A Supervised Learning Approach for Rainfall Detection From Underwater Noise Analysis

Andrea Trucco, *Senior Member, IEEE*, Roberto Bozzano, Emanuele Fava, Sara Pensieri, Alessandro Verri, and Annalisa Barla

***Abstract*—Underwater noise analysis allows estimation of parameters of meteorological interest, difficult to monitor with *in situ* devices, especially in very harsh environments such as polar waters. Rainfall detection is a fundamental step of acoustical meteorology toward quantifying precipitation and, indirectly, wind. To date, this task has been conducted with some success by using a few frequency bins of the noise spectrum and combining their absolute values and slopes into some inequalities. Unfortunately, these algorithms do not perform well when applied to spectra obtained by averaging multiple noise recordings made over the course of an hour. Supervised, machine learning models allow the use of all the frequency bins in the spectrum, exploiting relationships that are difficult for a human observer to identify. Among the different models tested, a binary classifier based on random forest performed well with moderate computational load. Using a dataset consisting of over 18 000 hourly averaged spectra (approximately 25 months of *in situ* recordings) and comparing the results with measurements from a surface-mounted rain gauge, the proposed system detects precipitations greater than 1 mm/h with 90% probability, keeping the false alarm probability below 0.5%. This system has demonstrated remarkable robustness as performance is achieved without intentionally excluding any spectra corrupted by sounds produced by other sources, such as naval traffic and wind blowing over the sea surface.**

***Index Terms*—Acoustical meteorology, machine learning, noise analysis, rainfall detection, supervised learning, underwater acoustics.**

I. Introduction

THE MEASUREMENT of rain and wind in the marine environment is an essential operation in understanding and monitoring natural phenomena, especially in relation to climate change and risk prevention [1]–[4]. Satellites for meteorological observation provide a valuable contribution, although the spatial and temporal resolutions they provide do not always meet monitoring requirements. This problem is particularly felt in the polar environment due to the reduced coverage that these satellites offer at higher latitudes [3], [5]. Weather surveillance radars, operating along the coast, and surface rain gauges and anemometers, installed on oceanographic fixed or mobile platforms, also present critical issues that make it difficult to deploy these devices on a large scale [3], [6], [7]. For these reasons, estimating wind speed and rainfall intensity using underwater acoustic noise is considered a crucial technique for a better understanding of the oceans, either as an alternative to or in support of satellites, coastal radar systems, and meteorological buoy networks [6], [8].

In recent years, acoustical meteorology has received considerable attention, demonstrating that wind and rain can be measured with satisfactory accuracy using low-cost underwater devices installed on fixed moorings or moving platforms, in a variety of ocean environments [1], [3], [7], [9]–[17]. However, several problematic issues are still present and need convincing answers to achieve proper operation of these devices in the field [2], [18], [19]. These issues include the possibility of performing wind and rain estimates when only acoustic data averaged over a significant period of time are available. On the one hand, the deployment in polar waters of a network of measurement devices that are capable of operating autonomously could impose a drastic reduction in processing and storage resources. Indeed, such devices might be in need of operating completely autonomously for one or more years. On the other hand, in mobile platform installations, it may be necessary to minimize the transmission resources to be committed during the rare and brief periods of surfacing. These savings requirements led Vagle *et.al.*, in their pioneering work [9], to propose a method for estimating wind speed using the average of the acoustic data acquired, at various times, over a period of one hour. Recently, a similar proposal has also been formulated and tested for rain monitoring [20].

In the literature concerning the acoustic measurement of rain, the intensity of precipitation is estimated through two distinct steps [3], [7], [10]–[13], [15], [18], [20], [21], [22]: the detection of rainfall and, if any, the estimation of its intensity. With the exception of [20], for both operations the input data are derived from acoustic signals gathered over a few seconds or, at most, a few minutes. This article investigates the possibility of detecting precipitation over a one-hour period by exploiting only the average of consecutive acoustic spectra acquired at

Manuscript received February 8, 2021; revised May 31, 2021; accepted June 16, 2021. This work was supported in part by the Italian framework Programma Nazionale di Ricerca in Antartide under Grant PNRA18_00154. *(Corresponding author: Andrea Trucco.)*

Associate Editor: J. Potter.

Andrea Trucco and Emanuele Fava are with the Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture, University of Genoa, 16145 Genova, Italy (e-mail: trucco@ieee.org; emafa97@gmail.com).

Roberto Bozzano and Sara Pensieri are with the National Research Council of Italy, Institute for the Study of Anthropic Impact and Sustainability in the Marine Environment, 16149 Genova, Italy (e-mail: roberto.bozzano@cnr.it; sara.pensieri@cnr.it).

Alessandro Verri and Annalisa Barla are with the Department of Informatics, Bioengineering, Robotics, and System Engineering, University of Genoa, 16146 Geneva, Italy (e-mail: alessandro.verri@unige.it; annalisa.barla@unige.it).

Digital Object Identifier 10.1109/JOE.2021.3091769

intervals of a few minutes during that hour, without performing any processing to filter out measurements potentially affected by noise sources other than rain. The detection also aims to reveal intermittent rain falling for a period shorter than the hour under examination. Furthermore, the adopted dataset consists of spectra acquired during different deployments of the acoustic sensor, seasons, and environmental conditions, covering about 25 months of operation, using the same platform. Over such an extended time interval, while precipitation detection is an essential step in quantifying rainfall, it can also be useful in estimating wind speed, due to the combined effect of these phenomena on underwater noise, and in monitoring of oceanographic parameters, such as sea surface salinity.

The methods mentioned above [1], [7], [9], [10], [13], [14], [21] that aim to perform rainfall detection using short-term acoustic data (short-term being the term adopted to indicate data gathered over some seconds or a few minutes) do not provide satisfactory performance when hourly averaged acoustic data are used as input. In addition, the detection is performed by decision rules that exploit only the values and slopes of the acoustic spectrum at a few predetermined frequencies (for this reason, these methods will be referred to as rule-based). To overcome this restriction, a recent paper [20] proposed machine learning methods to estimate rainfall intensity and wind speed using all the frequency bins of the underwater noise spectrum as input data to exploit implicit relationships that are not evident to the human observer. In [20], machine learning methods are also applied for rainfall detection, using hourly averaged spectra as input data. Unfortunately, detection is limited to precipitation intensities greater than 1 mm/h and the performance obtained over a one-year period is worse than that reported in [15], where a rule-based estimation method [21] was fed with short-term data collected by the same equipment and over the same time period used in [20].

This article is aimed at improving the performance of rule-based methods in rainfall detection, exploiting all the frequency bins of the spectrum within a scheme based on supervised machine learning models, already successfully applied to other detection problems of the underwater acoustic domain [23]–[25]. The new knowledge that this work introduces is twofold: the demonstration that hourly averaged spectra can be used to detect rainfall with better performance than that achieved by rule-based methods fed by short-term data; and, second, an in-depth analysis of the potential and limitations of the machine learning models adopted, made possible by experimentation on real data collected at sea over a period of more than two years. In addition, the detection scheme proposed here is capable of operating even with extremely light precipitation, being able to detect rainfall intensity of 0.1 mm/h, a value that represents the instrumental limit of most commercial rain gauges.

The rest of this article is organized as follows. Section II presents a brief state-of-the-art in rainfall detection from underwater noise. Section III describes the experimental setup and data used in this study, the detection algorithms from the prior literature, and the proposed approach, based on machine learning models. Section IV reports and compares the detection results obtained from rule-based methods, the proposed approach, and a weather radar system operating simultaneously in the area of interest. Finally, Section V concludes this article.

II. State-of-the-Art Overview

The methods proposed in past decades for rainfall detection through underwater noise analysis are based on spectral values and slopes at given frequencies, compared among them or against fixed thresholds. In [9], the spectral slopes between 3 and 8 kHz and between 3 and 19.5 kHz are compared to specific thresholds to achieve an indication of the precipitation presence. In [1], rainfall is classified in several categories depending on the difference between the average spectral levels in two bands: from 4 to 10 kHz and from 10 to 30 kHz. In [10], rainfall is detected when a set of inequalities, in which the spectral levels at 5, 8, and 25 kHz are linearly combined and compared against specific thresholds, are satisfied. Similarly, in [7], the spectral levels at 5.4, 8.3, and 21 kHz are adopted with different coefficients and thresholds. In addition, spectra corrupted by transient noise or by high wind are discarded, and a continuity check is applied to reduce the false detection rate: if no new rainfall detections occur within 10 min of the first detection, then such a detection is assumed to be false. At a later time, Nystuen proposed a new version of the detection algorithm [21] in which a higher number of inequalities combine the spectral levels at 5, 8, and 20 kHz (as is and squared) and the slopes between 2 and 8 kHz and between 8 and 15 kHz. In [13], this algorithm is further refined by updating a couple of threshold levels. Finally, in [14], a detection scheme is introduced in which the minimum and maximum spectral levels between 10 and 20 kHz are exploited. All these algorithms are set through the authors' observations of acoustic spectra collected in rainy and nonrainy conditions. Moreover, they are designed to process input data derived from short-term acoustic signals.

A statistical assessment of the detection results is provided only in [3], [7], and [15], whereas the other papers cited above evaluate the proposed algorithms only on a few selected cases. In [7], thanks to the removal of noisy samples and the introduction of a continuity check, the probability of false alarm (P_{fa}, i.e., the probability of detecting rain in the absence of precipitation) is 0.004. The probability of detection (P_d, i.e., the probability of detecting rain when precipitation occurs) is 0.6 for a rainfall intensity greater than 5 mm/h and 0.8 for a rainfall intensity greater than 10 mm/h. In [15], the authors applied the detection algorithm described in the work of Nystuen [21], including noisy sample removal and a continuity check, obtaining $P_{fa} = 0.0052$ and $P_d = 0.584$ for a rainfall intensity greater than 0.1 mm/h. P_d increases to 0.839 when only samples with rainfall intensity greater than 1 mm/h are considered. Finally, in [3], Yang *et al.* reported a $P_d = 0.7$ for a rainfall rate greater than 3 mm/h using the acoustic device described in [13].

