Article

Robust Vector BOTDA Signal Processing with Probabilistic Machine Learning

Abhishek Venketeswaran [1], Nageswara Lalam [1,2,*], Ping Lu [1,2], Sandeep R. Bukka [1,2], Michael P. Buric [3] and Ruishu Wright [1]

[1]   National Energy Technology Laboratory, 626 Cochrans Mill Road, Pittsburgh, PA 15236, USA
[2]   NETL Research Support Contractor, 626 Cochrans Mill Road, Pittsburgh, PA 15236, USA
[3]   National Energy Technology Laboratory, 3610 Collins Ferry Road, Morgantown, WV 26505, USA
[*]   Correspondence: nageswara.lalam@netl.doe.gov

Abstract: This paper presents a novel probabilistic machine learning (PML) framework to estimate the Brillouin frequency shift (BFS) from both Brillouin gain and phase spectra of a vector Brillouin optical time-domain analysis (VBOTDA). The PML framework is used to predict the Brillouin frequency shift (BFS) along the fiber and to assess its predictive uncertainty. We compare the predictions obtained from the proposed PML model with a conventional curve fitting method and evaluate the BFS uncertainty and data processing time for both methods. The proposed method is demonstrated using two BOTDA systems: (i) a BOTDA system with a 10 km sensing fiber and (ii) a vector BOTDA with a 25 km sensing fiber. The PML framework provides a pathway to enhance the VBOTDA system performance.

Keywords: optical fiber sensors; data analytics; deep neural networks; distributed fiber sensors; sensor data

1. Introduction

Brillouin-based distributed fiber-optic sensors have gained tremendous attention over the past two decades due to their capability of measuring both strain and temperature over tens of kilometers. A wide range of structural health monitoring applications on large civil infrastructure (oil and gas pipelines, bridges, and railways), power grids, and border security have been demonstrated [1,2]. Brillouin optical time-domain analysis (BOTDA) is based on stimulated Brillouin scattering, where pump pulses and probe continuous waves (CWs) injected into a fiber from both ends excite acoustic waves to facilitate power coupling between the two counter-propagating optical waves. The pulsed pump waves generate a gain or loss for the CW probe waves with downshifted or upshifted frequencies. A Brillouin gain/loss spectrum (BGS/BLS) can be obtained over the sensing fiber distance by sweeping frequencies around the Brillouin frequency shift (BFS), which depends on the strain and temperature of the fiber. Several vector BOTDA (VBOTDA) system configurations have been proposed to extract both the BGS and Brillouin phase spectrum (BPS) simultaneously. The BPS is independent of Brillouin gain and is unaffected by nonlocal effects. BPS exhibits low noise and is therefore a good target for deriving sensor information [3–6]. For instance, an IQ demodulation algorithm was proposed with heterodyne detection and measured both the BGS and BPS [7]. A double-frequency phase modulation method was proposed that provides measurements of both BGS and BPS [8]. Recently, VBOTDA was proposed by using a vector network analyzer to acquire BGS and BPS [9,10].

The incident pump pulses at a frequency ν_0 with a certain pulse width counter-propagates to a CW probe wave at a frequency $\nu_0 - \nu_B$. The input pulses generate a backscattered Brillouin signal, which is downshifted to the pump signal by the BFS ν_B (around 10–11 GHz for silica single-mode fibers). The probe experiences Brillouin amplification and Brillouin phase shift when the difference between the input pump and

probe wave frequency is around the BFS. A suitable signal processing technique must be employed to evaluate the BFS from the measured BGS/BPS over the fiber length. Several signal processing techniques have been proposed in the literature to estimate the BFS from BGS and/or BPS measurements. These techniques can be categorized as either (i) curve fitting (CF) or (ii) machine learning (ML)-based approaches. The latter approach is increasingly being favored for long-distance BOTDA systems for its capability to process data in real time. Several studies in the literature have reported that the ML-based models have provided more accurate BFS predictions than their CF counterparts in scenarios such as (i) low signal-to-noise ratio and (ii) coarsely resolved BGS/BPS measurements [11,12]. ML techniques can indeed be combined with Brillouin-based distributed fiber sensors to enhance their performance and to enable advanced signal processing and data analysis. These advancements open up opportunities for a wide range of intelligent monitoring applications, including structural health monitoring, environmental monitoring, and smart infrastructure management. In the literature, several efforts are being made to address these challenges and to develop robust and efficient ML algorithms that can accelerate the signal processing speed and handle the complexity and volume of data generated by distributed fiber sensors.

BOTDA systems are often deployed in harsh environments over long distances and are prone to several environmental and systemic factors that can increase sensor noise and can degrade the performance of the sensing system. Several factors reduce the SNR of a BOTDA system, such as the strength of the backscattered Brillouin signal (typically at nW), silica fiber double path loss (typically 0.4 dB/kmat 1550 nm), noise sources, and other nonlinear effects that limit the input power of pumps and probes [13]. Polarization noise, relative intensity noise (RIN), amplified spontaneous emission (ASE) noise, and thermal and shot noise are the major noise sources that degrade the SNR [14].

A data processing algorithm based on radial basis function neural networks (RBFNN) was proposed in [15], for BOTDR sensing. The RBFNN algorithm was shown to accelerate data processing compared with the Levenberg–Marquardt nonlinear least-squares algorithm. A recurrent neural network-based data processing framework was demonstrated [16] on data collected using a commercial Brillouin-based distributed temperature sensing system (AP Sensing, N4385B). This approach utilized an autoregressive input layer often found in time series analysis literature as autoregressive models. A ML-based algorithm using support vector machines [17] was used to extract BFS from both gain and phase data collected from a BOTDA sensing system. The proposed algorithm improved the data processing speed by 100 times compared with the conventional nonlinear least squares data processing technique. Deep neural networks (DNNs) with autoencoder architectures were introduced to extract BFS from BOTDA data [11,18]. A BOTDA sensing system using a 25 km long large-effective-area fiber (LEAF) sensing fiber coupled with stacked autoencoder-based data processing framework exhibited capabilities in extracting both strain and temperature measurements simultaneously [18]. Further, a denoising autoencoder architecture was shown to both denoise as well as extract temperature from BOTDA data in [11]. Another ANN-based signal processing framework was developed using radial basis functions (RBF) [19,20]. This framework utilizes a shallow neural architecture and requires a much smaller number of parameters compared with deep neural network architectures.

A major drawback of all the existing BOTDA ML models proposed in the literature is that they do not predict the confidence intervals (CI) of the predicted parameters. The reliability of the sensor data and the associated ML model can only be assessed by propagating the measurement noise in the BGS/BPS data to obtain estimates of prediction uncertainties in strain/temperature. Currently, the ability to predict BFS uncertainty from BGS data has only been demonstrated for the simple case of quadratic CF-based BGS processing [21].

We propose a novel *probabilistic machine learning* (PML) framework [22–24] for processing BOTDA data, which preserves the existing advantages of ML models while adding a new capability of providing simultaneous CI estimates. The core advantage of the pro-

posed PML framework in this work is the characterization of noise in the data in terms of confidence intervals. The mean and the standard deviation of peak frequency of BFS at each location on the fiber are computed by the PML framework. This enables us to obtain an estimate of the effect of noise in the system at each discrete point on the fiber. First, the mathematical frameworks of the CF and ML approaches are discussed in Section 2, in the context of BFS estimation from BGS and/or BPS measurements. Next, the mathematical framework of the proposed approach is introduced in Section 3. The model development and training procedure are outlined in Section 4. The experimental setup used to generate the data is described in Section 5. The capability of the proposed PML framework to provide simultaneous predictions of BFS and its CI is demonstrated in Section 6 from measurements obtained from two custom BOTDA systems, namely (i) BGS data collected from a 10 km range BOTDA system and (ii) BGS and BPS data collected from a 25 km range custom VBOTDA system. The performance of the PML approach is compared with the CF approach because only these two approaches provide estimates of the parameters and their CIs. Finally, the conclusions and the scope for future work are summarized in Section 7.

