Source Detection and Tracking for Underwater Distributed Acoustic Sensing

Konstantinos Theofilos Drylerakis*, Mohammad Belal†, Rafael Mestre*, Timothy J. Norman*, Christine Evers*
School of Electronics and Computer Science, University of Southampton, Southampton, UK*
Ocean BioGeosciences, National Oceanography Centre, Southampton, UK†
Dept. of Mathematical Sciences, University of Liverpool, Liverpool, UK†
Dept. of Physics, University of Southampton, Southampton, UK†
{ktd1g20, r.mestre, t.j.norman, c.evers}@soton.ac.uk*, mob@noc.ac.uk†

***Abstract*—Distributed Optical Fiber Sensing (DOFS) transforms conventional fiber optic cables into an extensive network of continuous sensors. It achieves this by exploiting the spectral, polarization and/or phase sensitivity of the propagating light to measurands of temperature, strain, pressure, vibrations etc. To harness the novel capabilities of optical fibers to remotely capture, process and coherently analyze ambient vibration (e.g., acoustic) fields, it is crucial to address the challenges of the diversity of noise introduced in DOFS measurements, in particular, within the under-explored submarine environment. This research introduces a comprehensive workflow for the detection of active (uncontrolled) acoustic sources, comprised of successive denoising steps that deal with the distinctive properties of such environments. Leveraging the spatio-temporal density of DOFS measurements, we develop a method based on data covariances for the automatic extraction of features in an unsupervised manner, together with additional features introduced to distinguish active source signals from noise. Consequently, this work takes the denoising of underwater DOFS data one step further through the application of a tracking algorithm on real, novel submarine DOFS data, laying the foundation for broader applications of DOFS data analysis in marine environmental sensing and monitoring.**

***Index Terms*—distributed acoustic sensing, machine learning**

I. Introduction

Dynamic (up to several kHz) differential strain sensing via Distributed Optical Fiber Sensing (DOFS), more commonly referred to as Distributed Acoustic Sensing (DAS), utilizes fiber optic cables to detect vibrations. In that, pulses of coherent light interact with imperfections in the glass, called scattering points, which move in accordance with external perturbations, e.g., acoustic vibrations. This movement alters the phase of backscattered light, which allows for differential optical phase change evaluation between spatially separated sections along the fiber optic cable. This enables real-time monitoring of environmental changes with dense spatio-temporal insights. DAS can simultaneously track characteristic event frequencies, e.g., marine-vessel motion, cetacean activity, seismic etc. Consequently, spatial, temporal, and, hence, spectral, diversity across acoustic sources can be leveraged to perform a variety of tasks, e.g., sound event detection, classification, source localization, tracking etc. [1].

This work was supported by UK Research and Innovation Centre for Doctoral Training in Machine Intelligence for Nano-electronic Devices and Systems [EP/S024298/1] and the Defence Science and Technology Laboratory.

State-of-the-art systems based on this approach, i.e., DAS, use data-driven machine learning (ML) algorithms either partially, e.g., as the final step of data clustering based on features extracted with more traditional signal processing techniques [2], or in an end-to-end manner, e.g., for self-supervised deep learning-based denoising of fiber data [3]. While classical ML models have been considered, e.g., support vector machines for train tracking [4], the majority of ML-based approaches for DAS are based on neural networks (NNs). The main advantage of NNs is the automatic feature extraction during the training stage, in comparison to the hand-crafted features required for classical ML models, making them more adaptive when moving to other sites or DAS systems [5]. A variety of NN-based approaches can be found in DAS literature, such as feature extraction via transfer learning [6], [7], [8], emulating signal processing workflows with NNs [9], classification based on supervised learning using labelled datasets [10], [11], as well as approaches for latent space exploratory data analysis or the simulation of DAS settings through unsupervised learning and generative modelling [5], [11].

However, these approaches depend on large, manually-annotated DAS datasets, or assumptions regarding the properties of noise, while little work has been done towards exploring the features extracted in the process. Additionally, the focus is primarily on detection or denoising tasks, with [4] comprising the only end-to-end framework for active source detection and tracking using DAS. Related work regarding DAS-based trajectory extraction can be found in [12] and [13]. The work in [12] relies on the Hough transform of z-score values extracted from DAS data, to detect road vehicles and extract trajectories based on the detection results. The work in [13] developed a new enhanced fiber optic acoustic sensor for DAS, which, combined with the technique of steered-response power phase transform, enables the detection and localization of drones. Preliminary evaluation of the capabilities of DAS for marine traffic detection and localization can also be found in [14] and [15]. However, extracting smooth trajectories from vessel track estimates using DAS data, remains to be examined.