The machine learning approach proposed in [20] for rainfall monitoring applies supervised models to hourly averaged acoustic spectra, extending the analysis to all the frequency bins instead of only a few frequencies and slopes. For the detection task, a binary classifier is built through the CatBoost algorithm, setting the lower bound for rainfall intensity equal to 1 mm/h.

When the detector is applied to the available one-year dataset (through the cross-validation scheme), a $P_{fa} = 0.0332$ and a $P_d = 0.811$ are obtained: a poorer performance than that obtained in [15] using the same dataset but exploiting short-term data in place of hourly averaged data.

III. Material and Methods

A. Experimental Measurements

The acoustic underwater noise and the rainfall intensity at sea surface were collected from June 17, 2011 to September 6, 2013 (with a few breaks, approximately 1.5 months overall) by apposite sensors installed on the meteooceanographic observatory W1M3A, moored on a deep-sea bed of 1200 m, about 80 km off the Ligurian coast, in the northwestern part of the Mediterranean Sea, as detailed in [15], [26], and [27].

The rainfall intensity was measured with a Vaisala RAINCAP Sensor, comprised in a Vaisala Weather Transmitter WXT520, placed on the upper part of the buoy trellis, at about 10 m above sea level. Precipitation measurements were acquired at high temporal resolution (5 s) and contribute to the measurements of the hourly cumulative rainfall intensity [15]. During the hours when the cumulative precipitation was acquired, the hourly average wind speed was also computed using measurements from a WindSonic 2-D anemometer installed on the same trellis on the observatory at 10 m above sea level.

The underwater acoustic noise was acquired by a dedicated oceanic recorder, based on passive aquatic listener (PAL) technology [13], [28], [29], clamped to the body of the platform at a depth of 36 m. This device is designed to operate unattended at sea for a long period of time powered by an internal battery, and to acquire an average of seven acoustic noise snapshots per hour. Each snapshot consists of a time series of 4.5 s, sampled at 100 kHz, which is processed on board to obtain a spectrum composed of 64 frequency bins, with a resolution of 0.2 kHz from 0.1 to 3 kHz and 1 kHz from 3 to 50 kHz. The spectra of the snapshots acquired in one hour (at an average interval of about 9 min from each other) were averaged, producing a mean spectrum that is included in the acoustic dataset used in this article for rainfall detection.

In the entire period of operation, 18 193 hourly averaged acoustic spectra were collected and are available for processing, amounting to about two and a half times those considered in Taylor *et al.* [20]. The concurrent measurements of the hourly rainfall intensity are also available and are assumed, in this study, as the ground truth. The rain gauge measured a precipitation greater than 0.1 mm/h in 876 of the 18 193 h considered. The maximum rainfall intensity measured was 51.5 mm/h, and the distribution of the observed intensities is shown in Fig. 1. The average wind speed ranged between 0.4 and 20.7 m/s with the distribution shown in Fig. 2. Finally, the tracks of the automatic identification system (AIS) used on ships reveal how many of them transited near the buoy in the period of data acquisition. Considering a circle with a radius of 5 km, centered at the position of the buoy, the number of hours in which at least one ship crossed the circle is 1999, of which 78 are characterized by the presence of rain.

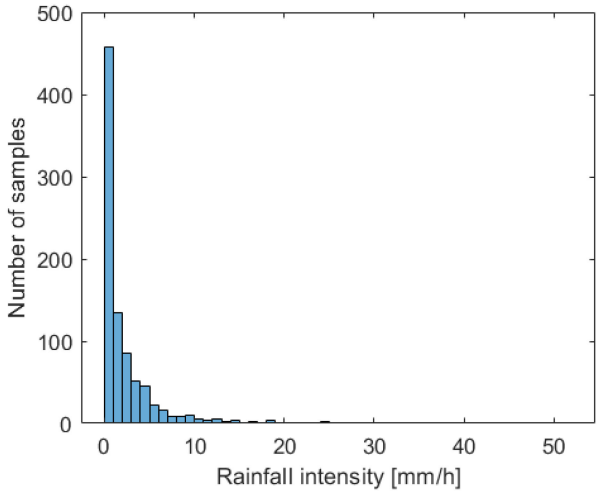


Fig. 1. Rainfall intensity distribution in the dataset samples.

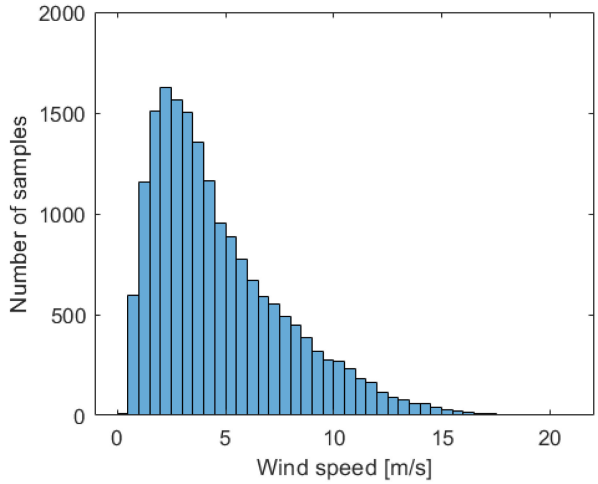


Fig. 2. Wind speed distribution in the dataset samples.

Fig. 3 provides an example of hourly averaged spectra acquired by the underwater acoustic system, comparing the spectra acquired over two consecutive hours in which the wind speed remained nearly constant. The difference between the two spectra is mainly due to rainfall, which was absent in the first hour and present in the second hour. The spectrum in Fig. 3(a), related to a wind speed of about 3.5 m/s, varies visibly despite the fact that the amount of rain accumulated in the second hour is only 0.1 mm/h. In Fig. 3(b), where the accumulated rain is 1.1 mm/h, the difference between the spectra, both related to a wind speed of about 2.1 m/s, increases significantly. The differences observed should not lead to the conclusion that rainfall detection is easy: spectra very similar to those collected when precipitation is present can be produced by an increase of wind speed, in the absence of rainfall. To show this fact, two spectra are introduced in Fig. 3(a) and (b), acquired a few hours after those already discussed, in the absence of rainfall, but with a higher wind speed: 5 m/s instead of 3.5 m/s and 10 m/s instead of 2.1 m/s, respectively. These ambiguities seem to confirm the need to

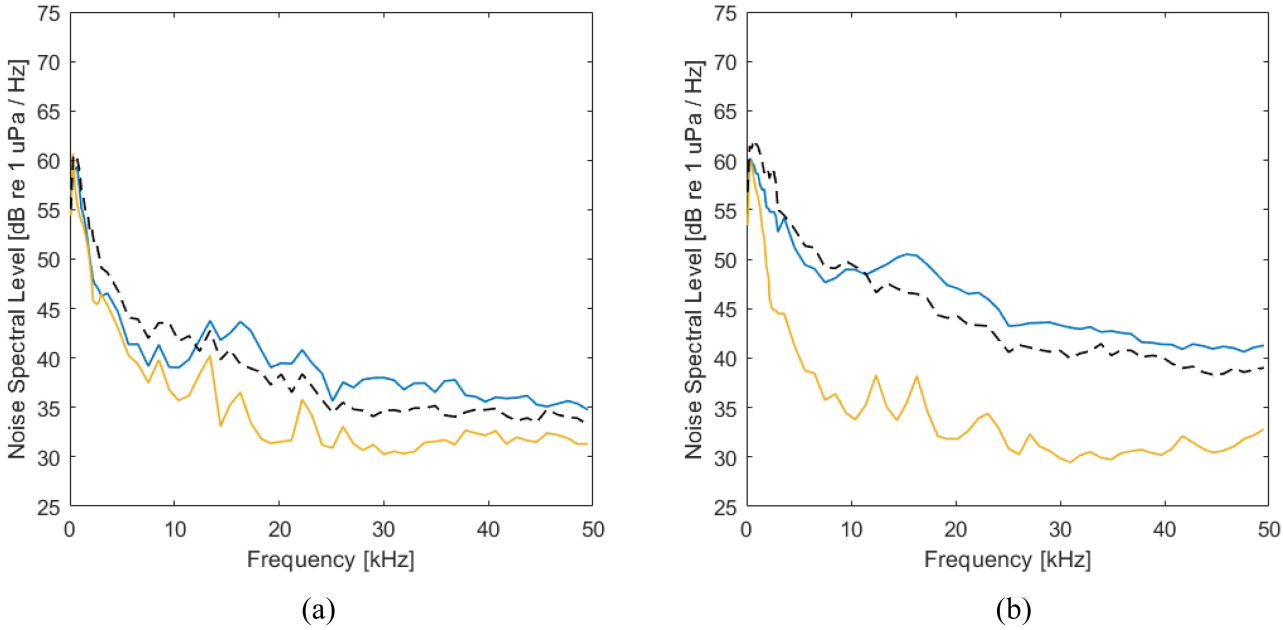


Fig. 3. Comparison of spectra acquired in two consecutive hours, with (light blue lines) and without (golden lines) rainfall. The wind speed is approximately the same. The dashed line is a spectrum acquired a few hours later, without rainfall, but with a higher wind speed. (a) Wind speed: 3.5 m/s for light blue and golden lines; 5 m/s for the dashed line. Rainfall intensity: 0.1 mm/h for the light blue line, 0 mm/h otherwise. (b) Wind speed: 2.1 m/s for light blue and golden lines; 10 m/s for the dashed line. Rainfall intensity: 1.1 mm/h for the light blue line, 0 mm/h otherwise.

