IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 22, 2021, accepted February 3, 2021, date of publication February 11, 2021, date of current version February 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058744

Metadata-Driven Universal Real-Time Ocean Sound Measurement Architecture

ENOC MARTÍNEZ[1], ALBERT GARCÍA-BENADÍ[1], DANIEL M. TOMA[1], ERIC DELORY[2], (Senior Member, IEEE), SPARTACUS GOMÁRIZ[1], (Member, IEEE), AND JOAQUÍN DEL-RÍO[1], (Senior Member, IEEE)

[1]SARTI-MAR Research Group, Electronics Department, Universitat Politècnica de Catalunya, 08800 Vilanova i la Geltrú, Spain
[2]Oceanic Platform of the Canary Islands, 35200 Telde, Spain

Corresponding author: Enoc Martínez (enoc.martinez@upc.edu)

This work was supported in part by the Spanish Government Project RESBIO under Grant TEC2017-87861-R, in part by the EU projects ENVRI-FAIR under Grant 824068, in part by the JONAS INTERREG Atlantic Area-funded project under Grant EAPA_52/2018, in part by the Generalitat de Catalunya SARTI-MAR Project under Grant 2017 SGR 371, and in part by the Scholarship Program under Grant FPI-UPC_2015.

ABSTRACT Underwater sound in the oceans has been significantly rising in the past decades due to an increase in human activities, adversely affecting the marine environment. In order to assess and limit the impact of underwater noise, the European Commission's Marine Strategy Framework Directive (MSFD) included the long-term monitoring of low-frequency underwater sound as a relevant indicator to achieve a good environmental status. There is a wide range of commercial hydrophones and observing platforms able to perform such measurements. However, heterogeneity and lack of standardization in both hydrophones and observing platforms makes the integration and data management tasks time-consuming and error-prone. Moreover, their power and communications constraints need to be addressed to make them suitable for long-term ocean sound monitoring. Measured underwater sound levels are challenging to compare because different measurement methodologies are used, leading to a risk of misunderstandings and data misinterpretation. Furthermore, the exact methodology applied is not always public or accessible, significantly reducing ocean sound data re-usability. Within this work, a universal architecture for ocean sound measurement is presented, addressing hydrophone integration, real-time *in situ* processing and data management challenges. Emphasis is placed on generic and re-usable components, so it can be seamlessly replicated and deployed in new scenarios regardless of the underlying hardware and software constraints (hydrophone model, observing platform, operating system, etc.). Within the proposed architecture, a generic implementation of an underwater sound algorithm based on underwater noise measurement best practices is provided. Standardized and coherent metadata with emphasis on strong semantics is discussed, providing the building blocks for FAIR (Findable, Accessible, Interoperable, Reusable) ocean sound data management.

INDEX TERMS Ocean sound, underwater acoustics, sensor web enablement, interoperability, real-time systems, data acquisition.

I. INTRODUCTION

There has always been underwater background sound in the oceans due to natural factors. However, human activities such as shipping, construction, sonar and seismic exploration have been raising the underwater ambient noise to unprecedented levels in the past decades [1]. Ocean sound is an important environmental factor for many species, especially to those using underwater sound for localization and communication [2]. Thus, the introduction of energy into the marine habitat in the form of acoustic noise needs to be properly monitored and studied to minimize its harmful impact on the ecosystem.

In order to achieve a good environmental status in European waters, the European Parliament approved the Marine Strategy Framework Directive (MSFD) which aims to protect the marine environment and ecosystem. This directive includes a set of indicators to measure the status of European waters, including maximum ocean underwater sound levels considered as acceptable (MSFD indicator 11.2.1).

The associate editor coordinating the review of this manuscript and approving it for publication was Yougan Chen.

This indicator requires the long-term measurement of Sound Pressure Level (SPL) at the 1/3 octave bands centered at 63 and 125 Hz. It has also been suggested to extend the monitored bands from 10 Hz up to 20 kHz in order to achieve a more accurate assessment of underwater noise [3].

The term ''underwater noise'' usually refers to the anthropogenic sound that has the potential to cause negative impacts on marine life, while ''underwater ambient noise'' usually refers to the background sound with no distinguishable sources [4]. Within this work, the generic term ''ocean sound'' is used to refer to the overall underwater sound level, regardless of its source or its impact on marine life.

A. OCEAN SOUND MEASUREMENT SYSTEMS

As concern about underwater noise increased, numerous studies have been presented analyzing ocean sound [5]–[8]. Commercial off-the-shelf hydrophones used in ocean sound monitoring usually provide raw acoustic data, whether as acoustic recordings or in streaming mode. Thus, acoustic data has to be post-processed to obtain meaningful underwater sound levels. Due to the high sampling rate used by acquisition or recording systems (usually from tens to hundreds of kHz), streaming raw acoustic data from an observing platform to a shore station is only possible with broadband communications. Cabled observatories equipped with hydrophones are an excellent observing platform for long-term ocean sound monitoring, since they do not have constraints regarding bandwidth and power [9], [10]. However, these infrastructures are scarce and costly to maintain.

Some surface buoys equipped with broadband radio communications are also capable of streaming acoustic data in real-time to shore station for real-time processing [11]. However, the autonomy of these systems is greatly constrained by their power availability. These buoys also need to be located close to shore, since broadband radio links have limited coverage range, satellite transmission being currently cost-prohibitive.

Moored autonomous recorders composed of one or several hydrophones and a recording unit have also been used in large-scale deployments for underwater noise assessment [6], [8], [12]. The autonomy of such devices is mainly reduced by two factors: power and storage capacity. The storage of acoustic data may require gigabytes per day depending on the sampling rate. Usually data acquired by these devices are only available after recovery, so they do not provide real-time capability.

Underwater gliders have also been used to sense the underwater soundscape [13], [14]. Although underwater gliders have intermittent communication link to shore stations during surface time, their satellite communications have very limited bandwidth, not suitable for raw acoustic data transmission.

Telemetry is usually one of the more power-demanding components of an autonomous system such as underwater gliders. Thus, reducing the data transmission (and its associated power consumption) is vital to extend their autonomy. If the acoustic data could be processed *in situ*, the amount of information to be transmitted would decrease by several orders of magnitude, allowing real-time measurements from autonomous platforms with satellite telemetry, such as gliders, profilers, moored buoys, etc.

However, processing acoustic data in real-time is not a trivial task due to the required computational power and intrinsic complexity of acoustic signals. Although there are hydrophone prototypes with embedded underwater sound level algorithms, they are not yet widely used and their embedded algorithms still have room for improvement [15].

B. STANDARDIZATION AND INTEROPERABILITY

Scientific instruments, including hydrophones, tend to use proprietary and non-standardized protocols. Thus, in order to integrate these instruments to observing platforms, *ad hoc* drivers are usually developed. Furthermore, observing platforms have a broad range of software and hardware architectures (operating systems, communications link, computational/power constraints), so these *ad hoc* drivers need to be developed for each instrument-platform combination. This lack of code re-usability is costly, time consuming and requires in-depth knowledge of both instrument and observing platform functionalities [16].

There have been several attempts to facilitate instrument integration by standardizing instruments protocols, such as IEEE 1451 [17]. However, instrument manufacturers did not embrace this approach and still use proprietary interfaces. Other approaches try to describe instrument interfaces using machine-understandable descriptions [16]. Using this approach, interoperability can be achieved even if manufacturers do not adopt a common standard.

Beyond sensor integration, data management and interoperability also prove challenging for the ocean observing community. Data and metadata formats and procedures usually vary significantly across domains and institutions, leading to information silos and preventing the data to be effectively shared across different scientific communities. This lack of standardization has been the driving force of the Sensor Web Enablement (SWE) set of standards, supported by the Open Geospatial Consortium (OGC) [18]. SWE provides a standard framework for data and metadata ingestion, archival and retrieval with a strong focus on robust semantics and machine-to-machine interactions. However, the integration of sensor data to the SWE framework is not trivial and several approaches have been proposed [19], [20], [21].

Within this paper, a universal architecture for *in situ*, real-time ocean sound monitoring is proposed, compliant with the needs of the ocean observing community (i.e. MSFD indicators) and following best practices on underwater sound measurement methodologies [22], [23]. This architecture addresses interoperability issues at both sensor integration and data management levels by proposing a solution based on the SWE framework. Emphasis is placed on generic and re-usable components, so the proposed architecture can be replicated and deployed in new scenarios regardless of the underlying hardware and software constraints (hydrophone

model, observing platform, operating system, etc.). This work relies on the Sensor Model Language (SensorML) and Observations and Measurements (O&M) for coherent and semantically tied data and metadata [24], [25]. The ultimate goal of the proposed architecture is to provide the building blocks for ocean sound data management following the FAIR principles (Findable, Accessible, Interoperable and Reusable) [26], [27].

The paper is structured as follows. In section II an ocean sound measurement architecture is presented. Section III discusses an underwater sound level algorithm and its implementation. In section IV the proposed interoperability and metadata solution is discussed. In section V three use cases covering different scenarios are presented. The paper closes with conclusions and an outlook to future work.

II. UNIVERSAL OCEAN SOUND MEASUREMENT ARCHITECTURE

Metadata plays an important role during the data life-cycle, providing vital context to measurements: what was measured, where it was measured, who led the deployment, etc. However, metadata is usually compiled after data is acquired and targets only the measurements themselves, providing little information of the overall acquisition chain.

Within this work, a step forward is taken, using the metadata not only to provide contextual information, but also to configure a generic data acquisition chain. In other words, the data acquisition, processing and storage processes are driven by the sensor's metadata. So, metadata is not added to the acquired data once gathered, but prepared beforehand and controls the acquisition process. Using this approach, metadata may not only reflect what has been measured, but also unambiguously define the whole acquisition chain, including sensor setup, signal processing, formatting, etc. Thus, the acquisition chain in a deployment can be easily replicated based on its metadata.

Ocean sound is a non-trivial variable which requires complex signal processing. Slight differences in the acquisition and signal processing may lead to significant differences in the result. Thus, it is important to state the exact method applied. Within this work the concept of a metadata-driven measurement architecture is applied to ocean sound. Using this approach, metadata contains unambiguous information about the whole acquisition chain, from the instrument's low-level configuration to the details of signal processing.

The dataflow of this metadata-driven architecture, depicted in Fig. 1, starts with a generic hydrophone acquiring raw acoustic data. Since all its low-level specifications will be abstracted, almost any hydrophone may be used. The next components in the architecture are the SWE Bridge and the hydrophone's metadata description in SensorML format [24]. The SWE Bridge is an open source, standards-based universal driver [19]. One of its key features is its ability to interface scientific instruments without any previous knowledge using a SensorML description, regardless of its vendor-specific protocols. It can manage sensor communications in almost any format, ranging from plain ASCII communications through serial port to Ethernet high frequency binary streams (e.g. hydroacoustic data).

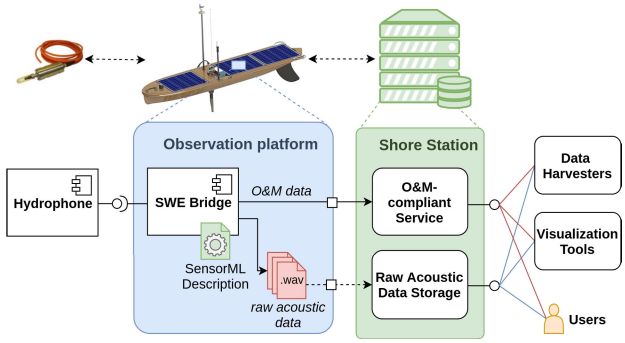

FIGURE 1. Proposed architecture for ocean sound monitoring. Hydrophone raw data is processed *in situ* by the SWE Bridge according to its SensorML metadata file. Then, two streams of data are sent to the shore station: underwater sound levels (sent in real-time) and acoustic recordings (real-time or delayed, depending on the telemetry used). At the shore station data is made available to users, data harvesters, visualization tools and others using standardized interfaces.

The SWE Bridge has been designed bearing in mind the heterogeneity of observing platforms and scientific instruments. Its re-usable and generic design facilitates the deployment in any platform. Considering power and communications constraints of observing platforms, stress has been put on an efficient and modular implementation using very limited computational resources and optimized telemetry. Its modularity and versatility have been proved in numerous deployments using various observing platforms, such as unmanned surface vehicles, underwater gliders, cabled observatories and autonomous buoys, among others.

The SWE Bridge includes an acoustics module that provides real-time underwater sound levels following the recommendations of the MSFD Task Group 11 on underwater noise [3]. Sound Pressure Level (SPL), Root Mean Squared pressure (RMS) and Sound Exposure Level (SEL) at 1/3 octave bands are calculated. Moreover, the acoustics module is not limited to the MSFD requirements and can be easily configured to monitor any band within the hydrophone's frequency range. The internal processing is discussed in detail in section III.

The metadata encoded in the SensorML description contains extensive information about the hydrophone's characteristics, deployment details, communication protocol, setup routines and more. The SWE Bridge is also configured through this metadata file, selecting different processing and formatting options. Thus, the SensorML file does not only contain hydrophone's metadata, it contains all the required information to replicate the deployment: from sensor configuration to underwater sound level algorithm details.

