG peak can be easily estimated by neighboring channels at the FBG vicinity [23], with possibility to improve the cost-effectiveness when coupled to a smart channel-selecting system assisted by deep learning [24]. Note that an FBG peak width is as low as a fraction of a nanometer, whereas an LPFG dip might be as wide as tenths of nanometers. Therefore, the channel density of AWGs allows its use for LPFG spectrum estimation [25].

However, other methods rely on edge filtering, with limited number of filters. To estimate an FBG position, for example, an LPFG could be used as edge filter [26], or other FBGs close to the sensing FBG [27]. Similarly, to interrogate LPFGs, one could use one or more FBGs as edge filters [28], [29], or a

This work was supported in part by the Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Instituto Nacional de Energia Elétrica (INERGE-UFJF), and Santo Antônio Energia.

F. O. Barino is with the Department of Circuits, Federal University of Juiz de For a, Brazil (e-mail: alexandre.bessa@engenharia.ufjf.br).

A. B. dos Santos is with the Department of Circuits, Federal University of Juiz de For a, Brazil (e-mail: felipe.barino@engenharia.ufjf.br).

The data that support the findings of this study are available from the corresponding author, F. Barino, upon reasonable request.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

sparse filter bank [30], [31].

It is important to note that the use of filters to measure the spectral characteristics of sensors requires a high degree of technical expertise, because a careful and well-designed experimental setup is necessary to obtain accurate and meaningful results. Once the filtered power is highly dependent on LPFG transmission characteristics, the results could be highly sensor-dependent, i.e., vary a lot for sensors with the same resonant wavelength, but different coupling efficiency, for example. Therefore, the interrogator is often tailored to a single sensor or to an extremely limited set of sensors. However, the use of a filter bank coupled to ML methods solved this limitation, by fusing information from a wide range of sensor transmission characteristics [30], [31].

Fig. 1 shows the scheme for a ML-assisted sparse filter interrogator. Note the intricate relationship between the power readings and the LPFG resonant dip position depends on several factors, such as LPFG full width at half maximum, coupling efficiency, and resonant dip asymmetries. Hence, the use of machine learning to process the optical power readings and output the LPFG position improved the signal processing due to the high level of abstraction of such models.

In this work, we evaluate the LPFG interrogators based on sparse filter bank coupled to ML models under a metrology perspective. We revisit the models and dataset presented in [31], to preset a study on the uncertainty and its propagation through the interrogator signal conditioning ML model.

Note that the use of ML models in the signal processing chain of optical fiber sensors is widely adopted. Machine learning models has been used for multiparameter discrimination [15], [11], [32], [33], for temperature profile reconstruction based on FBG array data [34], for tilted FBG sensor interrogation [35], [36], FBG data fusion and weight estimation [13]. However, those works lack an error evaluation method that follows the best practices of the instrumentation and measurement field.

Note that the machine learning models are difficult to interpret, thus the uncertainty characteristic and its propagation through the model is difficult to evaluate, especially when different ML models are being compared. In this work we present a framework for uncertainty characterization in a ML-based LPFG interrogator using the Monte Carlo method.

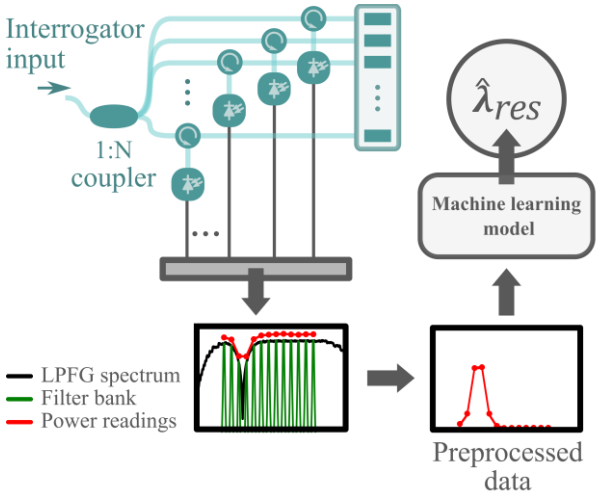


**Fig. 1.** Schematic of a sparse filtering and ML assisted LPFG interrogator.

The goal of this work is to inspect the uncertainty propagation through the ML models. Hence, we studied how minor fluctuations on the LPFG's resonant wavelength were processed by the ML models and inspected the models' noise suppression capability and how this feature impacted the interrogator's resolution. The models considered in this work were: Linear model, multilayer perceptron (MLP), fuzzy inference system with expert training (FIS expert), and fuzzy inference model with genetic algorithm optimization (FIS GA opt.) [31].

## II. Methods

### A. The sparse filter LPFG interrogator

The LPFG interrogator studied in this work is schematized in Fig. 1. The optical input, i.e., the light transmitted by an LPFG sensor enters a coupler to split the light between 13 FBG filters with equidistant Bragg wavelengths, forming the sparse filter. Circulators, photodetectors, and ADCs turn the filtered power into information for processing.

Hence, the filtered power is the ML model's input. The raw data was preprocessed to reduce the optical source influence and normalize its intensity to unit sum. First, the LPFG sensor's filtered power was subtracted from the filtered optical source (the raw data without a sensor), obtaining a 13 elements vector representing the LPFG transmission. Then each element was divided by the vector sum [31] . The resulting vector was called input strength, and each element represents how close an LPFG resonant dip is to its respective filter element.

An LPFG spectrum alongside the filtered power and input data can be seen in detail in Fig. 2. The model's target, i.e., desired output is the LPFG resonant wavelength and is also shown (~1552 nm for this example).

### B. The ML models

All models had 13 inputs (input strength) and one output (LPFG resonant wavelength). The models used were multilayer perceptron (MLP), fuzzy inference system with expert learning (FIS expert), and fuzzy inference system with genetic algorithm learning (FIS GA opt.). We also used a linear regression model as a simpler baseline model. All the models were pretrained, in fact, we used the same models as in [31], without further

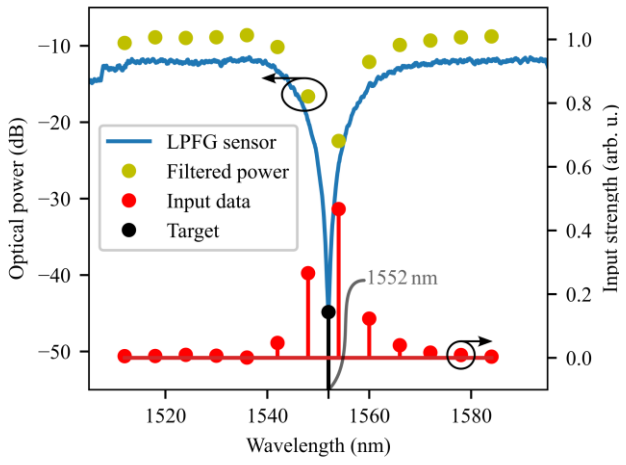


**Fig. 2.** Sparse filtering and input data acquisition.

optimization or training.