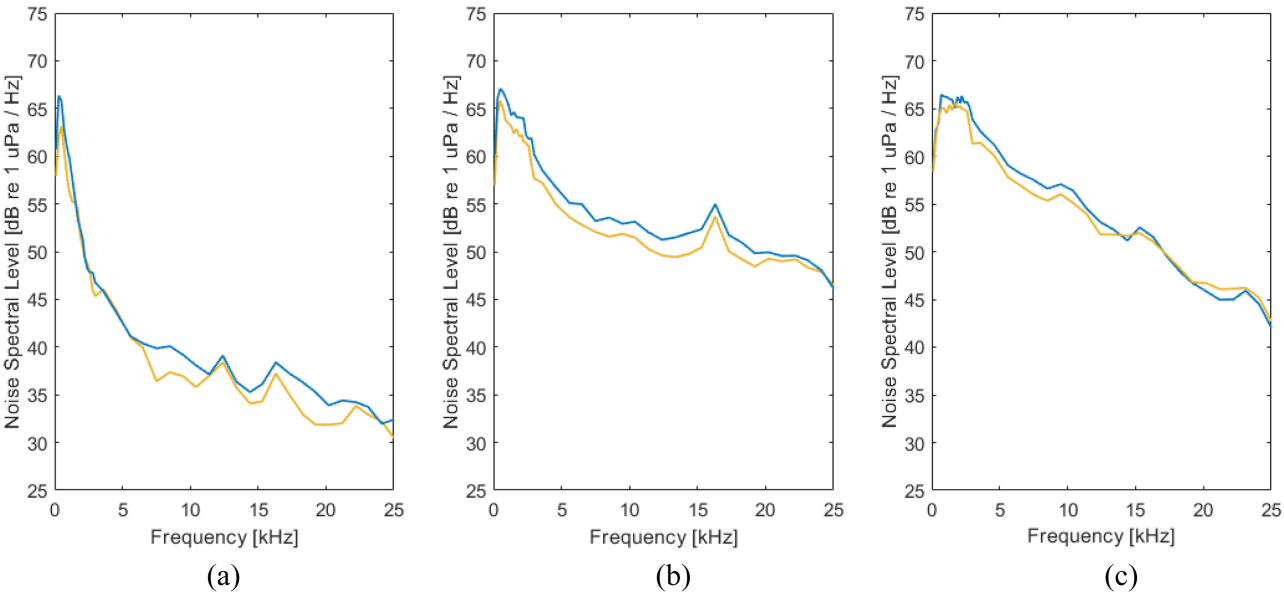


Fig. 4. Comparison of the spectrum acquired over the hour in which a ship passage occurred (light blue line) and that acquired in the hour before the passage (golden line). Wind and rainfall are almost constant. (a) Wind speed: 2.6 m/s; no rainfall. (b) Wind speed: 3.4 m/s; rainfall intensity: 7.5 mm/h. In (c) the spectra of two consecutive hours with constant conditions (wind speed: 15.5 m/s; rainfall intensity: 4.5 mm/h) and with no ship passage are compared.

exploit all the frequency bins of the spectrum and to learn the implicit relationships from data that allow one to successfully detect the rainfall. To address this need, the adoption of machine learning methods is an effective and viable solution.

Two examples of the ship-passage effects on hourly averaged spectra are shown in Fig. 4. Panels (a) and (b) compare the spectrum acquired in the hour before the passage with that of the hour in which the passage occurred. In Fig. 4(a), the average wind speed in the two consecutive hours was about 2.6 m/s and there was no rain. In Fig. 4(b) instead, the wind speed was about 3.4 m/s and the rainfall produced an accumulation of about 7.5 mm/h in both hours. For comparison, Fig. 4(c) shows the spectra of two consecutive hours when the wind remained at a speed of about 15.5 m/s and the rain produced an accumulation of

about 4.5 mm/h, with no ship passages. Although an increase in spectral values (on the order of a few decibels) is visible in the spectra related to the ship passages, in almost the entire 0–25-kHz band, a similar change is also visible in Fig. 4(c) although no ship passage induced it.

B. Rule-Based Algorithms for Rainfall Detection

The notation $S(f_{\mathrm{k}})$ is introduced to indicate the sound spectral level of underwater noise, measured in dB re 1 μPa^2 Hz^{-1}, at the frequency f_{k} expressed in kHz.

In [9], two rainfall detection rules are proposed, similar to each other, but with a different number of examined frequency bins. They are denoted by V1 and V2

$$\begin{aligned}\mathrm{V1}\quad & S(19.5)-S(3)>-13.82 \text{ OR } S(12.5)-S(3)>-10.54\\ & \text{OR } S(8)-S(3)>-6.82\\ \mathrm{V2}\quad & S(19.5)-S(3)>-13.25 \text{ OR } S(8)-S(3)>-6.82.\end{aligned} \tag{1}$$

In [7], rainfall is detected if at least one of the following three conditions is verified, the third condition being specific for drizzle

$$S(21)+2.35\,S(5.4)>194 \tag{2}$$

$$S(21)>48 \text{ AND } S(5.4)>53 \tag{3}$$

$$S(21)>44 \text{ AND } S(21)-0.7\,S(8.3)>14. \tag{4}$$

In addition, the removal of spectra corrupted by noise and the temporal continuity check are applied [7], as described in Section II. These operations lose their meaning when this algorithm is applied to hourly averaged spectra.

In [13] and [21], rainfall is detected if at least one of the following four conditions is verified, the third condition being specific for drizzle and the fourth for rain with high wind:

$$S(20)-0.75\,S(5)>5 \text{ AND } S(5)\leq 70 \tag{5}$$

$$S(8)>60 \text{ AND } Q(2,8)>\theta \text{ AND } S(20)>45 \tag{6}$$

$$\begin{aligned}&S(8)<50 \text{ AND } Q(8,15)>-5 \text{ AND } S(20)>35\\ &\quad \text{AND } S(20)>0.9\,S(8)\end{aligned} \tag{7}$$

$$\begin{cases} S(20)+0.1144\,S^2(8)-12.728\,S(8)>-307\\ \text{AND } Q(2,8)>\theta\\ \text{AND } S(20)+0.1\,S^2(8)-11.5\,S(8)<-281 \text{ AND}\\ 51<S(8)<64\end{cases} \tag{8}$$

where $Q(f_1, f_2)$ is the spectral slope, in dB/decade, between the frequencies f_1 and f_2 (expressed in kHz):

$$Q\ (f_1,f_2)=\frac{S(f_1)-S(f_2)}{\log_{10}(f_1)-\log_{10}(f_2)}. \tag{9}$$

The difference between the algorithms in [13] and [21] is only the value assigned to the constant θ: $\theta =$ –18 dB/decade in [21] and $\theta =$ –13 dB/decade in [13].

The algorithms proposed in [1], [10], and [14] are not included in this collection because the work of Black *et al.* [1] is mainly dedicated to the classification of rain, downstream of a detection carried out by other means; Nystuen and Selsor [10] represented a preliminary version of the algorithm proposed in [7]; Kuhner [14] presented a general idea based on the maximum spectral slope observed between 10 and 20 kHz, with no specific detection algorithm.

C. Rainfall Detection by Supervised Learning Models

An alternative to the rule-based detection algorithms described above is to exploit all the information available, looking for relationships between the spectrum frequency bins (none excluded) and the rainfall presence through machine learning models driven by experimental observations. A supervised binary classifier (the classes of which are rainy and nonrainy) that receives as input the frequency bins of an hourly averaged spectrum can successfully perform precipitation detection. In this study, standard supervised learning methods well suited to address binary classification in the presence of a large number of features are adopted: two linear classification techniques (linear discriminant analysis and logistic regression), a kernel-based method (support vector machine), and an ensemble learning method (random forest). Given the type and size of the available data, the spectrum of the methods adopted is considered sufficient to assess the advantage of employing a data driven approach to effectively detect rainfall.

1) Notation for Data and Performance: To discuss the characteristics of the models mentioned, the following notation will be used. A sample is an hourly averaged spectrum and is indicated by the vector $\boldsymbol{x}$, $\boldsymbol{x}\in\mathbb{R}^d$. The vector is composed of d frequency bins, called features. In this study, d is equal to 64 and the dataset contains 18 193 samples, a fraction of which are used for the training of the statistical models. The training set is indicated by $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^{L}$, where L is the number of samples used for the training phase; $\boldsymbol{x}_i$ is the ith training sample, associated with either rainy or nonrainy classes; and y_i is equal to ± 1 depending on membership of $\boldsymbol{x}_i$ to one of the two classes: $+1$ for the rainy case and -1 for the nonrainy case. The samples not belonging to the training set constitute the test set. The application of the trained model to a sample $\boldsymbol{x}$ taken from the test set allows us to assign such a sample to the $+1$ or -1 class. Since the actual rain condition is also known for the samples of the test set, P_d, P_{fa}, the overall accuracy (OA, i.e., the probability of correct classification) and the receiver operating characteristic (ROC) curve can easily be estimated. The ROC curve shows the possible tradeoffs between P_d and P_{fa} and can be traced by varying the threshold used to decide membership of the test sample $\boldsymbol{x}$ on the basis of the real-valued score produced by the trained model, when the model is applied to $\boldsymbol{x}$. The area under the ROC curve (*AUC*) is commonly used to quantitatively evaluate the detector positioning between the detector choosing at random (*AUC* = 0.5) and the ideal detector (*AUC* = 1.0).

