Research Article Vol. 57, No. 20 / 10 July 2018 / Applied Optics 5859

applied optics

NARX neural network model for strong resolution improvement in a distributed temperature sensor

Luís Cicero Bezerra da Silva,[1,*] Jorge Leonid Aching Samatelo,[1] Marcelo Eduardo Vieira Segatto,[1] João Paulo Bazzo,[2] Jean Carlos Cardozo da Silva,[2] Cicero Martelli,[2] and Maria José Pontes[1]

[1]Telecommunications Laboratory (LabTel), Electrical Engineering Department, Federal University of Espírito Santo, Vitória-ES 29075-910, Brazil
[2]Graduate Program in Electrical and Computer Engineering, Federal University of Technology—Parana, Curitiba-PR 80230-901, Brazil
*Corresponding author: luis.b.silva@aluno.ufes.br

Received 1 March 2018; revised 5 June 2018; accepted 5 June 2018; posted 6 June 2018 (Doc. ID 325182); published 10 July 2018

This paper proposes an approach to process the response of a distributed temperature sensor using a nonlinear autoregressive with external input neural network. The developed model is composed of three steps: extraction of characteristics, regression, and reconstruction of the signal. Such an approach is robust because it does not require knowledge of the characteristics of the signal; it has a reduction of data to be processed, resulting in a low processing time, besides the simultaneous improvement of spatial resolution and temperature. We obtain total correction of the temperature resolution and spatial resolution of 5 cm of the sensor.

OCIS codes: (140.3490) Lasers, distributed-feedback; (060.2420) Fibers, polarization-maintaining; (060.3735) Fiber Bragg gratings.

https://doi.org/10.1364/AO.57.005859

1. INTRODUCTION

Distributed temperature sensors (DTS) are important in monitoring large structures such as hydroelectric generators [1], dams [2,3], oil and gas platforms [4], subway tunnels [5], mines, and warehouses [6]. Temperature interrogation provides fundamental information about the current operating conditions of these structures, ensuring safety for employees in such locations [7].

Performance of DTS systems is determined by temperature resolution, characterized by how close the measured temperature provided by the sensor is the actual temperature. The performance is also given by the spatial resolution, defined as the smallest heated region that can be accurately detected by the sensor. Both spatial resolution (less than 1 m [8]) and temperature resolution (below to 3°C [9]) are strongly influenced by several factors, such as signal attenuation, dispersion, local losses, and laser signal pulse width [6].

Many authors have proposed different solutions in order to improve the performance of DTS systems [10–15]. However, the presented solutions generally improve either the spatial resolution or the temperature resolution, being scarce solutions to improve both simultaneously. For example, one can mention the use of signal processing techniques that focus exclusively on improvements in signal-to-noise ratio and uncertainties in temperature resolution [16–18]. Besides these reported works, [19,20] used a neural network to process the Brillouin time-domain trace to extract the temperature information along the fiber after the data acquisition process. However, the processing was applied directly on the signal containing the temperature information, which is nonlinear as Brillouin scattering and is influenced by several factors intrinsic to the acquisition process that distort the signal such as attenuation and dispersion.

In the results presented in [12], the signal processing technique was applied specifically to improve the spatial resolution of DTS equipment, with the objective of identifying temperature variations in regions smaller than 1 m. The algorithm used is based on the principle that signals with excessive and possibly erroneous details have a high total variation. Consequently, the reduction of the total signal variation tends to be a close match of the original signal, as it removes unwanted signal information, while preserving important details such as edges. Thus, the regularization by total variation allows preserving edges while smoothing noise in flat regions, even with low signal-to-noise ratio. The main advantage of the proposed method is the ability to correctly reconstruct hot regions on the optical fiber with dimensions up to 15 cm [12].

As an improvement, the present work proposes the use of a nonlinear autoregressive with an external input (NARX) neural network to provide accurate temperature measurements with a spatial resolution of 5 cm. The proposed model also presents a new approach to solve the problem of the large amount of data generated by a DTS, allowing a faster acquisition, or even in real-time (temperature) acquisition, depending on the desired

application. The validation of the model was accomplished considering a temperature profile with different sizes of the heated regions in the optical fiber.