The MLP had 13 input nodes, 5 neurons in the hidden layer, and a single output neuron. The activation function considered was the rectified linear unit (ReLU).

The fuzzy systems had three modified gaussian membership functions (equal for all 13 inputs). They represented a LPFG centered at the given input filter, near this filter, or far from it. The fuzzy sets were named CENTERED, NEAR, and FAR, respectively. The fuzzified variables were processed by a Mamdani inference engine with simple rules that related the fuzzy sets to the output sets. The output sets represented were the LPFG was positioned, and the rule base took into consideration all the inputs. If the input indicated that the LPFG was centered to this input but far from all others, the output set equivalent to the filter's Bragg wavelength was activated. Similarly, if two inputs indicated that the LPFG was near them, the output set centered between these two inputs was activated.

The output was defuzzyfied by the weighted average method, to estimate the LPFG position. Two FIS were considered, one trained with expert insights on data and another using genetic algorithm. The trainable parameters were simply the input membership functions.

The baseline model was fully vectorial, its output was calculated by the inner product between the input strength vector and the FBG position vector, then calibrated to the training data. Further details on these models can be found in [30], [31].

Previous work using these models showed promising results for machine learning methods. When compared to a regression model, the error of ML-based models was half the baseline regression model [31], see Table I. One can assign this improvement to the high abstraction capability of such models, with ability to learn complex and high dimensional relationships using examples. However, there are limitations, specifically regarding generalization and results for data far away from the training examples. Thus, is important to well characterize and address the uncertainty in ML models, especially for those in instruments signal conditioning.

### *C. Uncertainty in ML models*

The uncertainty in ML models is divided between two types: epistemic and aleatoric [37]. The first is the model uncertainty and is caused by the lack of knowledge regarding an input, therefore it occurs for points far from the training set representativeness. The second occurs due to the variability on the outcome of an event, or training data. For the instrumentation and measurement point of view, these are similar to the systematic and aleatoric (stochastic) errors found in measurements.

TABLE I – ML MODELS CONSIDERED AND THEIR STATISTICS, AS ESTIMATED IN [31].

| | Linear model | MLP | FIS expert | FIS GA opt. |
|---|---|---|---|---|
| Mean abs. error (nm) | 0.762 | 0.668 | 0.443 | 0.352 |
| Men sq. error ($nm^2$) | 0.979 | 0.737 | 0.317 | 0.242 |
| Bias (nm) | -0.031 | -0.083 | -0.124 | -0.101 |
| Standard variation (nm) | 0.989 | 0.854 | 0.549 | 0.481 |

To address the uncertainty in complex machine learning models, authors have proposed several methods for uncertainty quantification and modelling. These methods could be based on the Bayesian framework [38], on ensemble of models [39], or on the evidential framework [40]. In [41] the authors present a comprehensive review on the most relevant methodologies for uncertainty quantification in deep learning, exploring their advantages and disadvantages. These methodologies are model-focused and often aim to provide an output coupled to its uncertainty, to improve decision making, especially on sensitive cases.

The best practices in data-driven and machine learning modelling state that the dataset should be split. The train-test split ensures that the model is optimized (trained), using a set of data different from that used to evaluate the model [42], [43]. Hence, one could simulate the model behavior under new data and estimate the model's ability to extrapolate. Therefore, performance statistics of the models, such as mean squared error (MSE) and mean absolute error (MAE), must be estimated using the test set. However, the test-set evaluation performed in the lab might not translate to the field due to uncertain nature of in-field applications. Thus, it is important to characterize the challenges that might surge during the model effective use [44], to identify if eventual faults are due to physical problems (such as data acquisition) or model errors [45].

As consequence, in order to effectively train a model, a large amount of data is necessary. Additionally, to minimize the epistemic uncertainty, this data should be representative of the model use case. I.e., training data should be as similar to the data presented after model implementation as possible. Recently, researchers have approached this data size challenge using transfer learning [46], because it can be an workaround for the difficulty to acquire in-field data for model training, and also used synthetic data for training deep learning models [47]–[49].

In the context of LPFG signal conditioning, to effectively train an LPFG interrogator, the data should come from several sensors. It must cover the entire interrogator spectral window and present a wide range of spectral characteristics. This should ensure the interrogator's ability to estimate the resonant wavelength of any unknown LPFG sensor (with characteristics similar to the training set sensors), thus minimizing systematic error.

However, in this case, there is a tradeoff between spectral variability and single-sensor evaluation. A high degree of sensor variability might be difficult to achieve in practical terms, thus one could prioritize training with a limited number of sensors, with several measurements per sensor. On the other hand, the use of several sensors might impose limitations on single-sensor variability, due to limited time and resources. While the former might improve the model's resolution and the aleatoric uncertainty analysis using the test set, it lacks variability. Thus, the trained model might be unable to operate for unknown sensors. The latter, on the other hand, prioritizes

the model extrapolation ability (reduce epistemic uncertainty) at the cost of poor aleatoric uncertainty estimation.

Note that the ideal case is to provide an interrogator with minimal epistemic uncertainty and good aleatoric uncertainty characterization, so the instrument can be well specified and reliable.

The approach presented in [31] favors extrapolation and lacks a discussion on measurement uncertainty, i.e., how the interrogator might perform for fluctuating single-sensor measurements. An analysis of the models' performance under a wide variety of LPFGs was presented, in which the model bias and variance were estimated based on the residual statistics calculated using the test set. However, the work lacks the analysis of this single-sensor wavelength fluctuation.

Considering that the ML models used to perform the LPFG interrogation are simple, compared to the ML state-of-the-art, it is reasonable to assume that the epistemic uncertainty is low. Indeed, the models considered in this work have limited parameter space, compared to the training samples. Therefore, in this work we focus on the aleatoric uncertainty only.

One of the advantages of machine learning is the robustness and noise suppression ability, which was shown in [31] for several signal-to-noise ratio (SNR). However, for the instrumentation point-of-view this ability might affect the measurement. Note that an intelligent noise suppression system could reduce the instrument's resolution, due to excessive filtering of input fluctuations. We studied this effect by analyzing the propagation of LPFG uncertainty through the ML model. Ideally, all the fluctuations observed at the LPFG should be transferred to the ML model output, so there is no resolution loss.