2) Linear Discriminant Analysis: Linear discriminant analysis (LDA) is a traditional method [30], based on decision theory and Bayes theorem, in which the probability density functions for the samples belonging to the $+1$ and -1 classes

are assumed to be multivariate Gaussian with mean vectors μ_{+1} and μ_{-1}, respectively, and the same covariance matrix Σ. The knowledge of these class-conditional densities, $f_{\boldsymbol{x}|+1}(\boldsymbol{x})$ and $f_{\boldsymbol{x}|-1}(\boldsymbol{x})$, together with the prior probabilities for the two classes, $P(-1)$ and $P(+1)$, makes computation of the class posterior probabilities possible for a given sample $\boldsymbol{x}$, $P(+1 \mid \boldsymbol{x})$ and $P(-1 \mid \boldsymbol{x})$. Specifically, the probability for the $+1$ class given sample $\boldsymbol{x}$ is

$$P\ (+1 \mid \boldsymbol{x}) = \frac{f_{\boldsymbol{x}|+1}\,(\boldsymbol{x})\,P\,(+1)}{f_{\boldsymbol{x}|+1}\,(\boldsymbol{x})\,P\,(+1) + f_{\boldsymbol{x}|-1}\,(\boldsymbol{x})\,P\,(-1)}. \quad (10)$$

Sample $\boldsymbol{x}$ is assigned to the $+1$ class if this probability exceeds 0.5, to the -1 class otherwise. Although the 0.5-threshold is optimum in terms of overall classification accuracy, a different value can be set to modify the balance between P_d and P_{fa}, and, therefore, the tuning of the threshold allows the tracing of the detector's ROC curve. The samples belonging to the training set are used in LDA to estimate the mean vectors, the covariance matrix, and the prior probabilities mentioned above. LDA is a linear classification method because membership of sample $\boldsymbol{x}$ can be equivalently assigned working on the log-odds function (i.e., the logarithm of the ratio between $P(+1 \mid \boldsymbol{x})$ and $P(-1 \mid \boldsymbol{x})$), which is a linear equation in $\boldsymbol{x}$.

3) Logistic Regression: The logistic regression (LR) model assumes the log-odds function to be a linear function in $\boldsymbol{x}$ and derives the equations for the class posterior probabilities without introducing any assumption about the class-conditional density functions [30]. In the binary case, the probability for the $+1$ class given the sample $\boldsymbol{x}$ results

$$P\ (+1 \mid \boldsymbol{x}) = \frac{1}{1 + \exp\left(\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}\right)} \quad (11)$$

i.e., a sigmoid function whose parameters β_0 and $\boldsymbol{\beta}$ can computed by maximizing a conditional log-likelihood function. The maximization is achieved through an iterative procedure in which the training set samples are exploited and the Newton–Raphson algorithm is typically applied to find the root of the first derivative [30]. As in LDA, sample $\boldsymbol{x}$ is assigned to the $+1$ class if $P(+1 \mid \boldsymbol{x})$ is greater than 0.5, to the -1 class otherwise, but different threshold values can be used to trace the detector's ROC curve and change the balance between P_d and P_{fa}.

4) Support Vector Machine: A support vector machine (SVM) assigns sample $\boldsymbol{x}$ to one of the two classes based on the score of the discriminant function

$$h\ (\boldsymbol{x}) = \sum_{i=1}^{L} \alpha_i y_i K\,(\boldsymbol{x}_i, \boldsymbol{x})\ + b \quad (12)$$

where $K(\cdot,\cdot)$ is a kernel function and the coefficients α_i and b are optimized by solving a quadratic programming problem [30]. This optimization problem exploits the samples of the training set and includes a parameter C that bounds the range for $\alpha_i:\ 0 < \alpha_i < C,\ i = 1, 2, \ldots, L$. Sample $\boldsymbol{x}$ is assigned to the $+1$ class if $h(\boldsymbol{x})$ is positive to the -1 class otherwise. As for previous methods, different threshold values can be used to trace the detector's ROC curve and change the balance between P_d and P_{fa}. In the SVM literature, the most commonly adopted kernels are the linear function, polynomial function of order q and Gaussian radial basis function (RBF), defined, respectively, as

$$K\ (\boldsymbol{x}_i, \boldsymbol{x}) = \boldsymbol{x}^T\ \boldsymbol{x_i} \quad (13)$$

$$K\ (\boldsymbol{x}_i, \boldsymbol{x}) = \left(1 + \boldsymbol{x}^T \boldsymbol{x_i}\right)^q \quad (14)$$

$$K\ (\boldsymbol{x}_i, \boldsymbol{x}) = \exp\left(-||\boldsymbol{x} - \boldsymbol{x}_i||^2 / 2\sigma^2\right) \quad (15)$$

where σ^2 is a specific parameter of the RBF kernel.

The choice of C and σ^2 (if the case) requires specific attention and, possibly, an optimization stage [30]. In addition, although not strictly necessary, all features of the dataset samples are often preliminarily standardized, so that each of them has a zero mean and a unitary variance. This operation makes features insensitive to the scales on which they are measured and favors numerical stability in the solution of the quadratic programming problem mentioned above.

5) Random Forest: Random forest (RF) is an ensemble model that aggregates the predictions individually achieved by many decision trees, separately trained on a subset of samples randomly chosen from the training set [30]. A decision tree is an acyclic connected graph, where each node represents a decision rule (called split) related to a single feature that leads to the partition of data in two groups. To automatically set the structure and splits of a decision tree, classification and regression trees (CART) is a widely adopted algorithm in which a new node is created by identifying the feature that yields the best split in terms of a preselected metric. In an RF model, B trees are generated and trained in an independent and identically distributed way by performing, for each tree $\mathcal{T}_b$, $b = 1, \ldots, B$, the following steps [30]: a) a subset of L samples is drawn randomly from the training set, uniformly and with replacement (this means that some samples are taken more than once, others are not chosen at all); b) such a subset is used to grow the tree $\mathcal{T}_b$, for each node of which a pool of m features is selected (at random and uniformly from the d features) and used to identify the best feature and the best decision rule to split the node into two daughter nodes; c) the previous step is repeated until at least one of the predefined stopping criteria is satisfied. When all the B trees are generated, an unknown sample $\boldsymbol{x}$ is classified as follows: the sequence of decision rules of the bth tree is applied to $\boldsymbol{x}$ in such a way that the corresponding class prediction $\hat{y}^b(\boldsymbol{x})$ is reached (namely, $+1$ or -1); the predictions from all the trees of the RF are used to compute a score

$$g\ (\boldsymbol{x}) = \frac{1}{B} \sum_{b\,=\,1}^{B} \hat{y}^b\,(\boldsymbol{x}) \quad (16)$$

sample $\boldsymbol{x}$ is assigned to the $+1$ class if $g(\boldsymbol{x})$ is positive to the -1 class otherwise. Threshold values different from zero can be used to trace the detector's ROC curve and tune the balance between P_d and P_{fa}. Although the setting of B and m does not critically affect performance, it deserves some investigation, recalling that these two parameters affect the computational burden.

6) Cross-Validation: The assessment of a trained statistical model performance is a crucial task for which *K*-fold cross-validation represents an easy and extensively applied option

TABLE I
PROBABILITIES OF DETECTION P_d, AND FALSE ALARM P_{fa}, FOR THE RULE-BASED ALGORITHMS APPLIED TO HOURLY AVERAGED SPECTRA

Algorithm	P_d	P_{fa}
Vagle *et al.*, 1990 [9] – V1	0.900	0.630
Vagle *et al.*, 1990 [9] – V2	0.880	0.570
Ma *et al.*, 2005 [7]	0.300	0.001
Nystuen, 2011 [21]	0.671	0.116
Nystuen *et al.*, 2015 [13]	0.586	0.094

[30]. To exploit the available data for both training and testing a machine learning model, the dataset is split in *K* subsets (called folds), nonoverlapped, and of approximately equal size. Taking the *k*th subset aside, the model is trained using the other *K*–1 subsets of data, and the test is performed using the data of the *k*th subset. This operation is repeated for *k* ranging from 1 to *K*, in such a way that every sample is used, in turn, to train and test the supervised model. By combining the predictions performed at each step *k* on the data subset kept aside, a prediction for each sample of the entire dataset is finally available. Because the prediction consists of the probability (or score) for membership of the sample of a given class, after setting a threshold value, estimation of *OA*, P_d, and P_{fa} is possible. In addition, the tuning of such a threshold allows the generation of the cross-validated ROC curve [30].

To cope with the different cardinality of the two classes in the dataset, the dataset partition in *K* subsets can be performed by a stratified scheme according to which each subset maintains approximately the same class proportions as the original dataset.

IV. RESULTS AND DISCUSSION

To delineate the desired performance of the rain detector, it is necessary to recall that rainfall is present in 5% of the one-hour periods included in the dataset and that the precipitation limit, which distinguishes between rainy and nonrainy hourly averaged spectra is particularly low (i.e., 0.1 mm/h). In this scenario, it is strictly necessary that the false alarm probability be very low, while a detection probability not too close to one may be acceptable. Consequently, the performance of a detector cannot be considered acceptable if P_{fa} exceeds 0.01.

A. Performance of Rule-Based Algorithms

The application of the algorithms introduced in Section III-B to the dataset described in Section III-A provides the results summarized in Table I. It is important to recall that these algorithms were designed to detect rainfall using short-term acoustic spectra, whereas in this study they are applied to hourly averaged spectra.