Since the SWE Bridge can be deployed within any observation platform, the hydrophone's data stream is processed *in situ* and in real-time. Ocean sound levels are usually required for time windows of tens of seconds, so the processed data stream requires several orders of magnitude less

storage and bandwidth than the raw data. The main advantage of processing the data on-board of the observation platform is that processed data may be transmitted in real-time, even if the bandwidth is limited (e.g. satellite communications). On the contrary, raw acoustic data (acoustic recordings) has be stored internally until recovery (acoustic recorder, underwater glider, etc.), unless broadband communications are available (e.g. cabled observatory).

The SWE Bridge generates processed data following the O&M standard. Thus, underwater sound levels may be injected in real-time into an O&M-compliant service such as Sensor Observation Service (SOS) or SensorThings [28], [29]. Although the preferred format for underwater sound levels is the metadata-enriched O&M format, CSV (comma separated values) output may also be provided by the SWE Bridge. Due to its large volume and their specific nature, acoustic recordings can be archived in a generic data storage service such as File Transfer Protocol (FTP) or ERDDAP [30].

Once the data are injected to a cyber-infrastructure, data and metadata are managed and disseminated following the FAIR principles. Further services can query and access both data and metadata, including users, data visualization tools and data harvesters (e.g. data assembly centers and portals). At this stage, the critical importance of metadata arises, since users and machines interacting with the cyber-infrastructure may not know beforehand the acquisition and processing details. Robust semantics and standardized metadata formats facilitate the interpretation and contextualization of the acquired data to both human and machines.

III. UNDERWATER SOUND LEVEL ALGORITHM

Measured Underwater sound levels are sometimes difficult to compare because different measurement methodologies or acoustic metrics are used, leading to a risk of misunderstandings between scientists from different disciplines [23]. The goal of the presented architecture is to obtain underwater sound measurements using appropriate metrics, following best practices on the field and community-accepted procedures [3], [22], [23].

This work focuses on providing an implementation of ocean sound algorithms that can be seamlessly integrated into resource-constrained platforms to provide real-time, *in situ* measurements. One of the critical aspects for these platforms is the computational cost required to apply underwater sound level algorithms in real-time, while maintaining a reasonable frequency resolution (Δf). In order to select the algorithm's implementation a computational cost analysis of different techniques was performed, focusing on execution time and memory usage (see appendix A).

Ocean sound is usually measured using the Sound Pressure Levels (SPL) over a time window T at different 1/3 octave band frequencies [23]. Both the time window and the frequency bands may be adjusted by the operator depending on the deployment. Finally, the algorithm implementation within the SWE Bridge is discussed.

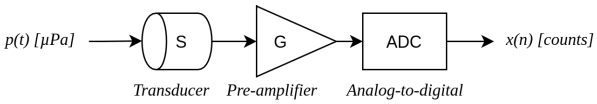

FIGURE 2. Generic hydrophone acquisition chain. Underwater pressure is converted to an analog voltage by the transducer. Its output is amplified by the preamplifier and finally it is converted to a discrete signal using an analog-to-digital converter.

A. SIGNAL CONDITIONING STAGE CHARACTERIZATION

In order to calculate SPL levels, the first step is to obtain the pressure signal. Acquisition systems or digital hydrophones usually do not provide the pressure signal, but the dimensionless output from the analog-to-digital converter (ADC). Thus, the signal conditioning stage has to be characterized. The typical signal conditioning stage of a hydrophone, depicted in Fig. 2, contains a transducer, a pre-amplifier and an ADC converter. The following signal conditioning properties need to be known:

- S: transducer sensitivity (dB re 1 V/μPa)
- G: pre-amplifier gain (dB)
- V_{ref}: ADC's reference voltage (V)
- M: ADC's resolution in bits (dimensionless)
- f_s: Sample Rate (Hz)

Once these parameters are known, i.e. encoded in the SensorML file, the instantaneous pressure (also named pressure signal) can be calculated from the raw dimensionless samples with (1). Note that S and G have been converted from dB to linear, S_{lin} and G_{lin}.

$$p(n) = \frac{x(n)}{S_{lin} G_{lin} \frac{1}{LSB}} [\mu Pa] \tag{1}$$

where $x(n)$ is the raw digital counts provided by the ADC and LSB^{-1} (least significant bit) is the ADC conversion coefficient in $counts/V$.

The LSB can be calculated from the ADC's specifications using (2):

$$LSB = \frac{V_{fs}}{2^M} = \frac{2V_{ref}}{2^M} \; [V/counts] \tag{2}$$

where M is the ADC's number of bits, V_{fs} is the ADC's full-scale voltage range and V_{ref} is ADC's reference voltage. In (2), it has been assumed that the hydrophone's ADC has a symmetric reference voltage ($V_{ref+} = -V_{ref-}$).

Some hydrophones already provide the total sensitivity S_{total} of the hydrophone considering the amplifier gain and the ADC's conversion coefficient. In this case the sensitivity is given in dB re count/μPa and the raw dimensionless samples may be directly converted to pressure using (3) (note that total sensitivity has been converted to linear units, S_{total}).

$$p(n) = \frac{x(n)}{S_{total}} [\mu Pa] \tag{3}$$

Once the input signal has been converted from dimensionless samples to a discrete pressure signal $p(n)$ in μPa,

the Sound Pressure Level (SPL) over a time interval T can be calculated applying (4) [23].

$$SPL_N = 10\log_{10}\left(\frac{1}{N}\sum_N \frac{p(n)^2}{p_0^2}\right)[dB\ re\ \mu Pa] \quad (4)$$

where N is the number of samples within the time window ($N = T \cdot f_s$) in the pressure signal's segment and the reference pressure in seawater is $p_0 = 1\ \mu$Pa.

In (4) the whole bandwidth of the pressure signal $p(n)$ is considered. However, usually it is not desirable to provide the overall SPL value since it is highly dependent on the hydrophone's bandwidth. In order to avoid the dependency on the sampling rate, SPL values are calculated over a frequency band, generally 1/3 octave bands as required by the MSFD indicators.

B. 1/3 OCTAVE BAND FREQUENCIES

The calculation of SPL values over a specific frequency band can be achieved by different means. Within this work the Fast Fourier Transform (FFT) is used to estimate the SPL over a set of frequency bands. Other methods such as a filter bank and Goertzel's algorithm were analyzed to compute SPL levels, but the FFT proved to be the most robust and computationally efficient (see annex A).

In order to calculate the SPL over a frequency band, the first step is to estimate the power spectral density (PSD) of the pressure signal by means of its periodogram:

$$P_{xx}(f) = \frac{1}{N}|P(f)|^2\ [\mu Pa^2/Hz] \quad (5)$$

where P_{xx} is the periodogram over a time interval $T = N/fs$, $P(f)$ is the Discrete Fourier Transform (DFT) of the discrete pressure signal $p(n)$ and f is the normalized frequency [31].

The DFT assumes that a signal is stationary and periodic, which is not the case in acoustics signals. This causes the DFT to ''see'' discontinuities around the edges of the time window (segment of length N), which leads to spectral leakage. The spectral leakage is the spread of power from one DFT bin to its adjacent bins. In order to reduce the spectral leakage, it is a common practice to apply a smoothing window function to the signal before calculating its DFT, as shown in (6) [32]:

$$p_w(n) = w(n) \cdot p(n) \quad (6)$$

where $w(n)$ is a window function of length N samples and $p_w(n)$ is the resulting windowed pressure signal. Applying a window reduces the spectral leakage, but it has severe implications on the resulting spectrum [32]. The resulting $p_w(n)$ signal amplitude is reduced around the edges of the time window, minimizing the discontinuities. However, since the amplitude is reduced, its overall energy diminishes. In order to correct this loss of amplitude the result of the DFT has to be scaled with the coherent gain (CG). This factor can be calculated as the sum of the window components (W) normalized by the number of samples (N), as showed in (7) and (8).

$$W = \sum_N w(n) \quad (7)$$

$$CG = \frac{W}{N} \quad (8)$$

From the spectral point of view, the amount of energy within each DFT bin is also modified by the window function. Each DFT bin can be understood as a very narrow band-pass filter with a Δf bandwidth. However, when a window is applied, the bandwidth of this hypothetical filter is slightly increased due to the aperture of the window's spectral response main lobe [33].To correct the extra power contribution in each DFT bin due to the bandwidth increment, the normalized equivalent noise bandwidth (*nenbw*) correction factor is used. The *nenbw* factor can be calculated using (9) [33].

$$nenbw = N\frac{\sum_N w(n)^2}{W^2} \quad (9)$$

In order to compensate the mentioned side-effects of windowing, the periodogram of a windowed signal (also known as modified periodogram) can be calculated using (10) [31].

$$P'_{xx}(f) = \frac{1}{N \cdot nenbw}\left|\frac{P_w(f)}{CG}\right|^2 = \frac{|P_w(f)|^2}{\sum_N w(n)^2} \quad (10)$$

Being $P_w(f)$ the DFT transform of the windowed pressure signal $p_w(n)$. Note that *nenbw* and CG corrections assume that the input signal is stationary and has a flat spectral response, i.e. white Gaussian noise. In real-world acoustic signals this is rarely the case, so a small error due to windowing side-effects is expected.

One of the weak points of the periodogram as a PSD estimator is its high variance. In order to reduce the variance of each estimate, several periodograms can be averaged using the Bartlett's or Welch's methods. However, when averaging periodograms there is a trade-off between the frequency resolution Δf and the variance reduction [34]. On the other hand, SPLs are usually calculated over large periods of time (usually tens of seconds), so averaging can help to reduce the computational costs in terms of memory and number of operations [35].

The Bartlett's method slices the signal with M samples into K non-overlapping, sequential segments ($M = K \cdot N$). Then it calculates the periodogram for each segment and averages the results (Fig. 3, top) [34]. This approach does not require CG nor *nenbw* corrections, but the effect of the spectral leakage is greater.

The Welch method is similar, but the slices are windowed and overlapped (Fig. 3, bottom) [34], [35]. This method has significantly less spectral leakage, but uses the CG and *nenbw* corrections, which may also induce errors due to the non-stationary and non-gaussian nature of real-world acoustic signals. Additionally, an increase of computational cost due to the FFT overlapping is expected.

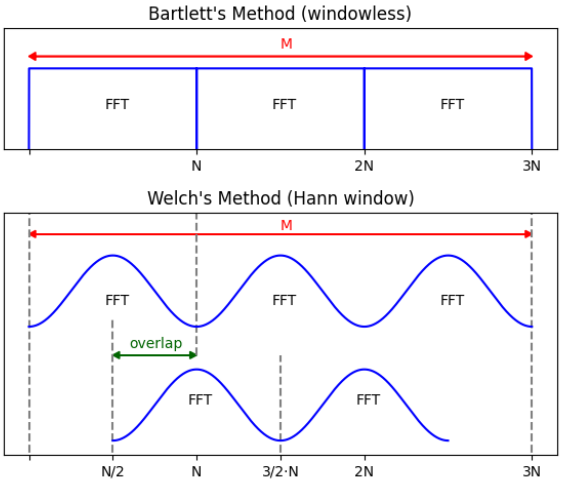

FIGURE 3. Welch's and Bartlett's methods for power spectral density estimation. The Bartlett's method uses sequential, non-overlapped, windowless data segments. The Welch method uses overlapped windowed data segments, in this example it uses a Hann window and a 50% overlap.

Then, the power spectral density can be estimated using the Welch's method or the Bartlett's method by averaging K periodograms:

$$\overline{P}_{xx}(f) = \frac{1}{K} \sum_{i=0}^{i=K-1} P'_{xx_i}(f) \tag{11}$$

where $\overline{P}_{xx}(f)$ is the power spectral density estimation, K is the number of periodograms being averaged and $P'_{xx_i}(f)$ is the periodogram of the ith data segment. The number of segments to be averaged depends on the length of each data segment N and on the overlapping [35].

Once the power spectral density has been estimated, the SPL value within a frequency band over a time interval $T' = M/f_s$ can be calculated with (12).

$$SPL_{f,N} = 10 \log_{10} \left(\frac{2}{N} \sum_{f_l}^{f_h} \overline{P}_{xx}(f) \right) [dB\ re\ \mu Pa] \tag{12}$$

Being f_l and f_h the low / high limit frequencies of the desired third-octave band and N is the number of samples within each FFT segment. Since the PSD of a real signal is symmetric, the negative frequencies are redundant and can be omitted [33]. However, to consider their energy contribution, a factor of 2 has to be applied to the sum of the positive frequency bins. Note that the result of (12) depends on multiple parameters, including the total number of samples M, the number of samples in each data segment N, the overlapping between data segments and the window function.

In (12) several adjacent bins are added to approximate a third-octave band. However, there is only a finite number of frequency bins and f_l and f_h may not coincide with them. For instance, the pressure signal's power spectrum calculated over a time period of 1 second has a bin width of 1 Hz. If the SPL in the third-octave band centered at 63 Hz is calculated, the edge frequencies are 56.6 Hz and 70.8 Hz, which are not aligned with the frequency bins. In order to correct this misalignment a bandwidth correction needs to be applied, as shown in (13) [22].

$$SPL'_{f,N} = SPL_{f,N} - 10 \log_{10} \frac{BW_{real}}{BW_{ideal}} \tag{13}$$

where $SPL'_{f,N}$ is the sound pressure level value with the bandwidth correction, BW_{ideal} is the ideal bandwidth of the third octave band (f_h - f_l) and BW_{real} is the real bandwidth summed in (12).

C. ALGORITHM IMPLEMENTATION

The algorithm previously presented has been integrated within the SWE Bridge software to provide a universal ocean sound monitoring tool.

Usually SPL values are computed in time windows of tens of seconds, e.g. 20 seconds [3]. Within the SWE Bridge this can be adjusted by the user to match different deployment scenarios.