### *D. The data*

All models used in this work was pretrained. Hence, in this work, only the test set was considered. However, the data came from 83 LPFG sensors and compromised of 528 spectra, randomly split into training and test sets. The training was 369 examples (LPFG spectra). The test set, on the other hand, had 159 LPFG spectra. All the data was as clean from noise as possible.

The noise was generated latter added to the test set, using the Monte Carlo method, after the simulations a total of 63600 test set input/output pairs were obtained regarding different noise characteristics.

### *E. The Monte Carlo simulation*

In this work, the models were evaluated using the Monte Carlo method, to estimate the uncertainty propagation through them. The equivalence between the Monte Carlo method for uncertainty calculation and the guide for uncertainty in measurement (GUM) methodology has been previously shown for a wide range of applications [50]–[52] and thus we followed

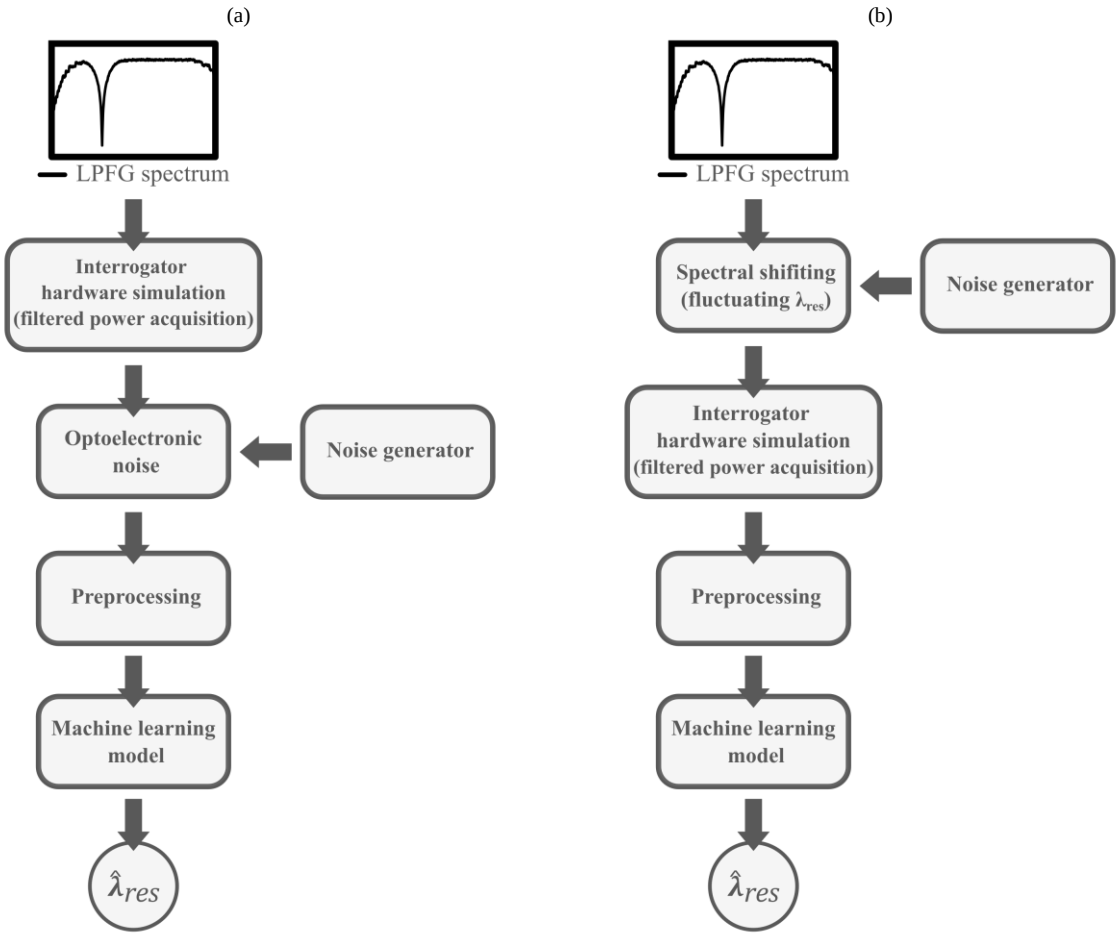


**Fig. 3.** Procedure for uncertain measurement generation for the Monte Carlo method: (a) uncertainty due to optoelectronic noise and (b) uncertainty due to LPFG sensor fluctuation.

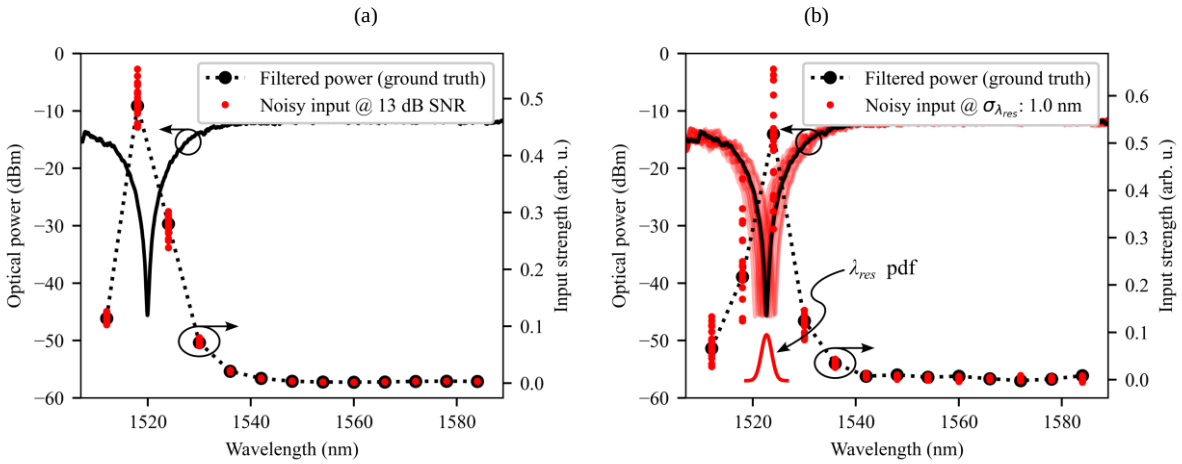


**Fig. 4.** Example of uncertainties at the interrogator input: (a) uncertainty due to optoelectronic noise and (b) uncertainty due to LPFG sensor fluctuation.

a similar approach in this work.