This paper is organized as follows: Section 2 shows the fundamental aspects of distributed temperature sensing and concepts of a NARX neural network. Section 3 presents the details of the developed model. Section 4 presents the results obtained by numerical simulations, where, after applying the proposed model, we reconstruct the sensor response signal for hot regions with only 5 cm of the resolution measured along the entire length of the fiber. Finally, the conclusions are presented in Section 5.

2. THEORETICAL ASPECTS

A. Distributed Temperature Sensor

DTS systems are devices for temperature measurement in large structures that use optical fibers as a sensor element, which means that no transducers are required along the entire length of the fiber to carry out measurements [5,21].

The principle of interrogation consists in coupling light with high power density into the core of the fiber in order to generate nonlinearities. One of the nonlinearities generated in this process is known as spontaneous Raman scattering (SRS), which has strong dependence on the temperature of the medium in which it was generated [22]. Therefore, by detecting the SRS and performing the processing of this signal, it is possible to raise the temperature profile along the entire length of the fiber, as shown in Eq. (1) [10]:

$$T(z) = \left[\frac{1}{T_{\text{ref}}} - \frac{k}{hcv'} ln\left(\frac{R(T)}{R(T_{\text{ref}})}\right)\right]^{-1}, \tag{1}$$

where $R(T)$ and $R(T_{\text{ref}})$ are the backscattering ratios (amplitudes of signal components) measured at the arbitrary section of the sensing fiber and the reference fiber spool maintained at a known temperature T_{ref}, respectively. v' is the wavenumber separation from the input wavelength. h is the Planck's constant. The constant c is the speed of light in the free space; meanwhile, k is the Boltzmann's constant, and z is the position in the fiber. The spatial localization of the backscattered light is determined by the previously known propagation speed inside the fiber [10].

It is important to emphasize that DTS systems can carry out measurements at several points along a fiber optic cable, respecting its spatial resolution. In addition, such sensors make use of the inherent advantages of optical fibers such as low loss, electromagnetic immunity, and signal multiplexing, among others [5,6,23].

B. Nonlinear Autoregressive Method with External Input

Artificial neural networks (ANN) are a set of learning models inspired by biological neural networks, which are mainly used to estimate a function dependent on a large number of inputs that are usually unknown. They are composed by a number of interconnected neurons. Each neuron receives an input signal from other neurons or external stimuli, processes that signal and produce a transformed output signal to other neurons or external outputs.

The characteristics of an ANN algorithm enable ANN to learn from examples and then to generalize examples that have never been seen before. ANN can easily identify and learn interconnected patterns between input data sets and corresponding target values. After training the data, ANN can be applied to a set of new independent input data to predict an outcome [24].

ANN learning basically consists of modifying the weights of the connections between the neurons, where the initial weights (synapses) are modified iteratively, by an algorithm. There is supervised learning, presented by a training set, consisting of inputs and corresponding desired outputs. And the unsupervised learning, which occurs when the network updates its weights without the use of desired input–output pairs and without indications about the adequacy of outputs produced [25]. In this work, a supervised model is used.

A recurrent dynamic neural network is one that contains feedback synaptic connections and delays that allow the flow of information between neurons of distinct layers. Recurrence is a type of short-lived memory mechanism that allows the network to recall information from a recent past. While delays provide exact values of past information at the current instant, feedback loops perform some kind of processing (filtering) on past information [26].

An important class of recurrent dynamic neural network are NARX, where the actual system outputs serve as input to the model. The dynamic behavior of the NARX model is described by [27]

$$y(n) = f[y(n-1), \ldots, y(n-d_y); u(n), u(n-1), \ldots, u(n-d_u+1)], \tag{2}$$

where $u(n)$ and $y(n)$ represent, respectively, the input and output of the model at time n, while $d_u < 0$ and $d_u \leq d_y$ are the orders of input memory and output memory.

