IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2020, accepted January 9, 2020, date of publication January 13, 2020, date of current version January 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966254

Analysis of Template-Based Detection Algorithms for Inshore Bryde's Whale Short Pulse Calls

OLAYINKA O. OGUNDILE[1,2], (Member, IEEE), AND DANIEL J. J. VERSFELD[1], (Member, IEEE)
[1]Department of Electrical and Electronics Engineering, Stellenbosch University, Stellenbosch 7600, South Africa
[2]Department of Physics and Telecommunication, Tai Solarin University of Education, Ijebu Ode 2118, Nigeria

Corresponding author: Olayinka O. Ogundile (ogundileoo@gmail.com)

This work was supported in part by the National Research Foundation of South Africa (NRF) under Grant 116036.

ABSTRACT Marine mammals use sound for communication and echolocation within their ecosystems. The detection of these sounds is an important aspect of signal processing, such that we can estimate the spatial position and direction of arrival of these mammals, and have an understanding of their ecology. Passive acoustic monitoring (PAM) is widely used to understand marine mammal movement and vocal repertoire. In PAM, datasets are accumulated over days, months or years. Thus, it is impracticable to manually analyse the datasets because it is very large. This motivated the development of automated sound detection techniques for marine mammals, which most often varies depending on the vocal duration, frequency range and call type. In this paper, continuous recordings of Bryde's whale (*Balaenoptera edeni edeni*) short pulse calls ($< 3.1s$ long) were collected on a weekly basis from December 2018 to April 2019 on sighting of the individual in a single site in the endmost South-West of South Africa. The sound, previously not documented off South Africa, was observed on visual confirmation of the presence of inshore Brydes's whale. In addition, the paper develops and analyses two automated template-based detection algorithms for this short pulse call, employing dynamic time warping (DTW) and linear predictive coding (LPC) techniques. These proposed template-based detectors are novel, as they have not being previously used in Bryde's whale sound detection in the literature. When applied to the continuous recordings of the short pulse calls, the DTW-based and LPC-based detection algorithms obtained a sensitivity of 96.04% and 97.14% respectively for high signal-to-noise ratio (about $10dB$ above the ambient sound). Otherwise, for low SNR, the DTW-based and LPC-based detection algorithms obtained a sensitivity of 94.98% and 96.00% respectively. These detection algorithms exhibit low computational time complexity and can be modified to analyse the movement of obscure but vocal marine species instead of manual identification.

INDEX TERMS Bryde's whale, DTW, LPC, PAM, pulse call, sound detection.

I. INTRODUCTION

Over the years, increased human marine activities such as fishery and shipping have threatened the ecosystems of marine mammals [1]–[5]. As a result of these anthropogenic impacts, it is difficult to make informative decisions about the movement and spatial position of marine species [3]. Also, since marine mammals spend most of their time below water, it is difficult to visually observe and monitor these marine mammals. Therefore, passive acoustic monitoring (PAM) provides a valuable modality for study of marine mammal movement and distribution because (1) animals are very vocal, and (2) sound propagates much further in water. Besides, PAM can be used to collect datasets in outlying areas over days, months or years. More importantly, PAM is used in unfavourable weather conditions and it is suitable for the tracking of highly mobile marine mammals such as cetaceans [3], [6].

Bryde's whales, also referred to as Eden's whales are species of the order Cetacea. They are currently grouped as a single species called *Balaenoptera edeni* (*B. edeni*), where two subspecies have been suggested: *Balaenoptera edeni edeni* (*B. e. edeni*) and *Balaenoptera edeni brydei* (*B. e. brydei*). The *B. e. edeni* is the small, coastal form

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

of the *B. edeni* while *B. e. brydei* refers to the large, offshore form [6]–[9]. Several studies have been recently carried out on the existence of Bryde's whales from different geographic location, such as the Gulf of Mexico [9]–[11], Gulf of California [12], [13], Hauraki Gulf, New Zealand [3], [14], [15], Eastern Tropical Pacific [6], [16], and Southern Brazil [17]–[20]. While some of these literatures have described the population, spatial distribution, genetic and phylogenetic features, attributed nomenclature, a few have described the potential vocal repertoire of this marine species. In [10] and [12], the vocal repertoire of Bryde's whales in the Gulf of Mexico and Gulf of California respectively are described. It is observed that the recorded Bryde's whale calls from this region ranges from 50Hz to 1200Hz. However, aside describing the Bryde's whale vocal repertoire, the authors in [10] and [12], do not identify method(s) to automatically detect these established calls as proposed in this paper.

Historically, in the extreme South-West of South Africa, *Olsen* [21] described a new species of whale. He named them Bryde's whale after Mr. Johan Bryde who was at that time the Norwegian consul to South Africa. Although, in *Olsen* [21], the Bryde's whales harvested off South Africa was labelled *B. brydei*, it was later revealed in *Best* [22] that there are two allopatric forms of Bryde's whale off South Africa. Subsequently, it was affirmed that the *B. brydei* described by *Olsen* [21] include characteristics from the inshore and offshore forms of *B. edeni* [23]–[25]. Much recently, different literature have described the genetic and phylogenetic features of Bryde's whales off South Africa [26], [27]. However, no work has previously documented the Brdye's whales calls off South Africa. In this paper, we analyse a continuous recording of inshore Bryde's whale (*B. e. edeni*) short pulse calls collected on sighting of the individual in a single site in the endmost South-West of South Africa. This observed pulse call is previously undocumented for *B. e. edeni* off South Africa and it can serve as an important contribution to the study of Bryde's whales vocal repertoire off South Africa. In addition, we present the characteristic of the recorded short pulse call. Similar to other studies [10], the call is identified by observation and matching the spectral and temporal features described in closely related studies such as in [12], which was conducted in the Gulf of California.