We considered the model's input as the main source of uncertainty. However, the input changes due to the interrogator's optoelectronics and to the LPFG position. So, we used semi-synthetic data to evaluate the model, by modifying the test set using simulations. The modifications were: addition of noise to the filtered power and LPFG resonant dip shifting, that represented the change due to optoelectronic noise and to LPFG position, respectively. The step-by-step procedure used to simulate a single uncertain measurement is shown in Fig. 3. Thus, in this work, these steps were repeated 200 times for each test-set spectrum, for each type of modification (optoelectronic noise and LPFG fluctuation).

To make a single measurement with noisy power measurements, the interrogator hardware was simulated using a test set spectrum to output the optical power filtered by each filter element. Then we added additive gaussian noise (AWGN) at the inputs, as described in [31]. The resulting array was then processed by the ML algorithm to estimate the LPFG resonant wavelength. See Fig. 3a for reference.

To make a single measurement with noisy LPFG, a test set spectrum was shifted (the resonant dip was moved) to a given resonant wavelength. Then the hardware was simulated to output the filtered power and the resulting array processed by the ML model to estimate the resonant wavelength. See Fig. 3b for reference.

Fig. 4 shows the results of such steps on an LPFG spectrum. In this figure, the black parameters are the static (ground truth), whereas the red values show the noise added during the Monte Carlo simulations, performed by the steps shown in Fig. 3. We repeated these steps 20 times for each case, simulating 20 measurements.

Whereas the effect of AWGN at the inputs was described in [31] to assess the hardware impact, the effect of slight deviations in the LPFG position has not been investigated yet.

To generate the noise, the function `numpy.random.randn` was used. These values were used to add the hardware noise, proportionally to the input filtered power (given the SNR) and to shift a given LPFG spectra, thus obtaining new optical power at each filtering element by simulating the hardware and preprocessing the filtered data array accordingly to the method described in [31].

The ML model input for the 20 noise simulations shown in Fig. 4 can be seen in Fig. 5, where the color intensity is related to the input strength. This figure illustrates well the input uncertainty due to: a) optoelectronic noise and b) FOS fluctuations. In this section and the following, the respective uncertainties will be called $u_a$ and $u_b$ and they are associated,

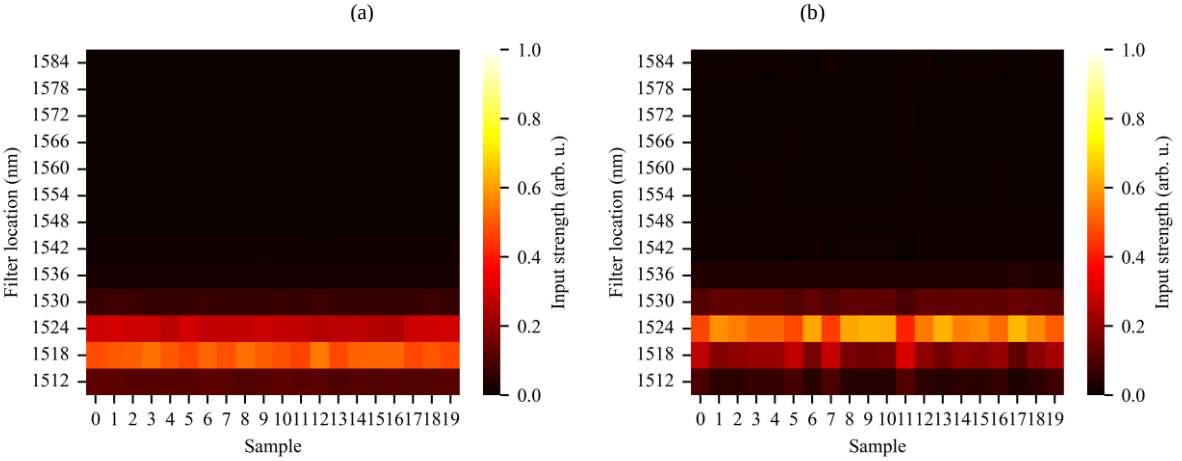


**Fig. 5.** Time series of the input values shown in Fig. 4.

respectively, to an input SNR and random LPFG noise of standard deviation $\sigma_{\lambda_{res}}$.

This process of acquiring 20 samples for each test-set spectrum was repeated 10 times for several input uncertainties, resulting in 200 simulations for each test set spectra. These represented 20 measurements over 10 experiments. The values used as $u_a$ were 6, 8, 11, 15, 20 dB, whereas the values for $u_b$ were 0.10, 0.15, 0.24, 0.37, 0.57, 0.88, 1.36, 2.10, 3.24, and 5.00 nm.

### *F. Evaluation metrics*

The results were evaluated as follows: the 159 test set spectra were processed by the ML models, at each iteration, the estimated resonant wavelength $\hat{\lambda}_{res}$was compared to the target $\lambda_{res}$, thus obtaining the residues $r = \lambda_{res} - \hat{\lambda}_{res}$. The residue bias and standard deviation were then estimated, for each iteration, based on the 20 samples per spectra per experiment. Finally, we evaluated how these metrics behaved for the 10 experiments, studied the uncertainty propagation, and the impact of the noise robustness on the interrogator resolution.

In the results section we considered the mean and standard deviation of the results obtained during the 10 experiments calculated.

## III. Results

### *A. Uncertainty due to optoelectronic noise*

First, let us analyze the effects of the optoelectronic noise, i.e., the input fluctuations of a fixed LPFG sensor. Fig. 6 shows the interrogator's bias as a function of $u_a$, for each model. Note that, generally, the bias doesn't present a clear trend with respect to the SNR, except for the Linear model under high noise, but with great variability among our tests. The standard deviation, on the other hand, shows a clear increasing trend correlated to the input noise, see Fig. 7. These results are in accordance with the results shown for the MSE in [31], where it was shown that the ML models, especially the FIS models, are less susceptible to noise.

Based on such results, we were able to infer that the input noise doesn't affect the bias of the more robust models, MLP and FIS, tested in this work. Therefore, we calculated the typical range (with 95% confidence interval) for the input bias for the $u_a$ values considered. We found that the Linear model bias lies in the range of 0.240 ~ 0.028 nm, the MLP in the 0.074 ~ 0.093 nm range, the FIS-ex in the 0.123 ~ 0.124 nm, and the FIS-ga in the 0.101 ~ 0.102 nm range. These values were also in accordance with the results found previously and are important to better characterize the possible models for the interrogator processing.