2. Mathematical Background

Firstly, the mathematical notations adopted in this work are introduced. Let $\mathcal{D}$ denote a single BGS and BPS measurement dataset (obtained at a single point along the optical fiber and resolved at n frequencies):

$$\mathcal{D} \equiv \{(\nu_i, g_i, \phi_i), \forall i = 1, \cdots, n\}, \tag{1}$$

where frequency, gain, and phase are denoted by (ν_i, g_i, ϕ_i), respectively. Let the BFS and Brillouin line width (full width at half maximum (FWHM)) be denoted by ν_B and w, respectively. Further, let $\lambda = [\lambda_1 = \nu_B, \lambda_2 = w]$ denote the vector of parameters of interest. From here on, a single measurement sample of BGS or BPS will be denoted by an input matrix $\mathbf{X}$, wherein

$$\mathbf{X} = \begin{bmatrix} \nu_1 & \cdots & \nu_n \\ g_1 & \cdots & g_n \end{bmatrix}^\mathsf{T}, \quad or \quad \mathbf{X} = \begin{bmatrix} \nu_1 & \cdots & \nu_n \\ \phi_1 & \cdots & \phi_n \end{bmatrix}^\mathsf{T}, \quad or \quad \mathbf{X} = \begin{bmatrix} \nu_1 & \cdots & \nu_n \\ g_1 & \cdots & g_n \\ \phi_1 & \cdots & \phi_n \end{bmatrix}^\mathsf{T}. \tag{2}$$

The schematic of the Brillouin gain and phase spectrum is outlined in Figure 1 for illustrative purposes. Next, the mathematical backgrounds of curve fitting and machine learning approaches are briefly discussed below for the case of BGS processing to extract BFS.
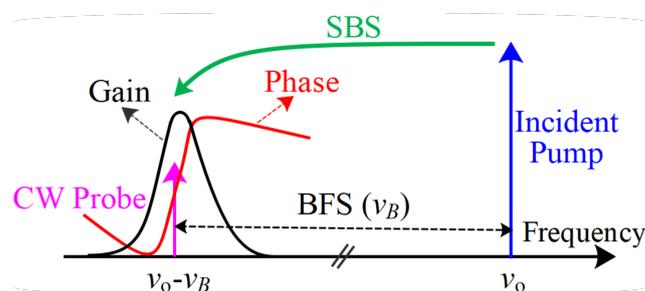
Figure 1. Schematic of Brillouin gain and phase spectrum.

2.1. Curve Fitting Approach

Curve fitting approaches model the BGS as a smooth parametric function (Lorentzian, Gaussian, and pseudo-Voigt) depending on the parameters of the incident pump light [25,26]. The BGS profile usually has a Lorentzian shape if the pump light pulse width is larger than 10 ns [27], and a Lorentzian fit is often adopted for long-distance BOTDA applications. The BGS profile resembles a Gaussian shape for short incident pulse widths (<10 ns) due to

the Doppler broadening effect. The most commonly utilized BGS functions in the literature are given in Table 1.

Table 1. Commonly used BGS [28].

Gain Spectrum	Function $f_c(\nu; \lambda)$	Parameters λ
Lorentzian	$f_c(\nu; \lambda) = \dfrac{g_0}{1 + 4\xi^2}, \quad \xi = \dfrac{(\nu - \nu_B)}{w}$	$\lambda \equiv [g_0, \nu_B, w]$
Gaussian	$f_c(\nu; \lambda) = g_0 \exp\left[-4 \ln 2\xi^2\right], \quad \xi = \dfrac{(\nu - \nu_B)}{w}$	$\lambda \equiv [g_0, \nu_B, w]$
Pseudo-Voigt	$f_c(\nu; \lambda) = g_0 \left[\dfrac{p}{1 + 4\xi^2} + (1 - p) \exp\left(-4 \ln 2\xi^2\right)\right]$	$\lambda \equiv [p, g_0, \nu_B, w]$

The BGS measurements are modeled by choosing a suitable BGS function $f_c(\cdot; \lambda)$ with parameters λ and adding measurement noise as given below:

$$g_i = f_c(\nu_i; \lambda) + \epsilon_i, \tag{3}$$

where ϵ_i denotes the measurement error. The parameters λ are estimated from the dataset $\mathcal{D}$ by using a nonlinear regression based optimization framework:

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} (\epsilon_i)^2 = \underset{\lambda}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} (g_i - f_c(\nu_i; \lambda))^2. \tag{4}$$

Linear curve fitting using measurements around the zero de-phase frequency region is commonly used to estimate BFS from BPS measurements [9].

The BFS uncertainty from BGS data has only been demonstrated for the simple case of a quadratic BGS profile [21]. However, for other BGS spectra (e.g., Lorentzian, Gaussian etc.), the CIs of the parameter estimates $\hat{\lambda}$ can be estimated using the asymptotic Gaussian distribution of the corresponding least-squares estimator [29]. This asymptotic distribution can be computed from the covariance estimates of the measurement errors (ϵ_i in Equation (3)).

2.2. Machine Learning Approach

Machine-learning-based BOTDA data processing algorithms that have been proposed in the literature construct a *direct vectorial functional mapping* $f_d(\cdot; \Theta)$ between the parameters λ and the BGS dataset $\mathbf{X}$:

$$\lambda = f_d(\mathbf{X}; \Theta) + \mathbf{e}, \tag{5}$$

where Θ denotes the parameters of the ML model and $\mathbf{e}$ denotes the *prediction* error. The training of the ML model requires an annotated dataset containing BGS measurements corresponding to different parameters λ. The ML model parameters are obtained using a suitable optimization framework:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \|\mathbf{e}_i\|^2 = \underset{\lambda}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \|\lambda_i - f_d(\mathbf{X}_i; \Theta)\|^2, \tag{6}$$

where $\left\{\left(\boldsymbol{\lambda}_j, \mathbf{X}_j\right), j = 1, \cdots, N\right\}$ denotes the training dataset of N BGS measurements and $\|\cdot\|$ denotes the Euclidean norm. Upon training, the parameter estimates $\hat{\boldsymbol{\lambda}}$ for a given BGS ($\mathbf{X}$) can be computed directly using

$$\hat{\boldsymbol{\lambda}} = f_d\left(\mathbf{X}; \hat{\boldsymbol{\Theta}}\right). \tag{7}$$

Several ML models have been proposed in the literature for the purpose of BFS extraction, including support vector machines [17], principal component regression [30], and artificial neural networks [20,31]. Recent studies in the literature have increasingly advocated for the use of deep neural networks [12,32] with varying architectures such as (i) stacked and denoising autoencoders [11,18], and convolutional neural networks [33,34] to extract BFS from BGS/BPS measurements for low signal-to-noise ratio, long range, and dynamic sensing scenarios.

2.3. Comparison of Curve Fitting vs. Machine Learning for BFS Extraction

- Data Processing Time: The computational advantage of the ML approach over the CF approach is evident when comparing the schematics of both approaches, as shown in Figure 2. The CF approach requires repeating the optimization iterations for each BGS, while the ML approach can predict the BFS and FWHM directly from the BGS measurements using Equation (7), once the ML model is trained offline.

- Interpretability: However, the CF approach is more interpretable since the function f_c can be chosen based on the underlying optical physics knowledge. On the other hand, no such reasoning exists to construct f_d in the ML approach and several different ML models have been proposed in the literature.

