Research Article | Vol. 30, No. 8 / 11 Apr 2022 / Optics Express 12484
Optics EXPRESS

Distributed humidity fiber-optic sensor based on BOFDA using a simple machine learning approach

CHRISTOS KARAPANAGIOTIS,* KONSTANTIN HICKE, ALEKSANDER WOSNIOK, AND KATERINA KREBBER

Bundesanstalt für Materialforschung und-Prüfung, Unter den Eichen 87, 12205 Berlin, Germany
christos.karapanagiotis@bam.de

Abstract: We report, to our knowledge for the first time, on distributed relative humidity sensing in silica polyimide-coated optical fibers using Brillouin optical frequency domain analysis (BOFDA). Linear regression, which is a simple and well-interpretable algorithm in machine learning and statistics, is utilized. The algorithm is trained using as features the Brillouin frequency shifts and linewidths of the fiber's multipeak Brillouin spectrum. To assess and improve the effectiveness of the regression algorithm, we make use of machine learning concepts to estimate the model's uncertainties and select the features that contribute most to the model's performance. In addition to relative humidity, the model is also able to simultaneously provide distributed temperature information addressing the well-known cross-sensitivity effects.

1. Introduction

Humidity and temperature are key environmental factors with their monitoring being of great importance for a wide range of civil and geotechnical applications, from structural health monitoring of dikes, dams, and embankments to long pipeline and subsea cable corrosion detection. Temperature distributed fiber optic sensors (DFOS) are well studied and they are attractive for many practical applications due to their ability to provide distributed sensing over long distances with high spatial resolution, to work remotely and in harsh environments and due to their immunity to electromagnetic interference [1].

Fiber optic humidity sensors have progressed significantly over the last years [2], however target-oriented investigations of truly distributed fiber optic humidity sensing are still a very recent development. Silica optical fibers (SOF) without coating are not sensitive to humidity and approaches such as humidity-induced strain based on the usage of hygroscopic coatings have been proposed [2–5]. Among the most prominent coating materials for humidity-induced strain is polyimide (PI) but we note that studies have shown that standard SOF with acrylate coating are also sensitive to humidity with their sensitivity depending highly on temperature [6]. Nevertheless, the sensitivity of the PI-coated SOF is significantly higher than that of acrylate-coated SOF due to the higher hygroscopic capacity. Other solutions based on polymer optical fibers (POF) were also examined [7] and tested in the field [8].

The challenge of distributed humidity monitoring in SOF does not only lay in the fact that they are not inherently sensitive to humidity. Even when the aforementioned humidity sensing solutions are used, the system's response is affected by the occurrence of cross-sensitivities such as temperature or tensile strain [9,10]. Strain cross-sensitivity effects in humidity sensing can be minimized by mechanically isolating an optical fiber, e.g., with a protective conduit. On the other hand, temperature influences, cannot be isolated and thus solutions that include the discrimination between temperature and humidity are required. A simple approach for distributed temperature and relative humidity sensing that has been proposed makes use of a two-fiber configuration, where one fiber is coated with acrylate (which is almost insensitive to humidity)

#453906 https://doi.org/10.1364/OE.453906
Journal © 2022 Received 27 Jan 2022; revised 3 Mar 2022; accepted 7 Mar 2022; published 29 Mar 2022

Optics EXPRESS

and one with PI which is sensitive to both measurands [3]. However, this method suffers from reduced spatial comparability and is impractical for many applications where only one fiber is available. Nevertheless, it is worth mentioning that simultaneous temperature and humidity measurements in optical fibers have been reported for point and quasi-distributed fiber Bragg grating sensors [11–14].

Brillouin optical frequency domain analysis (BOFDA) is a well-established method for temperature and strain distributed fiber sensing [15] and can be a promising candidate for distributed humidity sensing as well. BOFDA can be employed for many km-long optical fiber sensing and can provide a spatial resolution even on a centimeter scale [16,17]. Furthermore, its data acquisition does not involve the use of fast electronics rendering BOFDA a cost-effective solution [15].

To our knowledge, this is the first report on distributed humidity BOFDA sensing using a single PI-coated SOF which is also capable to discriminate temperature and humidity effects. First, we demonstrate a Brillouin optical frequency-domain analysis (BOFDA) setup including a laser diode with high frequency stability that is required when using commercially available fibers with low sensitivity to humidity changes. Humidity and temperature predictions are made by a linear regression algorithm, which is a simple and well-understood algorithm in machine learning and statistics [18]. The linear regression model is trained by using as features not only the Brillouin frequency shifts (BFS) but also the linewidths of the multi-peak Brillouin gain spectrum (BGS). The performance of the model is evaluated using a leave-one-out cross-validation-based method [19] to ensure that the model generalizes and interpolates well on new data. Furthermore, we evaluate the importance of every feature on the model's performance, and we estimate the minimum number of features that is required using backward feature elimination [18].