The high noise insensitivity of the FIS-based models could be attributed to the fact that the logic used in these models focus on the FBGs close to the LPFG's resonant wavelength. Indeed, in these models, only the data from FBGs close to the resonant wavelength are used. Thus, once the filtered power is fuzzified, the minor fluctuations on the far-from $\lambda_{res}$ FBGs doesn't affect the resonant wavelength estimation, because the changes aren't great enough to be considered part of the NEAR or CENTERED fuzzy sets (please see [31] for a detailed description of the fuzzy sets).

Even though the MLP model used the whole input data to estimate the resonant wavelength, similarly to the Linear model, its sensitivity to noise was much lower. Certainly, as MLP are well known for their noise immunity [53].

Regarding the Linear model, once the input is a function of the input strength and its position inner product, little noisy variations on FBGs very far from the LPFG dip (low input strength), impact a lot on the output. Indeed, since a contribution of a very far wavelength induce excessive error on the estimation. Moreover, note that as the noise grows, the bias becomes very difficult to estimate, because the impact of very far inputs become much more prominent. At the limit, the bias variability could be as big as the interrogator's full scale, since extreme filters might have high noise.

This effect illustrates the "attention" attribute of the ML-models tested. We believe that the low deviation found for those models might be due to the minimal effect of input strengths that filter far from the important ones (close to the target resonant wavelength) have on the estimation.

Note that the metrics' variance was low over the 200 simulations (20 measurements over 10 experiments), indicating that the number of Monte Carlo iterations were sufficient in this case. Except for the Linear model residuals' bias under high noise; however, as discussed, this effect was due to the model itself.

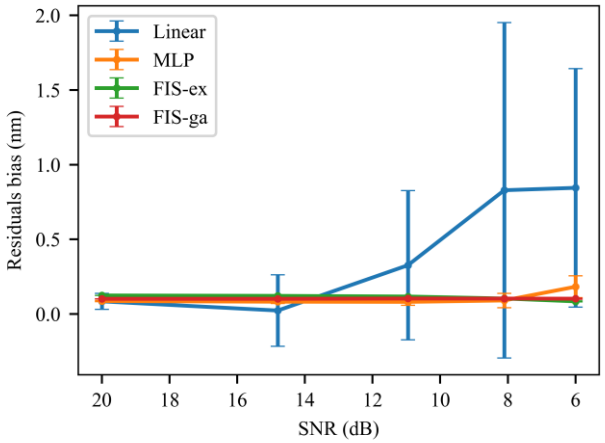


**Fig. 6.** Bias due to input power SNR.

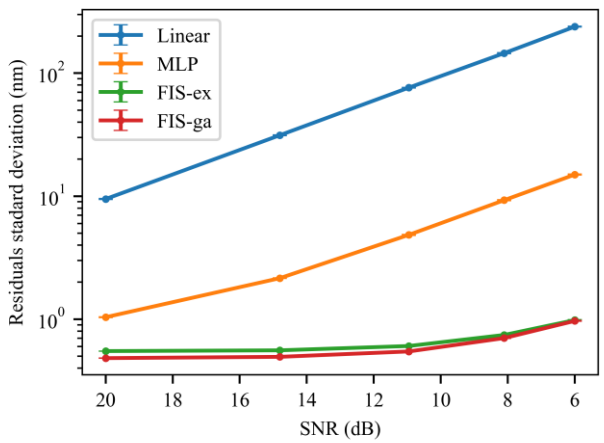


**Fig. 7.** Standard deviation due to input power SNR.

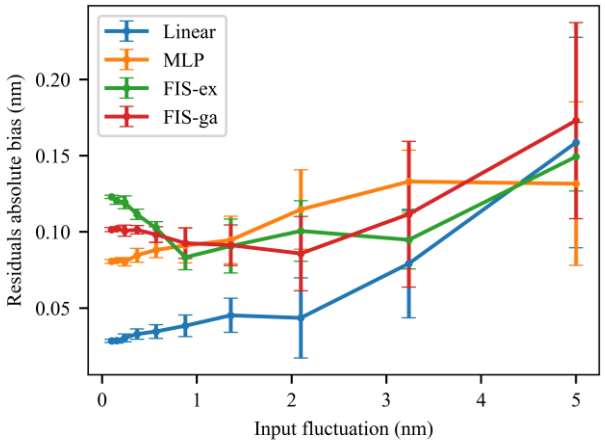


**Fig. 8.** Bias due to LPFG sensor fluctuation.

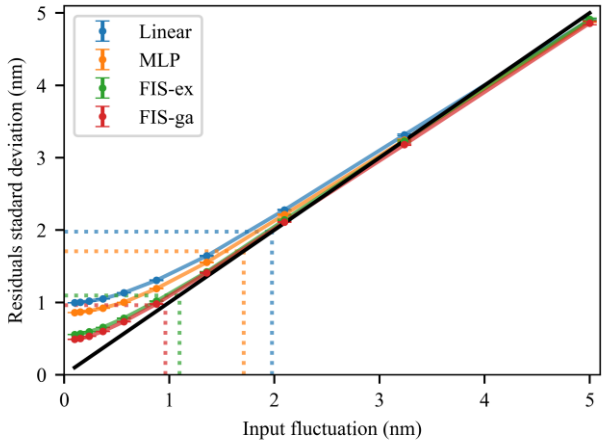


**Fig. 9**. Standard deviation due to LPFG sensor fluctuation.

### *B. Uncertainty due to the LPFG fluctuations*

However, note that the high noise insensitivity of the fuzzy-based models might rise a concern regarding the LPFG sensor fluctuation and the interrogator resolution. Indeed, as the models lessen the variations in the FBG's filtered power, the models might not transfer all LPFG fluctuations to $\hat{\lambda}_{res}$. In other words, the noise robustness might result in loss of interrogator resolution.

To investigate this effect, we studied the impact of minor fluctuations in the LPFG resonant wavelength ($u_b$) and their impact on the residual's bias and standard deviation, similarly to the previously shown.

Fig. 8 shows the bias as a function of $u_b$. The results show that the bias also doesn't correlate to the input fluctuation, thus we also evaluated the typical bias range for fluctuating LPFGs. We calculated the range 0.028 ~ 0.030 nm for the Linear model, 0.081 ~ 0.083 nm for the MLP, 0.120 ~ 0.122 nm for the FIS-ex, and 0.099 ~ 0.101 nm for the FIS-ga. These values illustrate the stability of the proposed interrogator, since they are in accordance with the previously shown, and provide a glimpse on the models' epistemic uncertainty.

Unlike the optoelectronic noise, that occurs all over the inputs, the LPFG fluctuation occurs at the resonant wavelength vicinity. Consequently, our results didn't show as high variations as before, especially for the Linear model. So, the "Attention" property of ML-models is less important for LPFG variation, indicating that the sensor's uncertainty propagates well trough the model.