To address these gaps, this work proposes a framework for decomposing DAS data, based on features extracted by leveraging only their covariances. The approach is based on the unsupervised dimensionality reduction technique of principal

component analysis (PCA) [16]. Furthermore, two features are defined that allow to identify appropriate subsets of principal components (PCs) that correspond to active sources, which enables the selective denoising of such datasets and, as a result, facilitates their use in tracking applications.

The proposed PCA-based method is applied to real, novel marine DAS data, and, with the aid of a tracking algorithm, the results demonstrate the capability of DAS systems to be utilized in real-time tracking of active sources. Furthermore, a feature-level comparison to the baseline method for DAS-based train tracking developed in [4], exhibits the effect that assumptions regarding noise can have on performance and, combined with feature visualization, demonstrates the advantages of the proposed approach.

II. Background and Data Acquisition

Dynamic-DOFS (e.g., DAS) applied from an end of a conventional optical fiber within a cable (e.g., seafloor), enables monitoring range of several (100-150) kilometers with a high spatial resolution (few meters) [1], [3]. This results in dense spatio-temporal datasets that can be regarded as $m \times n$ matrices. Each of the m rows corresponds to measurements along time for a specific location on the cable, also referred to as a channel, while each of the n columns corresponds to measurements along all such channels, collectively representing the entire cable at a specific timestamp.

Whilst noise from interrogation-instrument and Fresnel reflections due to prevalent use of flat optical connectors in legacy seafloor cable systems can be mitigated to an extent, noise contributions from nonlinear and non stationary ambient marine environment poses a formidable challenge. Especially, when attempting to disentangle dynamic signals of interest. Consequently, to date, only limited attempts have been made to tackle such datasets for the said interest [1], [14], [15].

The data used in this work were obtained using a DOFS system based upon differential change in backscattered Rayleigh phase (dΦ-DVS) [17], bespoke developed as part of the National Oceanography Centre (NOC) intelligent marine fiber sensing research program. The data were acquired from the shore end of a $5,650$ m offshore seafloor cable during field trials conducted by the NOC in November 2021, at the SmartBay cabled observatory, Galway Bay (Ireland). Besides monitoring the infrastructure, several experiments during the field campaign were carried out to evaluate the system's ability to intelligently identify and track anthropogenic activity, e.g., surface marine vessel movements. Consequently, these movements were orchestrated in various orientations relative to the cable, e.g., constant speed trajectory as much parallel to the cable as possible. An illustrative example is presented in Fig. 1.

We focused on a 30 Hz eigenfrequency associated with the vessel motion. Spatial sampling density was set to 1.275 m, with 200 Hz of sampling frequency, which resulted in a DOFS array with 4428 channels. Only a subset of the data is used to test the proposed framework, i.e., over channels 1801 to 4300 ($\sim 3,200$ m).

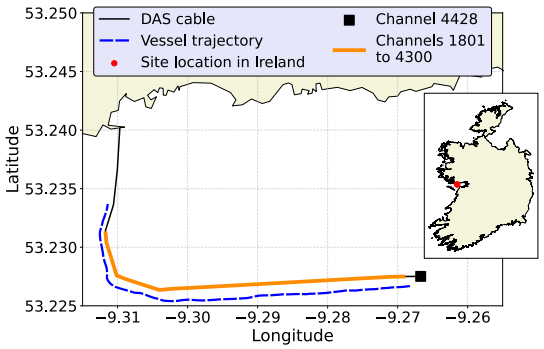

Fig. 1: Trajectory of a vessel moving parallel to the DAS cable.

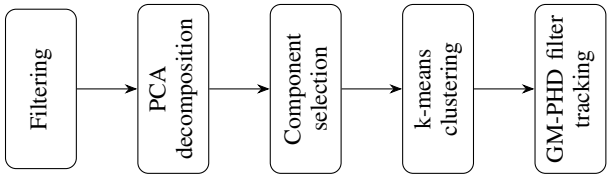

Fig. 2: Block diagram of the proposed framework.