The algorithms in [9] achieve high P_d, but this is accompanied by excessive P_{fa}. A bias in hydrophone sensitivity cannot be the cause of the problem, because the quantities compared with thresholds in (1) are subtractions between measurements. One option to make the algorithms more selective is to arbitrarily

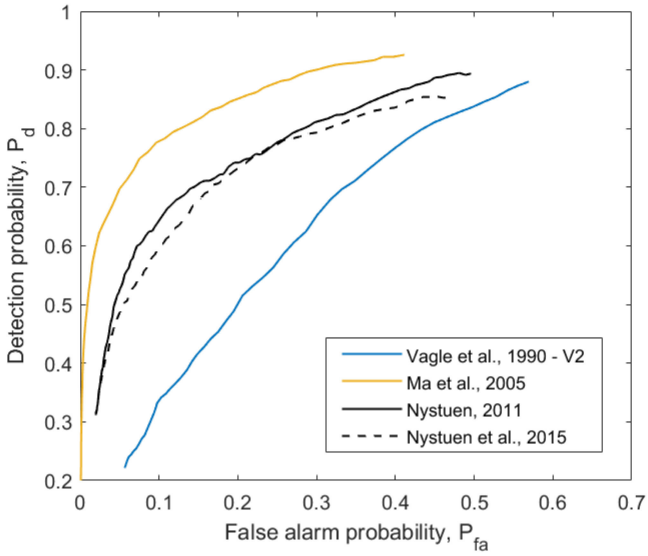


Fig. 5. ROC curves obtained by varying threshold values [9] and hydrophone sensitivity [7], [13], [21] in the rule-based algorithms listed in Table I.

increase the threshold values, modifying the V2 rule, as follows:

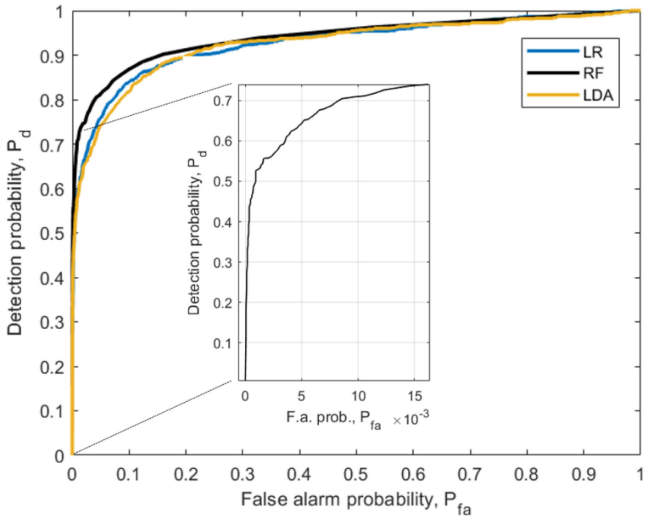
$$S(19.5) - S(3) > -13.25 + \delta \,\text{OR}\, S(8) - S(3) > -6.82 + \delta \quad (17)$$

where $\delta > 0$. Varying the value of δ between 0 and 5, the ROC curve in Fig. 5 is obtained. The P_{fa} reduction is obtained but, unfortunately, it is accompanied by a significant P_d decrease. When a similar modification is applied to the V1 rule, the performance is worse since the ROC curve is always below that shown in Fig. 5 for the V2 rule.

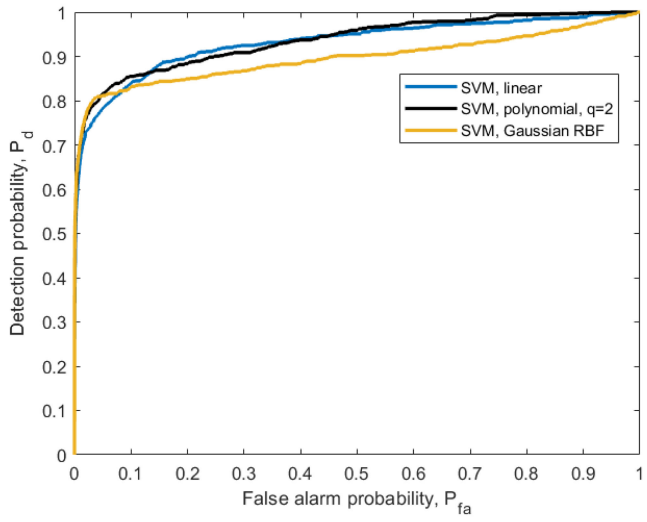
What is more, the algorithms in [7], [13], and [21] do not provide satisfactory performance, because P_d is too low, as for [7], or P_{fa} is too high, as for [13] and [21]. As discussed in [12], the performance of these algorithms can be optimized by considering potential errors in hydrophone sensitivity. To do this, the values $S(f_k)$ in equations from (2) to (8), for whatever f_k, are replaced by $S(f_k) + \varepsilon$, where ε is intended to compensate a sensitivity bias. Varying ε between –10 and 10 dB re 1 μPa^2 Hz^{-1}, the ROC curves shown in Fig. 5 are obtained. This comparison clearly evinces that the rule-based algorithm achieving the best performance (with the discussed correction) is the one proposed in [7]. In particular, for $\varepsilon = 2$ dB re 1 μPa^2 Hz^{-1}, a detection probability $P_d = 0.521$ is accompanied by $P_{fa} = 0.010$. The corrections introduced fail to reduce the false alarms of the other rule-based algorithms [9], [13], [21] to acceptable values: P_{fa} always remains significantly higher than 0.01.

B. Performance of the Supervised Learning Models

To assess and compare the detection performance of the statistical models, tenfold cross-validation with stratification in dataset partitioning is adopted. In addition, for the SVM approach feature standardization is applied, the constant C is set equal to 1.0, according to common practice, and the variance σ^2 of the Gaussian RBF kernel, after some tests, is set equal



(a)



(b)

Fig. 6. ROC curves for the supervised learning models listed in Table II. (a) LDA, LR, and RF, with a zoom for the RF model. (b) SVM with three kernel functions.

to 8.0. For the parameters of the RF, $B = 100$ trees and $m = 22$ features are used, although a change of these values in even rather broad ranges does not significantly affect the performance obtained. Data processing is performed using MATLAB and, in particular, the Statistics and Machine Learning toolbox. Additionally, it has been verified that results perfectly consistent with those shown below can be obtained using the scikit-learn library for the Python programming language.

The ROC curves obtained from the trained models are shown in Fig. 6 and their performance is summarized in Table II, where the P_d value for which $P_{fa} = 0.01$ is reported. In the cross-validation procedure, one of the K folds into which the dataset has been split, in turn, is not used for training but is instead used for testing. At the end of the procedure and after setting the decision threshold, it is possible to calculate P_d and P_{fa} in each fold used for the test. The data inserted in Table II are the average

TABLE II
DETECTION PROBABILITY, P_d, FALSE ALARM PROBABILITY, P_{fa}, OVERALL ACCURACY, *OA*, AND AREA UNDER THE ROC CURVE, *AUC*, FOR THE SUPERVISED LEARNING MODELS. FOR THE PROBABILITIES, THE AVERAGE ± THE STANDARD DEVIATION IS REPORTED. FOR EACH CLASSIFIER, THE THRESHOLD VALUE *TV*, USED TO OBTAIN THE AVERAGE P_{fa} EQUAL TO 0.01, IS REPORTED

Classifier	P_d	P_{fa}	*OA*	*AUC*	*TV*
LDA	0.583 ± 0.053	0.010 ± 0.003	0.970	0.926	0.38
LR	0.597 ± 0.058	0.010 ± 0.003	0.971	0.928	0.33
SVM, linear	0.667 ± 0.041	0.010 ± 0.003	0.974	0.931	-0.67
SVM, polynomial, q=2	0.702 ± 0.060	0.010 ± 0.002	0.976	0.936	-0.51
SVM, Gaussian RBF	0.703 ± 0.062	0.010 ± 0.003	0.976	0.897	-0.71
RF	0.708 ± 0.054	0.010 ± 0.002	0.977	0.941	-0.42

and standard deviation of the P_d and P_{fa} values calculated on each fold. The average P_{fa} is 0.01 since the threshold is set precisely to achieve this result. The threshold values used for each classifier are also included in Table II and should be read recalling that the optimal threshold values (i.e., those values that maximize OA) are: 0.5 for LDA and LR; 0 for SVM and RF.

The linear classifiers (i.e., LDA and LR) perform moderately better than the best rule-based algorithm, increasing the probability of detection to about 0.6. A further advantage is offered by the SVM and RF classifiers for which probability of detection exceeds 0.7. The change of the kernel function for the SVM classifier does not significantly alter the performance, although the linear case shows a lower detection ability and the Gaussian case reports the worst *AUC* figure. The *OA* values are all greater than 0.97, but this finding has little relevance because it is strongly influenced by the correct classification of nonrainy samples (probability 0.99, P_{fa} being 0.01) which are by far the most numerous. Overall, the best option among the models considered is the RF classifier because it achieves the best performance figures, shows a stability better than that of SVM classifiers with polynomial or RBF kernels, requires a computational load lower than that of such SVM classifiers, and is not appreciably affected by changes in the parameter setting. Accordingly, in the remainder of this section, further analysis and performance comparisons will be carried out with reference to the RF-based classifier. Indeed, the goal is not to identify the best model, but rather to demonstrate that the machine learning approach is well suited for rainfall detection also in case of drizzle phenomena, characterized by low rainfall intensity.

A portable computer with an Intel Core i7 CPU of 1.9 GHz and 16 GB of RAM memory trains the RF classifier, using a MATLAB routine and the entire dataset, in about 20 s. The execution of the detection task on more than 18 000 samples of the dataset requires less than 2 s.

C. In-Depth Analysis and Comparisons

In Fig. 6(a), the zoom of the ROC curve for the RF model demonstrates that P_d remains greater than 0.6 even if P_{fa} is reduced by as much as 0.0035. More precisely, the following

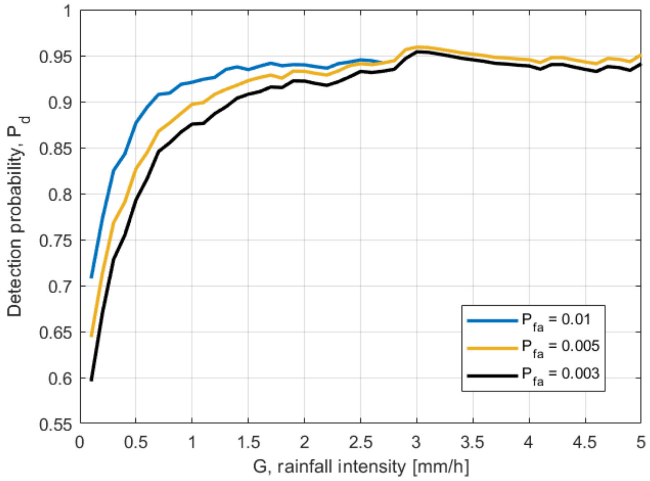


Fig. 7. Detection probability for the rainy samples with a rainfall intensity greater than or equal to G. Three RF-based classifiers, with different false alarm probabilities, are considered.

probability pairs, $\{P_d, P_{fa}\}$, lie on that curve: {0.661, 0.006}, {0.644, 0.005}, {0.623, 0.004}, {0.588, 0.003}.