Note that for the residuals' standard deviation, it is desired that all input fluctuations (that represent the resonant wavelength standard deviation) are passed through the model to the estimated resonant wavelength. So, ideally, a 1:1 ratio between $u_b$ and the residual standard deviation is desired. However, the measuring instruments presents a resolution, so not all minor fluctuations can be transferred to the interrogator's output. Indeed, the results showed that the residuals' standard deviation clips at a certain value, see Fig. 9.

The concern that the more robust-to-noise models (fuzzy-based models) might damp the LPFG fluctuations was discarded, since we didn't find the estimated standard deviation of the residuals were smaller than $u_b$ in any case. In fact, we found that for lower LPFG fluctuations, the estimation was worse, as expected for any measuring instrument. We also found that for values above the interrogator's estimated resolution (2σ of the values shown in Table I), the ratio between input fluctuations and output fluctuations is 1:1. Therefore, under normal circumstances, all input fluctuations ($u_b$) should be passed through the model to the estimated resonant wavelength. The estimated resolution for each model is shown in Fig. 8 as dotted lines.

As noted before, the metrics variance were low enough for us to consider that the number of iterations used in this work was sufficient.

## IV. Conclusions

In this work we discussed on the application of machine learning for the interrogation of LPFG-based fiber optic sensors. The interrogation technique discussed is based on the use of a sparse filter bank, composed by 13 FBGs. The discussion focuses on the performance evaluation and uncertainty propagation through the interrogator's signal processing chain, using previously reported methods and models. We considered the aleatoric uncertainty, caused by input uncertainty, both from optoelectronic noise and from the LPFG fluctuation.

Our results showed the models don't induce systematic error under uncertain scenarios, due to the stable residuals bias found in this work. We were also capable to verify the damping effect that the FIS models present concerning optoelectronic noise, with great insensitivity to the SNR, specially from 20 dB to 11 dB.

Furthermore, the results presented showed this damping effect doesn't impact on the propagation of LPFG fluctuations. I.e., the standard deviation calculated for the residuals are consistent with the standard deviation of the input sensor. Therefore, a deviation at the input $\lambda_{res}$ is passed to the estimated $\lambda_{res}$ in a 1:1 ratio, as long as the input deviation is above the interrogator's resolution.

Moreover, the methodology used in this work could be used to generate equations for the uncertainty propagation in ML-assisted FOS instruments and interrogators. Note that the proposed methodology is model-independent and can be applied to a wide variety of models with proper adjustment,

regarding the number of iterations, for example. Other types of uncertainties could also be considered, like the FBG Bragg's wavelength stability, for example. For that, the noise generation should be inserted into the interrogator hardware simulation to shift the Bragg wavelength before filtering the inputted LPFG spectrum.

To further improve the use and reliability of ML in FOS instruments and interrogators, we believe that addressing uncertainty in models using transfer learning and synthetic data is relevant, because acquiring large amount of data is a difficult and expensive in the context of optical instruments.

## REFERENCES

[1] T. A. Soomro *et al.*, "Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review," *IEEE Rev Biomed Eng*, 2022, doi: 10.1109/RBME.2022.3185292.

[2] M. M. Kumbure, C. Lohrmann, P. Luukka, and J. Porras, "Machine learning techniques and data for stock market forecasting: A literature review," *Expert Syst Appl*, vol. 197, p. 116659, Jul. 2022, doi: 10.1016/J.ESWA.2022.116659.

[3] D. T. Hoang and H. J. Kang, "A Motor Current Signal-Based Bearing Fault Diagnosis Using Deep Learning and Information Fusion," *IEEE Trans Instrum Meas*, vol. 69, no. 6, pp. 3325–3333, Jun. 2020, doi: 10.1109/TIM.2019.2933119.

[4] P. Coliaie *et al.*, "Machine Learning-Driven, Sensor-Integrated Microfluidic Device for Monitoring and Control of Supersaturation for Automated Screening of Crystalline Materials," *ACS Sens*, vol. 7, no. 3, pp. 797–805, Mar. 2022, doi: 10.1021/ACSSENSORS.1C02358/SUPPL_FILE/SE1C02358_SI_001.PDF.

[5] Z. Zhou *et al.*, "Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays," *Nature Electronics 2020 3:9*, vol. 3, no. 9, pp. 571–578, Jun. 2020, doi: 10.1038/s41928-020-0428-6.

[6] J. Guo, Y. Cheng, D. Luo, K. Y. Wong, K. Hung, and X. Li, "ODRP: A Deep Learning Framework for Odor Descriptor Rating Prediction Using Electronic Nose," *IEEE Sens J*, vol. 21, no. 13, pp. 15012–15021, Jul. 2021, doi: 10.1109/JSEN.2021.3074173.

[7] Q. Nazir and C. Shao, "Online tool condition monitoring for ultrasonic metal welding via sensor fusion and machine learning," *J Manuf Process*, vol. 62, pp. 806–816, Feb. 2021, doi: 10.1016/J.JMAPRO.2020.12.050.

[8] N. Bandari, J. Dargahi, and M. Packirisamy, "Miniaturized Optical Force Sensor for Minimally Invasive Surgery with Learning-Based Nonlinear Calibration," *IEEE Sens J*, vol. 20, no. 7, pp. 3579–3592, Apr. 2020, doi: 10.1109/JSEN.2019.2959269.

[9] D. Li *et al.*, "Machine Learning-Assisted Multifunctional Environmental Sensing Based on a Piezoelectric Cantilever," vol. 7, pp. 2767–2777, 2022, doi: 10.1021/acssensors.2c01423.

[10] S. Ali, T. Glass, B. Parr, J. Potgieter, and F. Alam, "Low Cost Sensor with IoT LoRaWAN Connectivity and Machine Learning-Based Calibration for Air Pollution Monitoring," *IEEE Trans Instrum Meas*, vol. 70, 2021, doi: 10.1109/TIM.2020.3034109.

[11] F. O. Barino, G. Ébias, J. Bittencourt, D. Discini, and A. B. Santos, "Two-dimensional long-period fiber grating sensor for touch applications," *Microw Opt Technol Lett*, vol. 63, no. 2, pp. 647–652, Feb. 2021, doi: 10.1002/mop.32599.

[12] M. A. Jucá, M. A. Jucá, I. V. C. Pereira, P. C. G. Spelta, and A. B. dos Santos, "Identification of external media using a long-period grating and optical time-domain reflectometry," *Applied Optics, Vol. 62, Issue 8, pp. C43-C48*, vol. 62, no. 8, pp. C43–C48, Mar. 2023, doi: 10.1364/AO.476282.