2. Sensing principle

2.1. Feature extraction and linear regression algorithm

In stimulated Brillouin backscattering, and particularly in BOFDA, two counterpropagating continuous optical waves, pump and probe (Stokes) with frequencies ν_p and ν_s, respectively, generate an optical beating with frequency $\nu = \nu_p - \nu_s$. Only if the frequency offset ν between the pump and probe matches the so-called BFS ν_b, the Stokes wave is amplified. A slight detuning of the frequency offset can still initiate Stokes amplification due to the medium's damping rate, which is related to the Brillouin linewidth w. Similar to time-domain Brillouin backscattering, the BGS is described by a sum of Lorentzian profiles [20]:

$$g(\nu) = g(\nu_p - \nu_s) = \sum_j g_{Bj} \frac{\left(\frac{w_j}{2}\right)^2}{(\nu - \nu_{bj})^2 + \left(\frac{w_j}{2}\right)^2} \tag{1}$$

where g_{Bj} is the gain of the j^{th} peak. The number of the Lorentzian profiles depends on the existing acoustic modes which, in turn, are affected by the doping concentration and the geometry of the optical fiber [20]. The spectral properties of the BGS, namely the BFS, the linewidth and the gain (i.e., the peak amplitude) can be determined by performing Lorentzian curve fitting (LCF). In this paper, the LCF is performed by using the Lmfit Python library for nonlinear least-squares minimization and curve-fitting problems [21].

In Brillouin backscattered distributed sensing the BFS depends on the refractive index and the velocity of the longitudinal acoustic waves, and it is shown that it is linearly proportional to temperature and tensile strain [15]. A linear relation should be assumed for humidity-induced strain as well. However, the linearity of BFS can be disturbed by humidity-induced hysteresis effects depending on the type of optical fiber [9]. In a previous paper, we examined such effects

Optics EXPRESS

in commercially-available PI-coated SOF with different coating thicknesses and we found that after annealing, hysteresis effects were negligible in those fibers [3].

The linewidth of the BGS could serve as an additional parameter with great significance for multiparameter sensing [13,22]. The linewidth is related to the lifetime of the acoustic phonons τ_B as follows [23] :

$$w \propto \frac{1}{\tau_B} \tag{2}$$

The phonon lifetime increases with temperature and hence the linewidth decreases. Due to the low linewidth temperature sensitivity, an experimental setup with a high signal-to-noise ratio (SNR) and laser stability is required. Studies on the influence of tensile strain on linewidth show that they are not related to each other [24] and the same can be assumed for humidity-induced strain. This will be studied in the results section of this paper.

The impact of temperature and relative humidity on BFS and linewidth can be described conventionally as follows:

$$Y = CX + C_0 \tag{3}$$

where

$$Y = \begin{bmatrix} \nu^{(1)} \\ w^{(1)} \\ \vdots \\ \nu^{(n)} \\ w^{(n)} \end{bmatrix}, \quad X = \begin{bmatrix} T \\ RH \end{bmatrix}, \quad C = \begin{bmatrix} C_{T,\nu}^{(1)} & C_{RH,\nu}^{(1)} \\ C_{T,w}^{(1)} & C_{RH,w}^{(1)} \\ \vdots & \vdots \\ C_{T,\nu}^{(n)} & C_{RH,\nu}^{(n)} \\ C_{T,w}^{(n)} & C_{RH,w}^{(n)} \end{bmatrix} \tag{4}$$

Equation (3) is a multi-output linear equation with the independent variables being the temperature (T) and the relative humidity (RH). $\nu^{(n)}$ and $w^{(n)}$ represent the BFS and linewidth of the nth peak, respectively. C is the matrix that describes the change of every BFS and linewidth if temperature and relative humidity are changed by one unit (1 °C and 1%RH respectively). $C_{T,\nu}^{(n)}$ and $C_{RH,\nu}^{(n)}$ are the temperature and humidity coefficients (sensitivity) of the nth BFS, respectively; while $C_{T,w}^{(n)}$ and $C_{RH,w}^{(n)}$ are the temperature and relative humidity coefficients of the nth linewidth. C_0 is the offset.

In this paper, we predict humidity and temperature by making use of a linear regression algorithm with the measurands of interest being the target variables (multi-target regression) and the BFS and linewidths of the multipeak BGS representing the features. In other words, the algorithm is trained to predict X by using Y and hence the Eq. (3) can be rewritten as:

$$X = C'Y + C_0' \tag{5}$$

The regression algorithm employs the least-square optimization procedure, which minimizes the residual sum of squares between the true target values and the predicted ones. For the training of the regression algorithm the state-of-the-art Python library called "Scikit-learn: Machine Learning in Python" is utilized [25].

2.2. Fiber configuration

Our approach is tested with a commercially available PI-coated optical fiber with a 15 µm coating thickness (Fibercore SM1250(10.4/125)P). Figure 1 shows a simplified illustration of the fiber configuration. The total length of the fiber is approximately 400 m, where the last 60 m are placed in a climate chamber. This part of the fiber exposed in controlled environmental conditions in the climate chamber is conventionally called fiber-under-test (FUT). The spatial sampling of the

Optics EXPRESS

BOFDA setup is set to 6 m. The set temperature range is from 40 °C to 60 °C with a step of 4 °C, while the relative humidity is set in the range from 20% to 80% with a step of 10%. This results in 42 combinations of temperature and humidity.