There is strong relationship with the number of points N and the computational cost of each FFT. As a compromise between frequency resolution and computational cost, the SWE Bridge performs by default an FFT for each second of data, achieving a resolution of $\Delta f = 1$ Hz, which is sufficient for ocean sound analysis [36]. However, the frequency resolution may be adjusted by the user. The use of segments which are not power of 2 may slightly increment the computational cost, but a well-known frequency resolution is achieved in exchange. The time window set by the user is used to control the number of periodograms being averaged in order to reduce their variance. So, the user can select both frequency resolution and time window.

Both averaging methods are implemented within the SWE Bridge, the windowless Bartlett's method and the Welch method, using a Hann window with 50% overlap. Both approaches are depicted in Fig. 3.

1) SOUND EXPOSURE LEVEL AND ROOT MEAN SQUARE

Although SPL is the main parameter when measuring ocean sound, some other parameters may also be of interest, such as the Sound Exposure Level (SEL) and the Root Mean Squared (RMS) pressure. Since both parameters use similar calculations as the SPL, they are also implemented within the SWE Bridge to provide extra information at very little computational cost.

SEL is a measure of the integral of the square of the sound pressure over a stated time interval or event expressed in decibels [23]. This parameter is closely linked with the SPL value, but making the time interval explicit. It can be easily derived from an SPL value using (14) [22].

$$SEL_T = SPL'_{f,N} + 10 \log_{10} \left(\frac{T'}{T_0} \right) [dB\ 1\ \mu Pa^2 s] \tag{14}$$

Being T' the time window over which the SPL value was calculated (proportional to M) and $T_0 = 1$ second is the time reference.

The P_N is the RMS value of the pressure signal over a segment of length N, calculated using (15):

$$P_N = \sqrt{\frac{1}{N}\sum_{N} p(n)^2}[Pa] \tag{15}$$

In order to maintain the same time window as SPL and SEL values, several RMS values are averaged, as stated in (16).

$$\overline{P}_{N,M} = \frac{1}{M}\sum_{M} P_N[Pa] \tag{16}$$

Being $\overline{P}_{N,M}$ the mean of M RMS values. The SWE Bridge provides the averaged mean of several RMS values for two reasons: to maintain the same time window used in SPL and SEL calculations, and to prevent the use of very large buffer which may result in memory overflow and excessive computational cost.

2) ACQUISITION AND TIMESTAMPING

In streaming mode, hydrophones usually send pressure samples grouped in packets. Within the SWE Bridge, the acquisition of incoming packets is performed through a configurable circular buffer. By default, the SWE Bridge defines a safe size for the buffer, but the user may adjust this parameter through the SensorML description to optimize the system's memory usage. A large circular buffer will waste a lot of memory, while a small buffer may result in overflows and loss of data.

Some hydrophones have an accurate internal real-time clock (RTC), providing precise timestamp information within the data stream's header, e.g. NeXOS A2 (see section V-C). The SWE Bridge can identify timestamps in the stream header and use them to maintain the time base. However, if a hydrophone does not provide timing information (e,g, NAXYS Hydrophone, see section V-A), incoming packets are timestamped upon arrival based on the platform's system clock. Thus, within the SWE Bridge the time base is maintained regardless of the hydrophone stream.

3) MISSING PACKETS

Most digital hydrophones stream their samples grouped in packets using the User Datagram Protocol (UDP), which does not guarantee that all packets will be received. Data loss (packets not delivered) is intrinsic to UDP communications and cannot be avoided. Most digital hydrophones include a packet counter in their header, so the acquisition software can detect missing packets.

Since the time base has to be maintained, these missing packets have to be filled with null values, e.g. a packet with all values set to zero. However, these null values may induce frequency glitches and digital broadband distortion. In order to avoid these glitches, the SWE Bridge periodically calculates the mean value of the pressure signal during the last second. When a missing packet is detected, instead of filling the signal with zeros, the mean value is used. The spectral implications of these approach are shown in Fig. 4. This is critical to avoid erroneous data in hydrophones that have an offset in their signal or are measuring very low frequency signals.

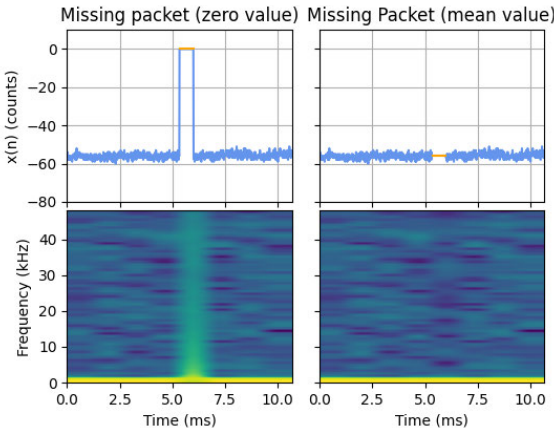

FIGURE 4. Spectral glitches due to missing packets. Left graphs show the pressure signal filled with zeros when a missing packet is detected (top) and its spectrogram (bottom), showing an artificial broadband glitch. Right graphs show the pressure signal filled with the mean value when a missing packet is received (top) and its spectrogram (bottom). It can be seen that the spectral glitch practically disappears when the mean value is used instead of zeros.

IV. METADATA AND INTEROPERABILITY

Interoperability can be defined as the ability of two systems to exchange information and use the information that has been exchanged. It can be separated in two different layers: syntactic and semantic. The syntactic interoperability is the ability of two systems to understand their formats and interfaces, allowing the flow of information, while semantic interoperability focuses on providing unambiguous meaning to the information that has been exchanged.

Within the proposed architecture, metadata plays a vital role, giving contextual information about what is being measured, when and how. It also describes the communication's interface, protocol and signal conditioning characteristics. In other words, metadata is key to obtain both syntactic and semantic interoperability.

Relevant metadata to achieve both syntactic and semantic interoperability in hydrophone deployments have been identified and structured within a Hydrophone SensorML Template, depicted in Fig. 5. A hydrophone SensorML description file, encoded in XML format, can be easily processed by the different software components in the architecture, so the metadata can be streamlined alongside data throughout the dataflow. In appendix B an example of a hydrophone SensorML description file can be found.

A. SYNTACTIC INTEROPERABILITY

The first step in any acquisition chain is to achieve syntactic interoperability between the acquisition software and the instrument, i.e. a hydrophone. Since manufacturers tend to use vendor-specific protocols and formats, achieving syntactic interoperability is not a trivial task. Usually an *ad hoc* driver is generated to interface a sensor to an acquisition

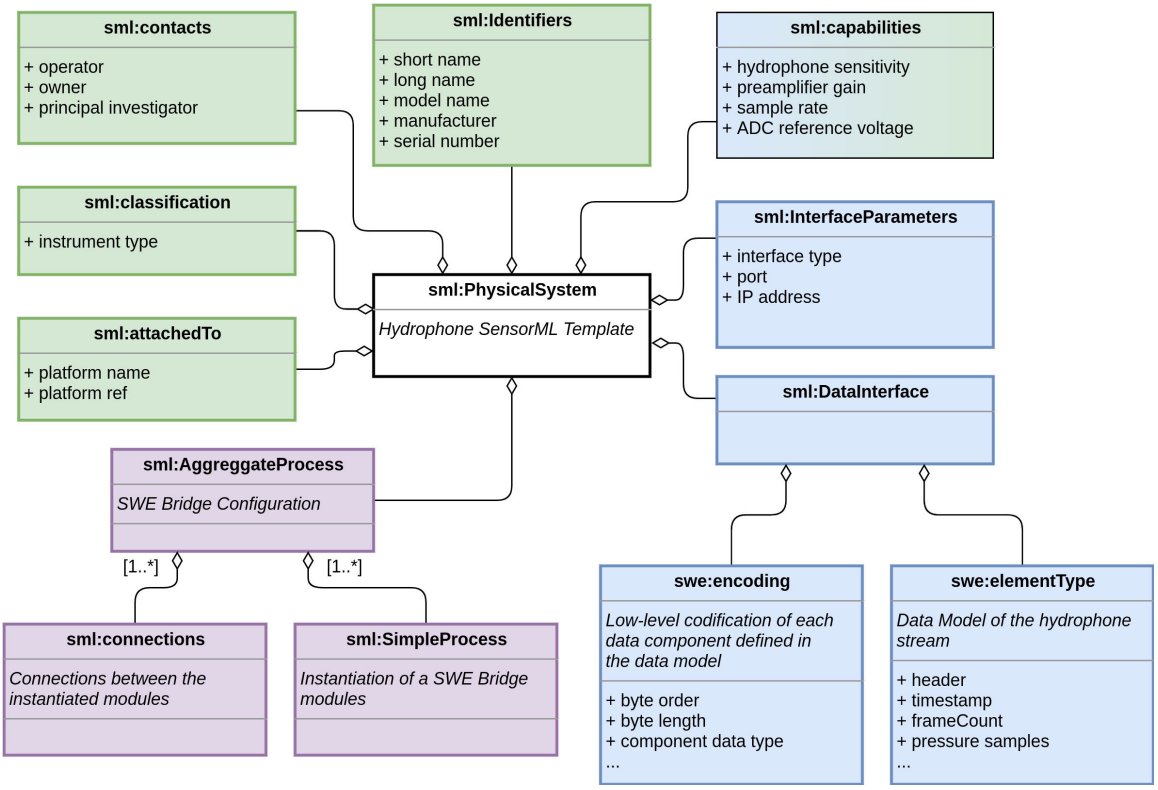

FIGURE 5. Diagram of the information contained within a Hydrophone SensorML Template. Green elements are used to provide contextual information for the measurements. Blue elements are structures targeting syntactic interoperability. The *sml:capabilities* element is used in both, semantically and syntactically. Purple elements are used to configure the acquisition and data workflow within the SWE Bridge.

system. However, within any driver there is a lot of implicit metadata to achieve syntactic interoperability, such as the communication protocol, the meaning of each field within a data stream, encoding, etc. All this information can be encoded in unambiguous way using the SensorML standard. A software component able to understand this standard could automatically configure its acquisition to achieve on-the-fly syntactic interoperability. Within the proposed architecture this component is the SWE Bridge.

However, hydrophone streams can be complex, containing headers, counters and acoustic data arranged in different ways. A detailed description of the stream has to be carefully organized and encoded. Signal conditioning characteristics such as transducer sensitivity, preamplifier gain and ADC parameters also need to be specified in order to transform dimensionless raw samples to pressure (see section III-A). Fig. 6 shows how the information within a hydrophone SensorML description is used by the SWE Bridge to configure the acquisition.

In order to communicate with a hydrophone, the first step is to define its communication interface: protocol (e.g. UDP or TCP), IP address and port number. Using this information, a communication link can be established and data streams start to flow. To process the incoming acoustic data, the low-level details of the stream must be known: packet length, byte order, sample width, etc.

At this point the SWE Bridge has interfaced the hydrophone on-the-fly and can acquire pressure samples for further processing. Its embedded underwater sound level algorithm can take incoming samples and provide relevant ocean sound measurements such as SPL, SEL and RMS values for each frequency band.

By default, the SWE Bridge applies this processing to be compliant with the MSFD descriptors following Task Group 11 recommendations. However, it is possible to adjust its parameters to fit different applications such as the frequency bands, time window and processing parameters. This configuration is also be included within the SensorML description file, so it contains both hydrophone metadata and information about how data is processed, from the transducer to the algorithm's output.

B. SEMANTIC INTEROPERABILTIY

Semantic interoperability focuses on making the meaning of data and metadata explicit, with emphasis on machine to machine interactions. The overall goal is to provide machine understandable, standardized contextual information about the data, following the FAIR principles [26], [27].

The Hydrophone SensorML Template defines a minimum set of metadata elements to be included in a hydrophone SensorML description in order to provide accurate contextual information for the correct production of ocean

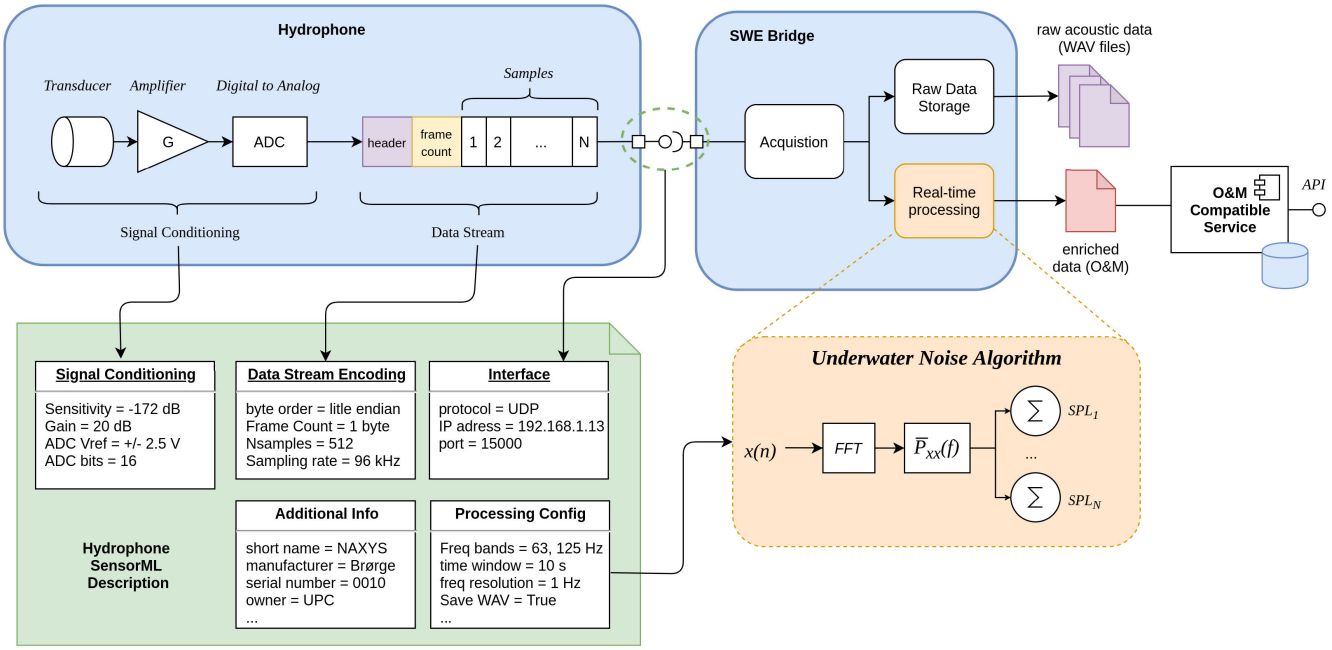

FIGURE 6. Metadata contained within a hydrophone SensorML deployment file and its role to achieve syntactic interoperability to establish a data-flow. It includes signal conditioning metadata (e.g. sensitivity, gain, reference voltage), data stream encoding (e.g. byte order, sampling rate), communications interface (protocol, address, port, etc.), processing and storage options and contextual metadata (e.g. name, manufacturer, deployment coordinates).