[13] D. Pal, A. Kumar, A. Gautam, and J. Thangaraj, "FBG Based Optical Weight Measurement System and Its Performance Enhancement Using Machine Learning," *IEEE Sens J*, vol. 22, no. 5, pp. 4113–4121, Mar. 2022, doi: 10.1109/JSEN.2022.3144173.

[14] K. Dey, N. Vangety, and S. Roy, "Simultaneous strain-temperature analysis by machine learning assisted FBG sensor," *Conference on Lasers and Electro-Optics (2022), paper ATh2K.5*, p. ATh2K.5, May 2022, doi: 10.1364/CLEO_AT.2022.ATH2K.5.

[15] S. Sarkar, D. Inupakutika, M. Banerjee, M. Tarhani, and M. Shadaram, "Machine Learning Methods for Discriminating Strain and Temperature Effects on FBG-Based Sensors," *IEEE Photonics Technology Letters*, vol. 33, no. 16, pp. 876–879, Aug. 2021, doi: 10.1109/LPT.2021.3055216.

[16] P. H. Chiu, Y. S. Lin, Y. C. Manie, J. W. Li, J. H. Lin, and P. C. Peng, "Intensity and Wavelength-Division Multiplexing Fiber Sensor Interrogation Using a Combination of Autoencoder Pre-Trained Convolution Neural Network and Differential Evolution Algorithm," *IEEE Photonics J*, vol. 13, no. 1, Feb. 2021, doi: 10.1109/JPHOT.2021.3050298.

[17] M. A. Soto, J. A. Ramírez, and P. D. Hernández, "Deep-Learning-Based Earthquake Detection for Fiber-Optic Distributed Acoustic Sensing," *Journal of Lightwave Technology, Vol. 40, Issue 8, pp. 2639-2650*, vol. 40, no. 8, pp. 2639–2650, Apr. 2022, doi: 10.1364/JLT.40.002639.

[18] D. F. Kandamali *et al.*, "Machine learning methods for identification and classification of events in $\phi$-OTDR systems: a review," *Applied Optics, Vol. 61, Issue 11, pp. 2975-2997*, vol. 61, no. 11, pp. 2975–2997, Apr. 2022, doi: 10.1364/AO.444811.

[19] Y. Shi, Y. Wang, L. Wang, L. Zhao, and Z. Fan, "Multi-event classification for Φ-OTDR distributed optical fiber sensing system using deep learning and support vector machine," *Optik (Stuttg)*, vol. 221, p. 165373, Nov. 2020, doi: 10.1016/J.IJLEO.2020.165373.

[20] P. Westbrook, "Big data on the horizon from a new generation of distributed optical fiber sensors," *APL Photonics*, vol. 5, no. 2, Feb. 2020, doi: 10.1063/1.5144123/1024146.

[21] J. López-Higuera, "Introduction to fibre optic sensing technology," in *Handbook of Optical Fibre*

*Technology*, J. López-Higuera, Ed., New York: Wiley, 2004, pp. 1–22.

[22] D. Tosi, "Review and Analysis of Peak Tracking Techniques for Fiber Bragg Grating Sensors," *Sensors 2017, Vol. 17, Page 2368*, vol. 17, no. 10, p. 2368, Oct. 2017, doi: 10.3390/S17102368.

[23] S. Weng, P. Yuan, W. Zhuang, D. Zhang, F. Luo, and L. Zhu, "SOI-Based Multi-Channel AWG with Fiber Bragg Grating Sensing Interrogation System," *Photonics 2021, Vol. 8, Page 214*, vol. 8, no. 6, p. 214, Jun. 2021, doi: 10.3390/PHOTONICS8060214.

[24] S. Chen *et al.*, "Cost-effective improvement of the performance of AWG-based FBG wavelength interrogation via a cascaded neural network," *Optics Express, Vol. 30, Issue 5, pp. 7647-7663*, vol. 30, no. 5, pp. 7647–7663, Feb. 2022, doi: 10.1364/OE.449004.

[25] H. Guo, G. Xiao, N. Mrad, and J. Yao, "Interrogation of a long-period grating sensor by a thermally tunable arrayed waveguide grating," *IEEE Photonics Technology Letters*, vol. 20, no. 21, pp. 1790–1792, 2008.

[26] S. Li, S. Ren, S. Chen, and B. Yu, "Improvement of Fiber Bragg Grating Wavelength Demodulation System by Cascading Generative Adversarial Network and Dense Neural Network," *Applied Sciences 2022, Vol. 12, Page 9031*, vol. 12, no. 18, p. 9031, Sep. 2022, doi: 10.3390/APP12189031.

[27] M. A. Jucá, D. B. Haddad, A. B. Santos, and A. B. dos Santos, "Interrogation system for optical sensor using filter bank and artificial neural network," *Microw Opt Technol Lett*, vol. 62, no. 12, pp. 4015–4020, Dec. 2020, doi: 10.1002/mop.32516.

[28] G. Sampaio, F. O. Barino, and A. B. dos Santos, "Long-period fiber grating sensor interrogation with single strain modulated FBG and harmonic analysis," *Optical Fiber Technology*, vol. 71, p. 102940, Jul. 2022, doi: 10.1016/j.yofte.2022.102940.

[29] H. J. Patrick, G. M. Williams, A. D. Kersey, J. R. Pedrazzani, and A. M. Vengsarkar, "Hybrid fiber Bragg grating/long period fiber grating sensor for strain/temperature discrimination," *IEEE Photonics Technology Letters*, vol. 8, no. 9, pp. 1223–1225, Sep. 1996, doi: 10.1109/68.531843.

[30] F. O. Barino, A. B. d. Santos, and A. B. dos Santos, "LPG Interrogator Based on FBG Array and Artificial Neural Network," *IEEE Sens J*, vol. 20, no. 23, pp. 14187–14194, Dec. 2020, doi: 10.1109/JSEN.2020.3007957.

[31] F. O. Barino, E. P. De Aguiar, L. De Mello Honório, V. N. H. Silva, A. P. López-Barbero, and A. B. Dos Santos, "A Fuzzy Approach to LPFG-Based Optical Sensor Processing and Interrogation," *IEEE Trans Instrum Meas*, vol. 71, 2022, doi: 10.1109/TIM.2022.3216390.

[32] K. Dey, N. Vangety, and S. Roy, "Machine learning approach for simultaneous measurement of strain and temperature using FBG sensor," *Sens Actuators A Phys*, vol. 333, p. 113254, Jan. 2022, doi: 10.1016/J.SNA.2021.113254.