The ability of the classifier to detect the precipitation can be analyzed as a function of the rainfall rate [7], [15], as shown in Fig. 7. In this case, P_d is estimated using the hourly samples in which the cumulated rainfall, measured by the rain gauge on the platform in one hour, is equal to or greater than a value G. The P_d curves shown in Fig. 7 are related to three choices of the threshold value, leading to different P_{fa}: 0.010, 0.005, and 0.003. P_d increases rapidly with G, reaching, respectively, 0.921, 0.897, and 0.876 for $G = 1$ mm/h. Although the probabilities of detection of the three detectors show significant differences for $G < 2$ mm/h, for rainfall intensities higher than this value the three detectors provide similar P_d. It is therefore possible to design acoustic detectors capable of detecting rainfall of intensity greater than 2 mm/h with a probability greater than 0.9, while keeping a false alarm probability of 0.003.

The sharp P_d increase with G observed in Fig. 7 shows that the missed detections are mainly related to drizzle phenomena characterized by low precipitation intensity. This relation is confirmed by the average of the rainfall intensities recorded by the surface rain gauge when the precipitation is detected or missed by the underwater acoustic device. Among the 876 acoustic samples collected in rainy conditions (with intensity greater than or equal to 0.1 mm/h), the RF-based classifier correctly detects 620 of them (70.8%) and misses the remaining 256 samples (29.2%). The average rainfall intensity measured for the detected samples is 2.98 mm/h, whereas the average intensity for the missed samples is 0.71 mm/h. The small fluctuations that the curves in Fig. 7 show, especially for G greater than 2 mm/h, are mainly due to the limited number of samples available to train the RF model (in the training phase) and to estimate the detection probability (in the test phase). The number of acoustic samples with rainfall greater than 3 mm/h is about 200, while the number of those with rainfall greater than 4 mm/h is reduced to less than 150.

Moving from missed detections to false alarms, an analysis of wind distribution provides some interesting insights. Fig. 8 shows the wind speed histograms for rainy samples correctly detected (620 samples), nonrainy samples correctly classified (17 145 samples), and nonrainy samples raising false alarms (172 samples, corresponding to $P_{fa} = 0.01$). The average wind speeds for these three categories are 8.5, 4.6, and 9.3 m/s, respectively. It is evident that the false alarm samples present a wind distribution more similar to that of the rainy samples than to that of nonrainy samples.

However, the histogram of nonrainy samples shows that there are over a thousand samples with wind speed greater than 10 m/s that are correctly classified. To analyze this issue in detail, Fig. 9 shows the estimated P_{fa} when the samples for which the wind speed is greater than W, $W \in [0.1, 10]$ m/s are considered. The three detectors already examined in Fig. 7 are included. Notwithstanding the considerable rise of P_{fa} with increasing wind speed, the probability of correct classification for nonrainy samples remains satisfactory (e.g., for $W = 10$ m/s, P_{fa} increases from 0.01 to 0.08, but the probability of correctly classifying a nonrainfall sample is still high: 0.92). Therefore, the wind-related similarity only partially explains why the detector is misled and false alarms occur.

Finally, the performance of the RF-based detector during the period of data collection is examined in Fig. 10, where the height of the bars indicates the rainfall intensity measured by the rain gauge and the colors distinguish samples correctly detected (light blue bars) from missed alarm samples (orange bars). Samples raising false alarms are inserted as white bars with black edges, and an arbitrary height of 2 is set for them. The two zoom panels show the typical behavior that characterizes the 25-month span of data collection with good uniformity.

As described in Section III-A, 1999 out of the 18 193 data set samples are characterized by the passage of a ship within 5 km of the platform during the observation hour. These samples are not discarded and are used, like all others, to train and test the statistical model. It is verified *a posteriori* that P_d and P_{fa} values estimated on these samples do not differ significantly from those already reported, thus supporting the robustness of the proposed detector. Fig. 11 shows a further zoom of a portion of the 25-month span introduced in Fig. 10, in which the one-hour intervals at which a ship passage occurred are indicated by a black diamond. Again, it can be verified that there is no correlation between the ship passages and false alarms or missed detections.

Since the rainy samples most susceptible to missed detection are those characterized by modest precipitation intensity, it is reasonable to expect that a P_d of 70.8% would allow the detection of rainy samples with the greatest impact in terms of cumulative precipitation. Fig. 12 compares the cumulative rainfall profiles over the 25-month span obtained by considering all rainfall events measured by the rain gauge (876 hourly samples) or only those detected by the proposed underwater acoustic system (620 hourly samples). It can be emphasized that although the detected rainy samples are 70.8% of the total, these samples, at the end of the observation period, gather 91.0%

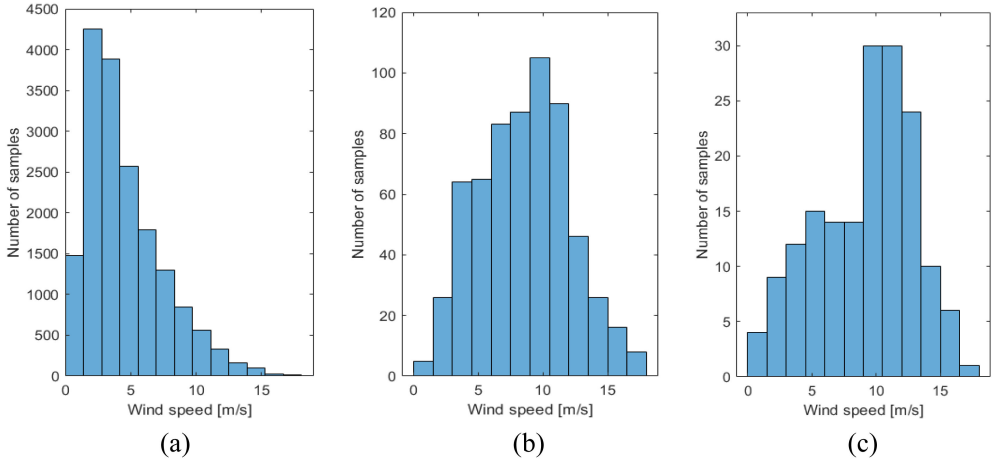


Fig. 8. Histograms of wind speed for: (a) nonrainy samples correctly classified; (b) rainy samples correctly detected; (c) nonrainy samples raising false alarms.

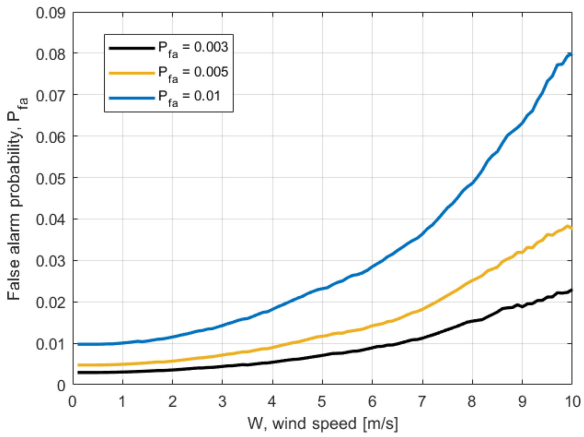


Fig. 9. False alarm probability for nonrainy samples with a wind speed greater than W. Three RF-based classifiers with different probabilities of false alarm (on the entire dataset) are considered.

of the cumulative precipitation (i.e., 1851 mm out of a total of 2035 mm).

The performance achieved by the RF-based detector acting on hourly averaged spectra can also be compared with those obtained by other underwater acoustic systems [3], [7], [15] acting on short-term spectra, summarized in Section II. By using data in Fig. 7, it is immediately possible to observe that the proposed system, at the same P_{fa} values and rainfall intensities, always provides a significantly higher detection capability. Moving from short-term spectra to hourly averaged spectra, according to the data presented in Section IV-A, the performance obtained from the detection algorithms used in [3], [7], and [15] worsen. As a result, the supervised learning models adopted in this study achieve a detection performance significantly better than those obtained from rule-based detection algorithms and better even than that obtained from the binary classifier proposed in [20].

Another useful comparison is with the rainfall detection capability of a weather radar installed on Mount Settepani, located at about 1400 m above sea level, about 87 km away from the buoy, covering the area of investigation and the data of which were used in [15]. In that study, rainfall detection by radar at the W1M3A observatory is characterized by $P_{fa} = 0.009$ accompanied by $P_d = 0.728$ for $G = 0.1$ mm/h and $P_d = 0.846$ for $G = 1$ mm/h. The data in Fig. 7 show that the performance of the proposed acoustic system is very close to that of radar: slightly worse for $G = 0.1$ mm/h and slightly better for $G = 1$ mm/h. However, it is important to emphasize the qualitative nature of this comparison because the radar performance refers to a time period of about 11 months [15], thus significantly shorter than the 25-month period considered in this study.

V. Conclusion

This study concerned the possibility of detecting precipitation, from drizzle phenomena to events of high intensity, using the underwater acoustic noise spectrum obtained from the average of the instantaneous spectra acquired, at various times, over the course of an hour. Since each sample is representative of an entire hour, to maintain sufficient temporal coverage, it was necessary to analyze all the spectra acquired, even those altered by the passage of ships, high wind, and other concurrent noises.