TABLE 1. Minimum set of terms from the NERC Vocabulary Server 2.0 used within the Hydrophone SensorML Template.

Vocabulary	Term	Label
W06	CLSS0002	instrument type
W07	IDEN0006	short name
W07	IDEN0002	long name
W07	IDEN0003	model name
W07	IDEN0012	manufacturer
W08	CONT0003	operator
W08	CONT0004	principal investigator
W08	CONT0002	owner
P07	CFSN0310	Sound Pressure Level in water

sound measurements. SensorML is a very flexible and versatile standard, however this can also prove an issue, since the same information can be encoded in multiple ways [37]. In order to provide a semantically-robust metadata, controlled vocabularies are used within the proposed architecture.

The NERC Vocabulary Server version 2.0 (NVS2.0) provides access to standardized terms covering a broad spectrum of disciplines relevant to the oceanographic and earth observing community [38]. These terms are organized collections of concepts, named vocabularies. Each concept has its own universal resource identifier (URI) that resolves, after content negotiation, in a self-descriptive resource description format (RDF) or an HTML page depending if the request was made by a human or a machine. NVS2.0 also holds a set of vocabularies specifically targeting SensorML terms, developed by the SWE Marine Profiles communities [39]. Table 1 shows the minimum set of terms from the NVS2.0 used in the hydrophone SensorML template.

TABLE 2. Minimum set of terms used from the Integrated Ocean Observing System Passive Acoustics Conventions vocabulary.

Vocabulary	Term
IOOS PAM Conventions	hydrophone_sensitivity
IOOS PAM Conventions	sample_rate
IOOS PAM Conventions	preamplifier_gain

NVS2.0 does not provide specific terms for some passive acoustics properties such as hydrophone sensitivity, pre-amplifier gain and sample rate. As an alternative for these terms the Integrated Ocean Observing System (IOOS) Passive Acoustic Monitoring (PAM) Conventions vocabulary has been adopted [40]. Table 2 shows the minimum set of terms from the IOOS PAM conventions vocabulary used within a hydrophone SensorML template. However, not all terms can be found in neither vocabularies, e.g. ADC's reference voltage.

The processed data generated by the SWE Bridge, alongside with the semantically-enhanced metadata from the SensorML are combined into O&M files, which can be injected in standard services such as SOS [28].

C. EMBEDDING METADATA IN WAV FILES

Alongside the semantically-enriched O&M data files, the SWE Bridge may also generate raw acoustic recordings using the Waveform Audio File Format (WAV). This format is an extension of the Resource Interchange File Format (RIFF) focusing on commercial audio.

Due to its versatility and flexibility, the WAV format has been broadly used in passive acoustics for hydrophone recordings. However, it does not consider any standardized

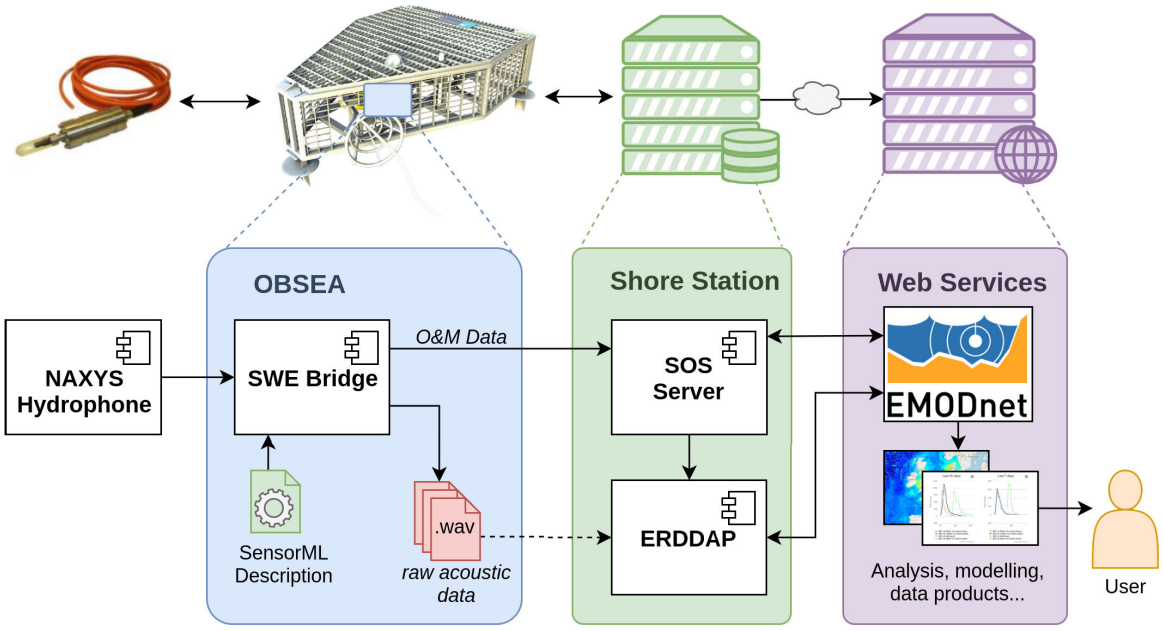

FIGURE 7. Ocean sound measurement dataflow at OBSEA. The SWE Bridge interfaces the NAXYS hydrophone based on the information on its SensorML description and generates two outputs: underwater sound levels encoded in O&M and acoustic recordings, sent to a SOS and ERDDAP servers respectively. Data is shared with data portals and repositories such as EMODnet Physics.

mechanism to embed metadata. A WAV file without its associated metadata may be useless since the calibration and contextual information are not known (hydrophone sensitivity, location, timestamp, etc.).

In order to overcome the lack of metadata, the SWE Bridge uses the ID3 tagging system. This informal standard takes advantage of the RIFF's chunk-based design, adding a chunk of metadata alongside the audio data while maintaining the format compatibility [41]. ID3 provides a list of items that can be included, such as author, song title, genre, composer, etc.

Although it focuses on commercial audio and it is not directly applicable to underwater acoustics, the ID3 tagging system includes the option to add user-defined tags. Within the SWE Bridge software, these user-defined tags are leveraged to include key-value pairs with all the relevant information contained within the SensorML hydrophone description, such as hydrophone name, model, sensitivity, deployment position, etc. Digital audio workstations and audio players can easily access to the embedded metadata. Using this workaround all the metadata described in section IV is embedded in the WAV file, maintaining its compatibility with WAV and RIFF formats.

V. USE CASES

A. NAXYS HYDROPHONE AT OBSEA

OBSEA Expandable Seafloor Observatory is a cabled, multi-parametric observing platform located 4 km off the coast of Vilanova i la Geltrú (Barcelona, Spain) at a depth of 20 meters. Since its deployment in 2009 it has been continuously acquiring data from numerous variables such as temperature, salinity, pressure, air temperature and underwater sound [42]. OBSEA is equipped with a NAXYS Ethernet 02345 hydrophone, streaming acoustic data. Since 2020 it is offering ocean sound measurements in real-time using the proposed ocean sound architecture, as shown in Fig. 7. It provides valuable data for the assessment of long-term underwater noise trends in the eastern Mediterranean Sea, compliant with the requirements of the MSFD.

TABLE 3. NAXYS Ethernet Hydrophone 02345 specifications.

Hydrophone Parameter	**Value**
sample rate	96000 Hz
hydrophone sensitivity	-192 dB re V/µPa
preamplifier gain	20 dB
ADC reference voltage	± 2.5 V
ADC number of bits	16
Processing Parameter	**Value**
frequency resolution	1 Hz
time window	10 s
overlap	50 %
window function	Hann
third-octave bands	63, 125, 2000 Hz

The SWE Bridge universal driver is deployed within the platform, using the information contained within the hydrophone's SensorML description file to setup the data acquisition. It automatically processes the incoming data stream, transforming from raw samples to pressure values. Its embedded underwater sound algorithm computes SPL, SEL and RMS values with a time window of 10 seconds. It uses the Welch method with a Hann window and an overlap of 50% (see section III-C). These parameters are calculated for the 63, 125 and 2000 Hz third-octave bands and the whole bandwidth (up to 48 kHz). The hydrophone's specifications and the processing setup is shown in table 3.

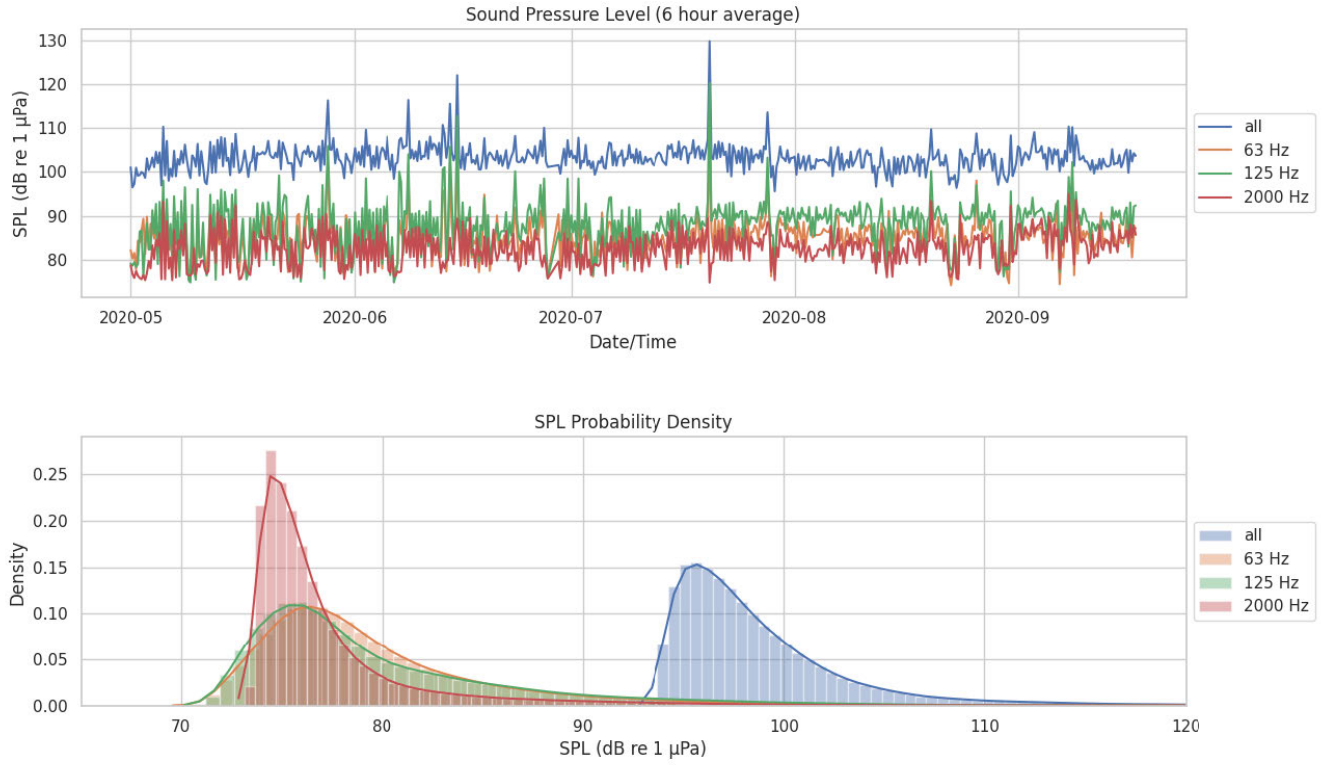

FIGURE 8. Sound Pressure Level (SPL) at full bandwidth, 63, 125 and 2000 Hz third-octave bands at OBSEA from May 1st to September 15th 2020, averaged in periods of 6 hours (top), and SPL histogram and estimated probability distribution function (bottom).

The algorithm result is stored in O&M files and sent to OBSEA's Sensor Observation Service. Data is available in real-time through the Helgoland Client at `http://sos.obsea.es/client` or directly through the SOS interface `http://sos.obsea.es/sos/`. The hydrophone's SensorML description is also available via SOS interface.