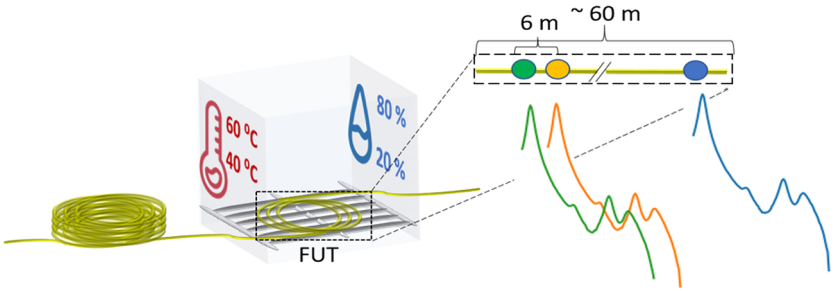

Fig. 1. Fiber configuration. The fiber-under-test (FUT) is placed in a climate chamber where the temperature and humidity range from 40 °C to 60 °C and from 20% to 80%, respectively. The length of the FUT is approximately 60 m and the spatial sampling is set to 6 m.

All the measurements are carried out using a Vötsch VCL4006 climate chamber, which allows simultaneous programmable and remote control of temperature and relative humidity. With the help of an external sensor we determine the minimum time that is needed to stabilize around the set temperature and relative humidity. The stabilization time is approximately 50 min, which depends on the change of the environmental conditions. Furthermore, when the stabilization time is elapsed the temperature and relative humidity uncertainty within a ten-minute time window is determined to be 0.08 °C and 0.2%, respectively.

2.3. Error estimation

Generalizability and error estimation is of great importance even for the simplest machine learning model. We evaluate the model's performance on its ability to generalize not simply on new data, but on different temperature and humidity values than those included in the training set. This approach will show the model's interpolation performance and a more realistic estimation of the errors that are expected. Specifically, we make use of a leave-one-out cross-validation method [19] by iteratively excluding from the training dataset a set of temperature and humidity and using it for testing. The mean error resulting from the testing errors from all the 42 sets of temperature and humidity is called variance [26]. The procedure is illustrated in Fig. 2.

2.4. Experimental setup

The experimental setup is shown in Fig. 3(a). A distributed feedback laser (DFB) that operates at 1550 nm is employed. Due to the low humidity sensitivity of the PI-coated fibers, a laser module with high frequency stability is required to minimize the statistical (aleatoric) uncertainty. In Fig. 3(b) we show in form of a violin plot the distribution of the absolute BFS deviation of the FUT resulting from subsequent measurements in stable environmental conditions in a climate chamber. We observe that the median value is around 40 kHz which indicates that for a FUT with humidity sensitivity equal to 100 kHz/%RH, the median estimated humidity error arising from the experimental setup is estimated to be around 0.4%RH.

A 20/80 polarization-maintaining optical coupler is utilized to split the laser optical power into two paths. The upper path (probe) is responsible for the laser frequency tuning, necessary for the stimulated Brillouin scattering, while the lower path (pump) is used for the acquisition of spatially resolved information. In the probe path, the frequency is tuned by a Mach–Zehnder modulator (MZM 1) driven by an RF signal generator (SG), which generates two sidebands and suppresses the carrier [27] as shown in Fig. 3(a). The lower sideband is filtered out by a fiber

Optics EXPRESS

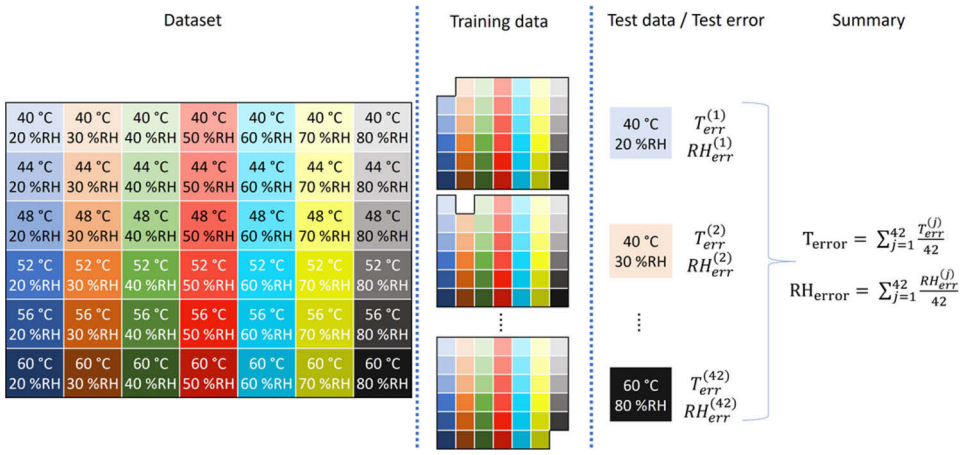

Fig. 2. Graphical representation of whole dataset and the pipeline of the error estimation based on leave-one-out cross-validation. The data were collected from measurements at n = 42 combinations of temperature and relative humidity.

