Automatic classification of whistles from coastal dolphins of the southern African subregion

Florence Erbs
Sea Search Africa, 4 Bath Road, Muizenberg, Cape Town 7945, South Africa

Simon H. Elwen[a]
Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, South Africa

Tess Gridley[a],[b]
*Centre for Statistics in Ecology, Environment and Conservation, Department of Statistical Sciences,
University of Cape Town, South Africa*

Passive acoustic monitoring (PAM) is commonly used to generate information on the distribution, abundance, and behavior of cetacean species. In African waters, the utilization of PAM lags behind most other continents. This study examines whether the whistles of three coastal delphinid species (*Delphinus delphis*, *Tursiops truncatus*, and *Tursiops aduncus*) commonly encountered in the southern African subregion can be readily distinguished using both statistical analysis of standard whistle parameters and the automated detection and classification software PAMGuard. A first account of whistles recorded from *D. delphis* from South Africa is included. Using PAMGuard, classification to species was high with an overall mean correct classification rate of 87.3%. Although lower, high rates of correct classification were also found (78.4%) when the two *T. aduncus* populations were included separately. Classification outcomes reflected patterns observed in standard whistle parameters. Such acoustic discrimination may be useful for confirmation of morphologically similar species in the field. Classification success was influenced by training and testing the classifier with data from different populations, highlighting the importance of locally collected acoustic data to inform classifiers. The small number of sampling populations may have inflated the classification success, therefore, classification trials using a greater number of species are recommended. © 2017 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4978000]

[WWA]   Pages: 2489–2500

I. INTRODUCTION

Passive acoustic monitoring (PAM), i.e., the recording of sound within a habitat (Merchant *et al.*, 2015), has been widely applied to generate information on the distribution, density, and abundance of numerous cetacean species (Zimmer, 2011). PAM is fast becoming the tool of choice in many areas of applied research to monitor anthropogenic activities in the marine environment and, more specifically, the effects of such activities on cetaceans (Weilgart, 2007; André *et al.*, 2011; Zimmer, 2011). The equipment and methodology used in PAM of cetaceans has developed rapidly over recent decades. Although the initial set up costs can be high, PAM can be cost effective in the medium to long term, generating detailed information over time, from inaccessible regions and across a range of spatial scales (e.g., Mellinger *et al.*, 2007; Van Parijs *et al.*, 2009). Therefore, PAM has wide ranging applications in Africa where there are large data gaps in our understanding of cetacean occurrence and finances are often limited.

There are several benefits of PAM beyond visual survey techniques for generating information on cetaceans. Notably, PAM enables data collection on species presence throughout the night and in poor weather conditions when visual data collection would not be possible. Archival marine acoustic recording units (ARUs) can be deployed in extreme or remote locations (Sousa-Lima *et al.*, 2013) where regular boat surveys would not be feasible. If deployed in arrays containing multiple devices, acoustic localization can be used to track the movements of individuals (Dunn and Hernandez, 2009). Distance sampling (Marques *et al.*, 2009) and mark-recapture approaches (Stevenson *et al.*, 2015) can be used to estimate animal density or abundance using acoustic data, and this field is growing rapidly (Marques *et al.*, 2013). Furthermore, combined visual and acoustic line transect surveys can be applied to better understand the distribution of cetacean species, for example, sperm whales (Barlow and Taylor, 2005), vaquita (Gerrodette *et al.*, 2011; Swift *et al.*, 2011), and Yangtze River dolphins (Turvey *et al.*, 2007; Richman *et al.*, 2014). If vocalizations are species specific and visual identification problematic (e.g., due to poor sightings conditions, brief or distant observations, or morphological similarity), acoustic identification can clarify the species (Oswald *et al.*, 2003; Oswald *et al.*, 2007; Gillespie *et al.*, 2013). Real-time species confirmation can assist in survey decisions, saving time and resources (Oswald *et al.*, 2007).