There are many applications for the NARX network. It can be used as a predictor, to predict the next value of the input signal [28], in standard classification problems [29–31], in online pattern recognition [32], and intelligent controller design [33]. It can also be used for nonlinear filtering, in which the target output is a noise-free version of the input signal [34]. In addition, NARX can be used in the modeling of nonlinear

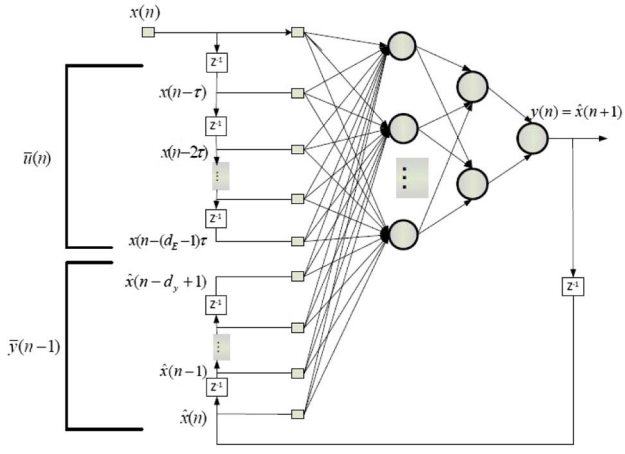

Fig. 1. NARX recurrent neural network architecture.

dynamic systems [35–37], being the output fed back to the input of the feedforward neural network as part of the standard NARX architecture because the true one is available during network training [24,27]. Or there is the possibility of creating a series-parallel architecture in which the true output is used instead of feeding back the estimated output [34].

Figure 1 shows the architecture of NARX.

In the context presented, this work sought to explore the ability of NARX to deal with nonlinear models and make the sensor response as accurate as possible.

3. DEVELOPED MODEL

In DTS systems, the signal that carries the temperature information is due to a nonlinearity, making it difficult to simultaneously analyze all the parameters that interfere with the temperature measurement.

In this context, the motivation of this work was to develop a model to improve the response of DTS systems without changing the internal structure of the same, thus exploring the NARX neural networks' ability to deal with nonlinear models.

The proposed model for improving the sensor response consists of three main parts: the pre-processing of the data, the regressor, and the reconstruction of the signal.

A. Pre-processing of the Data

The data used for training and validation of NARX is the experimental data obtained from a commercial DTS model AP Sensing N4385B. The DTS used on experiments has a spatial resolution of 1 m, acquisition time of 30 s, sample interval down to 15 cm, and temperature resolution of 0.04°C for fibers up to 2 km. In addition, the DTS unit employs a semiconductor laser with a wavelength of 1064 nm with maximum average output power of 17 mW, 0.2 numerical aperture, and a beam waist diameter of 50 μm.

The experiment to generate the temperature profiles consisted of measuring heating regions in fibers of different sizes provided by a thermal bath LAUDA ECO RE415G model, with stabilized temperature at 50°C, using the DTS to estimate the temperature along the fiber, as shown in Fig. 2.

Experimental data for training the algorithm composed a 202 × 89 matrix, where the first column represents the length of the fiber in meters, the 2–44 columns are the temperatures for different widths of the hot region in the fiber estimated by the DTS equipment, and the remaining columns are, respectively, reference temperatures (actual temperature). Therefore,

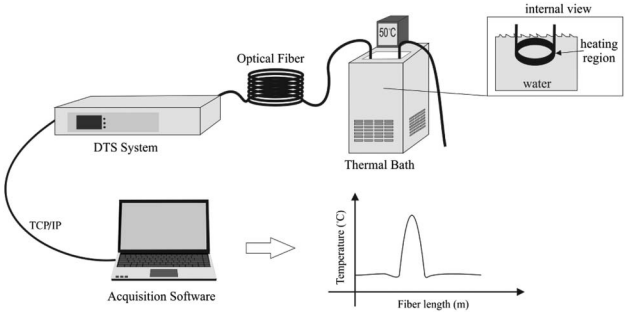

Fig. 2. Experimental setup used to obtain temperature profiles.

44 curves can be lifted that compare the actual temperature and that estimated by the DTS equipment.

Considering the algorithm proposed here intends to improve significantly the temperature and spatial resolution of measured data provided by a DTS, the first approach was training the NARX algorithm as a "calibration" process of the model.

For this training, just a few characteristics have been selected from each curve, such as the spatial resolution of the sensor W_{dts}, the actual region heated in the fiber W_{ref}, the maximum temperature measured by the DTS H_{dts}, the actual maximum temperature H_{ref}, and the minimum temperatures given by the DTS and the real temperature, B_{dts} and B_{ref}, respectively. W_{dts} and W_{ref} are the FWHM for each curve. Instead of using the "entire" curve in the analysis, selecting main characteristics of each curve drastically reduces the amount of data to be used as NARX input.

Figure 3 illustrates the selected feature selection.