III. Proposed Methodology

The stages of the proposed approach are summarized in Fig. 2. The rest of this section describes each stage in further detail.

A. Filtering

Using low, high, and band-pass Butterworth filters, several filtered versions of the DOFS data are generated based on uniformly spaced frequency bands. This filtering stage acts as a preprocessing step, following the standard practice in DOFS literature, and allows to focus on specific frequencies of interest. Any *a priori* knowledge regarding the acoustic sources can be used to determine the spacings between the frequency bands accordingly. While optional, in practice, this step can help isolate different types of noise that would otherwise overlap in the unfiltered data, and in turn affect the performance of PCA in the next step.

B. Data Decomposition

We propose to utilize PCA to decompose a DOFS dataset into a set of PCs such that each represents a group of acoustic sources that is statistically uncorrelated to any other such group. It involves transforming a dataset into a new coordinate system in such a way that the first axis, the first PC, captures the maximum variance in the data. Subsequent axes (PCs) capture decreasing amounts of variance in the data. PCA uses the eigendecomposition of the data covariance matrix to find a set of uncorrelated variables which are linear combinations of the original ones [16].

For a DOFS $m \times n$ data matrix, $\mathbf{D}$, of n temporal recordings made by m channels, the samples used for PCA are spatio-temporal areas of $\mathbf{D}$ of configurable size $r \times c$, that are extracted using a non-overlapping sliding window approach across both the temporal and spatial directions of $\mathbf{D}$. For simplicity, throughout this work it is assumed that such a

sliding window fits the relevant DOFS data matrix perfectly, e.g., no padding or cropping is needed for $\mathbf{D}$. Flattening each $r \times c$ spatio-temporal sample into a vector of $r \cdot c$ entries, the DOFS $m \times n$ data matrix, $\mathbf{D}$, can be mapped to a reshaped $(m \cdot n)/(r \cdot c) \times r \cdot c$ centered data matrix, $\hat{\mathbf{D}}$, which is transformed via PCA into $\hat{\mathbf{D}}_{PCA}$, where $\hat{\mathbf{D}}_{PCA} = \hat{\mathbf{D}} \times \hat{\mathbf{L}}$, with $\hat{\mathbf{L}}$ denoting the $r \cdot c \times r \cdot c$ matrix of maximum rank equal to $\min((m \cdot n)/(r \cdot c), r \cdot c)$ that contains in its columns the eigenvectors of the covariance matrix of $\hat{\mathbf{D}}$. The columns of $\hat{\mathbf{D}}_{PCA}$ are the PCs, the entries of which are called the PC scores, while the entries of each corresponding eigenvector in $\hat{\mathbf{L}}$ are called the PC loadings.

C. Dimensionality Reduction via Component Selection

While PCA can be used to decompose DOFS data into uncorrelated components, no information is provided regarding which components correspond to acoustic sources of interest. Therefore, this step identifies and maintains only the PCs that capture the presence of certain events in the DOFS dataset. In order to identify PCs that correspond to structured, coherent signals with a locally dense energy distribution, we propose to leverage two information theoretic measures, namely the Pearson correlation coefficient (PCC) and the Shannon entropy.

PCC is used to characterize the spatio-temporal correlations among the recordings of neighbouring DOFS channels. Given the size of the $r \times c$ sliding window used in section III-B, the eigenvector of $r \cdot c$ loadings returned for a PC, $\mathbf{v}$, of the PCA can be reshaped into an $r \times c$ matrix $\mathbf{L}(\mathbf{v})$, so as to coherently visualize the space-time expression of events. Then the feature, $a(\mathbf{L}(\mathbf{v}))$, of the average channel-wise PCC is defined as:

$$a(\mathbf{L}(\mathbf{v})) = \frac{\sum_{i=1}^{r-1} \sum_{j=i+1}^{r} PCC(L_i, L_j)}{C(r, 2)} \tag{1}$$

where $PCC(L_i, L_j)$ is the PCC between the i-th and j-th rows of matrix $\mathbf{L}(\mathbf{v})$, and $C(r, 2)$ is the number of all possible pairs formed by the rows of $\mathbf{L}(\mathbf{v})$. The assumption is that PCs that correspond to structured signals, are expected to have high values of a due to the spatial density of DOFS systems. On the contrary, PCs that correspond to spatio-temporally uncorrelated noise signals should have values of a closer to 0.