Although clearly advantageous, the application of PAM for research and monitoring of cetacean species in African

[a]Also at: Sea Search Africa, 4 Bath Road, Muizenberg Cape Town 7935, South Africa.
[b]Electronic mail: tessgridley@yahoo.co.uk

waters has lagged behind other more affluent regions. There have been some recent advances in this field (Hofmeyr-Juritz and Best, 2011; Gridley *et al.*, 2012; Gridley *et al.*, 2014; Gridley *et al.*, 2015), however, basic data on the call repertoire and vocal characteristics of most cetacean species encountered in southern Africa are still missing (Best, 2007; Elwen *et al.*, 2011). Such data are necessary for the implementation of successful PAM programs. A growing number of research groups operating in southern Africa are utilizing ARUs to monitor cetacean occurrence together with ambient noise conditions. In the absence of visual species confirmation, the success of such acoustic monitoring depends largely on how confidently species can be identified through their vocalizations. Although this has been demonstrated in other areas (e.g., Rendell *et al.*, 1999; Oswald *et al.*, 2003), no study has investigated species specific call characteristics of delphinids in southern Africa or the success of acoustic species discrimination using automatic classifiers. Intra-specific geographic variation in vocalization characteristics may also act to compromise discrimination, if classifiers are not trained using data from the region of interest (Gillespie *et al.*, 2013).

For dolphins, acoustic classification to species can be achieved using commonly produced vocalizations such as echolocation clicks (Soldevilla *et al.*, 2008; Roch *et al.*, 2011) or whistles (Rendell *et al.*, 1999; Oswald *et al.*, 2003; Oswald *et al.*, 2007; Gannier *et al.*, 2010), or a combination of these commonly emitted sounds types (Roch *et al.*, 2007). Whistles are narrow-band frequency modulated communication signals often used in social contexts (Herzing, 2000; Quick and Janik, 2008). Most often classification is based on discriminant function analysis (DFA; Rendell *et al.*, 1999), classification and regression trees (CART; Gannier *et al.*, 2010) or both methods (Oswald *et al.*, 2007) using standard parameters extracted from the whistle contour, although spectral analysis has also been employed (Roch *et al.*, 2007). Several whistle detection and classification programs have been developed (Oswald *et al.*, 2007; Gillespie *et al.*, 2013; Lin *et al.*, 2012; Lin and Chou, 2015). Some, such as ROCCA (Oswald *et al.*, 2007), require a human user to manually select high quality whistles suitable for classification, the fundamental frequency of which is then extracted and classified. Others, such as PAMGuard (Gillespie *et al.*, 2013), use an automated contour detection algorithm to identify contours for subsequent classification.

PAMGuard is an open-source software for the automated detection, localization, and classification of cetacean sounds (Gillespie *et al.*, 2009; Gillespie *et al.*, 2013). The program PAMGuard (version 1.13.04) differs from other whistle classification software by offering fully automated whistle detection through the integrated whistle and moan detector (WMD) and subsequent classification of whistle contour data (Gillespie *et al.*, 2013). As a whistle detector may only partially detect a whistle or break a whistle into several segments, the whistle classification method developed by Gillespie *et al.* (2013) has been carefully designed to be robust to the fragmentation taking place during the whistle detection process. This method therefore overcomes some of the commonly encountered issues with automated

whistle classification, and correct classification rates exceeding 94% have been reported (Gillespie *et al.*, 2013).

This study focuses on three commonly encountered whistling species found in the near-shore waters of the southern African subregion: common and Indo-Pacific bottlenose dolphins (*Tursiops truncatus* and *Tursiops aduncus*) and common dolphins, which were previously considered *Delphinus capensis* (Best, 2007) but now are considered *Delphinus delphis* (Cunha *et al.*, 2015). We describe the whistles of these species, including a novel description of *D. delphis* whistles, and test whether these species can be readily discriminated using standard whistle parameters. We investigate the performance of PAMGuard for automated classification of the same datasets. Of particular interest was how well PAMGuard could correctly distinguish between *T. aduncus* and *T. truncatus*—two closely related species that occur in sympatry or parapatry within the Indo-Pacific region and are notoriously difficult to identify from field observations. We examine whether classifiers are sensitive to macro- and micro-geographical variation in whistle characteristics using data from *T. aduncus* inhabiting the South West Indian Ocean and identify possible sources of error in the whistle classification process.

II. METHODS

A. Data collection

Acoustic recordings of free ranging dolphins were made from three widely separated sites within the southern African subregion: *T. truncatus* from Walvis Bay (Namibia), *T. aduncus* and *D. delphis* from Plettenberg Bay (South Africa), and *T. aduncus* from North and South of Unguja Island, Zanzibar Archipelago (Tanzania). The latter will be referred to as Zanzibar (Fig. 1).