[33] A. G. Leal-Junior, V. Campos, C. Díaz, R. M. Andrade, A. Frizera, and C. Marques, "A machine learning approach for simultaneous measurement of magnetic field position and intensity with fiber Bragg grating and magnetorheological fluid," *Optical Fiber Technology*, vol. 56, p. 102184, 2020, doi: https://doi.org/10.1016/j.yofte.2020.102184.

[34] A. Kokhanovskiy, N. Shabalov, A. Dostovalov, and A. Wolf, "Highly Dense FBG Temperature Sensor Assisted with Deep Learning Algorithms," *Sensors 2021, Vol. 21, Page 6188*, vol. 21, no. 18, p. 6188, Sep. 2021, doi: 10.3390/S21186188.

[35] E. D. Chubchev, K. A. Tomyshev, O. V. Butov, A. V. Dorofeenko, and I. A. Nechepurenko, "Machine Learning Approach to Data Processing of TFBG-Assisted SPR Sensors," *Journal of Lightwave Technology, Vol. 40, Issue 9, pp. 3046-3054*, vol. 40, no. 9, pp. 3046–3054, May 2022, Accessed: May 10, 2023. [Online]. Available: https://opg.optica.org/abstract.cfm?uri=jlt-40-9-3046

[36] Z. Cao *et al.*, "Improved Spectral Interrogation of Tilted Fiber Bragg Grating Refractometer Using Residual Convolutional Neural Networks," *Journal of Lightwave Technology*, vol. 40, no. 22, pp. 7403–7411, Nov. 2022, doi: 10.1109/JLT.2022.3200999.

[37] E. Hüllermeier and W. Waegeman, "Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods," *Mach Learn*, vol. 110, no. 3, pp. 457–506, Mar. 2021, doi: 10.1007/S10994-021-05946-3/FIGURES/17.

[38] J. Lampinen and A. Vehtari, "Bayesian approach for neural networks—review and case studies," *Neural Networks*, vol. 14, no. 3, pp. 257–274, Apr. 2001, doi: 10.1016/S0893-6080(00)00098-8.

[39] B. Lakshminarayanan, A. Pritzel, and C. B. Deepmind, "Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 6405–6416. doi: 10.5555/3295222.3295387.

[40] M. Sensoy, L. Kaplan, and M. Kandemir, "Evidential Deep Learning to Quantify Classification Uncertainty," *Adv Neural Inf Process Syst*, pp. 3179–3189, 2018.

[41] M. Abdar *et al.*, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Information Fusion*, vol. 76, pp. 243–297, Dec. 2021, doi: 10.1016/j.inffus.2021.05.008.

[42] S. Haykin, *Neural Networks and Learning Machines*. Prentice Hall, 2008.

[43] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 2008.

[44] V. Nasir and F. Sassani, "A review on deep learning in machining and tool monitoring: methods, opportunities, and challenges," *The International Journal of Advanced Manufacturing Technology 2021 115:9*, vol. 115, no. 9, pp. 2683–2709, May 2021, doi: 10.1007/S00170-021-07325-7.

[45] M. Taiebat and F. Sassani, "DISTINGUISHING SENSOR FAULTS FROM SYSTEM FAULTS BY UTILIZING MINIMUM SENSOR REDUNDANCY,"

*https://doi.org/10.1139/tcsme-2017-1033*, vol. 41, no. 3, pp. 469–487, 2018, doi: 10.1139/TCSME-2017-1033.

[46] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra, "Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0," *Information Fusion*, vol. 50, pp. 92–111, Oct. 2019, doi: 10.1016/J.INFFUS.2018.10.005.

[47] T. P. Merrifield *et al.*, "Synthetic seismic data for training deep learning networks," 2022, doi: 10.1190/INT-2021-0193.1.

[48] J.-F. Rajotte, R. Bergen, D. L. Buckeridge, K. El Emam, R. Ng, and E. Strome, "Synthetic data as an enabler for machine learning applications in medicine," *iScience*, vol. 25, p. 105331, 2022, doi: 10.1016/j.isci.

[49] A. Boikov, V. Payor, R. Savelev, and A. Kolesnikov, "Synthetic Data Generation for Steel Defect Detection and Classification Using Deep Learning," *Symmetry 2021, Vol. 13, Page 1176*, vol. 13, no. 7, p. 1176, Jun. 2021, doi: 10.3390/SYM13071176.

[50] G. M. Mahmoud and R. S. Hegazy, "Comparison of GUM and Monte Carlo methods for the uncertainty estimation in hardness measurements," *International Journal of Metrology and Quality Engineering*, vol. 8, p. 14, 2017, doi: 10.1051/IJMQE/2017014.

[51] A. Chen and C. Chen, "Comparison of GUM and Monte Carlo methods for evaluating measurement uncertainty of perspiration measurement systems," *Measurement*, vol. 87, pp. 27–37, Jun. 2016, doi: 10.1016/J.MEASUREMENT.2016.03.007.

[52] D. Theodorou, L. Meligotsidou, S. Karavoltsos, A. Burnetas, M. Dassenakis, and M. Scoullos, "Comparison of ISO-GUM and Monte Carlo methods for the evaluation of measurement uncertainty: Application to direct cadmium measurement in water by GFAAS," *Talanta*, vol. 83, no. 5, pp. 1568–1574, Feb. 2011, doi: 10.1016/J.TALANTA.2010.11.059.

[53] Y. Lee and S. H. Oh, "Input Noise Immunity of Multilayer Perceptrons," *ETRI Journal*, vol. 16, no. 1, pp. 35–43, Apr. 1994, doi: 10.4218/ETRIJ.94.0194.0013.

**Felipe Oliveira Barino** was born in Juiz de Fora, Brazil, in 1996. He received the B.S. and M.Sc. degrees in electronics from the Federal University of Juiz de Fora (UFJF), Juiz de Fora, in 2019 and 2021, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the same institution, with the Programa de Pós-graduação em Engenharia Elétrica (PPEE).

His current research interests are related to instrumentation, metrology, machine learning, transfer learning, optical fiber sensors, and sensor packaging.

**Alexandre Bessa dos Santos** was born in Juiz de Fora, Brazil, in 1975. He received the M.S. and Ph.D. degrees in electrical engineering from Pontifical Catholic University (PUC), Ponce, Brazil, in 2001 and 2005, respectively.

Since 2010, he has been an Associate Professor with the Department of Circuits with the Federal University of Juiz de Fora (UFJF), Juiz de Fora. His research interests include applied electromagnetism, optical sensors, computational methods, instrumentation, and metrology.