Raw acoustic data in WAV format is also generated by the SWE Bridge to allow further analysis and data validation. Since OBSEA is a cabled observatory and does not present communications constraints, the generated WAV files are sent in real-time to the OBSEA's ERDDAP server [30], available at `http://erddap.obsea.emso.eu`.

ERDDAP is a data server that provides a simple, consistent way to serve scientific data on the web. Alongside with its web interface it also provides a RESTful API allowing humans and/or machines to interact with data programmatically. Unlike SOS, it does not use strong semantics for metadata, but its ability to manage different files and formats makes it a perfect candidate to serve large datasets with heterogenous data. OBSEA's ERDDAP server contains WAV datasets and SPL timeseries. Using ERDDAP and SOS the data is shared with data portals and services, such as EMODnet Physics [43].

Fig. 8 shows ocean sound data at OBSEA from 1st of May until 15th of September 2020, acquired using the proposed metadata-driven architecture. SPL data is averaged in periods of 6 hours for visualization purposes.

TABLE 4. Values of Sound Pressure Level full bandwidth (all) and 1/3 octave bands centered at 63, 125 and 2000 Hz at OBSEA. All values are in dB re 1 μPa with a time window of 10 seconds.

metric	all	63 Hz	125 Hz	2000 Hz
mode	96.04	77.13	75.83	74.17
median	97.46	77.96	77.55	75.74
L90	103.81	86.79	88.42	81.89
mean	105.12	88.04	93.04	84.21

Ocean sound has a high variability, as it can be seen in the dataset. The arithmetic mean of large SPL dataset is heavily influenced by short, high-intensity events. In other words, it is highly sensible to outliers in the underwater sound level distribution [6], [44]. When assessing underwater sound levels, different averaging metrics can produce widely differing levels, which may result in misinterpretation of ocean sound data [45]. To complement the information provided by the arithmetic mean, additional metrics are provided at table 4. The mode and the median provide a better estimate of the expected SPL level during most of the time, and the L90 percentile shows the level which is not exceeded in 90% of the time. As it can be seen, the mean of the SPL values is higher than the L90 percentile (up to 95% in the 125 Hz third-octave band timeseries).

The spike observed on the 20th of July corresponds with a major maintenance operation at OBSEA using a large support vessel. The increment of the underwater sound level is caused by the proximity of the vessel's engine.

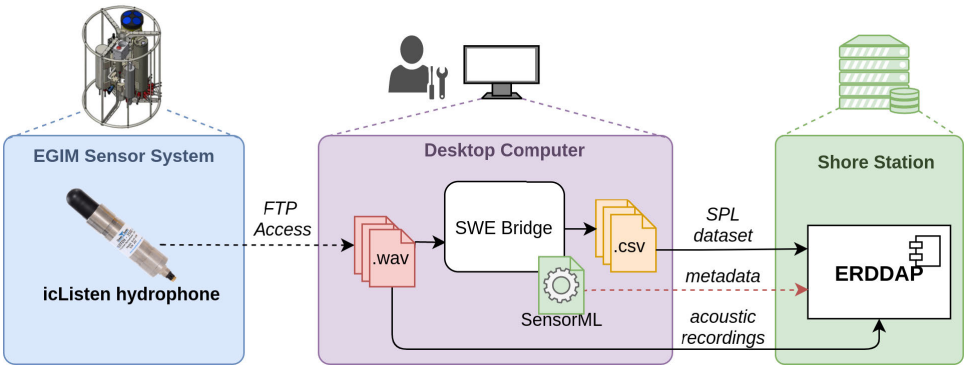

FIGURE 9. Ocean sound measurement dataflow of the EGIM sensor system at PLOCAN. Internally recorded data was stored on-board and downloaded via FTP. Afterwards the SWE Bridge processed the acoustic recordings, generating a Sound Pressure Level dataset in CSV format. The resulting dataset and its metadata is injected to an ERDDAP service.

Weekly periodic oscillations within the time series can be seen, especially in May and June. These oscillations are probably due to the increment of human activities, during workdays (e.g. fishing) and a decrease during weekends. However, in July and August these fluctuations are less evident due to the increase of recreational boating

From the probability density it can be seen that the sound levels at 2000 Hz 1/3 octave band are very stable, while low-frequency sound (63 and 125 Hz 1/3 octave bands) have a larger variability. This may be induced by the proximity of the port of Vilanova i la Geltrú, as low frequency noise produced by fishing vessels passing by are acquired. Higher frequencies are attenuated much more rapidly than lower frequencies, thus the 2000 Hz band is much less affected by distant events (e.g. shipping), presenting a narrower distribution.

B. EGIM PASSIVE ACOUSTICS RECORDINGS

Although real-time ocean sound processing is the main goal of the presented architecture, it is also capable of analyzing previously acquired data. The SWE Bridge can access acoustic recordings in WAV format and generate processed ocean sound datasets in O&M or CSV formats.

In order to illustrate this capability, recordings from a hydrophone deployed within an EGIM (EMSO Generic Instrument Module) have been processed [46]. The deployment was performed at the Plataforma Oceánica de Canarias (PLOCAN) in the Canary Islands, Spain, during June and July 2019. The EGIM sensor system contained an Ocean Sonics icListen HF hydrophone continuously recording data in its internal memory. The overall dataset was retrieved after instrument recovery and processed using the SWE Bridge. The hydrophone specifications (table 5) were encoded within a SensorML file and used by the SWE Bridge to convert the raw acoustic dataset to SPL, as shown in Fig. 9. The resulting dataset of SPL values was uploaded to PLOCAN's ERDDAP service, where it is publicly available (`http://erddap.plocan.eu/erddap`).

Fig. 10 shows the result of a very unstable underwater soundscape. Very intense low frequency noise (below 20 Hz) has been observed in the dataset. In order to reduce the spectral leakage from low frequencies the Welch method using a Hann window has been selected, with time window of 10 seconds and a Δf of 1 Hz and an overlap of 50% (see section III-C). The metrics of the overall dataset are shown in table 6.

TABLE 5. icListen HF hydrophone specifications.

Hydrophone Parameter	Value
sample rate	8000 Hz
hydrophone sensitivity	-169 dB re V/µPa
ADC reference voltage	± 3 V
ADC number of bits	24
Processing Parameter	**Value**
frequency resolution	1 Hz
time window	10 s
window function	Hann
overlap	50 %
third-octave bands	63, 125, 2000 Hz

TABLE 6. Metrics of the Sound Pressure Level full bandwidth (all) and third-octave-bands centered at 10, 63, 125 and 2000 Hz at PLOCAN during June and July 2019. All values are in dB re 1 μPa.

metric	all	10 Hz	63 Hz	125 Hz	2000 Hz
mode	115.99	89.96	84.36	84.38	86.19
median	116.33	90.40	85.15	85.28	85.82
L90	124.28	99.00	91.24	96.74	88.42
mean	123.37	107.53	92.80	92.53	89.17

PLOCAN's test site is approximately 1 km away from Taliarte's port, a relatively small fishing port. Small ships very close to the hydrophone may be the cause of spikes observed in the dataset. The low-frequency noise may be induced by shipping from the port of Las Palmas de Gran Canaria (a major commercial port), approximately 20 km away from the deployment site and possibly the logistics required around a wind energy converter deployed close to PLOCAN.

Although there is a clear influence of shipping, the changes in the trend can be clearly observed in the dataset in all frequency bands. Wind speed data collected from a weather station at PLOCAN also shows a correlation with the background trend (see Fig. 11), thus possibly stemming from the added sound energy generated by wind-driven wave-breaking processes.

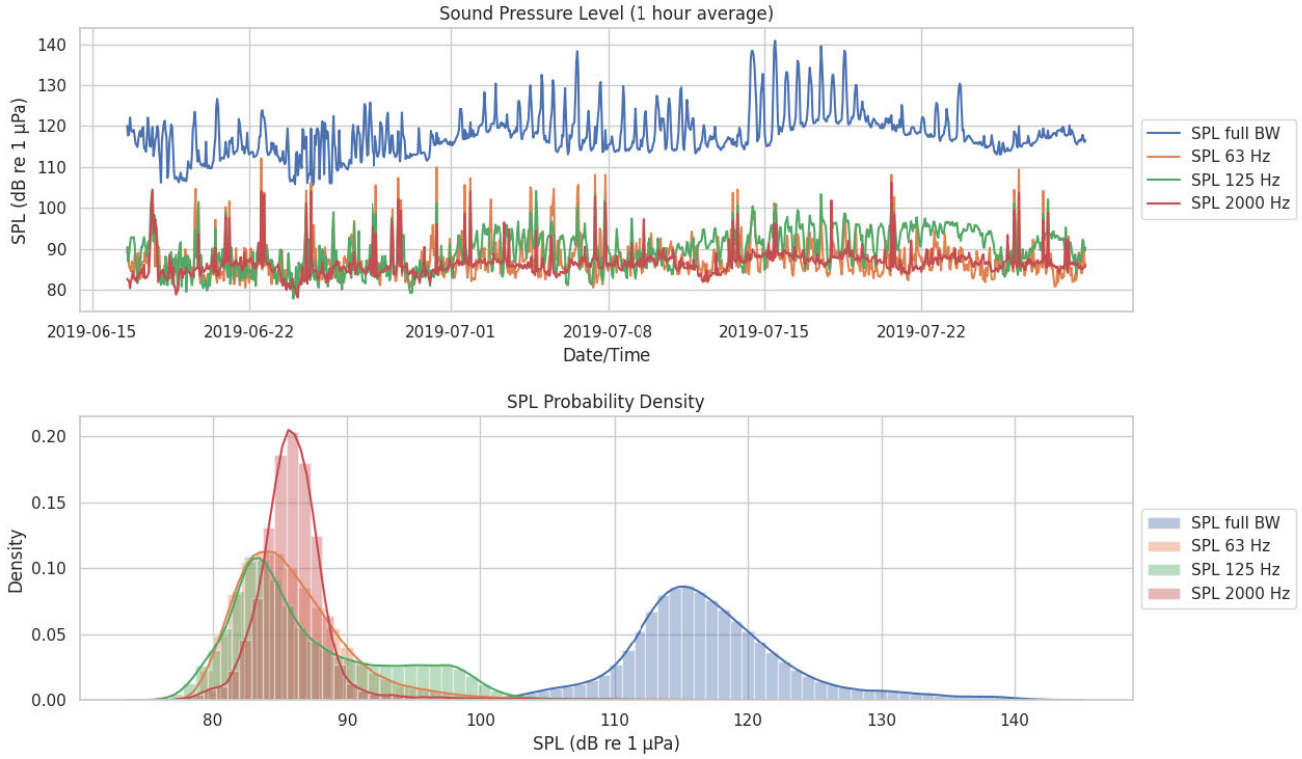

FIGURE 10. Sound Pressure Measurements at full bandwidth, 63, 125 and 2000 Hz third-octave bands at PLOCAN using an icListen hydrophone during June and July 2019, averaged in periods of 1 hour for visualization purposes (top), and SPL histogram and estimated probability distribution function (bottom).

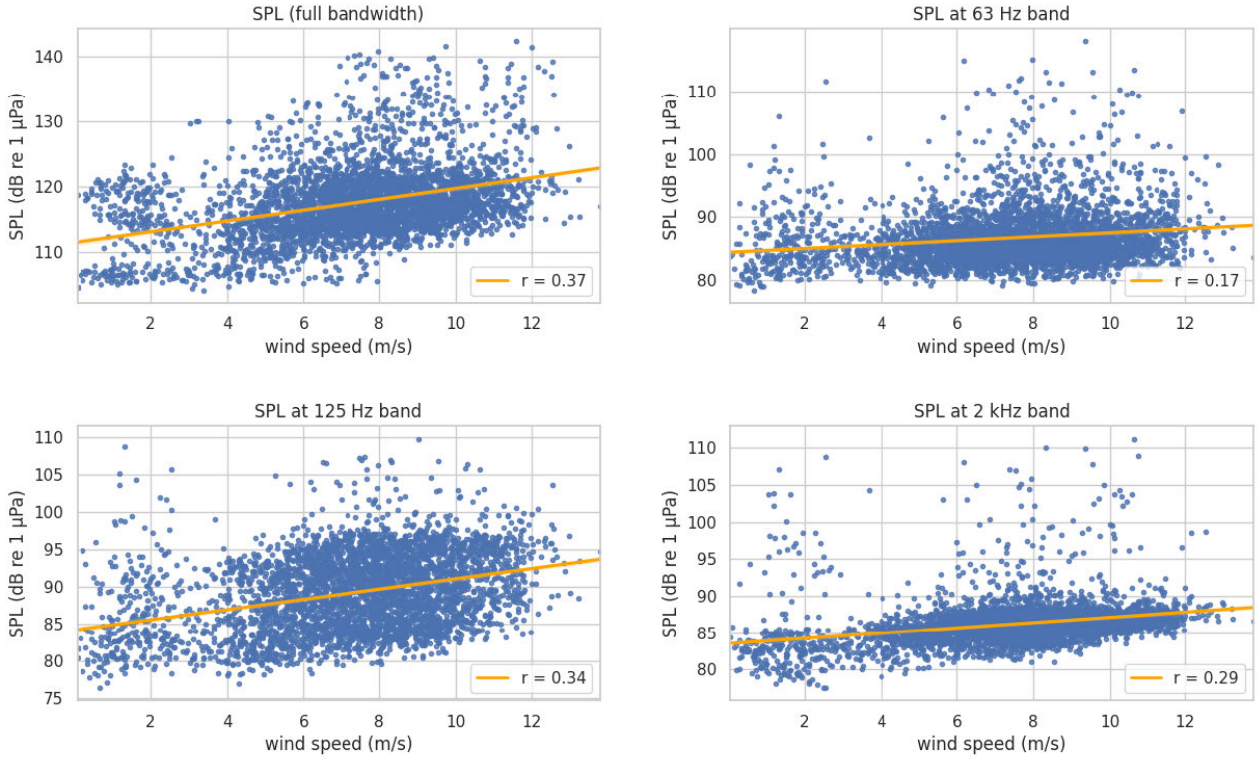

FIGURE 11. Linear regression between Sound Pressure Level (SPL) and wind speed and its correlation coefficient *r* at PLOCAN during June and July 2019. The linear regression analysis shows a clear correlation between both variables, although the outliers in SPL measurements induced by others sources (e.g. shipping) reduce the correlation coefficient *r*.