The acoustic recording information and periods when data were collected are summarized in Table I. In all cases, the HTI-96-MIN hydrophone (High Tech, Inc., Long Beach, MS, sampling frequency 96 kHz, flat frequency response between 2 Hz and 30 kHz $\pm$ 1 dB) was weighted with a 1 cm diameter steel chain and lowered between 1 and 6 m below the water surface from a small motorized research vessel, while the vessel was stationary, in idle, or moving slowly. During 2008–2009, acoustic recordings were digitized using an Edirol UA-25 sound card (Roland Corp., Shizuoka, Japan) and saved to a personal computer. Thereafter, a Zoom H4n digital recorder (Zoom Corp., Tokyo, Japan) was used to sample the data. Recordings were made over a range of behavioral states and group compositions typical for each species.

B. Descriptive statistics of whistles parameters

Acoustic files were visually and aurally inspected for the occurrence of whistles. Standard parameters were measured from whole whistle contours to determine the underlying differences between the four delphinid populations and two regions of Zanzibar. For this we defined a whistle contour as a continuous narrow-band sound >100 ms in duration and with at least part of the fundamental frequency exceeding 3 kHz (Gridley *et al.*, 2012; Gridley *et al.*, 2014).
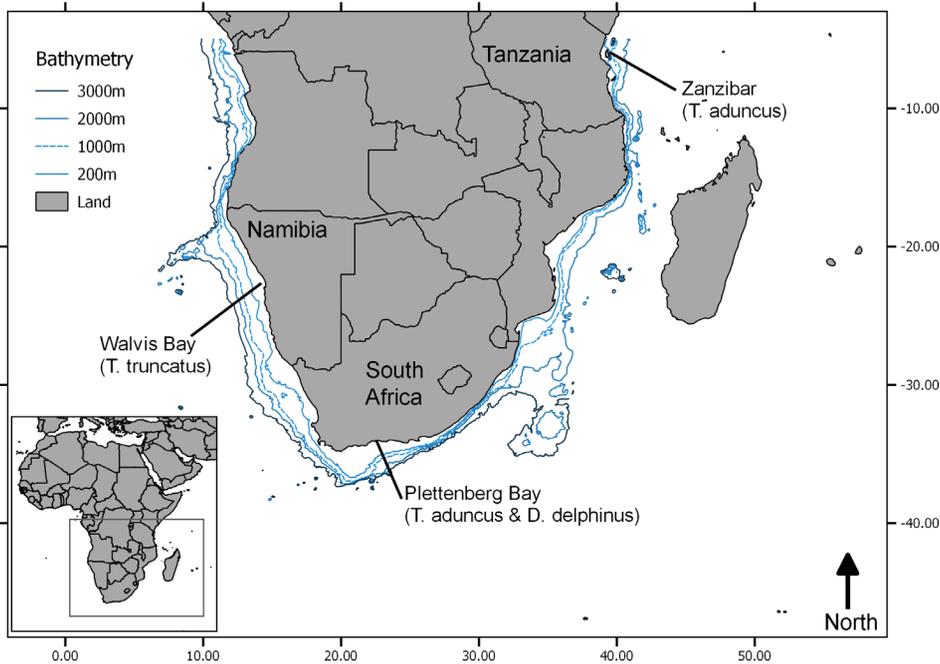
FIG. 1. (Color online) Map of the study area, the southern African subregion, showing the different locations where acoustic data were collected.

Whistle selection and measurement were conducted from the two *T. aduncus* populations as described in Gridley *et al.* (2012). Measurements of whistle contours from *T. truncatus* and *D. delphis* were derived as follows. Whistles were identified in the spectrogram display of Adobe Audition (Ver. 4.0 and Ver. 5.0, Adobe Systems Inc., San Jose, CA) and Raven Pro 1.4 (Cornell Bioacoustic Research Program, 2011) software using a Hanning window with a fast Fourier transform (FFT) length of 512 samples (time resolution 5.33 ms). Each whistle identified was visually assessed and graded based on the signal-to-noise ratio (SNR) as follows: (1) signal is faint but visible on the spectrogram, (2) signal is clear and unambiguous, (3) signal is prominent and dominates. High quality sounds were assessed as SNR 2 or 3, had a clear start/end, and were not masked by simultaneous sound. The number of inflection points, i.e., change in slope from positive to negative or vice versa was visually assessed and the duration was measured in Adobe Audition. Whistle characteristics, including the start, end, minimum, maximum, and frequency range, were measured in Raven using the selection function.