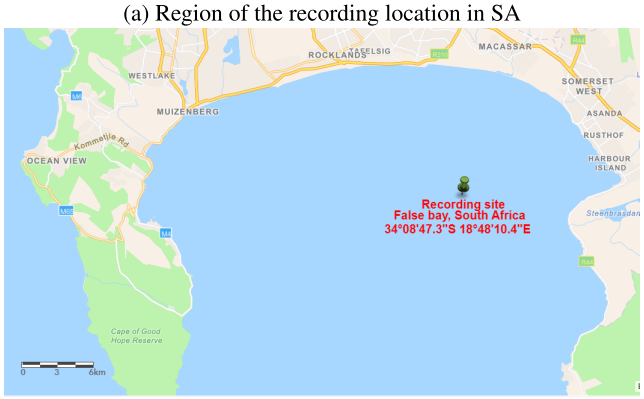
The datasets containing the *B. e. edeni* short pulse calls was accumulated over months. Therefore, it is impracticable to manually analyse all the collected datasets. In this regard, the paper develops two automated template-based detection algorithms for the short pulse call, employing dynamic time warping (DTW) [28] and linear predictive coding (LPC) [29], [30] techniques. Template-based detectors automatically recognise unknown sound signals only when a set of the signal is manually identified by an expert. This detection technique is widely used for sound detection in digital signal processing. As such, it has generally been adapted in animal vocalisation detection [31]–[33]. However, we emphasise that the template-based detection techniques have not been used in Brdye's whales sound detection in the literature. Thus, these proposed template-based algorithms for the observed Brydes's whale short pulse call is innovative and it produces good detection accuracy (sensitivity). As discussed in Section VI, when the proposed template-based detectors are applied to the continuous recordings of the Bryde's whale short pulse calls, the DTW and LPC template-based detectors obtained a sensitivity of 96.04% and 97.14% respectively for high signal-to-noise ratio (*snr*), depending on an empirically determined reliability value (Γ). On the other hand, the accuracy of the DTW and LPC template-based detectors decrease to 94.98% and 96.00% respectively as the background noise increase (the background noise is mostly due to bad weather conditions during recordings).

The rest of the paper is structured as follows. Section II describes the recording location, PAM set-up and the datasets preprocessing phase. In Section III, we explain the characteristics of the *B. e. edeni* short pulse call with standard parameters. Section IV briefly explains the two signal processing techniques used in the detection algorithms. The developed template-based detection algorithms are discussed in Section V. Section VI analyses the results of the detection algorithms for some specified parameters. The paper is concluded in Section VII.

II. RECORDINGS AND PREPROCESSING

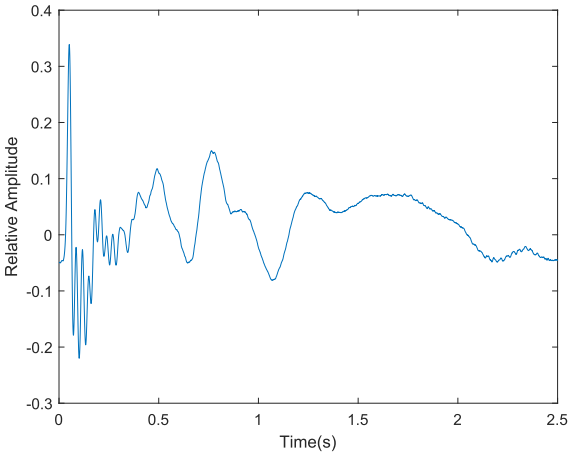
From December 2018 to April 2019, recordings were collected on a weekly basis to study *B. e. edeni* calls in an area of approximately $13km^2$, close to Gordon's bay harbour, False bay (34°08'47.3''S 18°48'10.4''E), South-West, South Africa, as shown in Fig. 1. In Fig. 1a, we show the region where the recordings is carried out in South Africa, while Fig. 1b shows the exact location where the recording is carried including the coordinates. The depth at the recording site was less than 30 meters. During these recordings, standard protocols were strictly adhered to as sanctioned by the Department of Environmental Affairs, South Africa. For instance, we kept the required minimum distance on sighting of the individual. The individual was identified each time it was sighted based on its features as discussed in [26]. Most times, recordings were carried out when a single individual is sighted. The individual was mature, but for all cases we could not identify the sex, whether it is a male or female. Note, some times, more than one individuals were sighted but no recording were carried out in such situations. The recordings is carried out using dipping hydrophones. In the set-up, a hydrophone (Aquarian Audio H2A-XLR Hydrophone with sensitivity $-180dB$ re: $1V/\mu Pa$ and frequency range from $10Hz$ to $100kHz$) was connected to a Zoom H1N recorder, operating at $96ksps$ at 24 bit resolution. Dataset was saved as raw samples (.wav format) in order to preserve the phase and amplitude data as best as possible. The deployment was to dip the hydrophone from a sail boat ($8m$ long with inboard engine), under varying conditions. That is, sailing (2-$4kts/h$), dropping the sails (less than $1kt/h$),

(a) Region of the recording location in SA

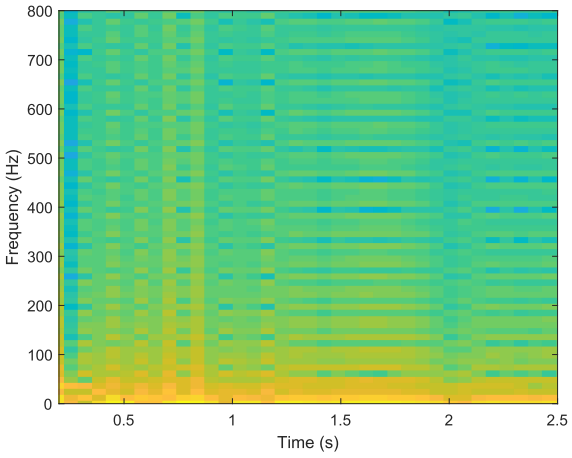


(b) Coordinates of the exact location in SA

FIGURE 1. Recording site in the South-West of South Africa (SA) (Map produced using http://www.arcgis.com.).



(a) Time series



(b) Spectrogram (parameters: f_s=96000, nfft=7750, overlap=75%)

FIGURE 2. Representation of the *B. e. edeni* pulse call.

heaving to (less than $2kts/h$). Before recordings, the engine of the boat is shut off. Also, recordings were done when the *B. e. edeni* was the only species sighted. Thus, the raw samples were filtered with a $3rd$ order Butterworth bandpass filter in MATLAB, R2018b. The filter eliminate frequencies under $10Hz$ and above $8000Hz$ in order to reduce background noise and the DC component. Fig. 2 depicts a filtered sample of the short pulse call (note, we presented the time series and spectrogram representation of the pulse call before filtering using Sonic Visualiser in Fig. 3, as the spectrogram of pulse calls can be more clearly viewed in Fig. 3 in comparison to Fig. 2b). Similarly, the resulting signal was analysed using MATLAB, where different characteristics explaining the main component of the pulse calls were extracted as discussed in Section III.