With the selection of features, the data were drastically reduced to only a 44 × 6 array, where the first three columns are the NARX input, and the last three columns are the output of NARX. The new data are shown in Table 1.

B. Regression Using NARX

Once the selection of characteristics was carried out, the next step was the regression that consisted in determining the weights of the NARX model, so that the error between the output variables (actual temperature) and the regressor input (sensor response) is minimized. However, the solution should correct not only the difference between the temperature estimated by the sensor and the actual temperature but also the spatial resolution.

In this context, we use a NARX as a regressor for the problem in question for three main reasons: first, a NARX handles well with nonlinear models, presenting an output with very little noise when compared with that of nonrecurrent neural networks [38,39]. Second, it is not necessary to know the factors that interfere with the sensor performance. And, finally, the prediction step is $O(n)^2$, with n equivalent to the number

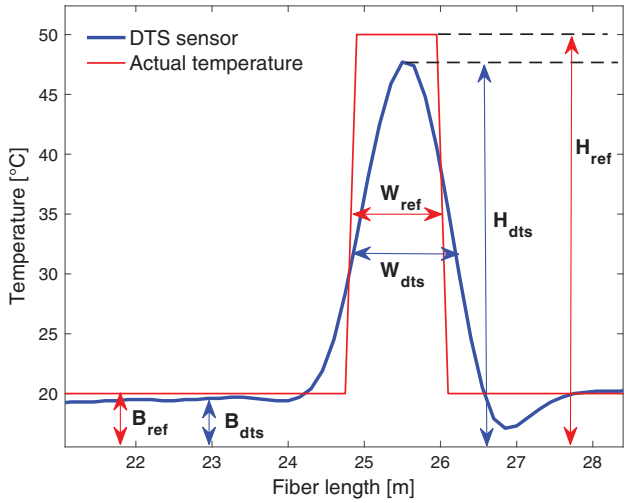

Fig. 3. Example of the selection of characteristics of one of the curves with a region heated in the fiber of 100 cm.

Table 1. NARX Input (W_{dts}, H_{dts}, B_{dts}) and Output (W_{ref}, H_{ref}, B_{ref})

W_{dts}	H_{dts}	B_{dts}	W_{ref}	H_{ref}	B_{ref}
$W_{dts,1}$	$H_{dts,1}$	$B_{dts,1}$	$W_{ref,1}$	$H_{ref,1}$	$B_{ref,1}$
$W_{dts,2}$	$H_{dts,2}$	$B_{dts,2}$	$W_{ref,2}$	$H_{ref,2}$	$B_{ref,2}$
⋮	⋮	⋮	⋮	⋮	⋮
$W_{dts,44}$	$H_{dts,44}$	$B_{dts,44}$	$W_{ref,44}$	$H_{ref,44}$	$B_{ref,44}$

of elements of the input sequence in the model (input+output −delay $\Rightarrow n = 3 + 3 - 1 = 5$). Without doing the characteristic extraction step, the complexity of the model is $O(N)^2$. Now, when doing the feature extraction step, the complexity becomes $O(n)^2$, with $n = 5 \ll N$. It is concluded that the fact of including a characteristic extraction step generates a drastic reduction in computational cost because this depends asymptotically on the number of inputs in the quadratic form model.

C. Reconstruction of the Signal

Finally, the reconstruction of the signal was performed from the NARX output. That is, with the new values of W_c, H_c, and B_c, we reconstruct the hot spot on the fiber, which is our region of interest in the fiber. Equation (3) represents how reconstruction of the signal was performed:

$$g = R\left(\frac{p_j - W_c}{2}, \frac{p_j + W_c}{2}, d\right) \times (H_c - B_c) + B_c, \quad \textbf{(3)}$$

where R represents a rectangular pulse function over the interval corresponding to the spatial resolution estimated by the NARX model, and p is the location in the fiber in the central reference position j. The value of p_j is found from the value of H_{dts}, that is, when

$$p_j = d(\max(H_{dts})), \quad \textbf{(4)}$$

where d is the vector corresponding to the position in the fiber, and W_c and H_c are the spatial resolution and temperature resolution corrected by the NARX model, respectively.

The reconstructed signal is presented in Section 4.