Entropy is used to examine whether a PC, when viewed as a spatial signal, has most of its energy concentrated at a certain location. Given the size of the $r \times c$ sliding window used in III-B and a PC, $\mathbf{v}$, of the PCA of the $m \times n$ DOFS data matrix, $\mathbf{D}$, the spatial entropy of $\mathbf{v}$ can be computed using:

$$se(\mathbf{v}) = -\sum_{i=1}^{m/r} p_i \cdot \log_{m/r}(p_i) \quad , \quad p_i = \frac{\sum_{j=1}^{n/c} v_{ij}^2}{\sum_{i=1}^{m/r} \sum_{j=1}^{n/c} v_{ij}^2} \tag{2}$$

where $v_{ij}, i \in \{1, \ldots, m/r\}, j \in \{1, \ldots, n/c\}$ is the relevant PC score on $\mathbf{v}$. The assumption is that PCs with low values of se will correspond to sources of limited spatial range in terms of the DOFS cable response, e.g., a vessel, and comprise more interesting targets in tracking scenarios (Fig. 3).

By selecting appropriate thresholds for these two features, the set of PCs of the PCA step in section III-B can be reduced to comprise only PCs of interest. Ultimately, this

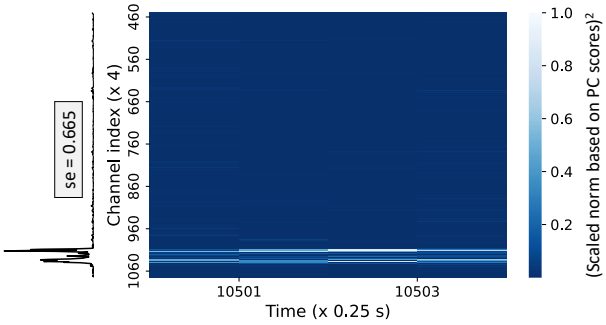

Fig. 3: The sparsity of energy across space during short periods can be measured via entropy, and help identify PCs of interest.

stage achieves the simultaneous denoising and dimensionality reduction of the input DOFS dataset. For a as defined in (1), theoretical limits can be drawn for the relevant threshold value based on a model for signal propagation. Such a limit prevents biases towards certain frequencies and is respected throughout this work, however, any deeper discussion exceeds the intended scope here. For se as defined in (2), the threshold choice depends on the acoustic source in question, and is an open problem for the proposed approach. For the purpose of this work, this threshold is selected experimentally.

D. Clustering

At this stage, the k-means [18] algorithm is used to cluster the dimensionally reduced DOFS data into 2 clusters, i.e., one for noise and one for active source signals. Given a set $S \neq \emptyset$ of PCs returned by the previous stage in section III-C, a spatio-temporal sample $\mathbf{d}$ from III-B is clustered based on its total energy $e(\mathbf{d})$ in the principal subspace defined by S as:

$$e(\mathbf{d}) = \sum_{\mathbf{v} \in S} {d_{\mathbf{v}}}^2 \tag{3}$$

where $d_{\mathbf{v}}$ is the score of $\mathbf{d}$ on PC $\mathbf{v}$. The idea behind the use of $e(\mathbf{d})$ for clustering is based on the fact that PCA, resulting in a rotation of the original coordinate system, preserves the Euclidean norm of the data upon transforming them. Therefore, $e(\mathbf{d})$ corresponds to the total energy in $\mathbf{d}$ of an activity that is represented by the relevant PC scores in the principal subspace defined by S. The noise cluster can then be selected to be the one with the lower centroid value.

E. Tracking

A Gaussian mixture probability hypothesis density (GM-PHD) filter [19] tracking algorithm completes the proposed framework, applied on the clustered data of III-D with the aid of a dynamical model of movement for the potential targets. Each single-target state consists of the projection of its position on the cable, and its velocity with respect to the cable's geometry. At each time step, the noisy measurements of the projection of positions of potential targets on the cable are obtained from the respective column of the binary matrix from III-D, using the row indices of this column's entries that belong to the cluster of active source signals.