III. FEATURES OF THE BRYDE'S WHALE SHORT PULSE CALL

During the five months recordings, one day per week of approximately two hours recordings, about fifteen different dataset were collected. A single recurrence call was observed which correspond to virtually all the vocalisation on sighting of the individual. Aside sighting the *B. e. edeni* whale physical features which corresponds to the descriptions in *Penry et al* [26], we are optimistic that this call is produced by the *B. e. edeni*. Firstly, Bryde's whales were historically sighted in this region by *Olsen* [21]. Also during recordings, no other whales or calves were cited within radius of the recording site. That is, no other cetaceans were in a radius of about 3 nautical miles. Besides, the call has been observed when we visually confirmed the presence of inshore Bryde's whales (that is, the Bryde's whale were in radius of $1NM$ when the sounds were observed). To further verify the call, we carried out comparative tests when Bryde's whales were not in the vicinity, and this "short pulse calls" were not observed. We assume that this call is probably used by the *B. e. edeni* for hunting or navigation since there are no other calves present during recordings. The pulse call has a small relative amplitude at the start, which increases to a maximum of 0.36 or minimum of -0.39, and decays rapidly as the call ends. The relative amplitude range of the Bryde's whale pulse calls is much larger than other major sources of biological noise in bays such as the snapping shrimp sound that can be misrepresented as the Bryde's whale sound. Also, in the frequency domain, the minimum frequencies of the pulse calls range between $0.07 \pm 0.02kHz$, while its highest

TABLE 1. Characteristics of *B. e. edeni* short pulse call recorded in the South-West of South Africa.

Call	F_{min} (kHz)	F_{max} (kHz)	A_{min}	A_{max}	C_d (s)
Short pulse	0.07 ± 0.02	1.0 ± 0.2	-0.39	0.36	$1.2 - 3.1$

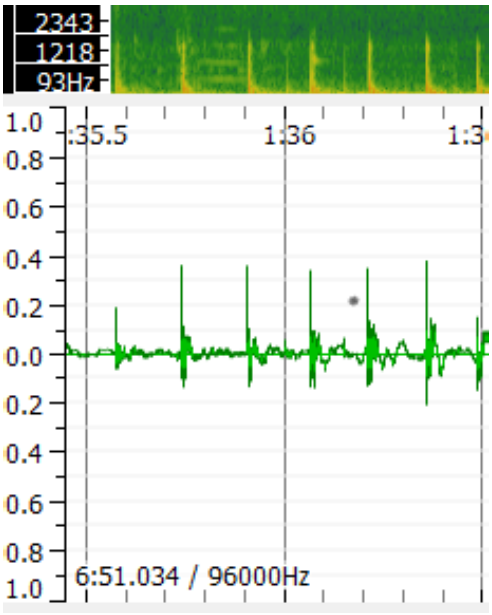


FIGURE 3. Time series and spectrogram representation of the *B. e. edeni* pulse call.

frequencies range between $1.0 \pm 0.2kHz$. Other characteristic describing the main component of the call is summarised in Table 1, where F_{min} and F_{max} are the minimum and maximum frequencies of the call respectively, A_{min} and A_{max} are the minimum and maximum relative amplitude respectively, and C_d is the call duration in seconds (s).

IV. DETECTION TECHNIQUES

A. DYNAMIC TIME WARPING

DTW introduced in *Itakura* [28] has been widely used in speech recognition, gesture recognition, medicine, data mining, and manufacturing [34]. Likewise, this technique have been used in marine mammal sound detection and classification [32], [35]. The DTW algorithm is used to find the best possible alignment between two time sequences, utilising the temporal distortions between the sequences. The time sequences are efficiently warped in a non-linear manner to match each other. For example, given two time series S_1 and S_2, of length i and j respectively, the DTW algorithm align the two time series by constructing an $i \times j$ matrix as [28]:

$$D = min \begin{pmatrix} D[i-1, j-1] \\ D[i-1, j] \\ D[i, j-1] \end{pmatrix} + |S_{1_i} - S_{1_j}|, \tag{1}$$

where each element in D represent the similarities between the two time series S_1 and S_2 at positions i and j respectively. In this paper, we define the difference between any two time series signal as $D_{i^{th},j^{th}}$. That is, the value of the element at the i^{th} and j^{th} position of the difference matrix D. For more information on the DTW technique, refer to [28], [34], [36].

B. LINEAR PREDICTIVE CODING

LPC, often referred to as inverse filtering have been successfully used for speech coding, synthesis and recognition. In addition, it has been used to analyse short length of marine mammal vocal signals [37]. The concept of LPC is to calculate an approximated value of the current speech sample $\hat{g}(\phi)$ by a linear combination of the preceding regenerated θ^{th} samples as [29], [30], [38]:

$$\begin{aligned} \hat{g}(\phi) &= \alpha_1 g(\phi - 1) + \alpha_2 g(\phi - 2) + \cdots + \alpha_\beta g(\phi - \theta) \\ &= \sum_{\beta,\theta=1}^{\psi} \alpha_\beta g(\phi - \theta), \end{aligned} \tag{2}$$

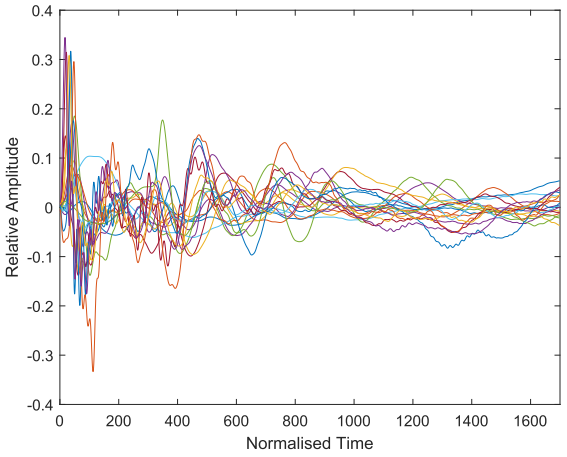
where α_β is the filter coefficients, $\hat{g}(\phi)$ is the approximated value of $g(\phi)$, $g(\phi - \theta)$ is the preceding θ^{th} samples, and $\psi = \beta^{th} = \theta^{th}$ is the polynomial order or the number of filter coefficients. These distinctive set of coefficients α_β can be calculated by minimising the sum of the squared differences between the linearly estimated samples and the original samples as defined in (3) [29], [30], [38]:

$$\epsilon(\phi) = g(\phi) - \hat{g}(\phi) = g(\phi) - \sum_{\beta,\theta=1}^{\psi} \alpha_\beta g(\phi - \theta), \tag{3}$$