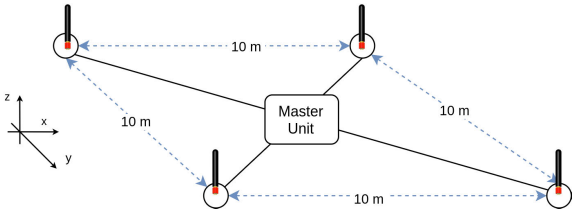

FIGURE 12. A2 hydrophone array diagram.

C. A2 HYDROPHONE ARRAY

The A2 Hydrophone Array is a digital passive acoustic transducer array designed for sound source localization (SSL) and tracking. It is composed of 4 slave hydrophones, called A2 hydrophones, streaming data to a master unit which processes the acoustic data in real-time. Therefore, the main capability of A2 is to provide directional sound source information for hydro-acoustic surveys [15]. Fig. 12 shows the block diagram of the A2 hydrophone array.

Time synchronization between the master unit and the slave units (A2 hydrophones) is accomplished by implementing the IEEE 1588 Precision Time Protocol (PTP) standard. This standard defines a network protocol enabling accurate and precise synchronization of the real-time clocks of devices in networked distributed systems, achieving a precision below the micro-second. The acoustic data is transmitted from the slave to the master by means of the Real-time Transport Protocol (RTP).

The A2 Hydrophone Array can be equipped with positioning sensors (pan, tilt, and compass) to allow the measurement of its geo-referenced position. The device can also receive relevant oceanographic parameters (sound velocity, temperature, depth, time) via Ethernet, in order to optimize the algorithms.

Within this work, the A2 hydrophone array's capabilities have been extended to provide ocean sound measurements alongside sound source localization. The proposed architecture for ocean sound measurement has been seamlessly integrated within the A2 hydrophone array, as depicted in Fig. 13. Taking advantage of the SWE Bridge's functionality, raw acoustic recordings for each hydrophone are also stored on-board for further analysis and data validation.

The A2 master unit, based on an embedded Linux single-board computer, receives real-time data from the slaves hydrophones. Thus, an instance of the SWE Bridge for each slave hydrophone has been deployed within the master unit.

The acoustic streams coming from the slave hydrophones are redirected to the SSL algorithm and to SWE Bridge instances. Each instance, processing data of a slave hydrophone, has its own SensorML hydrophone description file, used to setup the acquisition. Those files are identical with the exception of the communication's interface configuration, serial number and their calibration information. As output, four underwater sound levels and acoustic recordings datasets are generated, one per slave hydrophone.

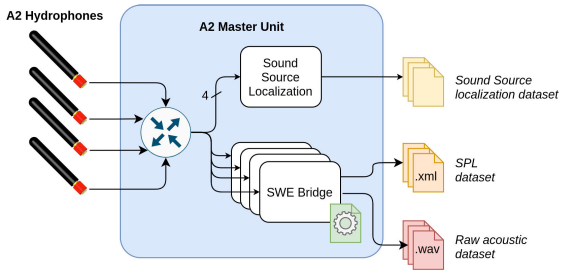

FIGURE 13. A2 hydrophone array extended with the SWE Bridge for ocean sound monitoring.

Successful lab tests were performed and a deployment is scheduled by the beginning of 2021.

VI. CONCLUSION AND FUTURE WORK

Within this work a universal architecture for ocean sound monitoring is proposed. A generic, real-time algorithm implementation for *in situ* underwater sound monitoring has been presented, following best practices on underwater noise measurement methodologies. The intrinsic constraints of state-of-the-art observation platforms have been carefully considered in order to achieve an efficient, generic and cross-platform implementation.

Open-source and open standards are used to ensure re-usability and universality of such architecture, regardless of the underlying components (hydrophone model, observations platform, etc.). Interoperability challenges at both syntactic and semantic level have been thoroughly analyzed and a solution based on the SensorML standard has been proposed. At the syntactic (operational) level, a characterization of the hydrophone leads to a seamless integration within the proposed architecture by abstracting the sensor characteristics. At the semantic level, relevant contextual information has been discussed and encoded using controlled vocabularies to achieve rich and coherent metadata, providing the building blocks for a FAIR data management.

The discussed metadata solution does not only provide contextual information, but effectively manages the acquisition system operation. Since all the acquisition and processing steps are explicitly encoded within the presented metadata solution, the whole acquisition and processing chain can be easily replicated.

In order to demonstrate its flexibility, the architecture has been applied to three different scenarios: real-time monitoring in a cabled observatory, the analysis of acoustic recordings and as the enhancement of a hydrophone array. In all cases the architecture effectively managed to abstract underlying hardware / software characteristics and underwater sound levels measurements were successfully acquired and processed.

One of the weakness of the proposed architecture is the need to generate large SensorML metadata files, which are soft-typed and based on a template. A future line of work would be to define a normative hydrophone description profile, so the generated metadata files could be validated against

a schema. Moreover, this schema could be integrated into existing SensorML editing tools to ease its generation and maintenance.

Future work in metadata could move beyond sensor deployment and focus on sensor calibration and operational history, providing traceability throughout the whole instrument's life-cycle.

As a future line of research, the proposed architecture could be expanded to other underwater acoustics applications. Since hydrophone integration and data pre-processing steps have been addressed within this work, new algorithms can be easily integrated, such as sound source recognition, marine species detection, bio-acoustics, etc.

APPENDIX A
ALGORITHM COMPARISON

Although the typical approach to calculate SPL at different band frequencies is the spectral analysis by means of FFT, there are other approaches that may achieve the same result. In this section, three methods were tested to obtain a computationally efficient SPL algorithm for an arbitrary number of 1/3 octave band frequencies. The tested algorithms were a filter bank, the FFT-based algorithm and the Goertzel's algorithm.

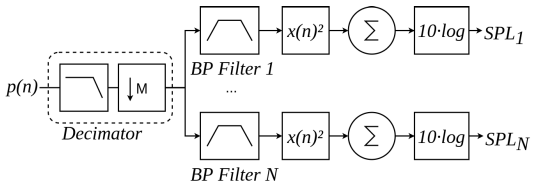

FIGURE 14. Block diagram of a SPL algorithm for multiple band frequencies using a filter bank.

The filter bank approach to calculate the third-octave SPL levels involves decimating and filtering, as shown in Fig. 14. Generally, hydrophones have a broad bandwidth, up to tens or hundreds of kHz. Thus, its sampling rate is usually much higher than the frequencies of interest. In order to save computational resources and to ensure the filter stability, the incoming signal needs to be decimated.

The decimation process involves low-pass filtering to avoid aliasing, and down-sampling. Since underwater sound level algorithms are usually calculated for low-frequency third-octave bands, the incoming signal may be over-sampled by several orders of magnitude over the Nyquist rate. In this case, decimating the signal in a single stage may result in a very costly and complex filtering arrangement, so it may be more efficient to decimate the incoming signal in multiple stages [47]. The optimal number of decimation stages needs to be calculated at run-time based on the hydrophone's sampling rate and the highest cut-off frequency of the third-octaves bands.

After being decimated, the pressure signal is filtered by a band-pass filter. Ideally such a filter would completely eliminate the energy in the rejection-band without affecting the power in the band-pass. In practice, it should have small ripple at the band-pass and a fast roll-off in the rejection bands to minimize its influence on the signal. When designing a filter there is a trade-off between the computational cost and the filter performance. Generally various frequency bands are desired; thus, a filter bank is required.

Once the signal has been filtered, the SPL value for a specific frequency band can be calculated using (4) over the filtered pressure signal.

On the contrary to the filter bank approach, the power spectral density of the pressure signal can be estimated by means of a Discrete Fourier Transform (DFT), as discussed in section III. To calculate the DFT, two different implementations where evaluated, the well-known Fast Fourier Transform (FFT) and the Goertzel's algorithm.

The FFT is an optimized implementation of the DFT for data segments with power of 2 length. On the contrary, Goertzel's algorithm is much slower DFT implementation, but it has the ability to independently calculate DFT terms. Goretzel's algorithm can be useful when only a small number of DFT bins M are required, i.e. $M < log_2 N$ [31]. In this particular application, the overall spectrum is not desired, just the bins containing frequencies within the bands of interest. So, it has been considered for ocean sound measurements.

All three SPL algorithms where implemented and tested to assess their performance: a filter bank (multi-stage decimation and 5th order elliptic band-pass filters), another FFT-based and the third one based on the Goertzel's algorithm. Several tests were executed to assess their performance in terms of required memory and execution time under different conditions, such as the number of samples in the signal and number of third-octave bands calculated. Although the MSFD only specifies two 1/3 octave bands to be monitored (63 and 125 Hz), it has been suggested to extend the monitoring range up to 20 kHz [3]. Thus, the computational impact of increasing the number of band frequencies calculated has been assessed.

The algorithms where implemented using the python3 programming language. Although it is a high-level scripted programming language and most of the complex operations are done under-the-hood, it is a good starting point to assess the feasibility of each algorithm and have a rough estimate of their computational cost. The results of these tests are shown in Fig. 15.

The filter bank memory usage grows rapidly when more third-octave bands are calculated (Fig. 15, top left). The filter bank is not heavily affected by the increase of the signal's length N (Fig. 15, bottom right), however it is significantly affected when the number of third-octaves bands are increased, since a different filter has to be applied for each band (top right).

The Goertzel's algorithm has a very limited memory usage and a very good performance when applied to small number of samples N. However, its execution time exponentially grows as N or the number of third-octave bands is increased. The dashed red line shows the Δf achieved with the DFT (both Goertzel and FFT), which is inversely pro-

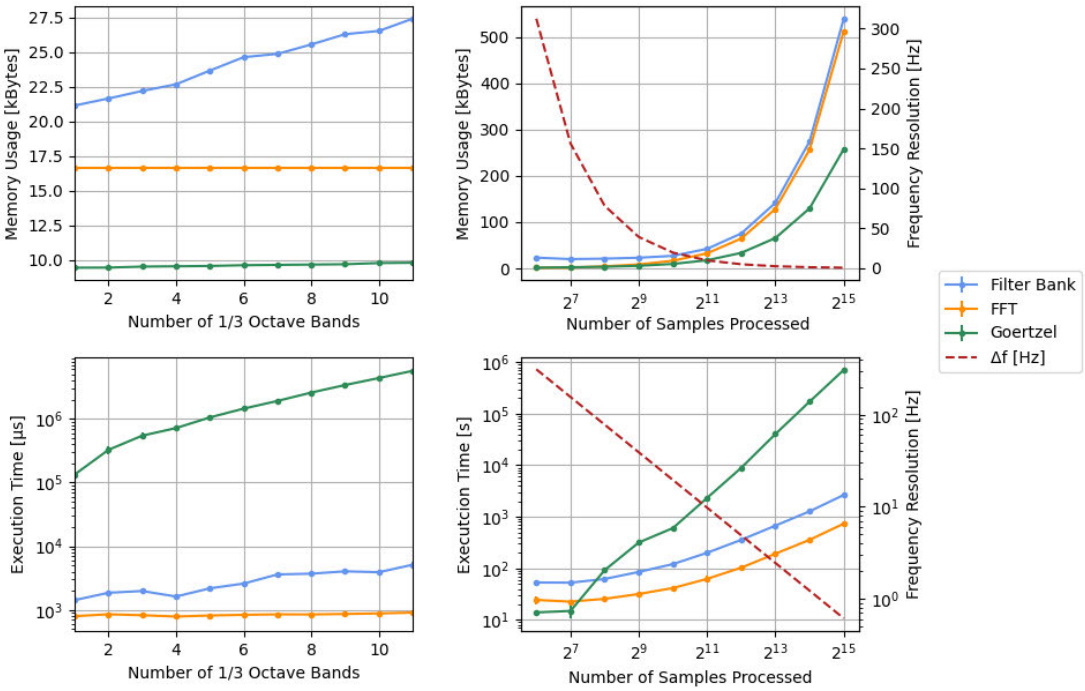

FIGURE 15. Performance test of three different SPL algorithms based on their 1/3 octave estimation method: filter bank, FFT and Goertzel's algorithm. The left graphs show the memory usage (top) and the execution time (bottom) depending on the number of third-octave bands calculated, starting from the band centered at 63 Hz (N is set to 4096). The right graphs show the memory usage (top) and execution time (bottom), depending on the number of samples in the time window (only 63 and 125 Hz third-octave bands). The dashed red line in rights graphs represent the frequency resolution Δf of the DFT for each value of N. The sampling frequency is set to 20 kHz.

portional to N. The Goertzel's algorithm is only faster than the other algorithms when Δf is in the order of tens of Hz, which is not acceptable for the intended applications. When accurate frequency resolution is required, the algorithm shows very poor performance (note the axis logarithmic scale). Thus, this algorithm is not suited for this particular application.