4. RESULTS AND DISCUSSION

Concerning the improvement of the DTS spatial resolution and temperature as well, we have considered the experimental data obtained from a commercial DTS as input to the NARX, and the NARX output (target) was the actual temperature that the DTS equipment should estimate. The samples were divided into training (70% of samples), validation (15% of samples), and test (15% of samples) data. The training data samples are provided to the neural network model during training, and the system is adjusted according to its error. The validation data is used to measure network generalization and to halt training of the neural network when the generalization stops improving, where MATLAB software was used during simulations with NARX.

In order to correctly detect hot regions in the fiber with dimensions of the order of 5 cm, it is sufficient to know the region in the fiber that such event occurred and the maximum temperature point in that region.

We use the output of the NARX model, which corresponds to these parameters (see feature selection in Fig. 3), and reconstruct the original DTS sensor signal for all temperature profiles.

Figure 4 shows that the proposed method processes the sensor response with the smallest possible error at a spatial resolution of 5 cm. We conclude that the NARX neural network model proposed here is more efficient than using total variation deconvolution in detecting small hot spots on the fiber. A significant improvement when compared with the result presented in a previous work (see Fig. 9) [12].

The result shown in Table 2 presents the robustness of the proposed model, correcting simultaneously not only the temperature resolution of the sensor but also the spatial resolution. However, it should be noted that the null error obtained in the temperature resolution is mainly due to the available measured data and tests performed. Different experimental schemes or signal processing methods would result in diverse experimental

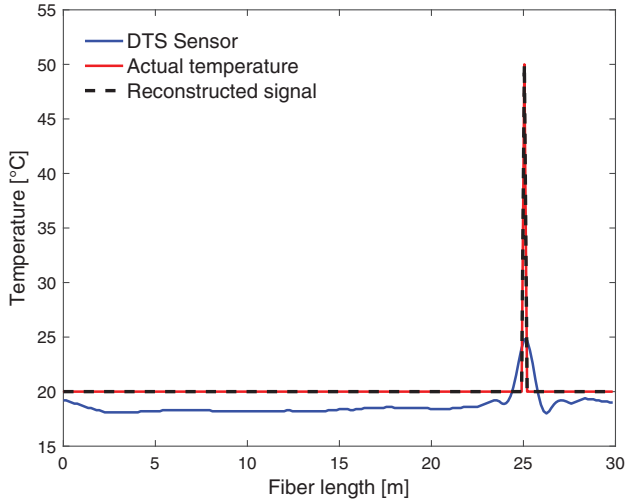

Fig. 4. Fully corrected DTS sensor response using NARX as a regressor for a fiber hot spot of only 5 cm.

Table 2. Error in Sensor Response Before and After Application of NARX for 10 Data Curves

Size Hot Spot	$Error_1$[a]	$Error_2$[b]	$Error_3$[c]
5 cm	0.983	25.20°C	4.553×10^{-7}
10 cm	0.943	24.00°C	6.806×10^{-7}
15 cm	0.857	21.60°C	6.124×10^{-7}
20 cm	0.802	20.20°C	1.623×10^{-6}
25 cm	0.754	18.20°C	1.026×10^{-6}
30 cm	0.705	16.50°C	2.871×10^{-7}
35 cm	0.654	14.90°C	5.401×10^{-7}
40 cm	0.622	14.10°C	1.193×10^{-4}
45 cm	0.610	12.40°C	1.122×10^{-4}
50 cm	0.616	11.60°C	1.512×10^{-6}

[a]$Error_1$-Error in spatial resolution without NARX;
[b]$Error_2$-Error in temperature resolution without NARX;
[c]$Error_3$-Error in spatial resolution with NARX.
Note: The error null obtained in the temperature resolution is mainly due to the available measured data and tests performed.

databases with different temperature resolutions, as discussed in [20,40].

In addition, the proposal presents a strong practical application, given the nature of the data processing performed by the DTS equipment, because these sensors usually have kilometers of extension, which correspond to a huge amount of data to be processed (nearly 1 TB/Day) [21].

Therefore, the reduction in the amount of data performed through the selection of characteristics constitutes a new approach to significantly improve the performance of DTS systems compared with traditional techniques that use the complete sensor response to provide the processed signal [12].