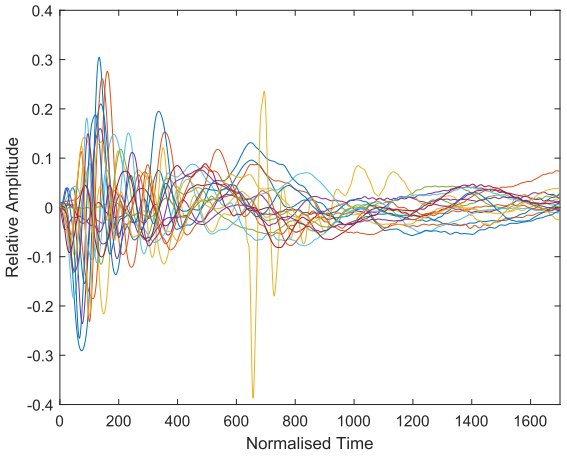
where $\epsilon(\phi)$ is the error between $g(\phi)$ and $\hat{g}(\phi)$. The coefficients α_β can be determined from (3) using the autocorrelation method. Typically, the number of coefficients ranges from 10-14. In this paper, we assume $\psi = \beta^{th} = \theta^{th} = 12$. This implies that the filter coefficients is a 12 order polynomial defined as $P(\alpha_\beta)$. See [29], [30] for more discussion on the LPC technique.

V. TEMPLATE-BASED DETECTION ALGORITHMS

In developing the detectors, some of the short pulse calls were manually identified from the datasets, recorded on different days to form the templates. These short pulse calls are identified from a small section of the dataset while the remaining section (the larger section) of the dataset is used to verify the performance of the detector. Fig. 4 shows some of the visually identified short pulse calls from two different days. The identified short pulse calls from each day are termed Template $\mathcal{A}$ $(T_\mathcal{A})$ and Template $\mathcal{B}$ $(T_\mathcal{B})$. Two templates were chosen to verify the performance of the developed detection algorithms for change in background noise. The recordings where the samples in $T_\mathcal{A}$ are identified has an average snr of $+3.84dB$ better in comparison to $T_\mathcal{B}$. Each of the template contains k number of samples of l varying lengths.



(a) $T_{\mathcal{A}}$



(b) $T_{\mathcal{B}}$

FIGURE 4. Signal waveform of *B. e. edeni* short pulse call.

From Table 1, observe that the short pulse call duration (C_d) is between $1.2 - 3.1s$. As such, the templates are specified to contain samples of different length l in the range of the C_d. Expressing each sample in the template as a row vector, we define the template as:

$$T_{\mathcal{A}||\mathcal{B}} = \begin{matrix} [t_{1,1}, t_{1,2}, \dots t_{1,C_d}] \\ [t_{2,1}, t_{2,2}, \dots t_{1,C_d}] \\ \vdots \quad \vdots \quad \vdots \\ [t_{k,1}, t_{k,2}, \dots t_{k,C_d}] \end{matrix}, \tag{4}$$

where t is the sampling point, k is the number of sample in a template. The value of k is chosen to be 18, 12 and 6 as shown in Section VI. The length of each sample in the template varies depending on the value of C_d. The template-based detectors is thus expatiated as follows.

A. DTW-BASED

As mentioned, the DTW algorithm is used to find the similarities between different time series waveform. Thus, in this detection algorithm, the template $T_{\mathcal{A}||\mathcal{B}}$ with k number of manually identified samples is warped with each other using (1) to form a $k \times k$ dissimilarities template matrix defined as:

$$\mathcal{T} = \begin{bmatrix} 0 & \mathcal{D}_{1,2} & \dots & \mathcal{D}_{1,k} \\ \mathcal{D}_{2,1} & 0 & \dots & \mathcal{D}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}_{k,1} & \mathcal{D}_{k,2} & \dots & 0 \end{bmatrix}, \tag{5}$$

where each element in $\mathcal{T}$ is the $D_{i^{th},j^{th}}$ similarities between each sample in the template. Subsequently, the algorithm finds the maximum entry of each column of $\mathcal{T}$ to form a $1 \times k$ row vector defined as:

$$\mathcal{T}_{max} = \left[\mathcal{T}_{max,1}\ \mathcal{T}_{max,2}\ \mathcal{T}_{max,3}\ \dots\ \mathcal{T}_{max,k}\right]. \tag{6}$$

The detection process then starts by sliding through the recordings $\mathcal{R}_{\mathcal{D}}$ with a defined window size (w) and overlapping size (o_v). Having known the C_d of most of the short pulse calls, we set $w = 1s$ and $o_v = w/2$. For each selected window, the algorithm calculates a relative energy of the waveform as defined as:

$$E = \sum_{1}^{C_d} (t_{C_d})^2, \tag{7}$$

where t is defined as above. Thus, a δ value is empirically set. Matching the values of δ and E, any selected window with an E lower than δ is not considered as the *B. e. edeni* pulse call. This significantly reduces the computational time complexity of the proposed detection algorithm. Onward, the similarities between the window frame with $E \geq \delta$, and the k samples in the template is computed using (1) to form a $1 \times k$ row vector defined as:

$$\mathcal{T}_w = \left[\mathcal{T}_{w_1} \quad \mathcal{T}_{w_2} \quad \mathcal{T}_{w_3} \quad \dots \quad \mathcal{T}_{w_k}\right], \tag{8}$$

where each element in $\mathcal{T}_w$ is the $D_{i^{th},j^{th}}$ similarities between the selected window frame and each sample in the template. Thereafter, the algorithm finds a count score (γ) by matching (8) and (6). That is, γ is computed by counting the number of times the value in each column of $\mathcal{T}_w$ is less or equal to the value in the corresponding column of $\mathcal{T}_{max}$. The value of γ is therefore compared with a predetermined reliability value Γ. This value of Γ determines the performance of the proposed detector. A small value of Γ increases the sensitivity of the detector with the price of increased false positive rate as is subsequently shown. With this in mind, a trade-off value should be defined for Γ based on observations. Results are presented in Section VI for different values of Γ. Note, Γ ranges between $0 - 1$. If $\gamma \geq \lfloor \Gamma * \gamma \rfloor$, the window size w is stored as the pulse call of the *B. e. edeni* whale.