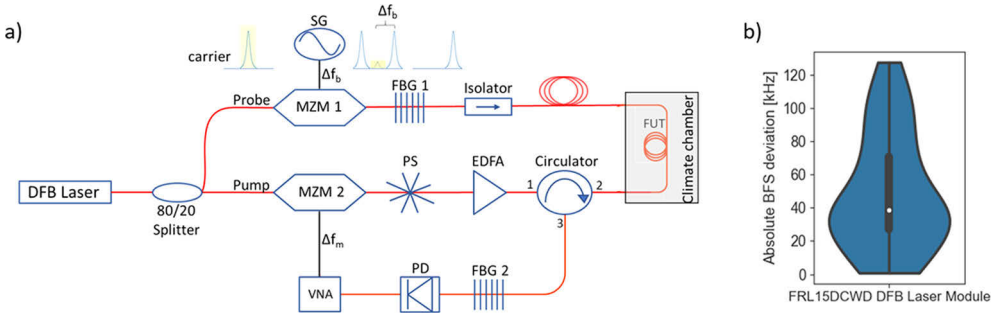

Fig. 3. a) BOFDA experimental setup. DFB: distributed feedback laser; MZM: Mach–Zehnder modulator; SG: signal generator; FBG: fiber Bragg grating; EDFA: erbium-doped fiber amplifier; PS: polarization scrambler; FUT: fiber under test; PD: photodiode; VNA: vector network analyzer. b) Distribution of the absolute BFS deviation at stable temperature and humidity conditions in a climate chamber. The white dot represents the median value; the black bar in the center of the plot the interquartile range; and the black lines stretched from the bar are the lower and upper adjacent values.

Bragg grating (FBG 1). The isolator is operated to protect the optical components in the probe path from the transmitted pump signal. The MZM in the pump path (MZM 2) is operated in the linear range of the transfer function and gradually modulates the amplitude of the continuous wave sinusoidally in a range of modulation frequencies that determines the possible measurement length and the spatial resolution [28]. A fiber squeezer-based polarization scrambler (PS) with an operation frequency at 700 kHz is used to alleviate the polarization fading [29]. After the PS, an erbium-doped fiber amplifier (EDFA) is employed to amplify the pump signal. The amplification needed depends mainly on the fiber length. The FBG 2 is used to filter out the Rayleigh component of the backscattered signal that arises from the fiber under test (FUT) letting only the Brillouin component to reach the photodiode. In the end, the VNA measures the system's response and modulates the MZM 2.

All the measurements are conducted with the same configuration in the experimental setup. The laser module's output power is set to 40 mW, the modulation step frequency Δf_m to 128

Optics EXPRESS

kHz and the Brillouin sweep step Δf_b to 4 MHz. The resolution bandwidth and the number of averages in the VNA to reach a high SNR are 1 kHz and 3, respectively. All these configurations result in a total measurement time of 5 min for the optical fiber at hand.

3. Results

3.1. Influence of relative humidity on the BFS

Similarly with acrylate-coated fibers, the BFS varies linearly with temperature and for the particular FUT at hand, the temperature coefficient is estimated to be 1.1 MHz/°C at 1550nm. In this section, we focus on the influence of humidity on the BFS.

The normalized data and the Lorentzian fitting of the fundamental peak at two humidity levels of 20% and 80% and at a constant temperature of 40 ° C is shown in Fig. 4(a). The data correspond to a random spatial position in the FUT. The manifest shift of the BGS with humidity towards higher frequencies shows that Brillouin backscattering in PI-coated optical fibers is sensitive to humidity variations. The good quality of Lorentzian fitting is attributed to the high SNR of the experimental setup.

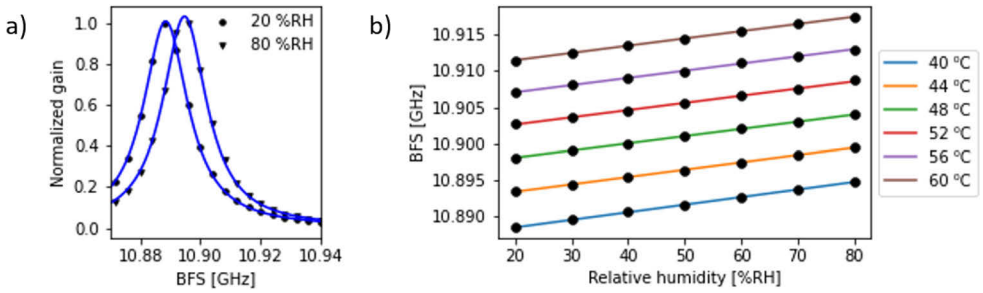

Fig. 4. a) Humidity-induced shift of the fundamental BGS in a PI-coated SOF. b) Linear fitting of the fundamental BFS with humidity at different values of temperature from 40 °C to 60 °.