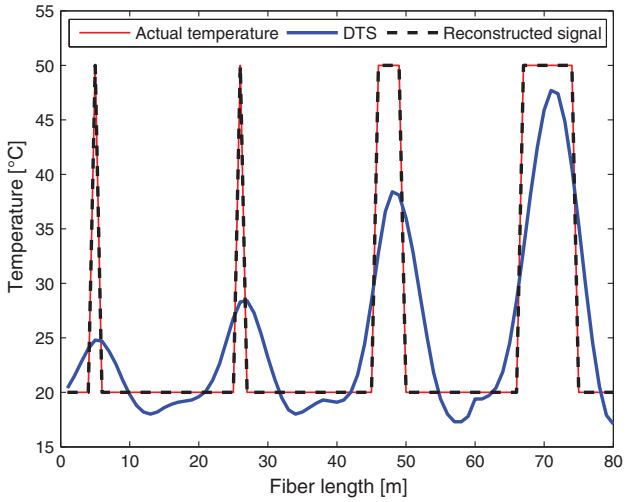

Fig. 5. Complete simultaneous correction of the DTS sensor response with NARX model for different hot spots in the fiber. Peaks 5, 15, 50, and 100 cm wide (left to right).

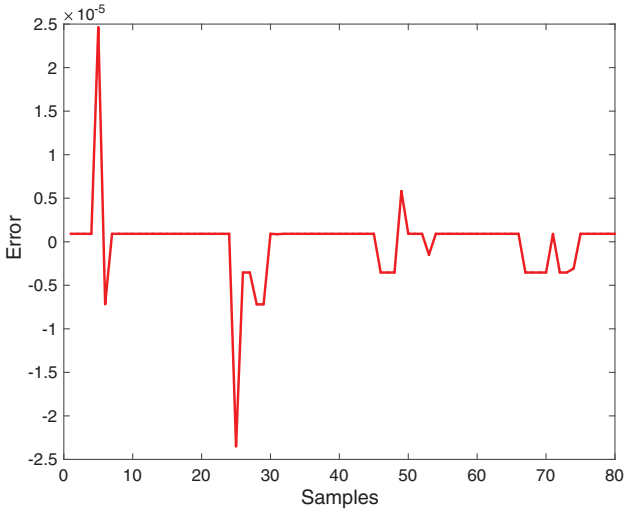

Fig. 6. Error obtained by the simultaneous correction of several peaks of temperature in the fiber, corresponding to the difference between the value estimated by the NARX model and the actual temperature.

Another advantage here is the fact that there are no losses due to compression of the data.

Table 2 shows the temperature profiles processed individually. However, in order to analyze the generality of the proposed model, we consider the situation where there are many temperature peaks with different widths. Figures 5 and 6 show the result obtained for this case and the obtained error, respectively.

In addition, other advantages of the method proposed in this paper are the possibility of application in any commercial DTS model, without the need for hardware changes that require modification in the optoelectronic circuit, which usually presents a high level of complexity besides high cost.

In summary, with a spatial resolution obtained of 5 cm, it favors the expansion of applications of the use of DTS systems to other areas focused on the thermal mapping of medium and small structures, such as pumps for water supply systems and motors in industrial plants [12].

5. CONCLUSIONS

In this work, we explored the ability of a NARX neural network to deal with nonlinear models to improve the response of a distributed temperature sensor. The proposed model was implemented in three basic steps, starting with the selection of sensor signal characteristics, next applying the regression using NARX, and, finally, performing the signal reconstruction. The method application led to excellent results by complete correction of the sensor's temperature, while the spatial resolution reached 5 cm.

A total of 44 different experimental temperature profiles available were analyzed, with different temperature peaks and heated regions in the fiber, thus enabling model validation. With the selection of features applied to the signal, the data size was dramatically reduced, consequently improving the overall processing time. It means to speed up the data acquisition process or even perform this action in real time. Thus, the extraction of characteristics from a DTS measured signal presented in this work constitutes a new proposal of how the problem of storing large amounts of data generated by DTS sensors can be solved.

Therefore, the developed model is robust with the potential to be implemented and process the response of DTS systems, without changes in the internal structure of the equipment.

Funding. Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); PETROBRAS.

REFERENCES

1. J. P. Bazzo, F. Mezzadri, E. V. da Silva, D. R. Pipa, C. Martelli, and J. C. C. da Silva, "Thermal imaging of hydroelectric generator stator using a DTS system," IEEE Sens. J. **15**, 6689–6696 (2015).
2. A. A. Khan, V. Vrabie, J. I. Mars, A. Girard, and G. D'Urso, "A source separation technique for processing of thermometric data from fiber-optic DTS measurements for water leakage identification in dikes," IEEE Sens. J. **8**, 1118–1129 (2008).
3. A. A. Khan, V. Vrabie, J. I. Mars, A. Girard, and G. D'Urso, "Automatic monitoring system for singularity detection in dikes by DTS data measurement," IEEE Trans. Instrum. Meas. **59**, 2167–2175 (2010).