From Table 1, the C_d ranges from 1.2-3.2s; thus, the algorithm synchronises every stored w in the range of the C_d. The algorithm achieves this by comparing the previously detected pulse call with the current one as:

$$\mathcal{O} = e_{w_p} - b_{w_c}, \tag{9}$$

where w_p and w_c is the previous and current detected pulse call respectively, e_{w_p} is last sampling point of w_p, and b_{w_c} is the first sampling point of w_c. If $\mathcal{O} \neq N_{o_v} - 1$, w_p is not synchronised with w_c (N_{o_v} is number of sampling points in o_v). Otherwise, it is synchronised as:

$$w_c = w_p + o_v. \tag{10}$$

The template-based DTW detector is hereby summarised in algorithm 1.

Algorithm 1 DTW Detection Algorithm

Input: $T_{\mathcal{A}||\mathcal{B}}$, k, $\mathcal{R}_{\mathcal{D}}$, w, o_v, δ, Γ, $w_p = 0$, $e_{w_p} = 0$
Output: w_c *as detected pulse call*
1: *build* $k \times k$ $\mathcal{T}$ *matrix based on (5)*
2: *find* $\mathcal{T}_{max}$ *based on (6)*
3: *choose window size, w,* o_v*: Slide through* $\mathcal{R}_{\mathcal{D}}$
4: *calculate E for each w*
5: $*$ ***if*** $E < \delta$
return to **3**
6: ***else***
7: *compute* $\mathcal{T}_w$ *based on (8)*
8: *match* $\mathcal{T}_{max}$ *and* $\mathcal{T}_w$ *to find* γ
9:• ***if*** $\gamma < \lfloor \Gamma * \gamma \rfloor$
return to **3**
10: ***else***
11: *store* w_c, *and* b_{w_c}
12: • ***end if***
13:$*$ ***end if***
14: ○ ***if*** $e_{w_p} - b_{w_c} \neq N_{o_v} - 1$
15: *Detect* w_c
16: ***else***
17: *Detect* $w_c = w_p + o_v$
18: *store* $w_p = w_c$, $e_{w_p} = b_{w_c}$
19: ○ ***end if***
return to **3**

Of note, in some cases, the duration of the manually identified calls and the automatic detected call differs in duration, such that the automatically detected call is slightly longer. In such situation, the automatic detected waveform duration can be synchronised to approximately match the manually identified waveform duration. The detected waveform can be divided into smaller windows s_w ($s_w \ll w$). Subsequently, (7) can be computed for these s_w waveforms. The result can be matched with a small δ (s_δ) ($s_\delta \ll \delta$), where s_δ is determined empirically. Doing this, the two ends of the detected waveform can be synchronised to fit an approximate of the manually identified call duration.

B. LPC-BASED

The LPC-based detector is developed from the filter coefficients α_β produced using (3). The detector first find the *roots* of the filter coefficient polynomial $P(\alpha_\beta)$ to obtain a $1 \times \psi - 1$ row vector as:

$$\mathcal{R} = \left[\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \ldots, \mathcal{R}_{\psi-1}\right], \tag{11}$$

where the elements of $\mathcal{R}$ are often complex numbers. Note that the *roots* of $P(\alpha_\beta)$ can be derived using different mathematical methods as presented in *Jia* [39]. Therefore, $\mathcal{R}$ is computed for the template $T_{\mathcal{A}||\mathcal{B}}$ to obtain a $k \times \psi - 1$ matrix defined as:

$$\mathcal{R}_{T_{\mathcal{A}||\mathcal{B}}} = \begin{bmatrix} \mathcal{R}_{1,1} & \mathcal{R}_{1,2} & \cdots & \mathcal{R}_{1,\psi-1} \\ \mathcal{R}_{2,1} & \mathcal{R}_{2,2} & \cdots & \mathcal{R}_{2,\psi-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}_{k,1} & \mathcal{R}_{k,2} & \cdots & \mathcal{R}_{k,\psi-1} \end{bmatrix}. \tag{12}$$

Each entry in $\mathcal{R}_{T_{\mathcal{A}||\mathcal{B}}}$ is subsequently compared with a defined complex reference point r_p ($r_p = 0 + i0$) to find the euclidean distance. This forms a corresponding $k \times \psi - 1$ distance template matrix defined as:

$$\mathrm{T} = \begin{bmatrix} \mathcal{D}_{\mathcal{R}_{1,1}} & \mathcal{D}_{\mathcal{R}_{1,2}} & \cdots & \mathcal{D}_{\mathcal{R}_{1,\psi-1}} \\ \mathcal{D}_{\mathcal{R}_{2,1}} & \mathcal{D}_{\mathcal{R}_{2,2}} & \cdots & \mathcal{D}_{\mathcal{R}_{2,\psi-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}_{\mathcal{R}_{k,1}} & \mathcal{D}_{\mathcal{R}_{k,2}} & \cdots & \mathcal{D}_{\mathcal{R}_{k,\psi-1}} \end{bmatrix}. \tag{13}$$

Thus, the algorithm finds the maximum entry of each column of T to form a $1 \times \psi$-1 row vector defined as:

$$\mathrm{T}_{max} = \left[\mathcal{T}_{max,1} \quad \mathrm{T}_{max,2} \quad \mathrm{T}_{max,3} \quad \ldots \quad \mathrm{T}_{max,\psi-1}\right]. \tag{14}$$

The detection process continues in a similar way as in the DTW detector using the same set of parameters (w, o_v, δ, Γ). However, (8) used in the DTW detector is computed in this case by comparing the *roots* of the selected sample of window w to r_p. In this way, we obtain a $1 \times \psi - 1$ distance sample row vector defined as:

$$\mathrm{T}_w = \left[\mathcal{T}_{w_1} \quad \mathrm{T}_{w_2} \quad \mathrm{T}_{w_3} \quad \ldots \quad \mathrm{T}_{w_\psi - 1}\right]. \tag{15}$$

The template-based LPC detector is summarised in Algorithm 2.