- Robustness: A key feature of a robust signal processing algorithm is its ability to accurately quantify the confidence/uncertainty in predictions. As mentioned earlier, curve fitting approaches use regression to estimate parameters and can yield CIs of the parameter estimates. However, the ML approach (Equation (5)) can not be used directly to estimate the CIs since the term $\mathbf{e}$ does not represent the sensor measurement error. Instead, $\mathbf{e}$ in Equation (5) should be interpreted as *prediction error* with an unknown probability distribution due to the nonlinear nature of f_d. Subsequently, the ML approach provides BFS estimates without providing a measure of uncertainty/confidence (i.e., confidence intervals or error bars) of the BFS predictions. With the increasing adoption of deep neural networks for BOTDA processing, it is even more crucial that the BFS be estimated along with its confidence level, in order to avoid over-fitting.
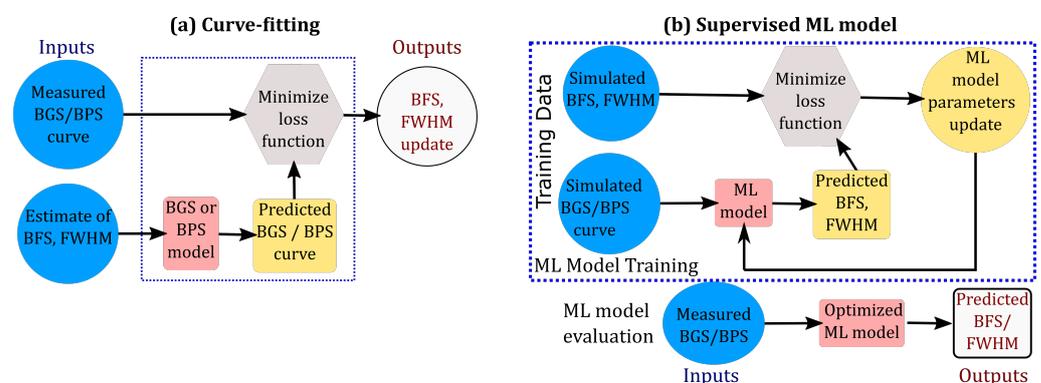
Figure 2. Schematic of (**a**) curve-fitting and (**b**) ML approach to the estimation of BFS and FWHM from BOTDA measurements.

3. Proposed Probabilistic Machine-Learning-Based BFS Extraction

Firstly, the BFS (ν_B) and FWHM (w) are modeled as statistically independent Gaussian random variables:

$$\boldsymbol{\lambda} \equiv \begin{bmatrix} \nu_B \\ w \end{bmatrix} = \boldsymbol{\mu}_{\boldsymbol{\lambda}}(\mathbf{X}) + \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}(\mathbf{X})\mathbf{n}, \quad \mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \quad n_1, n_2 \sim \mathcal{N}(0,1) \tag{8}$$

where the random variables $\mathbf{n}_1$, $\mathbf{n}_2$ are independent and identically distributed normal/Gaussian random variables. The representation in Equation (8) is often referred to as a *reparameterization trick* in the ML literature and is used in several deep learning architectures such as variational autoencoders [35]. The probabilistic model given by Equation (8) provides a pathway to achieve two objectives simultaneously:

1. Mean vector $\boldsymbol{\mu}_{\boldsymbol{\lambda}}$ directly predicts the means of BFS (ν_B) and FWHM (w) from BGS/BPS measurements $\mathbf{X}$ using a suitable ML model
2. Standard deviation matrix $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$ quantifies the uncertainty in estimates of BFS (ν_B) and FWHM (w) due to the noise in underlying measurements $\mathbf{X}$

Subsequently, Equation (8) can be expanded as follows:

$$\boldsymbol{\mu}_{\boldsymbol{\lambda}}(\mathbf{X}) = \begin{bmatrix} \mu_{\nu_B}(\mathbf{X}) \\ \mu_w(\mathbf{X}) \end{bmatrix}, \quad \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}(\mathbf{X}) = \begin{bmatrix} \sigma_{\nu_B}(\mathbf{X}) & 0 \\ 0 & \sigma_w(\mathbf{X}) \end{bmatrix} \tag{9}$$

where means of BFS and FWHM are denoted by $\boldsymbol{\mu}_{\boldsymbol{\lambda}} = (\mu_{\nu_B}, \mu_w)$ and standard deviations of BFS and FWHM are represented in the matrix $\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$ with diagonal elements $(\sigma_{\nu_B}, \sigma_w)$. When measurement noise in BGS/BPS data $\mathbf{X}$ is minimal, we expect the standard deviations to diminish proportionally ($\sigma_{\nu_B} \to 0$, $\sigma_w \to 0$) and the means to converge to their respective *true* values ($\mu_{\nu_B} \to \nu_B$, $\mu_w \to w$). In a realistic scenario, wherein there is considerable measurement noise in the Brillouin spectra, we expect the mean values to give the best estimate of the parameter values. At the same time, the standard deviations can be utilized to estimate the level of confidence that could be assigned to the estimated values. Substituting Equation (9) in Equation (8), we obtain

$$\boldsymbol{\lambda} \equiv \begin{bmatrix} \nu_B \\ w \end{bmatrix} = \begin{bmatrix} \mu_{\nu_B}(\mathbf{X}) + \sigma_{\nu_B}(\mathbf{X})n_1 \\ \mu_w(\mathbf{X}) + \sigma_w(\mathbf{X})n_2 \end{bmatrix}. \tag{10}$$

The means and standard deviations (henceforth denoted as $\boldsymbol{\Lambda}$) are modeled as functions of BGS and/or BPS measurements, using a suitable ML model:

$$\boldsymbol{\Lambda} \equiv \begin{bmatrix} \mu_{\nu_B} & \mu_w & \sigma_{\nu_B} & \sigma_w \end{bmatrix}^\mathsf{T} = f_d(\mathbf{X}; \boldsymbol{\Theta}). \tag{11}$$

wherein the parameters of the ML model are denoted by $\boldsymbol{\Theta}$. Using Equation (11) in Equation (10), it can be realized that the parameters of interest $\boldsymbol{\lambda}$ are represented using a *probabilistic machine learning* (PML) model.

The proposed procedure to estimate the PML model parameters $\boldsymbol{\Theta}$ is depicted in Figure 3. We utilize an offline training framework to train the PML model. A synthetic/simulated training dataset is obtained by using Lorentzian BGS and/or BPS curves for various BFS, FWHM parameters, and noise levels. Let $\{(\boldsymbol{\lambda}_j, \mathbf{X}_j), j = 1, \cdots, N\}$ denotes the training dataset of N BGS/BPS measurements corresponding to various values of BFS, FWHM, and noise values. It should be noted that each $\mathbf{X}_j$ is a matrix of n frequencies and corresponding gain and/or phase values, as given in Equation (2). The PML parameters cannot be estimated using deterministic loss functions as employed by CF (Equation (4)) and ML approaches (Equation (6)). This work proposes the use of the probabilistic objective function known as the log-likelihood function:

$$\hat{\boldsymbol{\Theta}} = \underset{\boldsymbol{\Theta}}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} \log[p(\boldsymbol{\lambda}_i; \boldsymbol{\Lambda} = f_d(\mathbf{X}_i; \boldsymbol{\Theta}))] \tag{12}$$

where $p(\lambda_i|\mathbf{\Lambda})$ denotes the joint probability density function of BFS and FWHM for a given set of parameters $\mathbf{\Lambda}$. From Equation (11), we can write the joint Gaussian probability density function as follows:

$$p(\lambda) \equiv p(\nu_B, w) = \frac{1}{2\pi} \det(\mathbf{\Sigma}_\lambda)^{-0.5} \exp\left(-\frac{1}{2}(\lambda - \mu_\lambda)^\mathsf{T} \mathbf{\Sigma}_\lambda^{-1} (\lambda - \mu_\lambda)\right). \tag{13}$$