Although FFT has an average memory usage with respect to N, it shows very good performance in terms of execution time. Moreover, its memory usage is constant regardless of the number of third-octave bands calculated. Thus, the FFT-based SPL algorithm has been selected and implemented within this work.

APPENDIX B
HYDROPHONE SensorML DESCRIPTION

This section provides an example of a hydrophone SensorML description, using OBSEA's NAXYS hydrophone as an example (see section V-A). The following XML snippets provide examples on how to define specific parts of the document. URLs are shortened to improve readability. The entire document is available online at `http://sos.obsea.es/sensorml/naxys.xml`.

Table 7 shows the metadata included within the NAXYS SensorML description. Some of the components such as hydrophone data stream and SWE Bridge configuration are not included in the table, but are discussed later in this section.

TABLE 7. Metadata included in the NAXYS Hydrophone SensorML description.

Term	Value
sml:classification	
instrument type	SDN:L05::369 [hydrophone]
sml:identifiers	
short name	NAXYS Hydrophone
long name	NAXYS Hydrophone at OBSEA
model name	NAXYS Ethernet Hydrophone 02345
manufacturer	Bjørge, AS
serial number	0010
sml:contacts	
operator	Enoc Martínez
owner	Universitat Politècnica de Catalunya
principal investigator	Joaquín del Río
sml:capabilities	
hydrophone sensitivity	-192 (dB re 1 V/µPa)
preamplifier gain	20 (dB)
sample rate	96000 (Hz)
ADC reference voltage	2.5 (V)
sml:InterfaceParameters	
interface type	UDP
port number	15000
IP	192.168.1.113
sml:attachedTo	
title	OBSEA
reference	http://sos.obsea.es/sensorml/obsea.xml
sml:position	
latitude	41.1819 (degree north)
longitude	1.7527 (degree east)
depth	20 (m)

A. IDENTIFICATION, CLASSIFICATION AND CAPABILITIES

Identification and classification sections provide a URI to a controlled vocabulary (definition attribute in the example

below), a label with a human-readable description of the term and a value. This value can be either a text string (e.g. NAXYS Hydrophone) or another URI to a controlled vocabulary where the info is described (e.g. description of the sensor type in NVS2.0's L05 vocabulary). The following snippet shows how to define the short name parameter:

```
<sml:identifier>
  <sml:Term definition="http://vocab.../IDEN0006/">
    <sml:label>short name</sml:label>
    <sml:value>NAXYS Hydrophone</sml:value>
  </sml:Term>
</sml:identifier>
```

In the capabilities section the electrical properties of the hydrophone are described. Using the SWE Common Data Model standard it is possible to provide unambiguous definition of the data type and the associated units. The following snippet shows how to describe the hydrophone sensitivity:

```
<sml:capability name="hydrophone_sensitivity">
  <swe:Quantity
     definition="http://mmisw.../hydrophone_sensitivity">
     <swe:label>hydrophone sensitivity</swe:label>
     <swe:uom code="dB_re_uPa"
       xlink:href="http:/vocab..../UDBL"/>
     <swe:value>-192</swe:value>
  </swe:Quantity>
</sml:capability>
```

B. CONTACTS

The contacts section provides information about the people and organization involved. SensorML does not describe a specific set of terms for these, but it uses the ISO 19115 to provide information about the contacts and their role.

C. INTERFACE PARAMETERS

In order to define unambiguously the interface details such as IP address and port number the *sml:interfaceParamaters* within a *sml:parameter* is used.

```
<sml:interfaceParameters>
  <swe:DataRecord>
    <swe:field name="portType">
      <swe:Category>
        <swe:label>port type</swe:label>
        <swe:value>UDP</swe:value>
      </swe:Category>
    </swe:field>
    <swe:field name="portNumber">
      <swe:Count>
        <swe:label>port number</swe:label>
        <swe:value>15000</swe:value>
      </swe:Count>
    </swe:field>
    <swe:field name="IP">
      <swe:Count>
        <swe:label>IP address</swe:label>
        <swe:value>192.168.1.113</swe:value>
      </swe:Count>
    </swe:field>
  </swe:DataRecord>
</sml:interfaceParameters>
```

D. ATTACHED TO

The *sml:attachedTo* element is used to provide information about the observing platform where the sensor is installed.

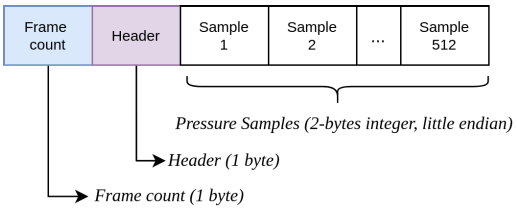

FIGURE 16. NAXYS Hydrophone stream, composed by a frame counter byte (0-255), a header byte and 512 pressure samples encoded as 2-byte integers in little endian.

According to the SensorML standard, the *xlink:title* attribute is used to specify the identifier of the hosting system while the *xlink:href* should point to a SensorML description of such platform:

```
<sml:attachedTo xlink:title="OBSEA"
  xlink:href="http://sos.obsea.es/sensorml/obsea.xml"/>
```

E. DATA STREAM

One of the most complicated aspects of a hydrophone SensorML description is complex streams in binary format. Although it is not trivial, the SWE Common Data Model Standard provides a good framework to encode such streams [48]. The NAXYS hydrophone sends streams periodically, each one with 1026 bytes. As depicted in Fig. 16, each stream has a frame counter to detect missing packets, a header byte containing configuration information (sampling rate and preamplifier gain) and 512 pressure samples, arranged in signed 2-byte integers. The following SensorML excerpt shows how the NAXYS hydrophone stream has been modelled using the *swe:DataStream* element:

```
<swe:DataStream>
  <swe:elementType name="dataModel">
    <swe:DataRecord>
      <swe:field name="frameCount">
        <swe:Count/>
      </swe:field>
      <swe:field name="header">
        <swe:Count/>
      </swe:field>
      <swe:field name="array">
        <swe:DataArray>
          <swe:elementCount xlink:href="#arrayCount"/>
          <swe:elementType name="samples">
            <swe:Count
              definition="http://vocab.../CFSN0310/"/>
          </swe:elementType>
        </swe:DataArray>
      </swe:field>
    </swe:DataRecord>
  </swe:elementType>
  <swe:encoding>
    <swe:BinaryEncoding byteOrder="littleEndian"
      byteEncoding="raw" byteLength="1026">
      <swe:member>
        <swe:Component
          dataType="http://www.opengis.net/.../unsignedByte"
          ref="dataModel/header"/>
      </swe:member>
      <swe:member>
        <swe:Component
          dataType="http://www.opengis.net/.../unsignedByte"
          ref="dataModel/frameCount"/>
      </swe:member>
```

```
    <swe:member>
      <swe:Component
        dataType="http://www.opengis.net/.../signedShort"
        ref="dataModel/array/samples"/>
    </swe:member>
   </swe:BinaryEncoding>
  </swe:encoding>
  <swe:values/>
</swe:DataStream>
```

The *swe:DataStream* element is used to encode the stream of the hydrophone. It contains two main components: the data model (*swe:elementType*) and the low-level encoding details (*swe:encoding*). Within the data model three fields are declared, the frame count, the header byte and an array of pressure samples. The number of pressure samples within the array is declared as a parameter outside the data stream and referenced in the data model:

```
<sml:parameter name="arrayCountParameter">
  <swe:Count id="arrayCount">
    <swe:value>512</swe:value>
  </swe:Count>
</sml:parameter>
```

The *swe:encoding* defines the low-level encoding details for the stream. These details include the byte order (little or big endian), the length of the overall stream and the encoding of each one of the fields declared in the data model. Each field encoding has an associated *swe:Component* which only has two attributes: data type and ref. Data type points to an online resource defining its computer number format (i.e. unsigned byte, 32-bit integer, 64-bit floating point, etc.). The ref element links this component to a data model's field following the SensorML standard referencing rules.

F. SWE BRIDGE CONFIGURATION FOR NAXYS HYDROPHONE

The previous elements of the NAXYS SensorML description file are focused on properties related directly to the hydrophone. However, to setup an acquisition chain some additional information may be required, such as output format, recording time, duty cycle, etc. In the proposed architecture all the metadata is embedded within the hydrophone SensorML description, thus the configuration of the acquisition itself is also managed from the SensorML file.

The SWE Bridge includes a set of modules, each one targeting a specific task: process incoming data from the communication's interface, calculate sound pressure levels, generate WAV files, generate O&M files and access GPIO (general purpose input output) among others [19]. These modules are described using the SensorML standard and can be easily configured adding a specific section to a hydrophone SensorML description file. Using the *sml:typeOf* inheritance method, a SWE Bridge module invoked by referencing its identifier. With the *sml:Settings* element it is possible to change the parameters values. If not specified, the default value is used. The following XML excerpt shows how to configure the Sound Pressure Level module:

```
<sml:SimpleProcess gml:id="SoundPressureLevel" >
  <sml:typeOf
    xlink:title="swebridge:modules:soundPressureLevel"/>
  <sml:configuration>
    <sml:Settings>
      <sml:setValue ref="parameters/integrationTime">
        10</sml:setValue>
      <sml:setValue ref="parameters/frequencyBands">
        full 63 125 2000</sml:setValue>
      <sml:setValue ref="parameters/soundExposureLevel">
        true</sml:setValue>
      <sml:setValue ref="parameters/rootMeanSquare">
        true</sml:setValue>
    </sml:Settings>
  </sml:configuration>
</sml:SimpleProcess>
```

As its name indicates, the Sound Pressure Level module calculates SPL values. Each module has a set of parameters which can be set using the *sml:setvalue* elements. These parameters reflect non-trivial acquisition aspects that are not directly related to the sensor, but the user may need to adjust to fine-tune the acquisition. In the previous XML excerpt, the following parameters of the SPL algorithm have been set:

- **integration time**: Time window of the SPL measurements.
- **frequency bands**: Third-octave bands to calculated (*full* corresponds to the whole hydrophone's bandwidth)
- **sound exposure level**: Flag to determine if the SEL level should be calculated
- **root mean square**: Flag to determine if the root mean squared (RMS) level should be calculated

In order to generate a workflow within the SWE Bridge the modules need to be connected. The following XML excerpts shows how to connect two SWE Bridge modules:

```
<sml:connection>
  <sml:Link>
    <sml:source
      ref=".../HighFreqStream/.../dataOut"/>
    <sml:destination
      ref=".../SoundPressureLevel/.../dataIn"/>
  </sml:Link>
</sml:connection>
```

ACKNOWLEDGMENT

Researchers want to acknowledge the support of the Associated Unit Tecnoterra composed by members of Universitat Politècnica de Catalunya (UPC) and the Consejo Superior de Investigaciones Científicas (CSIC). This work used the EGI infrastructure with the dedicated support of INFN-CATANIA-STACK.

REFERENCES

[1] J. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," *Mar. Ecology Prog. Ser.*, vol. 395, pp. 5–20, Dec. 2009.

[2] H. Slabbekoorn, N. Bouton, I. van Opzeeland, A. Coers, C. ten Cate, and A. N. Popper, "A noisy spring: The impact of globally rising underwater sound levels on fish," *Trends Ecol. Evol.*, vol. 25, no. 7, pp. 419–427, Jul. 2010.

[3] R. Dekeling, M. L. Tasker, A. J. Van der Graaf, M. A. Ainslie, M. H. Andersson, M. André, J. F. Borsani, K. Brensing, M. Castellote, D. Cronin, and J. Dalen, "Monitoring guidance for underwater noise in European seas, Part II: Monitoring guidance specifications," Publications Office Eur. Union, Luxembourg, Tech. Rep. JRC88045, 2014, doi: 10.2788/27158.