Algorithm 2 LPC Detection Algorithm

Input: $T_{\mathcal{A}||\mathcal{B}}$, k, $\mathcal{R}_{\mathcal{D}}$, w, o_v, δ, Γ, $w_p = 0$, $e_{w_p} = 0$, r_p, ψ
Output: w_c *as detected pulse call*
1: *compute* $\mathcal{R}$ *using (2), (3) and (11)*
2: *compute* $\mathcal{R}_{T_{\mathcal{A}||\mathcal{B}}}$
3: *Build* $k \times \psi - 1$ T *matrix based on* r_p
4: *find* T_{max} *based on (14)*
5: *similar process as in Algorithm 1, steps 3-18*

VI. TEST RESULTS AND DISCUSSION

In this section, the proposed template-based detectors were applied to recognise continuous recordings of the short pulse calls. In the results presented, we verified the performance of the detectors for different values of k. The reliability Γ is also a factor we used in the result comparisons. In addition, we evaluate the quality of detection algorithms by evaluating the detection sensitivities $\mathcal{S}$, false positive rates F_p and failure rates $\mathcal{F}$ of both methods as defined in (16) [40]:

$$\mathcal{S} = \frac{\eta}{\eta + \rho}, \quad F_p = \frac{\tau}{\eta + \tau} \text{ and } \mathcal{F} = 1 - \mathcal{S}, \tag{16}$$

TABLE 2. Detectors performance as a function of k: $T_{\mathcal{A}}$, $\Gamma = \frac{4}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	89.12	91.18	1.88	1.64	10.88	8.82
12	92.74	92.98	1.88	1.21	7.26	7.02
18	93.10	94.25	1.88	1.21	6.90	5.75

TABLE 3. Detectors performance as a function of k: $T_{\mathcal{B}}$, $\Gamma = \frac{4}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	88.25	90.00	3.45	3.12	11.75	10.00
12	91.00	91.00	3.20	2.63	9.00	9.00
18	91.25	92.50	3.20	2.63	8.75	7.50

TABLE 4. Detectors performance as a function of k: $T_{\mathcal{A}}$, $\Gamma = \frac{3}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	90.30	92.36	4.22	3.94	7.64	9.70
12	93.42	94.00	3.96	3.03	7.26	6.00
18	94.83	95.76	3.96	3.03	5.17	4.24

TABLE 5. Detectors performance as a function of k: $T_{\mathcal{B}}$, $\Gamma = \frac{3}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	90.42	91.84	5.32	4.92	9.58	8.16
12	92.00	92.24	5.06	4.44	8.00	7.76
18	93.28	94.66	5.06	4.44	6.90	5.34

TABLE 6. Detectors performance as a function of k: $T_{\mathcal{A}}$, $\Gamma = \frac{2}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	93.00	94.42	8.88	8.48	7.00	5.58
12	94.78	96.04	8.66	7.80	5.22	3.96
18	96.04	97.14	8.66	7.80	3.96	2.86

TABLE 7. Detectors performance as a function of k: $T_{\mathcal{B}}$, $\Gamma = \frac{2}{6}$.

k	$\mathcal{S}$ (%)		F_p (%)		$\mathcal{F}$ (%)	
	DTW	LPC	DTW	LPC	DTW	LPC
6	91.94	93. 10	9.96	9.04	8.06	6.90
12	93.32	93.32	9.18	8.46	6.68	6.68
18	94.98	96.00	9.18	8.46	5.02	4.00

where η is the number of times the manually detected short pulse call matches the output of the automatic detectors, ρ is the number of times the proposed detectors missed the manually detected pulse calls, and τ is the number of times the proposed detectors wrongly recognised a signal as the pulse call. In all cases, a high value of $\mathcal{S}$ is desirable to rate the accuracy of any sound detection technique. This in turns indicate a small value of $\mathcal{F}$ as shown in (16). More so, the smaller the value of F_p, the more dependable is the detection algorithm.

As earlier mentioned, we verified the performance of the detectors for varying weather conditions (background noise). Template $\mathcal{A}$ ($T_{\mathcal{A}}$) comprises of less noisy samples while $T_{\mathcal{A}}$ contain samples with more background noise. Table 2-7, shows the performance of the detectors as a function of k for different empirically selected values of Γ. In Table 2 and 3, a $\Gamma = \frac{4}{6}$ was used in both detection algorithms. Firstly, as k increases, the performance of the detectors improve linearly. The performance of the detectors for the two templates differ with $T_{\mathcal{A}}$ (Table 2) obtaining a superior performance in comparison to $T_{\mathcal{B}}$ (Table 3). This implies that the lower the noise in the identified samples used in the template, the better the performance of the detectors. Filtering can be a better way of reducing the noise as done during preprocessing but it cannot eliminate all noise components. Moreover, as shown in Table 2 and 3, the LPC-based detectors is a more robust recogniser as compared to the DTW-based detector as it offers better performance in terms of $\mathcal{S}$, F_p and $\mathcal{F}$.

In Table 4 and 5, the value of Γ ($\Gamma = \frac{3}{6}$) was reduced in both detection algorithms. As shown, the $\mathcal{S}$ of the algorithms increase with corresponding reduction in $\mathcal{F}$. However, F_p increase with a reduction in the value of Γ. Likewise in Table 6 and 7, as Γ ($\Gamma = \frac{2}{6}$) reduces further, the $\mathcal{S}$ increases while F_p also increase in both algorithms. An increase in F_p implies that the value of τ will increase as a result of a decrease in Γ (that is, $\Gamma \propto \frac{1}{\tau}$). Although η also increase as the value of Γ decreases, this increase is not as significant in comparison to the increase in the false positive (τ) calls detected. In real time, the false positive rate F_p of the detector must be as low as possible while maintaining a high level of detection accuracy (sensitivity). This means that there is a trade-off between $\mathcal{S}$ and F_p in empirically determining the value of Γ. Thus, in real time, Γ should be chosen depending on application requirements. Summarily, irrespective of the value of Γ, $T_{\mathcal{A}}$ (Table 2, 4 and 6) performs better than $T_{\mathcal{B}}$ (Table 3, 5 and 7) respectively, and the LPC-based detector offered superior performance in comparison to the DTW-based detector. Both detection algorithms exhibit low computational time complexity of order $O(\mathcal{R}_{\mathcal{D}})$ and can be used in real time to analyse the movement of obscure but vocal marine species instead of using traditional methods. In addition, the developed detectors can be modified to recognise different marine mammal sounds, where parameters such as w, o_v, δ, and Γ can be selected based on observation of the sound waveforms and application requirements.