By maximizing the log-likelihood function, we obtain a joint probability distribution that maximizes the probability of λ_i for a given $\mathbf{X}_i$. Using Equation (13) in Equation (12), we obtain a least-squares objective function:

$$\hat{\mathbf{\Theta}} = \underset{\mathbf{\Theta}}{\mathrm{argmax}} \frac{-1}{2N} \sum_{i=1}^{N} \left[2\log(2\pi) + \log\sigma_{\nu_B} + \log\sigma_w + \left(\frac{\nu_{B_i} - \mu_{\nu_B}}{\sigma_{\nu_B}}\right)^2 + \left(\frac{w_i - \mu_w}{\sigma_w}\right)^2 \right] \tag{14}$$

Upon training by maximizing Equation (14), the estimates of the means and standard deviations of BFS and FWHM can be computed by evaluating the PML model as depicted in Figure 3:

$$\hat{\mathbf{\Lambda}} \equiv \begin{bmatrix} \hat{\mu}_{\nu_B} & \hat{\mu}_w & \hat{\sigma}_{\nu_B} & \hat{\sigma}_w \end{bmatrix}^\mathsf{T} = f_d(\mathbf{X}; \hat{\mathbf{\Theta}}). \tag{15}$$
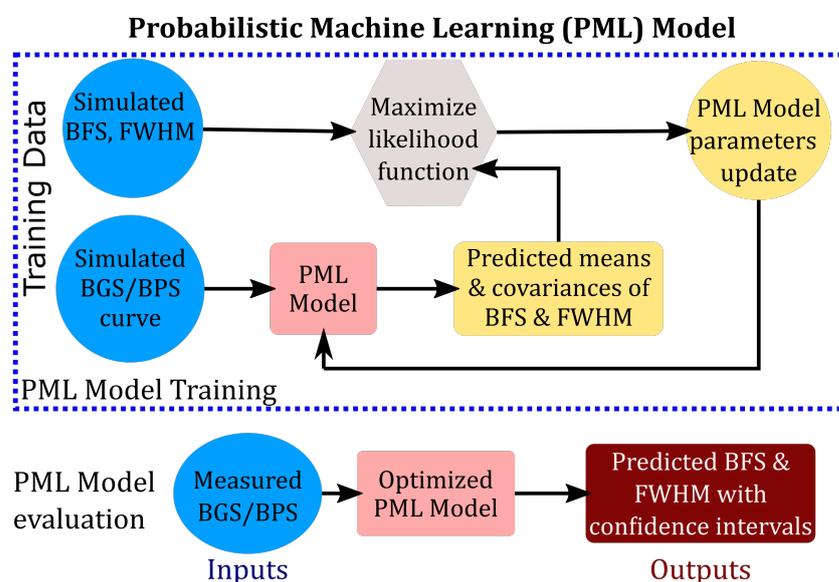
The point estimates and $100(1-\alpha)\%$ CI estimates of BFS and FWHM can be computed from Equation (15) as follows:

$$\hat{\lambda} = \begin{bmatrix} \hat{\mu}_{\nu_B} & \hat{\mu}_w \end{bmatrix}^\mathsf{T} \tag{16}$$

$$\hat{\mu}_{\nu_B} - s_\alpha \hat{\sigma}_{\nu_B} \le \nu_B \le \hat{\mu}_{\nu_B} + s_\alpha \hat{\sigma}_{\nu_B} \tag{17}$$

$$\hat{\mu}_w - s_\alpha \hat{\sigma}_w \le w \le \hat{\mu}_w + s_\alpha \hat{\sigma}_w \tag{18}$$

where $s_\alpha = \Phi\left(1 - \frac{\alpha}{2}\right)$ is the statistic computed from a standard normal cumulative distribution function Φ.

Figure 3. Schematic of PML approach to the estimation of BFS and FWHM with confidence intervals from BOTDA measurements.

This probabilistic framework has computational benefits in comparison with other PML approaches [22] such as dropout [36], bootstrapping [37] and weight randomization [38]. The PML approach has several advantages over CF and ML approaches, as listed below:

1. **Robustness**: The PML approach prevents overfitting that arises when using ML and DNN models to represent f_d.

2. **Flexibility**: It is compatible with the various ML models (e.g., neural networks [39] and support vector regression [40]).

3. **Speed**: It inherits the computational advantages of the ML approach and enables fast processing of BOTDA data with simultaneous assessment of prediction uncertainties.

4. PML Model Development and Training

A deep neural network (DNN) with a combination of convolutional and dense layers is chosen to represent f_d in Equation (11). It is evident from Equations (10) and (11) that the outputs of the DNN are the means and standard deviations of BFS and FWHM. Consequently, the DNN will henceforth be known as a *probabilistic deep neural network (PDNN)*. The PDNN was programmed in Python and the code is available in Code 1 [41]. This PML framework can be easily implemented with another ML or DNN model with minor modifications in the code.

4.1. PML Model Training

This work utilizes a simulated dataset of BGS and BPS measurements corresponding to a uniform grid of BFS, FWHM, and noise values. All of the parameters are sampled from uniform probability distributions $\left(y \sim \mathcal{U}(y_l, y_u), p(y) = \frac{1}{y_u - y_l}, \forall y_l \leq y \leq y_u\right)$ between lower and upper bounds, as given below:

$$s \sim \mathcal{U}\left(\frac{1}{18}, \frac{1}{5}\right), \quad \nu_B \sim \mathcal{U}(10, 50)\, MHz, \quad w \sim \mathcal{U}(5, 50)\, MHz \tag{19}$$

It should be noted that in addition to varying the BFS and FWHM values in a uniform grid, we also vary noise amplitude levels. The variation of BFS and FWHM is essential to predict the mean vector $\boldsymbol{\mu}_\lambda$, and the variation of noise amplitude is essential to predict $\boldsymbol{\Sigma}_\lambda$.

We generate the training dataset by first choosing a vector of frequency values to compute BGS and/or BPS: $\nu_1, \cdots, \nu_K$ and training sample size N. We have considered a large training sample size of $N = 50{,}000$.

For $j = 1 \cdots N$:

1. Uniformly sample s, ν_B, and w from the bounds in Equation (19) to obtain $\left(s_j, \nu_{Bj}, w_j\right)$

2. Simulate gain and phase values for each of the n frequencies and for $\left(\nu_{Bj}, w_j\right)$ using a suitable spectrum model.

$$g_j(\nu_i) = g\left(\nu_i; \nu_{Bj}, w_j\right), \forall i = 1 \cdots K \tag{20}$$

$$\phi_j(\nu_i) = \phi\left(\nu_i; \nu_{Bj}, w_j\right), \forall i = 1 \cdots K \tag{21}$$

This work has chosen Lorentzian BGS and BPS [7] (Equations (22) and (23)) given by the following:

$$g(\nu; \nu_B, w) = g_0 \frac{w^2}{4(\nu - \nu_B)^2 + w^2} \tag{22}$$

$$\phi(\nu; \nu_B, w) = g_0 \frac{2w(\nu - \nu_B)}{4(\nu - \nu_B)^2 + w^2} \tag{23}$$

where g_0 is the Brillouin gain amplitude, w is the Brillouin linewidth or FWHM, and ν_B is the BFS.