The relation of the BFS to relative humidity at different temperature values in the range from 40 °C to 60 °C is depicted in Fig. 4(b). The black dots represent the mean BFS of all positions in the FUT at stable temperature and humidity conditions. We observe that the BFS depends linearly on humidity at each temperature value and the R-squared is calculated equal to 0.999. The mean BFS humidity coefficient at every temperature is 100kHz/%RH with a standard deviation of 2 kHz. The low standard deviation indicates that the humidity coefficient is practically independent of temperature.

3.2. Relative humidity sensing and temperature cross-sensitivity

Although the fundamental BFS is related to the humidity, it can not address the temperature cross-sensitivity and thus more features that are not highly correlated are needed. These features can be the linewidths and the secondary BFS of the extended BGS. The BGS of the PI-coated fiber is characterized by multiple Brillouin peaks providing additional features that can be used to decouple temperature and humidity effects. A characteristic normalized logarithmic BGS of the FUT is presented in Fig. 5. The experimental data are shown with blue dots while the result of multiple LCF is depicted with the red curve. The BGS consists of four distinct peaks but the LCF algorithm makes use of five Lorentzian components since the last peak is very broad indicating that most likely two peaks are overlapping. In Fig. 5 all the Lorentzian components are illustrated with different colors as the legend indicates.

Optics EXPRESS

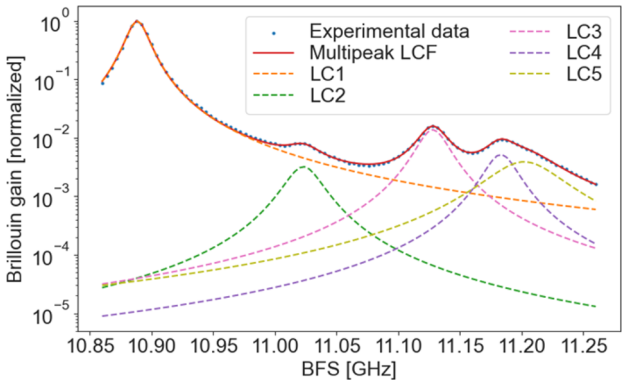

Fig. 5. Characteristic experimental logarithmic Brillouin gain spectrum (blue dots) consisting of 5 Lorentzian peaks. The multiple LCF is shown with a red curve while these of the individual Lorentzian components (LC) with dashed curves.

A linear regression model is trained as described in section 2.1 using the BFS and linewidths of all the Lorentzian components and validated as described in section 2.3. As an example, in Fig. 6 we present some characteristic distributed relative humidity predictions along the FUT at different set temperature conditions. As described in section 2.2 the FUT is around 60 m and the spatial sampling 6 m. The dots represent the relative humidity predictions extracted by the linear regression algorithm, while the dashed lines show the set relative humidity in the climate chamber. We note that the uncertainty of the relative humidity in the climate chamber is expected to be approximately 0.2%, as explained in section 2.2.

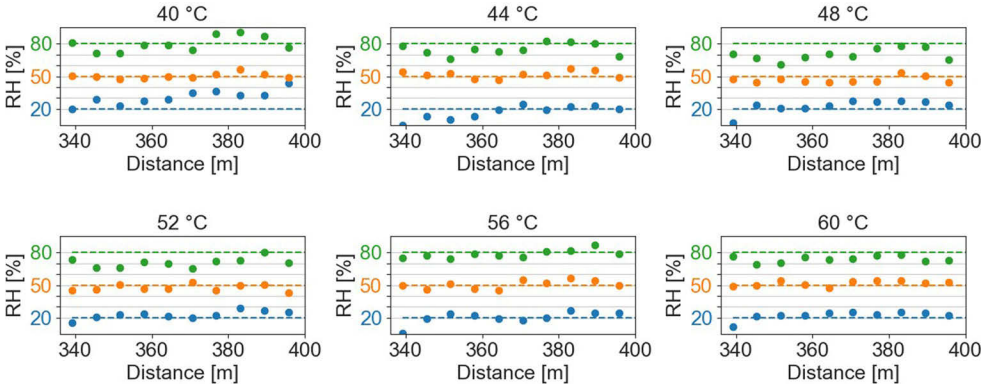

Fig. 6. Distributed relative humidity (RH) at different set temperatures along the FUT extracted by the linear regression algorithm. The dots represent the distributed RH predictions while the dashed lines the set RH in the climate chamber.

The model's performance is estimated in terms of mean absolute error (MAE). The total MAE and its standard deviation values are 0.8 °C and 0.4 °C, respectively; while those of the relative humidity are 8.7% and 4.3%, respectively. These errors correspond to the model's variance and shows the model's ability to generalize and interpolate on new data.