VII. CONCLUSION

The paper identified a short pulse call of a *B. e. edeni* whale off South-West South Africa. The behavioural patterns of the *B. e. edeni* whale is quite difficult to obtain. Thus, the recognition of this call is a noteworthy contribution to the knowledge of this species off South Africa and the world at large. In addition, two template-based detection algorithms were developed for this identified short pulse call, employing DTW and LPC techniques. Both algorithms were shown to demonstrate high sensitivity with reduced false positive rate. But, the LPC-based detector is a more robust recogniser as compared to the DTW-based detector. Besides, the developed

detection algorithms can be used in real time vocalisation detection because they both offer low computational time complexity. Moreover, the algorithms can be modified to analyse the movement of different obscure but vocal marine species instead of manual identification.

REFERENCES

[1] B. S. Halpern, "A global map of human impact on marine ecosystems," *Science*, vol. 319, no. 5865, pp. 948–952, 2018.

[2] R. L. Putland, N. D. Merchant, A. Farcas, and C. A. Radford, "Vessel sound cuts down communication space for vocalising fish and marine mammals," *J. Acoust. Soc. Amer.*, vol. 143, no. 3, p. 1898, Mar. 2018.

[3] R. Putland, L. Ranjard, R. Constantine, and C. Radford, "A hidden Markov model approach to indicate Bryde's whale acoustics," *Ecol. Indicators*, vol. 84, pp. 479–487, Jan. 2018.

[4] M. Soldevilla, J. Hildebrand, K. Frasier, L. A. Dias, A. Martinez, K. Mullin, P. Rosel, and L. Garrison, "Spatial distribution and dive behavior of Gulf of Mexico Bryde's whales: Potential risk of vessel strikes and Fisheries interactions," *Endangered Species Res.*, vol. 32, pp. 533–550, Jun. 2017.

[5] C. L. Moloney, S. T. Fennessy, M. J. Gibbons, A. Roychoudhury, F. A. Shillington, B. P. Von Der Heyden, and K. Watermeyer, "Reviewing evidence of marine ecosystem change off South Africa," *Afr. J. Mar. Sci.*, vol. 35, no. 3, pp. 427–448, Sep. 2013.

[6] R. Constantine, T. Iwata, S. L. Nieukirk, and G. S. Penry, "Future directions in research on Bryde's whales," *Frontiers Mar. Sci.*, vol. 5, p. 333, Sep. 2018.

[7] Committee on Taxonomy, Society of Marine Mammalogy. (2017). *List of Marine Mammal Species and Subspecies*. Accessed: Apr. 2019. [Online]. Available: https://www.marinemammalscience.org

[8] F. Kershaw, M. S. Leslie, T. Collins, R. M. Mansur, B. D. Smith, G. Minton, R. Baldwin, R. G. Leduc, R. C. Anderson, R. L. Brownell, and H. C. Rosenbaum, "Population differentiation of 2 forms of Bryde's whales in the Indian and Pacific Oceans," *J. Heredity*, vol. 104, no. 6, pp. 755–764, 2013.

[9] P. Rosel and L. Wilcox, "Genetic evidence reveals a unique lineage of Bryde's whales in the northern Gulf of Mexico," *Endangered Species Res.*, vol. 25, no. 1, pp. 19–34, Jul. 2014.

[10] A. N. Rice, K. J. Palmer, J. T. Tielens, C. A. Muirhead, and C. W. Clark, "Potential Bryde's whale (Balaenoptera edeni) calls recorded in the northern Gulf of Mexico," *J. Acoust. Soc. Amer.*, vol. 135, no. 5, pp. 3066–3076, May 2014.

[11] A. Širović, H. R. Bassett, S. C. Johnson, S. M. Wiggins, and J. A. Hildebrand, "Bryde's whale calls recorded in the Gulf of Mexico," *Mar. Mammal Sci.*, vol. 30, no. 1, pp. 399–409, 2014.

[12] L. Viloria-Gómora, E. Romero-Vivas, and J. R. Urbán, "Calls of Bryde's whale (Balaenoptera edeni) recorded in the Gulf of California," *J. Acoust. Soc. Amer.*, vol. 138, no. 5, pp. 2722–2725, Nov. 2015.

[13] S. M. Kerosky, A. Širović, L. K. Roche, S. Baumann-Pickering, S. M. Wiggins, and J. A. Hildebrand, "Bryde's whale seasonal range expansion and increasing presence in the Southern California Bight from 2000 to 2010," *Deep Sea Res. I, Oceanograph. Res. Papers*, vol. 65, pp. 125–132, Jul. 2012.

[14] R. Constantine, M. Johnson, L. Riekkola, S. Jervis, L. Kozmian-Ledward, T. Dennis, L. G. Torres, and N. A. De Soto, "Mitigation of vessel-strike mortality of endangered Bryde's whales in the Hauraki Gulf, New Zealand," *Biol. Conservation*, vol. 186, pp. 149–157, Jun. 2015.

[15] S. Izadi, M. Johnson, N. A. de Soto, and R. Constantine, "Nightlife of Bryde's whales: Ecological implications of resting in a Baleen whale," *Behav. Ecol. Sociobiol.*, vol. 72, no. 5, pp. 72–78, 2018.

[16] C. Castro, "Bryde's whale Balaenoptera edeni occurrence and movements in coastal areas off Ecuador, Peru and Panama. A preliminary report," Int. Whaling Commission, Cambridge, U.K. Tech. Rep. SC/67A/SH/15, Apr. 2017.

[17] L. D. Figueiredo and S. M. Simão, "Bryde's whale (Balaenoptera edeni) vocalizations from southeast Brazil," *Aquatic Mammals*, vol. 40, no. 3, pp. 25–231, 2014.

[18] L. Lodi, R. H. Tardin, B. Hetzel, I. S. Maciel, L. D. Figueiredo, and S. M. Simão, "Bryde's whale (Cetartiodactyla: Balaenopteridae) occurrence and movements in coastal areas of southeastern Brazil," *Zoologia*, vol. 32, no. 2, pp. 171–175, May 2015.