3. Sample $e \sim \mathcal{N}(0, 1)$ and add Gaussian noise corresponding to the noise amplitude s_j to obtain training dataset sample $\mathbf{X}_j$

$$\mathbf{X}_j = \begin{bmatrix} \nu_1 & \cdots & \nu_K \\ g_j(\nu_1) + s_j e & \cdots & g_j(\nu_K) + s_j e \\ \phi_j(\nu_1) + s_j e & \cdots & \phi_j(\nu_K) + s_j e \end{bmatrix}^{\mathsf{T}} \tag{24}$$

4.2. PML Model Architecture

The PDNN architecture is depicted in Figure 4. It has a combination of convolutional and dense layers. The input layer consists of two channels. The first 1D convolutional layer (Conv1D) has a kernel size of 15 and (Rectified Linear Unit) ReLU activation. The second Conv1D has a kernel size of 5, and both MaxPool1D layers have kernel size = 3. The dense layer has 32 units, with ReLU and L_2 regularization.
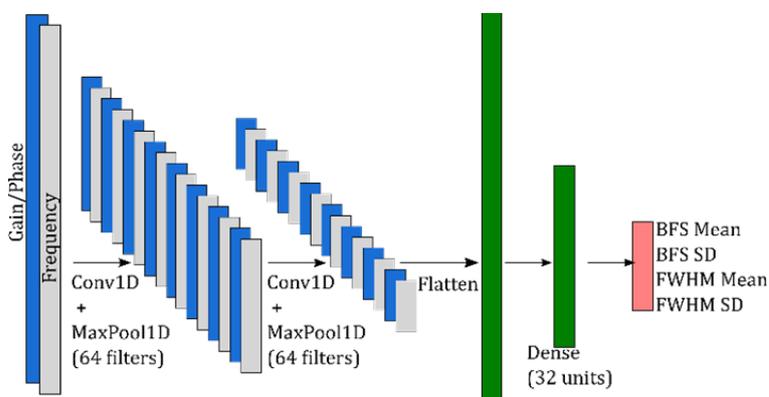
Figure 4. PML model architecture.

The PDNN is trained over this dataset using a training/validation split of 70/30. Further, a learning rate schedule and early stopping criterion are implemented to automate the model training.

5. Experimental Setup

The experimental setup of the vector Brillouin optical time domain analysis (VBOTDA) system is illustrated in Figure 5. In [10], we demonstrated a VBOTDA system, where a vector network analyzer was used to extract both the amplitude and phase spectrum of the Brillouin interaction over the sensing fiber.
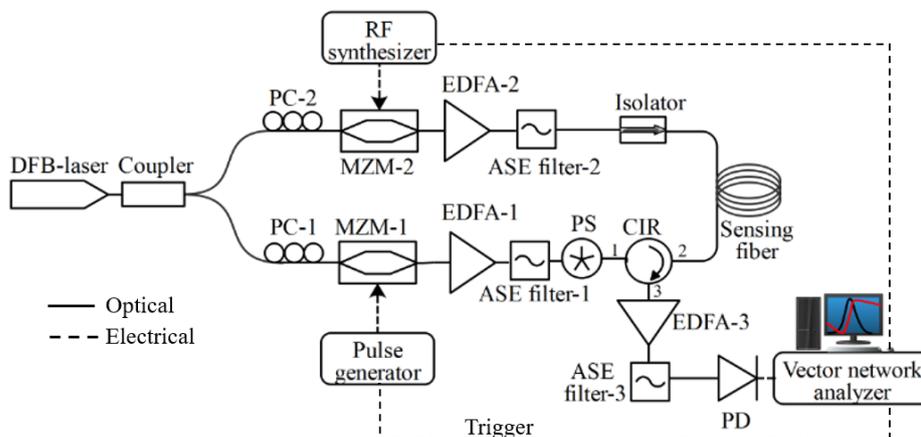

Figure 5. Experimental setup of VBOTDA system. DFB-laser: distributed feedback-laser, PC: polarization controller, MZM: Mach–Zehnder modulator, EDFA: Erbium-doped fiber amplifier, ASE: amplified spontaneous emission, PS: polarization scrambler, CIR: circulator, PD: photo-detector).

A distributed feedback (DFB) laser with a wavelength of 1550 nm was used as a laser source. The laser output signal was split into two branches using a 50/50 3 dB fiber coupler. One branch was for generating input pump pulses and the other branch was for the probe signal to make stimulated Brillouin operation over the fiber distance. A Mach–Zehnder modulator (MZM-1) was used to modulate the optical pulses using a pulse generator. The pulse peak-to-peak amplitude set at 4 Vpp, and the pulse width was set at 10 ns, corresponding to 1 m spatial resolution, whereas the pulse repetition frequency was 4 kHz, which has enough round-trip time to travel the full length of the fiber. In order to compensate the high lossy operation of MZM modulators, a polarization controller (PC) was employed at the input of each MZM to alter the polarization state of the signal and to reach maximum optical power after both MZMs. The output signal was then amplified using an erbium-doped fiber amplifier (EDFA-1). A narrow band-pass filter bandwidth: 0.8 nm was employed to reduce the amplified spontaneous emission (ASE) noise originating from the EDFA. To avoid polarization noise due to the intense polarization sensitivity of the stimulated Brillouin scattering mechanism, a high-speed polarization scrambler (PS) was used to reduce the polarization dependence of Brillouin spectra. The second MZM (MZM-2) was modulated with a frequency close to the BFS of the sensing fiber (10.82 GHz) via an external RF synthesizer. The signal was amplified by EDFA-2 and then sent to an isolator, which allows signal transmission in one direction. The input pump peak power (20 dBm) and the counter-propagating probe power (6 dBm) were sent to the fiber under test. The stimulated Brillouin signal from the circulator (CIR) port 3 was amplified using EDFA-3 and sent to ASE filter-3 to remove the ASE noise. Thereafter, the signal was detected using a photodetector (PD, bandwidth: 125 MHz) and analyzed with a vector network analyzer. In order to obtain the BGS spectra over the fiber distance, the RF synthesizer frequencies were swept around fiber under test BFS, which were 10.78 GHz to 10.9 GHz with a frequency step of 1 MHz. The RF synthesizer and pulse generator were synchronized with the vector network analyzer. For the fiber under test, we used two different lengths of sensing fibers, 10 km and 25 km, and obtained 3D BGS specta over the sensing fiber distance at various trace averages. The sensing fiber was kept at room temperature and under strain-free condition throughout all the measurements.

6. Results and Discussions

The PML framework is illustrated to process data using (i) a custom BOTDA system (BGS data) with a 10 km long sensing fiber with trace averages of 10 and 100 independently [42] and (ii) a custom VBOTDA system (BGS, and BPS data) with a 25 km range with 1000 trace averages. The results for both cases are presented and analyzed below:

6.1. Custom BOTDA System Using 10 km Long Sensing Fiber

The three-dimensional BGS was constructed with a sweep frequency step of 1 MHz with different trace averages of 10 and 100. The resultant BGS spectra are illustrated in Figure 6. The mean $\hat{\mu}_{\nu_B}$ and standard deviations $\hat{\sigma}_{\nu_B}$ of the BFS along the 10 km fiber length obtained from the two datasets are plotted in Figure 7. The two datasets are obtained for various trace averages (10 traces for Figure 7a,b and 100 traces for Figure 7c,d). It should be noted that the first few tens of meters have high standard deviations due to the dead zones [43] in Figure 7. As the fiber distance increases, the standard deviations escalate due to fiber attenuation, resulting in diminishing signal-to-noise ratio. It is also clear from Figure 7b,d that the PDNN's BFS uncertainty estimates reflect the well-documented phenomenon [9] that BFS uncertainty reduces with an increase in the number of traces. The proposed approach with confidence interval outputs at every location on the fiber quantitatively estimates the performance metrics of the fiber optic sensor. This information can be used to mitigate any defects and to enhance the performance of the fiber optic sensor along the length of the pipeline.
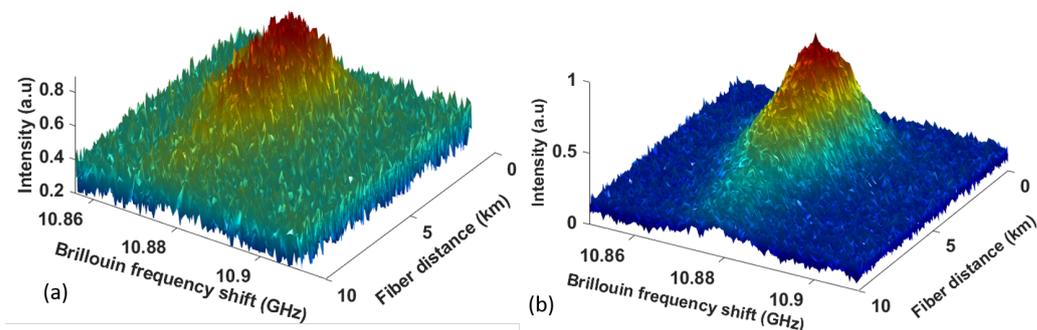
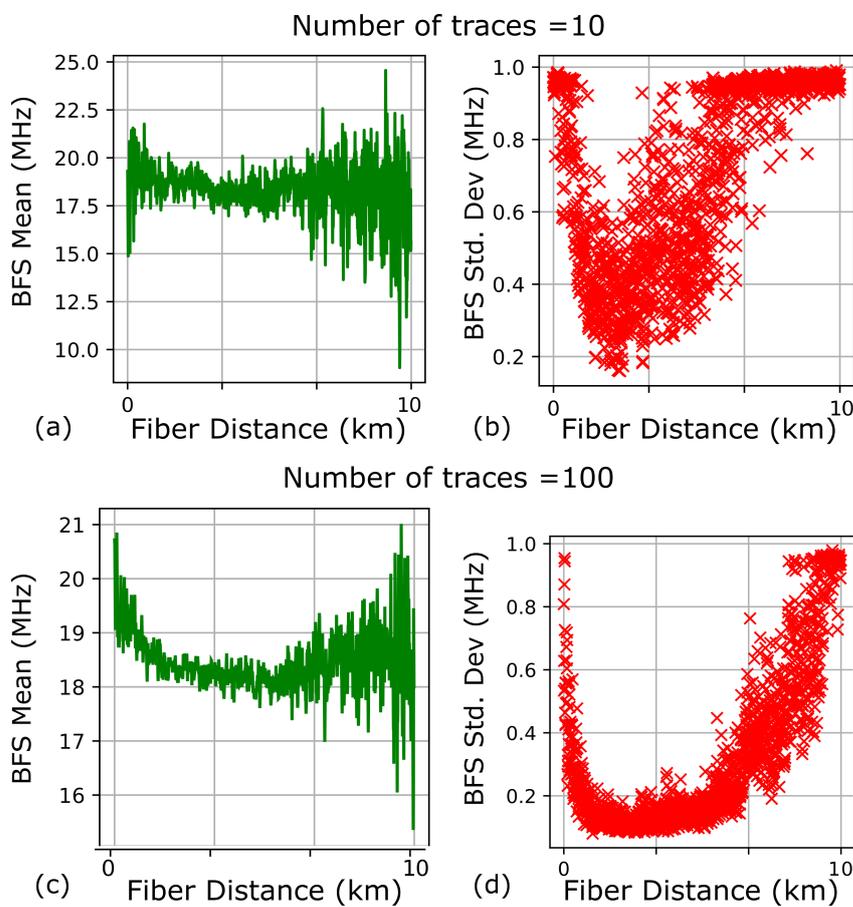**Figure 6.** Measured 3D BGS spectra using (**a**) 10 trace averages and (**b**) 100 trace averages.

Figure 7. PDNN estimates of mean (**a**,**c**) and standard deviations (Std. Dev) (**b**,**d**) from two 10 km BOTDA datasets. (**a**,**b**) PDNN estimates from BGS data averaged over 10 traces. (**c**,**d**) PDNN estimates from BGS data averaged over 100 traces.

6.2. Custom VBOTDA System Using 25 km Long Sensing Fiber

The BGS and BPS obtained over the fiber length are available as Dataset 1 [44] and Dataset 2 [45] respectively. Both are plotted in Figure 8.
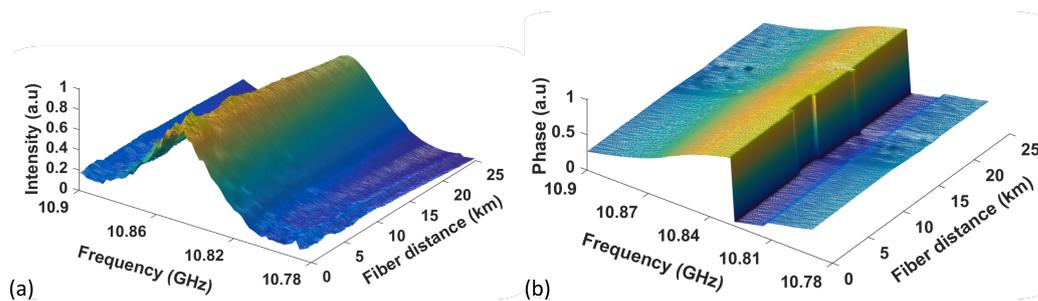
Figure 8. Measured 3D (**a**) BGS and (**b**) BPS over 25 km fiber.

The mean BFS (computed with respect to 10.8 GHz) is predicted using the PDNN model over the entire fiber length from the BGS and BPS spectra and is plotted with $\pm3\sigma_{\nu_B}$ (99.7%) confidence intervals in Figure 9a,c. These predictions are compared with a least-squares (LS) Lorentzian curve fitting model, as shown in Figure 9b,d. FWHM predictions from both models are shown in Figure 10.
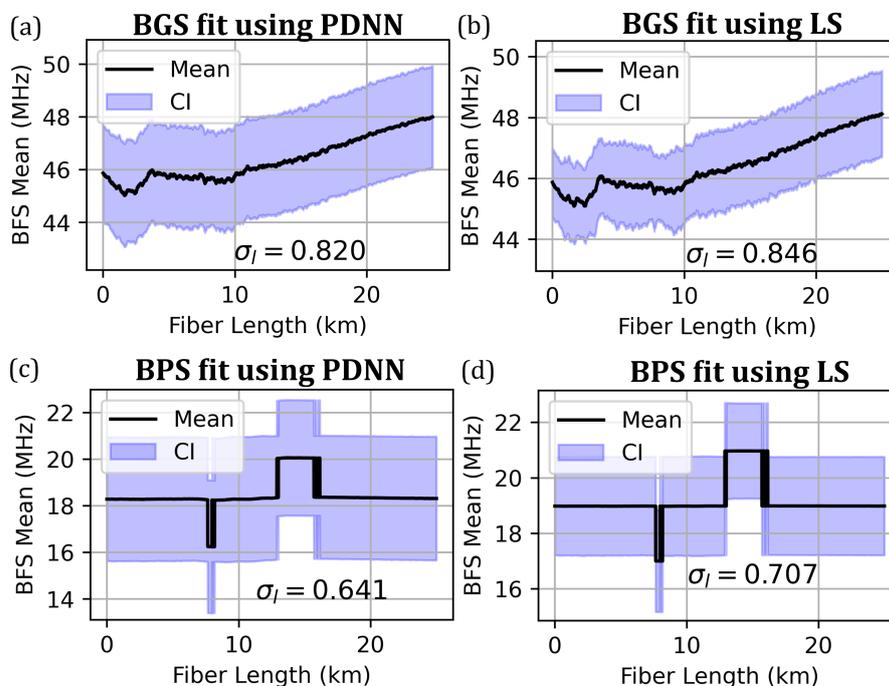

Figure 9. Comparison of BFS extracted using PDNN (**a**,**c**) and LS fit (**b**,**d**) from BGS or BPS measurements. The unit of confidence interval σ_l is MHz.
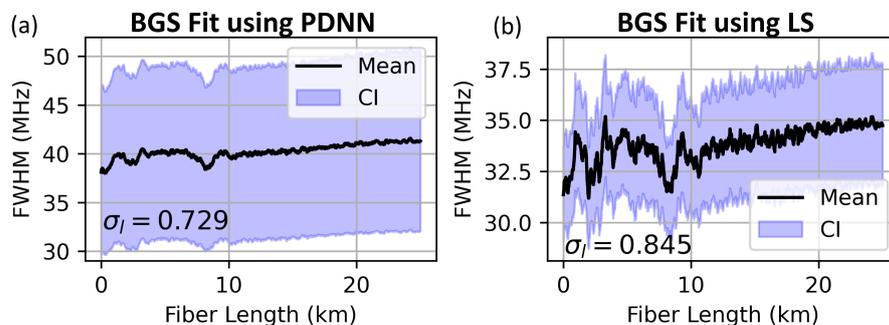

Figure 10. FWHM predictions using (**a**) PDNN and (**b**) LS fit from BGS data. The unit of confidence interval σ_l is MHz.