IV. Experimental setup

The DOFS NOC data described in Section II are filtered into 5 bands, using 10th order Butterworth filters. Band 1 corresponds to a 20 Hz low-pass filter; bands 2, 3, and 4, correspond to band-pass filters between 20-40 Hz, 40-60 Hz, and 60-80 Hz respectively; while band 5 corresponds to an 80 Hz high-pass filter. The sliding window used for PCA as described in III-B is chosen to cover 4 channels and 50 timesteps. The minimum value of $a(\mathbf{L}(\mathbf{v}))$ in (1) is set to 0.5, while the maximum value for the spatial entropy, $se(\mathbf{v})$, in (2) is set to 0.88. The data are processed in an online manner, where each input covers a 1 s period. For the GM-PHD filter, the surveillance region is set to the cable segment corresponding to channels 1801-4300 (see sect. II). Each target has a survival probability of 0.95 (selected empirically) and follows a linear Gaussian dynamical model [19]:

$$F_k = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}, \quad Q_k = \sigma_v^2 \begin{bmatrix} \Delta^4/4 & \Delta^3/2 \\ \Delta^3/2 & \Delta^2 \end{bmatrix} \tag{4}$$

as well as $H_k = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $R_k = \sigma_\epsilon^2$, where $\Delta = 0.25$ s is the sampling period (equivalently, the sampling rate is now 4 Hz, reduced from 200 Hz originally, due to the $\times 50$ dimensionality reduction along the temporal direction), $\sigma_v = 10$ m/s^2 is the standard deviation of the process noise and $\sigma_\epsilon = 20$ m is the standard deviation of the measurement noise. Up to 2 targets are born uniformly over the surveillance region with direction either towards the shore or towards the cable end, and parallel to the cable's geometry. No spawning model is implemented. The probability of detection is set to 0.5 with 10^{-3} clutter returns, uniformly distributed along the surveillance region.

V. Results

A. Performance Evaluation

Fig. 4 shows the performance of the proposed framework for a band 2 (20-40 Hz) subset of the NOC data during which the vessel completes two traversals along the cable. The relevant metrics with respect to the ground truth data are evaluated in Table I. For the rest of the bands, the outputs exhibit either sparse clutter or correspond to attenuated versions of the signal in band 2 due to unavoidable frequency leakage of the Butterworth filters. It is important to note that the GM-PHD filter output can be tailored according to the desired trade-off between accuracy and recall for the relevant application. This work emphasizes on accuracy, so as to estimate the vessel track as best as possible, and therefore the GM-PHD filter results of Fig. 4 and Table I correspond to accurate detections. Nevertheless, such results are only indicative of the DAS system's capabilities, due to possible inaccuracies in the relevant ground truth stemming from the inherent limitation in underwater environments of reliably estimating parameters such as the speed of sound in water, the cable location etc.

B. Baseline Comparison

As the classifier in [4] is trained using signals measured along terrestrial train tracks, it is unlikely to generalize to data acquired in underwater environments. Moreover, the classifier cannot be re-trained using the NOC data due to the inherent uncertainty affecting labels in underwater environments. As such, a direct comparison of our approach and [4] is not possible. Nevertheless, a feature-level comparison and data-related observations can be made for the early stages of the two approaches that are both based on PCA. Contrary to the data used in [4], the signal of interest is not the main source of energy in the NOC data. Upon further analysis, it was found that signals in the low frequencies up to 20 Hz were the most dominant, leading to uninformative features extracted by the method in [4] due to assuming that the signal of interest should correspond to the first two PCs. This becomes evident via measuring the uniformity of energy concentration across the cable for the two approaches via the spatial entropy defined in (2), for the visualizations in Fig. 3 and Fig 5. Such a comparison then demonstrates the effect that certain assumptions regarding the intensity of the signal of interest can have on performance. It must be noted that, ideally, a data decomposition technique should be able to separate the different sources without any pre-filtering, something that could not be achieved by PCA in this work for all data regions. Therefore, future work will investigate alternative and possibly non-linear models to enhance performance.