[4] A. J. Van der Graaf, M. A. Ainslie, M. André, K. Brensing, J. Dalen, R. P. A. Dekeling, S. Robinson, M. L. Tasker, F. Thomsen, and S. Werner, "European marine strategy framework directive good environmental status (MSFD-GES): Report of the technical subgroup on underwater noise and other forms of energy," Publications Office Eur. Union, Brussels, Belgium, Tech. Rep., Feb. 2012.
[5] J. L. Miksis-Olds and S. M. Nichols, "Is low frequency ocean sound increasing globally?" *J. Acoust. Soc. Amer.*, vol. 139, no. 1, pp. 501–511, Jan. 2016.
[6] N. D. Merchant, K. L. Brookes, R. C. Faulkner, A. W. J. Bicknell, B. J. Godley, and M. J. Witt, "Underwater noise levels in UK waters," *Sci. Rep.*, vol. 6, no. 1, p. 36942, Dec. 2016.
[7] M. van der Schaar, M. A. Ainslie, S. P. Robinson, M. K. Prior, and M. André, "Changes in 63Hz third-octave band sound levels over 42months recorded at four deep-ocean observatories," *J. Mar. Syst.*, vol. 130, pp. 4–11, Feb. 2014.
[8] M. Mustonen, A. Klauson, M. Andersson, D. Clorennec, T. Folegot, R. Koza, J. Pajala, L. Persson, J. Tegowski, J. Tougaard, M. Wahlberg, and P. Sigray, "Spatial and temporal variability of ambient underwater sound in the baltic sea," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 13237.
[9] G. Zhang, T. N. Forland, E. Johnsen, G. Pedersen, and H. Dong, "Measurements of underwater noise radiated by commercial ships at a cabled ocean observatory," *Mar. Pollut. Bull.*, vol. 153, Apr. 2020, Art. no. 110948.
[10] S. Viola *et al.*, "Continuous monitoring of noise levels in the gulf of catania (Ionian Sea). Study of correlation with ship traffic," *Mar. Pollut. Bull.*, vol. 121, nos. 1–2, pp. 97–103, Aug. 2017.
[11] H. R. Kolar, E. P. McKeown, M. E. Purcell, P. J. Gaughan, A. G. Westbrook, M. G. Barry, A. M. Akhriev, J. P. Hayes, A. Castelfranco, G. Nolan, D. J. Murray, K. P. Adlum, and D. F. Glynn, "The design and deployment of a real-time wide spectrum acoustic monitoring system for the ocean energy industry," in *Proc. MTS/IEEE OCEANS Bergen*, Jun. 2013, pp. 1–4.
[12] J. Tegowski, R. Koza, I. Pawliczka, K. Skóra, K. Trzcińska, and J. Zdroik, "Statistical, spectral and wavelet features of the ambient noise detected in the Southern Baltic sea," in *Proc. 23rd Int. Congr. Sound Vibrat.*, Henrico, VA, USA: International Institute of Acoustics and Vibration (IIAV), 2016, pp. 1–6.
[13] L. Liu, L. Xiao, S.-Q. Lan, T.-T. Liu, and G.-L. Song, "Using petrel II glider to analyze underwater noise spectrogram in the south China sea," *Acoust. Aust.*, vol. 46, no. 1, pp. 151–158, Apr. 2018.
[14] A. Dassatti, M. van der Schaar, P. Guerrini, S. Zaugg, L. Houegnigan, A. Maguer, and M. Andre, "On-board underwater glider real-time acoustic environment sensing," in *Proc. OCEANS IEEE Spain*, Jun. 2011, pp. 1–8.
[15] D. M. Toma, I. Masmitja, J. del Río, E. Martinez, C. Artero-Delgado, A. Casale, A. Figoli, D. Pinzani, P. Cervantes, P. Ruiz, S. Memè, and E. Delory, "Smart embedded passive acoustic devices for real-time hydroacoustic surveys," *Measurement*, vol. 125, pp. 592–605, Sep. 2018.
[16] J. del Rio, D. M. Toma, T. C. O'Reilly, A. Broring, D. R. Dana, F. Bache, K. L. Headley, A. Manuel-Lazaro, and D. R. Edgington, "Standards-based plug & work for instruments in ocean observing systems," *IEEE J. Ocean. Eng.*, vol. 39, no. 3, pp. 430–443, Jul. 2014.
[17] E. Song and K. Lee, "Understanding IEEE 1451-networked smart transducer interface standard—What is a smart transducer?" *IEEE Instrum. Meas. Mag.*, vol. 11, no. 2, pp. 11–17, Apr. 2008.
[18] A. Bröring, J. Echterhoff, S. Jirka, I. Simonis, T. Everding, C. Stasch, S. Liang, and R. Lemmens, "New generation sensor Web enablement," *Sensors*, vol. 11, no. 3, pp. 2652–2699, Mar. 2011.
[19] E. Martínez, D. Toma, S. Jirka, and J. del Río, "Middleware for plug and play integration of heterogeneous sensor resources into the sensor Web," *Sensors*, vol. 17, no. 12, p. 2923, Dec. 2017.
[20] A. Bröring, P. Maué, K. Janowicz, D. Nüst, and C. Malewski, "Semantically-enabled sensor plug & play for the sensor Web," *Sensors*, vol. 11, no. 8, pp. 7568–7605, Aug. 2011.
[21] J. del Rio, D. M. Toma, E. Martinez, T. C. O'Reilly, E. Delory, J. S. Pearlman, C. Waldmann, and S. Jirka, "A sensor Web architecture for integrating smart oceanographic sensors into the semantic sensor Web," *IEEE J. Ocean. Eng.*, vol. 43, no. 4, pp. 830–842, Oct. 2018.
[22] U. K. Verfuß, M. Andersson, T. Folegot, J. Laanearu, R. Matuschek, J. Pajala, P. Sigray, J. Tegowski, and J. Tougaard, "Bias standards for noise measurements. Background information, guidelines and quality assurance," Baltic Sea Inf. Acoustic Soundscape (BIAS) Project, Tech. Rep., 2015.
[23] S. P. Robinson, P. A. Lepper, and R. A. Hazelwood, "Good practice guide for underwater noise measurement," Nat. Meas. Office, Marine Scotland, Crown Estate, Teddington, U.K., Tech. Rep. 133, 2014.
[24] M. Botts and A. Robin, "OGC SensorML: Model and XML Encoding Standard 2.0," Open Geospatial Consortium, Wayland, MA, USA, Tech. Rep. OGC 12-000, 2014. [Online]. Available: http://www.opengis.net/doc/IS/SensorML/2.0
[25] S. Cox, "Observations and Measuremets-XML Implementation," Open Geospatial Consortium, Wayland, MA, USA, Tech. Rep. OGC 10-025r1, 2011. [Online]. Available: http://www.opengis.net/doc/IS/OMXML/2.0
[26] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, and J. Bouwman, "The FAIR guiding principles for scientific data management and stewardship," *Sci. Data*, vol. 3, no. 1, pp. 1–9, Mar. 2016.
[27] T. Tanhua, S. Pouliquen, J. Hausman, K. O'brien, P. Bricher, T. De Bruin, J. J. Buck, E. F. Burger, T. Carval, K. S. Casey, and S. Diggs, "Ocean FAIR data services," *Frontiers Mar. Sci.*, vol. 6, p. 440, Sep. Aug. 2019.
[28] A. Bröring, C. Stasch, and J. Echterhoff, "OGC sensor observation service," Open Geospatial Consortium, Wayland, MA, USA, Tech. Rep. OGC 12-006, 2012. [Online]. Available: http://www.opengis.net/doc/IS/SOS/2.0
[29] S. Liang, C.-Y. Huang, and T. Khalafbeigi. (2016). *OGC SensorThings API Part 1: Sensing*. [Online]. Available: http://www.opengis.net/doc/is/sensorthings/1.0
[30] R. A. Simons. (2019). *ERDDAP*. Monterey, CA, USA. [Online]. Available: https://coastwatch.pfeg.noaa.gov/erddap
[31] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, 4th ed. Upper Saddle River, NJ, USA: Pearson, 2007.
[32] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.
[33] G. Heinzel, A. Rüdiger, and R. Schilling, "Spectrum and spectral density estimation by the discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows," Max-Planck-Institut fur Gravitationsphysik, Hannover, Germany, Tech. Rep., 2002.
[34] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Pearson, 2005.
[35] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoustics*, vol. 15, no. 2, pp. 70–73, Jun. 1967.
[36] K. Betke, T. Folegot, R. Matuschek, J. Pajala, L. Persson, J. Tegowski, and M. Wahlberg, "BIAS standards for signal processing. Aims, processes and recommendations. Amended version," Baltic Sea Inf. Acoustic Soundscape (BIAS) Project, Tech. Rep., 2015.
[37] T. Paolo, F. Cristiano, O. Alessandro, and C. Paola, "Semantic profiles for easing SensorML description: Review and proposal," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 8, p. 340, Jul. 2019.
[38] NERC. (2016). *The NERC Vocabulary Server: Version 2.0*. [Online]. Available: http://vocab.nerc.ac.uk/
[39] A. Kokkinaki, L. Darroch, J. Buck, and S. Jirka, "Semantically enhancing SensorML with controlled vocabularies in the marine domain," in *Proc. Geospatial Sensor Webs Conf.*, 2016, pp. 29–31.
[40] S. Guan, H. Moustahfid, A. Milan, and J. Mize. (2013). *Metadata Conventions for IOOS Passive Acoustic Recordings (v1.0)*. [Online]. Available: https://mmisw.org/ont/ioos/passive_acoustic_metadata
[41] M. Nilsson. (2000). *ID3*. [Online]. Available: https://id3.org/
[42] J. Del-Rio, M. Nogueras, D. M. Toma, E. Martinez, C. Artero-Delgado, I. Bghiel, M. Martinez, J. Cadena, A. Garcia-Benadi, D. Sarria, J. Aguzzi, I. Masmitja, M. Carandell, J. Olive, S. Gomariz, P. Santamaria, and A. M. Lazaro, "Obsea: A decadal balance for a cabled observatory deployment," *IEEE Access*, vol. 8, pp. 33163–33177, 2020.
[43] A. Novellino, P. D'Angelo, G. Benedetti, G. Manzella, P. Gorringe, D. Schaap, S. Pouliquen, and L. Rickards, "European marine observation data network—EMODnet physics," in *Proc. OCEANS Genova*, May 2015, pp. 1–3.
[44] N. D. Merchant, P. Blondel, D. T. Dakin, and J. Dorocicz, "Averaging underwater noise levels for environmental assessment of shipping," *J. Acoust. Soc. Amer.*, vol. 132, no. 4, pp. EL343–EL349, Oct. 2012.
[45] N. D. Merchant, T. R. Barton, P. M. Thompson, E. Pirotta, D. T. Dakin, and J. Dorocicz, "Spectral probability density as a tool for ambient noise analysis," *J. Acoust. Soc. Amer.*, vol. 133, no. 4, pp. EL262–EL267, Apr. 2013.
[46] N. Lanteri, J. Legrand, B. Moreau, J. R. Lagadec, and J. F. Rolin, "The EGIM, a generic instrumental module to equip EMSO observatories," in *Proc. OCEANS - Aberdeen*, Jun. 2017, pp. 1–5.

[47] T. J. Rouphael, *RF and Digital Signal Processing for Software-Defined Radio*, 1st ed. Oxford, U.K.: Elsevier, 2009.
[48] A. Robin, "OGC SWE common data model encoding standard," Open Geospatial Consortium, Wayland, MA, USA, Tech. Rep. OGC 08-094r1, 2011. [Online]. Available: http://www.opengis.net/doc/IS/SWE/2.0

ENOC MARTÍNEZ received the B.S. degree in telecommunications engineering and the M.S. degree in electronic engineering from the Universitat Politècnica de Catalunya (UPC), where he is currently pursuing the Ph.D. degree with the SARTI Research Group, Electronics Department. His research interests include interoperability, sensor web standards, underwater noise measurements, and underwater acquisition systems.

ALBERT GARCÍA-BENADÍ received the bachelor's degree in physical sciences from the University of Barcelona, in 2000. He is currently the Technical and Quality Manager of the UPC Metrology and Calibration Laboratory accredited under ISO 17025. His current research interests include acoustics measurement systems, atmospheric parameterization, meteorology radars, and oceanic environment.

DANIEL M. TOMA received the M.Sc. degree in electrical engineering from the Gheorghe Asachi Technical University, Iasi, Romania, in 2008, and the Ph.D. degree in electronic engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2012. His current research interests include marine instrumentation, ocean observing systems, wireless ad hoc networks, interoperability in sensor networks, synchronization and scheduling in sensor networks, energy harvesting techniques, and passive acoustics. He is currently a member of the Remote Acquisition Systems and Data processing (SARTI), Electronic Engineering Department, UPC.

ERIC DELORY (Senior Member, IEEE) has worked for 25 years in research and development of monitoring systems, including signal processing, instrumentation and machine learning in biomedical applications, and environmental monitoring. He joined the Oceanic Platform of the Canary Islands in 2010 as the Head of the observatory, where his main activities have consisted in new systems developments and integration. He was the coordinator of NeXOS, a European collaborative project developing compact cost-effective sensors for the monitoring of ocean variables from autonomous platforms. He is involved in several European research infrastructure integration initiatives (EMSO, JERICO, and GROOM), spanning fixed and mobile ocean observing platforms, to observe the ocean but also develop, test, and validate new sensor technologies. He is currently an Associate Editor of the IEEE JOURNAL OF OCEANIC ENGINEERING and *Frontiers in Marine Science* and has recently co-edited the book *Challenges and Innovations in Ocean In Situ Sensors: Measuring Inner Ocean Processes and Health in the Digital Age* (Elsevier, 2018).

SPARTACUS GOMÁRIZ (Member, IEEE) received the M.Sc. and Ph.D. degrees in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. He is currently an Associate Professor with the Department of Electronic Engineering, UPC, and a member of the Remote Acquisition Systems and Data Processing (SARTI). His research interests include linear and nonlinear control theory, gain scheduled control, fuzzy control, design of navigation, guidance, and control systems for underwater vehicles. He is also a member of IEEE Oceanic Engineering Society and Spanish Committee of Automation (CEA) of International Federation of Automatic Control (IFAC).

JOAQUÍN DEL-RÍO (Senior Member, IEEE) was born in Catalonia, Spain, in 1976. He received the B.Sc. and M.Sc. degrees in telecommunication engineering and electronic engineering from the Universitat Politecnica de Catalunya (UPC), Barcelona, Spain, in 1999 and 2002, respectively. Since 2001, he has been an Assistant Professor with the Electronic Engineering Department, UPC. He is currently a member of the Remote Acquisition Systems and Data Processing (SARTI), focusing his research on electronic instrumentation, interoperability in marine sensor networks, and wireless sensor networks.