4. C. E. Campanella, G. Ai, and A. Ukil, "Distributed fiber optics techniques for gas network monitoring," in *IEEE International Conference on Industrial Technology (ICIT)* (2016), pp. 646–651.
5. X. Bao and L. Chen, "Recent progress in distributed fiber optic sensors," Sensors **12**, 8601–8639 (2012).
6. A. Ukil, H. Braendle, and P. Krippner, "Distributed temperature sensing: review of technology and applications," IEEE Sens. J. **12**, 885–892 (2012).
7. A. Bahadori, *Hazardous Area Classification in Petroleum and Chemical Plants: A Guide to Mitigating Risk* (CRC Press, 2013).
8. M. Höbel, J. Ricka, M. Wüthrich, and T. Binkert, "High-resolution distributed temperature sensing with the multiphoton-timing technique," Appl. Opt. **34**, 2955–2967 (1995).
9. J. Park, G. Bolognini, D. Lee, P. Kim, P. Cho, F. D. Pasquale, and N. Park, "Raman-based distributed temperature sensor with simplex coding and link optimization," IEEE Photon. Technol. Lett. **18**, 1879–1881 (2006).
10. H. S. Pradhan and P. K. Sahu, "Characterisation of Raman distributed temperature sensor using deconvolution algorithms," IET Optoelectron. **9**, 101–107 (2015).
11. I. Laarossi, M. A. Quintela, R. Ruiz-Lombera, J. Mirapeix, D. Lima, D. Solana, and J. M. Lopez-Higuera, "High-temperature distributed sensor based on Raman and multimode standard telecom fiber," in *Novel Optical Materials and Applications* (Optical Society of America, 2016).
12. J. P. Bazzo, D. R. Pipa, C. Martelli, E. V. da Silva, and J. C. C. da Silva, "Improving spatial resolution of Raman DTS using total variation deconvolution," IEEE Sens. J. **16**, 4425–4430 (2016).
13. N. Van De Giesen, S. C. Steele-Dunne, J. Jansen, O. Hoes, M. B. Hausner, S. Tyler, and J. Selker, "Double-ended calibration of fiber-optic Raman spectra distributed temperature sensing data," Sensors **12**, 5471–5485 (2012).
14. Y. S. Muanenda, M. Taki, T. Nannipieri, A. Signorini, C. J. Oton, F. Zaidi, I. Toccafondo, and F. Di Pasquale, "Advanced coding techniques for long-range Raman/BOTDA distributed strain and temperature measurements," J. Lightwave Technol. **34**, 342–350 (2016).
15. M. A. Soto, A. Signorini, T. Nannipieri, S. Faralli, G. Bolognini, and F. Di Pasquale, "Impact of loss variations on double-ended distributed temperature sensors based on Raman anti-Stokes signal only," J. Lightwave Technol. **30**, 1215–1222 (2012).
16. Z. Wang, J. Chang, S. Zhang, S. Luo, C. Jia, B. Sun, S. Jiang, Y. Liu, X. Liu, G. Lv, and Z. Liu, "Application of wavelet transform modulus maxima in Raman distributed temperature sensors," Photon. Sens. **4**, 142–146 (2014).
17. M. K. Saxena, S. Raju, R. Arya, R. Pachori, S. Ravindranath, S. Kher, and S. Oak, "Raman optical fiber distributed temperature sensor using wavelet transform based simplified signal processing of Raman backscattered signals," Opt. Laser Technol. **65**, 14–24 (2015).
18. A. Bahrampour, A. Moosavi, M. Bahrampour, and L. Safaei, "Spatial resolution enhancement in fiber Raman distributed temperature sensor by employing forward deconvolution algorithm," Opt. Fiber Technol. **17**, 128–134 (2011).
19. A. K. Azad, L. Wang, N. Guo, C. Lu, and H. Y. Tam, "Temperature sensing in BOTDA system by using artificial neural network," Electron. Lett. **51**, 1578–1580 (2015).
20. A. K. Azad, L. Wang, N. Guo, H.-Y. Tam, and C. Lu, "Signal processing using artificial neural network for BOTDA sensor system," Opt. Express **24**, 6769–6782 (2016).
21. Silixa, "Manuals, whitepapers," http://silixa.com/.
22. M. A. Farahani and T. Gogolla, "Spontaneous Raman scattering in optical fibers with modulated probe light for distributed temperature Raman remote sensing," J. Lightwave Technol. **17**, 1379–1391 (1999).