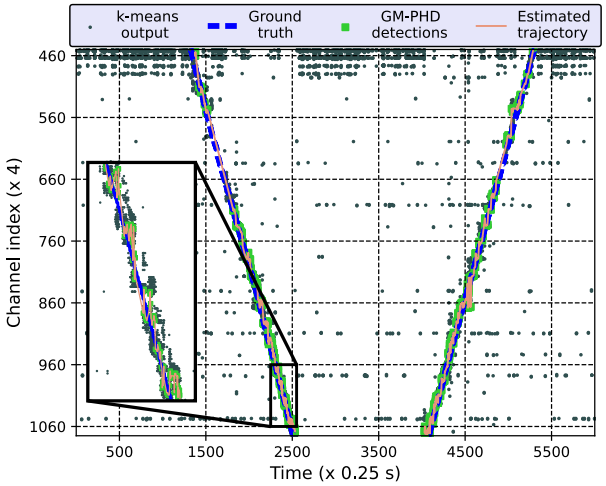

Fig. 4: Output of the proposed framework for the band 2 (20-40 Hz) NOC data, containing the tracks of the vessel trajectory.

TABLE I: Performance metrics for the results of Fig. 4.

Recall (%)	Error (m)	Error standard deviation (m)
45	39	31.7

C. Feature Visualization

Fig. 6 provides an illustrative example of PC loadings for a PC capturing a vessel. These loadings are the weights with which the 4×50 spatio-temporal DOFS samples are multiplied, based on the Frobenius inner product, to compute their value on the relevant feature (here corresponding to vessel presence). The frequency of the sinusoidal waves across the 4 channels in

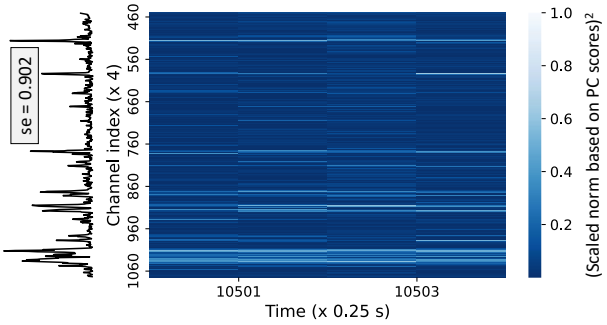

Fig. 5: Output of the PCA step in [4] for the data of Fig. 3.

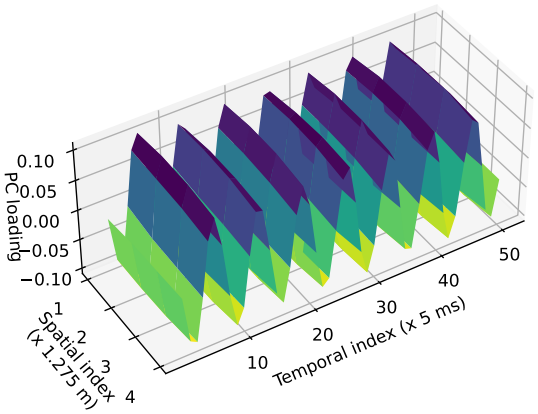

Fig. 6: Example of a set of PC loadings ($\mathbf{L}(\mathbf{v})$ in section III-C) for the band 2 NOC data of Fig. 4 visualized in the original DOFS data space, capturing a frequency around 30 Hz.

Fig. 6 is estimated to be around 30 Hz, confirming that phase oscillations in the data caused by the source (active around 30 Hz) are captured by features computed as detailed above.

VI. Conclusion

DOFS transforms conventional fiber optic cables into extensive arrays of acoustic sensors, allowing continuous and remote monitoring with dense spatio-temporal attributes. However, to fully exploit its potential, advanced data processing workflows are necessary, especially in challenging submarine environments. This work proposed a novel and comprehensive framework for analyzing marine DOFS data, employing PCA and two features in the PC space for tracking applications. Through successive filtering steps, the approach effectively denoised marine DOFS data, enabling their utilization in tracking scenarios with the GM-PHD filter tracking algorithm. The framework's blind denoising and unsupervised data processing capabilities, advance underwater DOFS data analysis, helping tackle new challenges in underwater real-time monitoring.

Acknowledgment

Data were acquired by NOC as part of the Trans National Access (Jerico S3 Project). NOC thanks the staff at SmartBay observatory, Dr. Arthur Hartog and Mr. Alexis Constantinos, for their help and support during the field campaign.

References

[1] A. H. Hartog, M. Belal, and M. A. Clare, "Advances in distributed fiber-optic sensing for monitoring marine infrastructure, measuring the deep ocean, and quantifying the risks posed by seafloor hazards," *Marine Technology Society Journal*, vol. 52, no. 5, pp. 58–73, 2018.