[19] L. R. Gonçalves, M. Augustowski, and A. Andriolo, "Occurrence, distribution and behaviour of Bryde's whales (Cetacea: Mysticeti) off south-east Brazil," *J. Mar. Biol. Ass.*, vol. 96, no. 4, pp. 943–954, Jun. 2016.

[20] R. Tardin, Y. Chun, S. Simão, and M. Alves, "Modeling habitat use by Bryde's whale Balaenoptera edeni off southeastern Brazil," *Mar. Ecol. Prog. Ser.*, vol. 576, pp. 89–103, Aug. 2017.

[21] Ø. Olsen, "On the external characters and biology of Bryde's whale *BalaeoEnopterabrydei*, a new rorqual from the coast of South Africa," *Proc. Zoolog. Soc. London*, vol. 83, no. 4, pp. 1073–1089, 1913.

[22] P. B. Best, "Two allopatric forms of Bryde's whale off South Africa," *Rep. Int. Whale Commn.*, vol. 1, pp. 10–38, 1977.

[23] P. Best, "Distribution and population separation of Bryde's whale Balaenoptera edeni off Southern Africa," *Mar. Ecol. Prog. Ser.*, vol. 220, pp. 277–289, Sep. 2001.

[24] N. Kanda, M. Goto, H. Kato, M. V. McPhee, and L. A. Pastene, "Population genetic structure of Bryde's whales (*Balaenopterabrydei*) at the interoceanic and trans-equatorial levels," *Conservation Genet.*, vol. 8, no. 4, pp. 853–864, 2007.

[25] T. Yamada, "Middle sized balaenopterid whale specimens in the Philippines and Indonesia," *Memoirs Nat. Sci. Museum*, vol. 45, pp. 75–83, Mar. 2008.

[26] G. S. Penry, P. S. Hammond, V. G. Cockcroft, P. B. Best, M. Thornton, and J. A. Graves, "Phylogenetic relationships in southern African Bryde's whales inferred from mitochondrial DNA: Further support for subspecies delineation between the two allopatric populations," *Conservation Genet.*, vol. 19, no. 6, pp. 1349–1365, Dec. 2018.

[27] G. S. Penry, K. Findlay and P. Best, "A conservation assessment of Balaenoptera edeni," in *The Red List of Mammals of South Africa, Swaziland and Lesotho*, M. F. Child, L. Roxburgh, E. D. L. San, D. Raimondo, and H. T. Davies-Mostert, Eds. Pretoria, South Africa: South African National Biodiversity Institute, 2016, p. 8.

[28] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-23, no. 1, pp. 67–72, Feb. 1975.

[29] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.

[30] L. Rabiner and R. Shafer, *Digital Processing of Speech Signals*. London, U.K.: Prentice-Hall, 1978.

[31] S. E. Anderson, A. S. Dave, and D. Margoliash, "Template-based automatic recognition of birdsong syllables from continuous recordings," *J. Acoust. Soc. Amer.*, vol. 100, no. 2, pp. 1209–1219, Aug. 1996.

[32] J. R. Buck and P. L. Tyack, "A quantitative measure of similarity for tursiops truncatus signature whistles," *J. Acoust. Soc. Amer.*, vol. 94, no. 5, pp. 2497–2506, Nov. 1993.

[33] K. Kaewtip, A. Alwan, C. O'Reilly, and C. E. Taylor, "A robust automatic birdsong phrase classification: A template-based approach," *J. Acoust. Soc. Amer.*, vol. 140, no. 5, pp. 3691–3701, Nov. 2016.

[34] E. J. Keogh and M. J. Pazzani, "Derivative dynamic time warping," Dept. Inf. Comput. Sci., Univ. California, Irvine, CA, USA, Tech. Rep., Apr. 2019, pp. 1–11.

[35] J. C. Brown and P. J. O. Miller, "Automatic classification of killer whale vocalizations using dynamic time warping," *J. Acoust. Soc. Amer.*, vol. 122, no. 2, pp. 1201–1207, Aug. 2007.

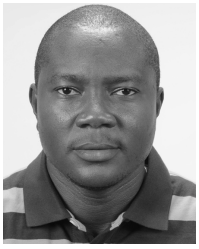
[36] C. Myers, L. Rabiner, and A. Rosenberg, "Performance tradeoffs in dynamic time warping algorithms for isolated word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 6, pp. 623–635, Dec. 1980.

[37] F. Pace, "Comparison of feature sets for Humpback whale song classification," M.S. thesis, Univ. Southampton, Southampton, U.K., 2008.

[38] C.-H. Min and A. Tewfik, "Automatic characterization and detection of behavioural patterns using linear predictive coding of accelerometer sensor data," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug./Sep. 2010, pp. 220–223.

[39] Y.-B. Jia, "Roots of polynomials," Iowa State Univ., Ames, IA, USA, Tech. Rep., Apr. 2019, pp. 1–11.

[40] T. Yack, J. Barlow, S. Rankin, and D. Gillespie, "Testing and validation of automated whistle and click detectors using PamGuard 1.0 NOAA technical memorandum NMFS," U.S. Dept. Commerce, Washington, DC, USA, Tech. Rep. NOAA-TM-NMFS-SWFSC-443, Apr. 2019.



OLAYINKA O. OGUNDILE (Member, IEEE) received the B.Eng. degree in electrical engineering from the University of Ilorin, Kwara, Nigeria, in 2007, the M.Sc. degree in communication engineering from The University of Manchester, U.K., in 2010, and the Ph.D. degree from the University of the Witwatersrand, Johannesburg, South Africa, in 2016. He is currently a Lecturer with the Department of Physics and Telecommunication, Tai Solarin University of Education, Nigeria. He is also a Postdoctoral Research Fellow with the Department of Electrical and Information Engineering, Stellenbosch University, South Africa. His research interests include digital communication, digital transmission techniques, digital signal processing/machine learning, channel estimation, forward error correction, and wireless sensor networks.

DANIEL J. J. VERSFELD (Member, IEEE) received the B.Eng. degree in electronic engineering and the M.Eng. degree in electronic engineering from North-West University, South Africa, in 1999 and 2001, respectively, and the Ph.D. degree from the University of Johannesburg, South Africa, in 2011. He is currently an Associate Professor with the Department of Electrical and Electronics Engineering, Stellenbosch University, Stellenbosch, South Africa. His research interests include algebraic coding for digital communications and signal processing applied to communications.