A dataset composed of more than 18 000 h of measurements at sea allowed an in-depth experimentation of different rainfall detection methods. Although the rainfall detection by rule-based algorithms taken from the literature have not provided satisfying performance on this type of spectrum, machine learning methods have shown that the detection can be carried out successfully. In this analysis, kernel-based and ensemble-learning models have demonstrated the best performance among the experimented supervised classifiers. In particular, the RF-based binary classifier has shown a satisfactory balance between computational burden and performance, reaching a detection probability greater than 90% when precipitation exceeds 0.7 mm/h and P_{fa} is 1% or, alternatively, when precipitation exceeds 1.4 mm/h and P_{fa} is 0.3%. This level of performance is slightly better than that obtained by a weather radar operating in the experiment area, and therefore the proposed method represents a promising alternative to obtain an estimate of rainfall intensity in areas where

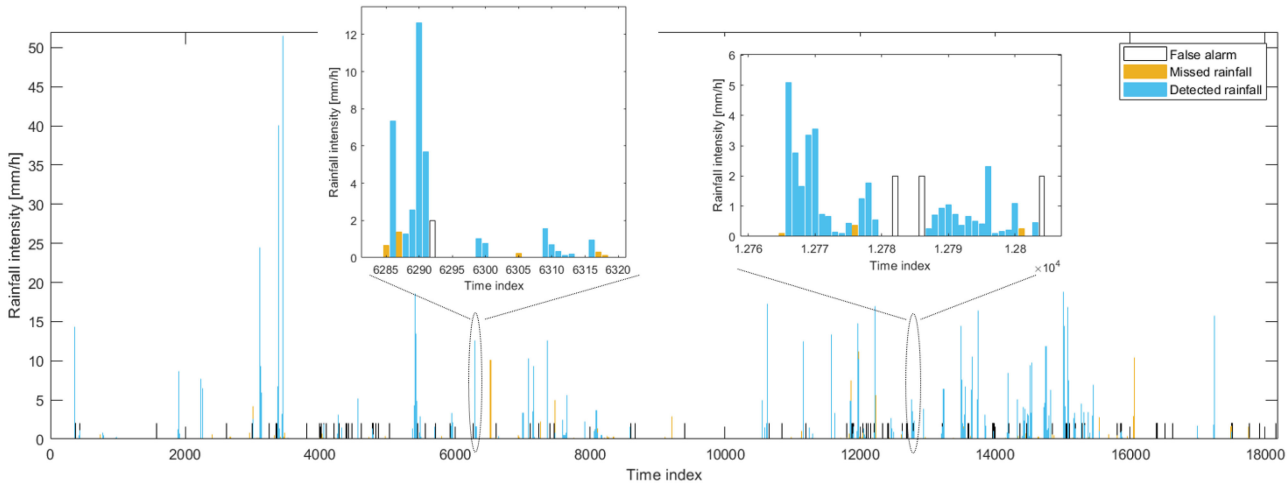


Fig. 10. Rainfall intensity during the 18 193 h of observation (one sample per hour; about 25 months of data collection), with indication of detected rainy samples (620 h), missed rainy samples (256 h), and false alarm samples (172 h). The zoom panels show two examples of the occurrence of the three cases on a fine scale.

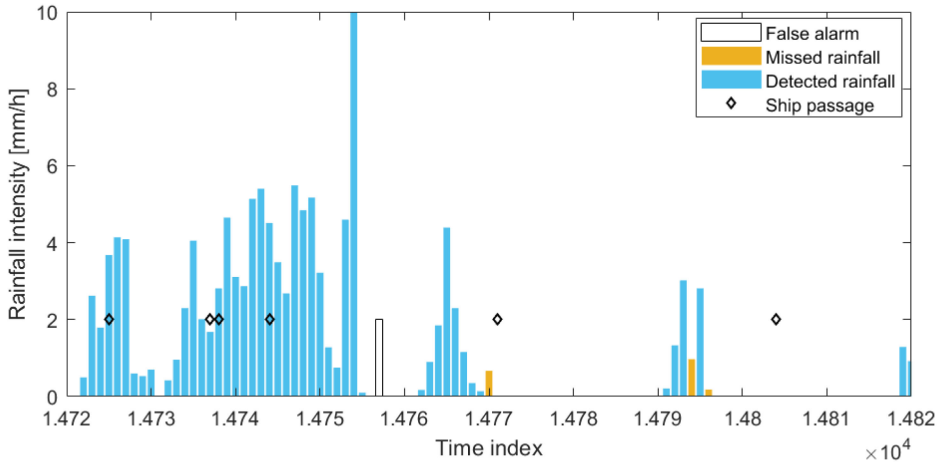


Fig. 11. Zoom of a portion of the 25-month span shown in Fig. 10, with the indication (black diamond) of the one-hour intervals at which a ship passage occurred.

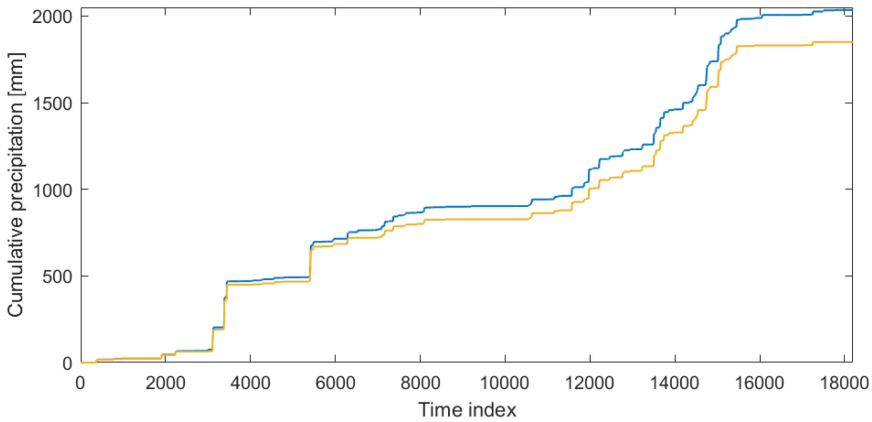


Fig. 12. Cumulative rainfall profiles over the 25-month span obtained by considering all rainfall events measured by the rain gauge (light blue line) or those detected by the underwater system (golden line).

environmental constraints do not allow the installation of rain gauges or radar systems. This is even more noteworthy in polar areas, where global warming is changing the hydrological cycle of those regions, thus increasing rainfall with respect to snow precipitation [31].

While the presence of high wind, especially above 10 m/s, induced a noticeable increase in the probability of false alarm, the performances did not undergo significant alterations in the hours in which a ship transited in the area where the underwater measurement device was placed. Similarly, no fluctuations in performance were observed on a seasonal basis, attributable to varying underwater propagation conditions. It is worth recalling that rainfall detection was based on the amount of precipitation accumulated over the course of an hour, and it is not possible to determine whether this amount is due to transient, intermittent, or continuous rain.

Although very promising, supervised learning models require a training phase that necessitates extensive collection of underwater acoustic spectra, accompanied by concomitant precipitation measurements to be used as ground truth. On the other hand, this is also partially necessary for rule-based algorithms that need specific calibrations to account for geographic location and hydrophone sensitivity. The possibility of using the trained detector in geographic areas other than the one in which the training data were collected is a topic for future investigation. However, it is reasonable to assume that in similar environmental settings, a trained detector can continue to operate successfully.

The performance obtained working on averaged spectra suggests that machine learning models may also be advantageous for rain detection using short-term acoustic spectra. This future research development is accompanied by a farther-reaching one: to design statistical learning models that act as regressors for accurate estimation of precipitation intensity and wind speed, making best use of information contained in multiyear time series of underwater acoustic noise.

ACKNOWLEDGMENT

This work is dedicated to the memory of Jeffrey A. Nystuen who started the pioneering research on the underwater sound produced by rain with Walter Munk at Scripps Institute of Oceanography/UCSD in the early 1980s and provided the authors with one of the devices he developed at the Applied Physical Laboratory, University of Washington, Seattle, WA to initiate underwater noise collection in the Ligurian Sea. The dataset and the Matlab scripts used in this work can be made available to researchers wishing to start scientific collaboration on these topics. Interested researchers are invited to send a request to sara.pensieri@cnr.it.

REFERENCES

[1] P. G. Black, J. R. Proni, J. C. Wilkerson, and C. E. Samsury, "Oceanic rainfall detection and classification in tropical and subtropical mesoscale convective systems using underwater acoustic methods," *Monthly Weather Rev.*, vol. 125, no. 9, pp. 2014–2042, 1997.

[2] S. C. Riser, J. Nystuen, and A. Rogers, "Monsoon effects in the Bay of Bengal inferred from profiling float-based measurements of wind speed and rainfall," *Limnol. Oceanogr.*, vol. 53, no. 5, pp. 2080–2093, 2008.

[3] J. Yang, S. C. Riser, J. A. Nystuen, W. E. Asher, and A. T. Jessup, "Regional rainfall measurement: Using passive aquatic listener during SPURS field campaign," *Oceanography*, vol. 28, no. 1, pp. 124–133, 2015.

[4] F. J. Wentz *et al.*, "Evaluating and extending the ocean wind climate data record," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2165–2185, May 2017.

[5] A. Y. Hou *et al.*, "The global precipitation measurement mission," *Bull. Amer. Meteorological Soc.*, vol. 95, no. 5, pp. 701–722, 2014.

[6] E. Amitai, J. A. Nystuen, L. Liao, R. Meneghini, and E. Morin, "Uniting space, ground, and underwater measurements for improved estimates of rain rate," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 2, pp. 35–38, Apr. 2004.

[7] B. B. Ma and J. A. Nystuen, "Passive acoustic detection and measurement of rainfall at sea," *J. Atmospheric Ocean. Technol.*, vol. 22, no. 8, pp. 1225–1248, 2005.

[8] E. Amitai, J. A. Nystuen, E. N. Anagnostou, and M. N. Anagnostou "Comparison of deep underwater measurements and radar observations of rainfall," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 406–410, Jul. 2007.

[9] S. Vagle, W. G. Large, and D. M. Farmer, "An evaluation of the WOTAN technique of inferring oceanic winds from underwater ambient sound," *J. Atmospheric Ocean. Technol.*, vol. 7, no. 4, pp. 576–595, 1990.