The mean BFS predictions from PDNN (for BGS and BPS) are in excellent agreement with those from LS fits. The PDNN processes 4500 spectra in 1.1 s (for both BGS and

BPS), while the CF approach takes 13.9 s and 18.5 s, respectively. The total computational speedup achieved with the PDNN approach is around 30 times more compared with the curve fitting approach. This quantum of speedup enables us to process the data in real time in the field as we collect sensor data all along the pipeline. Clearly, the PDNN model greatly reduces the data processing time. Furthermore, the BPS-based predictions exhibit lesser spatial uncertainty compared with that of BGS-based predictions.

7. Conclusions

We proposed and demonstrated a novel robust signal processing framework using PML to estimate the BFS from BOTDA systems. The PML is capable of processing BGS and BPS spectra in real time. Further, unlike ML models, which do not propagate uncertainties, our proposed PML model can provide estimates of predictive uncertainties. We compared the predictions obtained from the proposed PML with the conventional CF model and evaluated the BFS uncertainty and data processing time for both methods. The proposed PML model offers greater tolerance to measurement noise found in real-time strain and temperature extraction for a longer sensing range. Hence, the PML framework can be used to build a robust signal processing system and provides a pathway to enhance the VBOTDA system performance. Future advancements in ML techniques integrated with BOTDA sensor technology will continue to drive the evolution of distributed fiber optic sensors, enabling even more sophisticated and intelligent sensing capabilities. We plan to incorporate efficient data denoising techniques within the proposed PML framework in the future.

Author Contributions: Conceptualization, A.V., N.L., and P.L.; methodology, A.V. and N.L; software, A.V., S.R.B.; validation, A.V., N.L. and P.L.; formal analysis, A.V., S.R.B.; investigation, A.V. and N.L.; resources, R.W., and M.P.B.; data curation, A.V. and N.L.; writing—original draft preparation, A.V.; writing—review and editing, M.P.B. and R.W.; visualization, A.V.; supervision, M.P.B. and R.W.; project administration, R.W.; funding acquisition, R.W. and M.P.B. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data presented in this paper are available as Datasets 1-2 [44,45] and the code is available in Code 1 [41].

Conflicts of Interest: The authors declare no conflict of interest. A.V. and N.L. contributed equally to this paper. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, make any warranty, expressor implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

1. Bao, X.; Chen, L. Recent progress in Brillouin scattering based fiber sensors. *Sensors* **2011**, *11*, 4152–4187. [CrossRef] [PubMed]
2. Lu, P.; Lalam, N.; Badar, M.; Liu, B.; Chorpening, B.T.; Buric, M.P.; Ohodnicki, P.R. Distributed optical fiber sensing: Review and perspective. *Appl. Phys. Rev.* **2019**, *6*, 041302. [CrossRef]
3. Zhao, C.; Tang, M.; Wang, L.; Wu, H.; Zhao, Z.; Dang, Y.; Wu, J.; Fu, S.; Liu, D.; Shum, P.P. BOTDA using channel estimation with direct-detection optical OFDM technique. *Opt. Express* **2017**, *25*, 12698–12709. [CrossRef]

4.   He, H.; Yan, L.; Qian, H.; Li, Z.; Zhang, X.; Luo, B.; Pan, W. Efficient demodulation of Brillouin phase spectra and performance enhancement in BOTDA incorporating phase noise elimination. *J. Light. Technol.* **2019**, *37*, 4308–4314. [CrossRef]

5.   Li, Y.; An, Q.; Li, X.; Zhang, L. High-accuracy Brillouin frequency shift measurement system based on stimulated Brillouin scattering phase shift. *Opt. Eng.* **2017**, *56*, 056102. [CrossRef]

6.   Kadum, J.E.; Feng, C.; Schneider, T. Characterization of the Noise Induced by Stimulated Brillouin Scattering in Distributed Sensing. *Sensors* **2020**, *20*, 4311. [CrossRef] [PubMed]

7.   Tu, X.; Sun, Q.; Chen, W.; Chen, M.; Meng, Z. Vector Brillouin optical time-domain analysis with heterodyne detection and IQ demodulation algorithm. *IEEE Photonics J.* **2014**, *6*, 1–8. [CrossRef]

8.   Dossou, M.; Bacquet, D.; Szriftgiser, P. Vector Brillouin optical time-domain analyzer for high-order acoustic modes. *Opt. Lett.* **2010**, *35*, 3850–3852. [CrossRef]

9.   Lopez-Gil, A.; Soto, M.A.; Angulo-Vinuesa, X.; Dominguez-Lopez, A.; Martin-Lopez, S.; Thévenaz, L.; Gonzalez-Herraez, M. Evaluation of the accuracy of BOTDA systems based on the phase spectral response. *Opt. Express* **2016**, *24*, 17200–17214. [CrossRef]

10.   Lu, P.; Lalam, N.; Liu, B.; Buric, M.; Ohodnicki, P.R. Vector Brillouin optical time-domain analysis with Raman amplification and optical pulse coding. In Proceedings of the Photonic Instrumentation Engineering VI. International Society for Optics and Photonics, San Francisco, CA, USA, 5–7 February 2019; Volume 10925, p. 1092512.

11.   Wang, B.; Guo, N.; Wang, L.; Yu, C.; Lu, C. Robust and fast temperature extraction for Brillouin optical time-domain analyzer by using denoising autoencoder-based deep neural networks. *IEEE Sens. J.* **2019**, *20*, 3614–3620. [CrossRef]

12.   Venketeswaran, A.; Lalam, N.; Wuenschell, J.; Ohodnicki, P.R., Jr.; Badar, M.; Chen, K.P.; Lu, P.; Duan, Y.; Chorpening, B.; Buric, M. Recent Advances in Machine Learning for Fiber Optic Sensor Applications. *Adv. Intell. Syst.* **2021**, *4*, 2100067. [CrossRef]

13.   Lalam, N.; Ng, W.P.; Dai, X.; Wu, Q.; Fu, Y.Q. Performance improvement of Brillouin ring laser based BOTDR system employing a wavelength diversity technique. *J. Light. Technol.* **2018**, *36*, 1084–1090. [CrossRef]

14.   Urricelqui, J.; Soto, M.A.; Thévenaz, L. Sources of noise in Brillouin optical time-domain analyzers. In Proceedings of the 24th International Conference on Optical Fibre Sensors, SPIE, Curitiba, Brazil, 28 September–2 October 2015; Volume 9634, pp. 377–380.