In linear regression it happens very often that some of the features are highly correlated without contributing to the model's performance or they are independent of the label. For this reason, we identify the minimum number of features that are necessary for temperature and humidity discrimination.

Optics EXPRESS

The BFS coefficients for temperature and humidity are depicted in Fig. 7(a). The blue bars correspond to the temperature coefficients while the orange ones represent the humidity BFS coefficients. The error bars illustrate the standard deviation of the coefficients. We notice that all temperature and humidity coefficients are very close to each other and have a mean value of 1.1 MHz/°C and 100 kHz/%RH, respectively. The error is higher at higher-order peaks. This indicates that all these features are highly correlated to each other and thus one can assume that the usage of more than one BFS does not improve the linear regressor's performance. For this reason, only one BFS is kept, and in particular the BFS of the fundamental peak due to the lower errors.

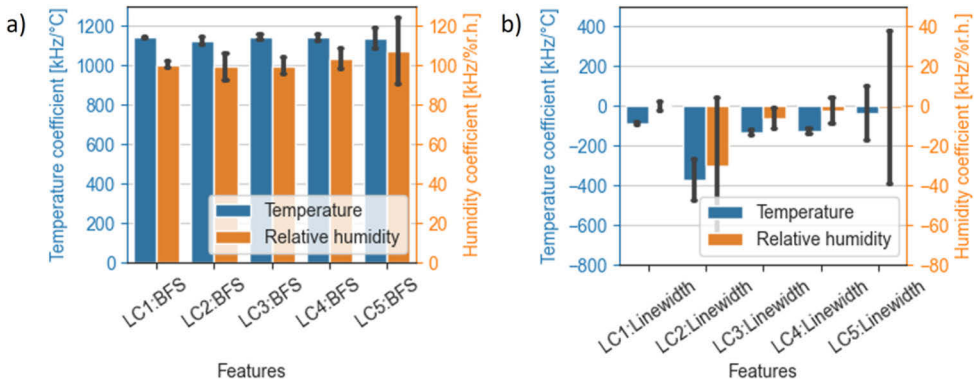

Fig. 7. Humidity (orange) and temperature (blue) coefficients of the BFS (a) and the linewidths (b).

Figure 7(b) shows the temperature and humidity linewidth coefficients. We observe that except for the second peak, the linewidth humidity coefficients are very close to zero, which indicates that there is no relation between them. Apart from the low coefficient values, their errors are too high to offer useful information. On the other hand, the linewidth clearly decreases with the temperature and hence the temperature coefficients have the potential to be very informative. The coefficient of the last peak can be safely excluded due to the large errors.

A linear regressor is trained again with the five remaining features providing a similar performance compared to that trained by using all the features. This proves that indeed the excluded features do not improve the performance. Afterwards, backward feature elimination

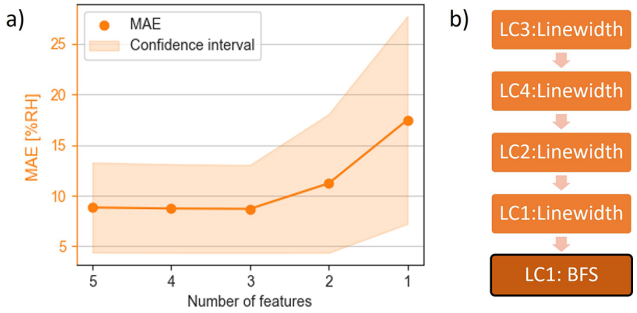

Fig. 8. a) MAE of the relative humidity vs the number of features left after each iteration of the backward feature elimination. The confidence interval represents the standard deviation of the prediction errors. b) Feature elimination sequence. The last feature is the one left when the algorithm is finished.

Optics EXPRESS

is applied to remove additional features that do not contribute significantly. Starting with five features, the algorithm iteratively removes the one that contributes least to the model's final performance until only one feature is left (Fig. 8). We notice that the model's performance starts decreasing only after the removal of the third feature which manifests that only three features are required, and in particular the BFS and the linewidth, respectively, of the fundamental peak and the linewidth of the first secondary peak. The feature elimination sequence is shown in Fig. 8(b).

4. Discussion

We have shown that distributed BOFDA sensing can be employed for humidity measurements in PI-coated fibers and can be used to decouple the temperature and humidity effects. The sensitivity of the conventionally used BFS of the fundamental peak of the investigated PI-coated fiber is around 100 kHz/%RH and its standard deviation in the temperature range between 40 °C and 60 °C is only 2 kHz. The latter renders the influence of temperature on the humidity coefficients negligible which allows the use of a very simple linear regression model as described in section 2.1.