[10] J. A. Nystuen and H. D. Selsor, "Weather classification using passive acoustic drifters," *J. Atmospheric Ocean. Technol.*, vol. 14, no. 3, pp. 656–666, 1997.

[11] J. A. Nystuen, E. Amitai, E. N. Anagnostou, and M. N. Anagnostou, "Spatial averaging of oceanic rainfall variability using underwater sound: Ionian sea rainfall experiment 2004," *J. Acoust. Soc. Amer.*, vol. 123, no. 4, pp. 1952–1962, 2008.

[12] J. A. Nystuen, S. E. Moore, and P. J. Stabeno, "A sound budget for the southeastern Bering Sea: Measuring wind, rainfall, shipping, and other sources of underwater sound," *J. Acoust. Soc. Amer.*, vol. 128, no. 1, pp. 58–65, 2010.

[13] J. A. Nystuen, M. N. Anagnostou, E. N. Anagnostou, and A. Papadopoulos, "Monitoring Greek seas using passive underwater acoustics," *J. Atmospheric Ocean. Technol.*, vol. 32, no. 2, pp. 334–349, 2015.

[14] J. Kuhner, "Automating the detection of precipitation and wind characteristics in navy ocean acoustic data," in *Proc. MTS/IEEE OCEANS Conf.*, 2018, pp. 1–7.

[15] S. Pensieri, R. Bozzano, J. A. Nystuen, E. N. Anagnostou, M. N. Anagnostou, and R. Bechini, "Underwater acoustic measurements to estimate wind and rainfall in the Mediterranean sea," *Adv. Meteorol.*, vol. 2015, 2015, Art. no. 612512.

[16] S. C. Riser, J. Yang, and R. Drucker, "Observations of large-scale rainfall, wind, and sea surface salinity variability in the eastern tropical pacific," *Oceanography*, vol. 32, no. 2, pp. 42–49, 2019.

[17] D. Cazau, J. Bonnel, and M. Baumgartner, "Wind speed estimation using acoustic underwater glider in a near-shore marine environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2097–2106, Apr. 2019.

[18] G. D. Quartly, T. H. Guymer, K. G. Birch, J. Smithers, K. Goy, and I. Waddington, "Listening for rain: Theory and practice," in *Proc. 5th Eur. Conf. Underwater Acoust.*, 2000, pp. 1–6.

[19] M. N. Anagnostou, J. A. Nystuen, E. N. Anagnostou, E. I. Nikolopoulos, and E. Amitai, "Evaluation of underwater rainfall measurements during the Ionian Sea rainfall experiment," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2936–2946, Oct. 2008.

[20] W. O. Taylor, M. N. Anagnostou, and D. Cerrai, E. N. Anagnostou, "Machine learning methods to approximate rainfall and wind from acoustic underwater measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 2810–2821, Apr. 2021.

[21] J. A. Nystuen, "Quantifying physical processes in the marine environment using underwater sound," in *Proc. 4th Underwater Acoust. Meas. Conf.*, 2011, pp. 20–24.

[22] J. Yang, W. E. Asher, and S. C. Riser, "Rainfall measurements in the North Atlantic ocean using underwater ambient sound," in *Proc. IEEE/OES China Ocean Acoust.*, 2016, pp. 1–4.

[23] A. Trucco, "Detection of objects buried in the seafloor by a pattern-recognition approach," *IEEE J. Ocean. Eng.*, vol. 26, no. 4, pp. 769–782, Oct. 2001.

[24] C. Barngrover, R. Kastner, and S. Belongie, "Semisynthetic versus real-world sonar training data for the classification of mine-like objects," *IEEE J. Ocean. Eng.*, vol. 40, no. 1, pp. 48–56, Jan. 2015.

[25] N. H. Klausner and M. R. Azimi-Sadjadi, "Manifold-based classification of underwater unexploded ordnance in low-frequency sonar," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 1034–1044, Jul. 2020.

[26] E. Canepa, S. Pensieri, R. Bozzano, M. Faimali, P. Traverso, and L. Cavaleric, "The ODAS Italia 1 buoy: More than forty years of activity in the Ligurian sea," *Prog. Oceanogr.*, vol. 135, pp. 48–63, 2015.
[27] R. Bozzano *et al.*, "The M3A network of open ocean observatories in the Mediterranean sea," in *Proc. MTS/IEEE OCEANS Conf.*, Bergen, Norway, 2013, pp. 10–14.
[28] M. N. Anagnostou, J. A. Nystuen, E. N. Anagnostou, A. Papadopoulos, and V. Lykousis, "Passive aquatic listener (PAL): An adoptive underwater acoustic recording system for the marine environment," *Nucl. Instrum. Methods Phys. Res. Sect. A, Accelerators Spectrometers Detectors Assoc. Equip.*, vol. 626, pp. S94–S98, 2011.
[29] S. Pensieri, R. Bozzano, M. N. Anagnostou, E. N. Anagnostou, R. Bechini, and J. A. Nystuen, "Monitoring the oceanic environment through passive underwater acoustics," in *Proc. MTS/IEEE OCEANS Conf.*, Bergen, Norway 2013, pp. 1–10.
[30] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Berlin, Germany: Springer, 2009.
[31] R. Bintanja, "The impact of arctic warming on increased rainfall," *Sci. Rep.*, vol. 8, no. 1, pp. 1–6, 2018.

Andrea Trucco (Senior Member, IEEE) was born in Genoa, Italy, on February 10, 1970. He received the M.Sc. and Ph.D. degrees in electronic and computer engineering from the University of Genoa, Genoa, Italy, in 1994 and 1998, respectively.

Since 1999, he has been with the University of Genoa, where, in 2016, he was appointed a Professor in telecommunications. He has been the Vice-Rector for International Relations (2015–2020) and the Vice-Director of the Department of Biophysical and Electronic Engineer (2008–2012). He was also a Senior Researcher with the Italian Institute of Technology (2011–2015). He has been the Principal Investigator in many research projects, funded by the EU, national agencies, and industrial companies. His fields of interest include acoustics, signal processing, pattern recognition, sonar systems, medical ultrasound, fault diagnosis. He is an author of more than 200 scientific papers (90 in international journals) and inventor in seven international patents.

Prof. Trucco is an Associate Editor for IEEE JOURNAL OF OCEANIC ENGINEERING and IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL. He received the student paper awards at competitions organized by the 9th International Symposium Unmanned Untethered Submersible Technology (1995) and the MTS/IEEE Oceans'97 International Conference (1997).

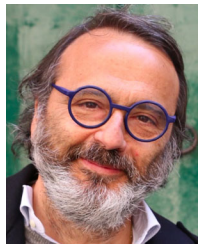
Roberto Bozzano received the M.Sc. degree in electronic engineering from the University of Genoa, Genoa, Italy, in July 1993.

Since 1996, he has been working with the National Research Council of Italy. His theoretical and experimental research activity concerns the field of complex and multisensory technologies for environmental monitoring with particular reference to the marine-maritime sector. Specifically, the activity concerns the design of complex multifunctional measurement systems and their remote management, the development of acoustic methodologies for the measurement of environmental parameters, the study of the physical processes of energy exchanges at the interface-sea, and, in general, the development of systems for the acquisition and processing of signals on real-time embedded platforms and architectures. He is the principal investigator of the offshore observatory W1M3A, part of ERIC-EMSO and ERIC-ICOS. He is an active member of regional and global alliances such as MonGOOS and OceanSITES.

Emanuele Fava was born in Novi Ligure, Italy, on July 9, 1997. He received the high school degree from Liceo Scientifico Tradizionale Amaldi, Novi Ligure, in 2016. He studied Ingegneria Elettronica e Tecnologie dell'Informazione at the University of Genoa, Genoa, Italy, from 2016 to 2019 and received the Bachelor's degree with a score of 110/110 con Lode. Since 2019, he has been working toward the Master's degree in internet and multimedia engineering with the University of Genoa.

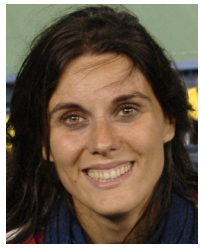
Sara Pensieri received the M.Sc. degree in telecommunication engineering and the Ph.D. degree in electronic and computer engineering, robotics, and telecommunications from the University of Genoa, Italy, in 2005 and 2013, respectively.

Since 2005, she has been a Researcher with the National Research Council of Italy in the Institute for the study of the anthropic impacts and the sustainability of the marine environment in Genoa. She is involved in the scientific and technological development of the W1M3A observatory. Her expertise spans from operational oceanography to the underwater acoustics and the analysis of physical processes at the air-sea interface. She was involved in several European Projects (MERSEA, EuroSITES, MyOCEAN, PERSEUS, FixO³, MINKE) aiming at the development of innovative methodologies for the marine environmental monitoring from unmanned platforms. She has experience in data-management and in the design of field experiment at sea during oceanographic cruises.



Alessandro Verri was born in Genoa, Italy, in 1960. He received the Ph.D. degree in physics from the University of Genoa, Genoa, Italy, in 1989.

He has been a Professor of computer science with the University of Genoa, since 2000. He visited MIT multiple times, ICSI, Heriot Watt University, and INRIA-IRISA. He published more than 150 papers in peer-reviewed journals and proceedings of international conferences. His research interest ranges from learning theory and algorithms to applications in the domain of computer vision and biomedical data analysis. He supervised more than 20 Ph.D. students and 120 M.Sc. students.



Annalisa Barla received the M.Sc. degree in physics and the Ph.D. degree in computer science from the University of Genoa, Genoa, Italy, in 2001 and 2005, respectively, working on kernel functions engineering for regularization methods in machine learning applied to image content understanding.

She is an Associate Professor of computer science with the University of Genoa. Her main areas of interest are in the field of machine learning spanning from the study of robust and reproducible variable selection methods with sparse regularization to the understanding and visualization of complex structured network data. She is an author of more than 60 peer-reviewed journal and conference papers.