23. IEEE, "Guide for temperature Monitoring of Cable Systems," Std 1718-2012 (2012), pp. 1–35.
24. S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. (1999).
25. J. M. Zurada, *Introduction to Artificial Neural Systems* (1992), Vol. **8**.
26. S. Qin, J. Feng, J. Song, X. Wen, and C. Xu, "A one-layer recurrent neural network for constrained complex-variable convex optimization," in *IEEE Transactions on Neural Networks and Learning Systems* (2016), pp. 1–11.
27. I. Leontaritis and S. A. Billings, "Input-output parametric models for non-linear systems Part I: deterministic non-linear systems," Int. J. Control **41**, 303–328 (1985).
28. H. Li, Y. Zhu, J. Hu, and Z. Li, "A localized NARX neural network model for short-term load forecasting based upon self-organizing mapping," in *IEEE 3rd International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017— ECCE Asia)* (2017), pp. 749–754.
29. J. Zhang, Z. Yin, and R. Wang, "Nonlinear dynamic classification of momentary mental workload using physiological features and NARX-model-based least-squares support vector machines," IEEE Trans. Human-Mach. Syst. **47**, 536–549 (2017).
30. R. T. Fleifel, S. S. Soliman, W. Hamouda, and A. Badawi, "LTE primary user modeling using a hybrid ARIMA/NARX neural network model in CR," in *IEEE Wireless Communications and Networking Conference (WCNC)* (2017), pp. 1–6.
31. X. H. Zhou, G. B. Bian, X. L. Xie, Z. G. Hou, and J. L. Hao, "Prediction of natural guidewire rotation using an sEMG-based NARX neural network," in *International Joint Conference on Neural Networks (IJCNN)* (2017), pp. 419–424.
32. D. A. Bonilla C, N. Nedjah, and L. de Macedo Mourelle, "Online pattern recognition for Portuguese phonemes using multi-layer perceptron combined with recurrent non-linear autoregressive neural networks with exogenous inputs," in *IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (2016), pp. 1–6.
33. D. D. Braga, R. Tanscheit, and M. M. B. R. Vellasco, "Neural network nonlinear plant identification as a tool in intelligent controller design," in *International Joint Conference on Neural Networks (IJCNN)* (2017), pp. 1472–1479.
34. MathWorks, "Design Time Series NARX Feedback Neural Networks," https://www.mathworks.com/help/nnet/ug/design-time-series-narx-feedback-neural-networks.html.
35. S. Bouaddi, I. Ahmed, and O. A. Mensour, "Modeling and prediction of reflectance loss in CSP plants using a nonlinear autoregressive model with exogenous inputs (NARX)," in *International Renewable and Sustainable Energy Conference (IRSEC)* (2016), pp. 706–709.
36. H. Ahmed, R. Ushirobira, D. Efimov, D. Tran, M. Sow, P. Ciret, and J. C. Massabuau, "Monitoring biological rhythms through the dynamic model identification of an oyster population," IEEE Trans. Syst. Man Cybern. **47**, 939–949 (2017).
37. A. Zibafar, S. Ghaffari, and G. Vossoughi, "Achieving transparency in series elastic actuator of sharif lower limb exoskeleton using LLNF-NARX model," in *4th International Conference on Robotics and Mechatronics (ICROM)* (2016), pp. 398–403.
38. T. Lin, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," IEEE Trans. Neural Netw. **7**, 1329–1338 (1996).
39. T.-N. Lin, C. L. Giles, B. G. Horne, and S.-Y. Kung, "A delay damage model selection algorithm for NARX neural networks," IEEE Trans. Signal Process. **45**, 2719–2730 (1997).
40. Y. Muanenda, "Recent advances in distributed acoustic sensing based on phase-sensitive optical time domain reflectometry," J. Sens. **2018**, 3897873 (2018).