15.   Zhang, Y.; Fu, G.; Liu, Y.; Bi, W.; Li, D. A novel fitting algorithm for Brillouin scattering spectrum of distributed sensing systems based on RBFN networks. *Opt.-Int. J. Light Electron Opt.* **2013**, *124*, 718–721. [CrossRef]

16.   da Silva, L.C.B.; Samatelo, J.L.A.; Segatto, M.E.V.; Bazzo, J.P.; da Silva, J.C.C.; Martelli, C.; Pontes, M.J. NARX neural network model for strong resolution improvement in a distributed temperature sensor. *Appl. Opt.* **2018**, *57*, 5859–5864. [CrossRef]

17.   Wu, H.; Wang, L.; Guo, N.; Shu, C.; Lu, C. Support vector machine assisted BOTDA utilizing combined Brillouin gain and phase information for enhanced sensing accuracy. *Opt. Express* **2017**, *25*, 31210–31220. [CrossRef] [PubMed]

18.   Wang, B.; Wang, L.; Guo, N.; Zhao, Z.; Yu, C.; Lu, C. Deep neural networks assisted BOTDA for simultaneous temperature and strain measurement with enhanced accuracy. *Opt. Express* **2019**, *27*, 2530–2543. [CrossRef] [PubMed]

19.   Venketeswaran, A.; Lalam, N.; Lu, P.; Ohodnicki, P.R.; Chen, K.P. Enhanced Signal Processing of Distributed Brillouin Fiber Sensors using a Decoupled Radial Basis Function Network. In *Proceedings of the Optical Fiber Sensors, Washington, DC, USA, 8–12 June 2020*; Optica Publishing Group: Washington, DC, USA, 2020; p. T3-51.

20.   Lalam, N.; Lu, P.; Venketeswaran, A.; Buric, M.; Ohodnicki, P.R. Raman-assisted BOTDA performance improvement with the differential pulse-width pair technique and an artificial neural network based fitting algorithm. In *Proceedings of the Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020, Virtual, 27 April–8 May 2020*; International Society for Optics and Photonics: Bellingham, WA, USA, 2020; Volume 11415, p. 1141503.

21.   Soto, M.A.; Thévenaz, L. Modeling and evaluating the performance of Brillouin distributed optical fiber sensors. *Opt. Express* **2013**, *21*, 31347–31366. [CrossRef]

22.   Lakshminarayanan, B.; Pritzel, A.; Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. *arXiv* **2016**, arXiv:1612.01474.

23.   Ovadia, Y.; Fertig, E.; Ren, J.; Nado, Z.; Sculley, D.; Nowozin, S.; Dillon, J.V.; Lakshminarayanan, B.; Snoek, J. Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift. *arXiv* **2019**, arXiv:1906.02530.

24.   Gal, Y. Uncertainty in Deep Learning. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2016.

25.   Xu, Z.; Zhao, L. Key parameter extraction for fiber Brillouin distributed sensors based on the exact model. *Sensors* **2018**, *18*, 2419. [CrossRef]

26.   Xu, Z.; Zhao, L.; Qin, H. Selection of spectrum model in estimation of Brillouin frequency shift for distributed optical fiber sensor. *Optik* **2019**, *199*, 163355. [CrossRef]

27.   Alem, M.; Soto, M.A.; Tur, M.; Thévenaz, L. Analytical expression and experimental validation of the Brillouin gain spectral broadening at any sensing spatial resolution. In Proceedings of the 2017 25th Optical Fiber Sensors Conference (OFS), Jeju Island, Republic of Korea, 24–28 April 2017; pp. 1–4.

28.   Lopez-Gil, A.; Angulo-Vinuesa, X.; Soto, M.A.; Dominguez-Lopez, A.; Martin-Lopez, S.; Thévenaz, L.; Gonzalez-Herraez, M. Gain vs phase in BOTDA setups. In Proceedings of the Sixth European Workshop on Optical Fibre Sensors. International Society for Optics and Photonics, Limerick, Ireland, 31 May–3 June 2016; Volume 9916, p. 991631.

29.   Seber, G.A.; Wild, C.J. *Nonlinear Regression*; John Wiley Sons: Hoboken, NJ, USA, 2003; Volume 62, p. 63.

30.   Azad, A.K.; Khan, F.N.; Alarashi, W.H.; Guo, N.; Lau, A.P.T.; Lu, C. Temperature extraction in Brillouin optical time-domain analysis sensors using principal component analysis based pattern recognition. *Opt. Express* **2017**, *25*, 16534–16549. [CrossRef]

31. Azad, A.K.; Wang, L.; Guo, N.; Tam, H.Y.; Lu, C. Signal processing using artificial neural network for BOTDA sensor system. *Opt. Express* **2016**, *24*, 6769–6782. [CrossRef]
32. Liehr, S. Artificial neural networks for distributed optical fiber sensing. In *Proceedings of the Optical Fiber Communication Conference, Washington, DC, USA, 6–11 June 2021*; Optical Society of America: Washington, DC, USA, 2021; p. Th4F-2.
33. Wu, H.; Wan, Y.; Tang, M.; Chen, Y.; Zhao, C.; Liao, R.; Chang, Y.; Fu, S.; Shum, P.P.; Liu, D. Real-time denoising of Brillouin optical time domain analyzer with high data fidelity using convolutional neural networks. *J. Light. Technol.* **2019**, *37*, 2648–2653. [CrossRef]
34. Zheng, H.; Yan, Y.; Wang, Y.; Sben, X.; Lu, C. Deep learning enhanced long-range fast BOTDA for vibration measurement. *J. Light. Technol.* **2021**, *40*, 262–268. [CrossRef]
35. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
36. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
37. Efron, B. Bootstrap methods: Another look at the jackknife. In *Breakthroughs in Statistics*; Springer: Berlin/Heidelberg, Germany 1992; pp. 569–593.
38. Cao, W.; Wang, X.; Ming, Z.; Gao, J. A review on neural networks with random weights. *Neurocomputing* **2018**, *275*, 278–287. [CrossRef]
39. Nix, D.A.; Weigend, A.S. Estimating the mean and variance of the target probability distribution. In Proceedings of the 1994 IEEE International Conference on Neural Networks (ICNN'94), Orlando, FL, USA, 28 June–2 July 1994; Volume 1, pp. 55–60.
40. De Brabanter, K.; Karsmakers, P.; De Brabanter, J.; Suykens, J.A.; De Moor, B. Confidence bands for least squares support vector machine classifiers: A regression approach. *Pattern Recognit.* **2012**, *45*, 2280–2287. [CrossRef]
41. Venketeswaran, A.; Lalam, N.; Lu, P.; Buric, M. Jupyter notebook containing the code for BFS and FWHM estimation using PDNN. Figshare. 2021. Available online: https://osapublishing.figshare.com/s/0fff0771379de177d5b3 (accessed on 9 May 2023).
42. Chen, T.; Xu, X.; Lalam, N.; Ng, W.P.; Harrington, P. Multi point strain and temperature sensing based on Brillouin optical time domain reflectometry. In Proceedings of the 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), Budapest, Hungary, 18–20 July 2018; pp. 1–6.
43. Muanenda, Y.; Taki, M.; Pasquale, F.D. Long-range accelerated BOTDA sensor using adaptive linear prediction and cyclic coding. *Opt. Lett.* **2014**, *39*, 5411–5414. [CrossRef]
44. Venketeswaran, A.; Lalam, N.; Lu, P.; Buric, M. Dataset for gain spectra for 25km long fibre. Figshare, 2021. Available online: https://osapublishing.figshare.com/s/2e08c4082b7653f665fb (accessed on 9 May 2023).
45. Venketeswaran, A.; Lalam, N.; Lu, P.; Buric, M. Dataset for phase spectra for 25km long fibre. Figshare, 2021. Available online: https://osapublishing.figshare.com/s/1ca66af63c3245a9c986 (accessed on 9 May 2023).