We applied feature selection techniques to minimize the number of features that are needed to acquire a low humidity error. We observed that mostly the features resulting from the first two Lorentzian components contribute to the system's performance. This manifests that instead of measuring the whole BGS, one could acquire the part of the spectrum that contains the two most informative Lorentzian peaks. Observing the BGS in Fig. 5, it is clear that only a sweep range of 0.2 GHz is required to obtain those peaks. This results in a twofold reduction of the measurement time because the latter is proportional to the Brillouin sweep range. Specifically, in this work, the measurement time could be shortened from 5 min to 2.5 min. The shortening of the measurement time is of particular interest in BOFDA, which is usually characterized by time-consuming measurements [17]. We note that recently we proposed a method based on convolutional neural networks that shortened the measurement time significantly opening the way for BOFDA to meet new applications, where faster monitoring is required [28].

Temperature and humidity inhomogeneities within the defined spatial resolution can affect the BGS and thus the extracted values of BFS and linewidth, respectively. Specifically, such inhomogeneities would result in a broader BGS with a BFS corresponding to the mean value of temperature and humidity levels. To avoid such effects, the spatial sampling rate can be increased so that no significant changes of humidity and temperature exist within the distance defined by the spatial resolution.

Our reported approach was applied in data recorded using a BOFDA setup. Since the algorithm's input is the BFS and the linewidths which can be also extracted from other Brillouin based techniques as well (e.g. BOTDA and BOCDA) the approach is not limited for BOFDA. However, due to the different working principle of the various Brillouin-based techniques, variations in the humidity and temperature errors, respectively, are expected."

Apart from the potential implementation of our approach in other Brillouin-based techniques, we note that our methodology can be applied for temperature and strain discrimination as well. Humidity variations cause the hygroscopic coating of the fiber to swell which in turn applies strain to the optical fiber. This is the humidity-induced strain effect that we described in the introduction. As long as we are able to detect humidity-induced strain, we can assume that tensile strain will be measurable as well. We note that for simultaneous temperature and strain discrimination a standard acrylate-coated SOF could be also used.

In recent years, non-linear machine learning algorithms have continuously revealed their untapped potential for advanced data analysis in DFOS [24,28,30,31]. A future work will investigate the use of decision trees and artificial neural networks in distributed humidity sensing.

Optics EXPRESS

5. Conclusions

In this paper, we demonstrated, to our knowledge for the first time a BOFDA humidity distributed sensor that, importantly, addresses the sensor's cross-sensitivity to temperature. Our approach is based on linear regression which is a simple algorithm in machine learning and statistics. The learning algorithm is trained by the BFS and linewidths of the multipeak BGS to predict relative humidity and additionally to decouple the effects of relative humidity and temperature. The model's performance was estimated in terms of MAE of relative humidity and temperature and found to be below 9%RH and 1 °C, respectively. Feature selection strategies revealed that only a few of the Brillouin peaks contribute significantly which manifests that the measurement time could be shortened because only a limited Brillouin sweep range is required to record the most informative part of the BGS. Our approach can pave the path towards distributed fiber-optic humidity monitoring in applications of civil and geotechnical engineering and, in particular, in structural health monitoring of infrastructures like, e.g., long subsea cables and for pipeline corrosion detection.

Funding. PhD program of Bundesanstalt für Materialforschung und -Prüfung (BAM).

Acknowledgments. The authors would like thank Marcus Schukar for the technical support. C.K. would also like to thank Christoph Völker and Sabine Kruschwitz for the fruitful discussions.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. A. H. Hartog, *An introduction to distributed optical fibre sensors* (CRC Press, 2017).
2. J. Ascorbe, J. M. Corres, F. J. Arregui, and I. R. Matias, "Recent Developments in Fiber Optics Humidity Sensors," Sensors **17**(4), 893 (2017).
3. P. Stajanca, K. Hicke, and K. Krebber, "Distributed Fiberoptic Sensor for Simultaneous Humidity and Temperature Monitoring Based on Polyimide-Coated Optical Fibers," Sensors **19**(23), 5279 (2019).
4. P. J. Thomas and J. O. Hellevang, "A high response polyimide fiber optic sensor for distributed humidity measurements," Sensors and Actuators B: Chemical **270**, 417–423 (2018).
5. T. Neves, R. Magalhães, L. Scherino, S. Martin-Lopez, H. F. Martins, P. Petagna, and L. Thévenaz, "Humidity Effect on Acrylate- and Polyimide-Coated Fibres for Distributed Sensing Applications," in *Optical Fiber Sensors Conference 2020 Special Edition* (2020), paper T3.73.
6. F. P. N. Tiago, Z. Li, Y. Fan, H. T. Kenny, P. Paolo, and T. Luc, "A kilometre-range distributed relative humidity sensor," in *Proc.SPIE* 11199, 1119922 (2019).
7. A. Schreier, A. Wosniok, S. Liehr, and K. Krebber, "Humidity-induced Brillouin frequency shift in perfluorinated polymer optical fibers," Opt. Express **26**(17), 22307–22314 (2018).
8. X. Lu, K. Hicke, M. Breithaupt, and C. Strangfeld, "Distributed Humidity Sensing in Concrete Based on Polymer Optical Fiber," Polymers **13**(21), 3755 (2021).
9. C. Galindez, F. J. Madruga, and J. M. Lopez-Higuera, "Influence of Humidity on the Measurement of Brillouin Frequency Shift," IEEE Photon. Technol. Lett. **20**(23), 1959–1961 (2008).
10. F. Farahi, D. J. Webb, J. D. C. Jones, and D. A. Jackson, "Simultaneous Measurement of Temperature and Strain - Cross-Sensitivity Considerations," J. Lightwave Technol. **8**(2), 138–142 (1990).
11. C. He, S. Korposh, R. Correia, L. Liu, B. R. Hayes-Gill, and S. P. Morgan, "Optical fibre sensor for simultaneous temperature and relative humidity measurement: Towards absolute humidity evaluation," Sensors and Actuators B: Chemical **344**, 130154 (2021).
12. Y. Wang, Q. Huang, W. Zhu, and M. Yang, "Simultaneous Measurement of Temperature and Relative Humidity Based on FBG and FP Interferometer," IEEE Photon. Technol. Lett. **30**(9), 833–836 (2018).
13. Y. P. Xu, X. Zhao, Y. F. Li, Z. G. Qin, Y. X. Pang, and Z. J. Liu, "Simultaneous measurement of relative humidity and temperature based on forward Brillouin scattering in polyimide-overlaid fiber," Sensors and Actuators B: Chemical **348**, 130702 (2021).
14. J. Mathew, Y. Semenova, and G. Farrell, "Fiber Optic Hybrid Device for Simultaneous Measurement of Humidity and Temperature," IEEE Sens. J. **13**(5), 1632–1636 (2013).
15. D. Garcus, T. Gogolla, K. Krebber, and F. Schliep, "Brillouin optical-fiber frequency-domain analysis for distributed temperature and strain measurements," J. Lightwave Technol. **15**(4), 654–662 (1997).
16. R. Bernini, A. Minardo, and L. Zeni, "Distributed Sensing at Centimeter-Scale Spatial Resolution by BOFDA: Measurements and Signal Processing," IEEE Photonics J. **4**(1), 48–56 (2012).