[2] E. R. Martin, F. Huot, Y. Ma, R. Cieplicki, S. Cole, M. Karrenbach, and B. L. Biondi, "A seismic shift in scalable acquisition demands new processing: Fiber-optic seismic signal retrieval in urban areas with unsupervised learning for coherent noise removal," *IEEE Signal Processing Magazine*, vol. 35, no. 2, pp. 31–40, 2018.

[3] M. van den Ende, I. Lior, J.-P. Ampuero, A. Sladen, A. Ferrari, and C. Richard, "A self-supervised deep learning approach for blind denoising and waveform coherence enhancement in distributed acoustic sensing data," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[4] C. Wiesmeyr, M. Litzenberger, M. Waser, A. Papp, H. Garn, G. Neunteufel, and H. Döller, "Real-time train tracking from distributed acoustic sensing data," *Applied Sciences*, vol. 10, no. 2, p. 448, 2020.

[5] L. Shiloh, A. Eyal, and R. Giryes, "Efficient processing of distributed acoustic sensing data using a deep learning approach," *Journal of Lightwave Technology*, vol. 37, no. 18, pp. 4755–4762, 2019.

[6] A. L. Stork, A. F. Baird, S. A. Horne, G. Naldrett, S. Lapins, J.-M. Kendall, J. Wookey, J. P. Verdon, A. Clarke, and A. Williams, "Application of machine learning to microseismic event detection in distributed acoustic sensing data," *Geophysics*, vol. 85, no. 5, pp. KS149–KS160, 2020.

[7] P. D. Hernández, J. A. Ramírez, and M. A. Soto, "Deep-learning-based earthquake detection for fiber-optic distributed acoustic sensing," *Journal of Lightwave Technology*, vol. 40, no. 8, pp. 2639–2650, 2022.

[8] A. F. Peña Castro, B. Schmandt, M. G. Baker, and R. E. Abbott, "Tracking local sea ice extent in the beaufort sea using distributed acoustic sensing and machine learning," *The Seismic Record*, vol. 3, no. 3, pp. 200–209, 2023.

[9] S. Kowarik, M.-T. Hussels, S. Chruscicki, S. Münzenberger, A. Lämmerhirt, P. Pohl, and M. Schubert, "Fiber optic train monitoring with distributed acoustic sensing: Conventional and neural network data analysis," *Sensors*, vol. 20, no. 2, p. 450, 2020.

[10] S. Jakkampudi, J. Shen, W. Li, A. Dev, T. Zhu, and E. R. Martin, "Footstep detection in urban seismic data with a convolutional neural network," *The Leading Edge*, vol. 39, no. 9, pp. 654–660, 2020.

[11] V. Dumont, V. R. Tribaldos, J. Ajo-Franklin, and K. Wu, "Deep learning for surface wave identification in distributed acoustic sensing data," in *2020 ieee international conference on big data (big data)*. IEEE, 2020, pp. 1293–1300.

[12] P. J. Thomas, Y. Heggelund, I. Klepsvik, J. Cook, E. Kolltveit, and T. Vaa, "The performance of distributed acoustic sensing for tracking the movement of road vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[13] J. Fang, Y. Li, P. N. Ji, and T. Wang, "Drone detection and localization using enhanced fiber-optic acoustic sensor and distributed acoustic sensing technology," *Journal of Lightwave Technology*, vol. 41, no. 3, pp. 822–831, 2022.

[14] L. Thiem, S. Wienecke, K. Taweesintananon, M. Vaupel, and M. Landrø, "Ship noise characterization for marine traffic monitoring using distributed acoustic sensing," in *2023 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea)*. IEEE, 2023, pp. 334–339.

[15] D. Rivet, B. de Cacqueray, A. Sladen, A. Roques, and G. Calbris, "Preliminary assessment of ship detection and trajectory evaluation using distributed acoustic sensing on an optical fiber telecom cable," *The Journal of the Acoustical Society of America*, vol. 149, no. 4, pp. 2615–2627, 2021.

[16] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016.

[17] A. H. Hartog, *An introduction to distributed optical fibre sensors*. CRC press, 2017.

[18] S. Lloyd, "Least squares quantization in pcm," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[19] B.-N. Vo and W.-K. Ma, "The gaussian mixture probability hypothesis density filter," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4091–4104, 2006.