Optics EXPRESS

17. T. Kapa, A. Schreier, and K. Krebber, "63 km BOFDA for Temperature and Strain Monitoring," Sensors **18**(5), 1600 (2018).
18. K. P. Murphy, *Machine learning : a probabilistic perspective* (MIT Press, 2012), Chap. 7.
19. Y. L. Zhang and Y. H. Yang, "Cross-validation for selecting a model selection procedure," J Econometrics **187**(1), 95–112 (2015).
20. Y. Koyamada, S. Sato, S. Nakamura, H. Sotobayashi, and W. Chujo, "Simulating and designing brillouin gain spectrum in single-mode fibers," J. Lightwave Technol. **22**(2), 631–639 (2004).
21. M. Newville, T. Stensitzki, D. B. Allen, M. Rawlik, A. Ingargiola, and A. Nelson, "Lmfit: Non-Linear Least-Square Minimization and Curve-Fitting for Python," Astrophysics Source Code Library (2016), https://ui.adsabs.harvard.edu/abs/2016ascl.soft06014N/abstract .
22. T. Schneider, *Nonlinear optics in telecommunications* (Springer 2004), Chap. 11.
23. R. W. Boyd, *Nonlinear Optics* (Elsevier, 2008), Chap. 9.
24. R. Ruiz-Lombera, A. Fuentes, L. Rodriguez-Cobo, J. M. Lopez-Higuera, and J. Mirapeix, "Simultaneous Temperature and Strain Discrimination in a Conventional BOTDA via Artificial Neural Networks," J. Lightwave Technol. **36**(11), 2114–2121 (2018).
25. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É Duchesnay, "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res. **12**, 2825–2830 (2011).
26. P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, "A high-bias, low-variance introduction to Machine Learning for physicists," Phys. Rep. **810**, 1–124 (2019).
27. M. Nikles, L. Thevenaz, and P. A. Robert, "Brillouin gain spectrum characterization in single-mode optical fibers," J. Lightwave Technol. **15**(10), 1842–1851 (1997).
28. C. Karapanagiotis, A. Wosniok, K. Hicke, and K. Krebber, "Time-Efficient Convolutional Neural Network-Assisted Brillouin Optical Frequency Domain Analysis," Sensors **21**(8), 2724 (2021).
29. A. Zadok, E. Zilka, A. Eyal, L. Thévenaz, and M. Tur, "Vector analysis of stimulated Brillouin scattering amplification in standard single-mode fibers," Opt. Express **16**(26), 21692–21707 (2008).
30. S. Liehr, "Artificial neural networks for distributed optical fiber sensing (Invited)," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, 2021), pp. 1–4.
31. A. Venketeswaran, N. Lalam, J. Wuenschell, P. R. Ohodnicki Jr, M. Badar, K. P. Chen, P. Lu, Y. Duan, B. Chorpening, and M. Buric, "Recent Advances in Machine Learning for Fiber Optic Sensor Applications," Advanced Intelligent Systems **4**(1), 2100067 (2022).