In most cases, the parameter data were non-normally distributed (Wilk-Shapiro test: $p < 0.05$) and the within-population variance was non-homogenous (Fligner-Killeen test of homogeneity of variances: $p < 0.05$). Statistical differences between the whistle parameters for the populations and species were therefore investigated using non-parametric Kruskal-Wallis tests conducted in the R statistical software (R Core Team, 2015). *Post hoc* Dunn's tests were applied to determine where significant differences lay following Zar (2010). The results were adjusted with the Holm multiple pair-wise correction method (Holm, 1979; Aickin and Gensler, 1996). The Holm procedure is a multistage test that adjusts the level of significance (α') according to the number of null hypotheses remaining to be tested (c) with $\alpha' = \alpha/c$, in order to control the familywise error rate inherent

TABLE I. Summary details of acoustic data collection and dataset used in the PAMGuard Whistle and Moan Detector (WMD) from four populations of delphinids recorded in Namibia (NAM), South Africa (SA), and Tanzania (TZ). Values in parentheses for Zanzibar North are the subsampled dataset used in analysis III.

Species names	Location	Latitude and Longitude	Period	Number of encounters	Number of files	Duration analyzed (hh:mm:ss.000)	Number of segment contours identified by PAMGuard WMD
Delphinus delphis	Plettenberg Bay, SA	34° 1' S 23° 25' E	March 2009 April 2014	2	11	00:53:21.750	14 777
Tursiops truncatus	Walvis Bay, NAM	22° 57' S 14° 30' E	2009 2011–2013	46	46	09:38:17.392	38 230
Tursiops aduncus	Plettenberg Bay, SA	34° 1' S 23° 25' E	March 2009 April 2014	8	47	05:30:40.797	28 767
	Zanzibar North	6° 9' S 39° 12'E	February–March 2008	14 (14)	85 (40)	08:36:08.182 (04:13:43.022)	10 316 (4952)
Tursiops aduncus	Zanzibar South, TZ			17	65	09:07:26.525	4743

to multiples comparisons tests. Regional variation between North and South Zanzibar was tested separately using non-parametric two-way analyses of variance (ANOVAs; Mann-Whitney U test) as the data failed the assumption necessary for parametric testing (Wilk-Shapiro test: $p < 0.05$).

C. Automated whistle detection and classification

The investigation of PAMGuard whistle classification was run as five analyses. (i) We investigated species level discrimination in PAMGuard by comparing whistles from the three study species. Here all data from *T. aduncus* were pooled. (ii) We investigated the discriminatory ability of PAMGuard in a classification test using data from *T. aduncus* from Zanzibar and Plettenberg Bay (∼4000 km apart) included separately. Previous studies have demonstrated macro-geographic variation in the whistles of these populations using standard whistle parameters such as end frequency (Gridley *et al.*, 2012). (iii) We investigated the discriminatory ability of PAMGuard in a classification test using data from *T. aduncus* from two sites around Zanzibar Island, which may potentially differ at a micro-geographic scale, North and South Zanzibar (∼80 km apart). In their study based on the analysis of mitochondrial deoxyribose nucleic acid (mtDNA), Särnblad *et al.* (2011) have demonstrated significant genetic differences between *T. aduncus* from the North and South of Zanzibar, indicative of population structuring. However, photo-identification data have demonstrated a degree of individual movement between these two areas (Gridley, 2011). (iv) We tested for bias associated with training the classifier with a dataset from a different geographical region to that where the test data were collected by running the training and classification analysis as in analysis II but only using one *T. aduncus* population (i.e., data from either Plettenberg Bay *or* Zanzibar). The classification outcomes under this scenario were then compared to a classification test using the alternate population, which had not been used in training. We hypothesized that the correct classification scores for *T. aduncus* would fall under this scenario compared to if the classifier was trained and tested with data collected from the same region. (v) We investigated the potential for false detections of burst pulse (BP) sounds to influence classification success.

Automated whistle detection and classification in PAMGuard is described in detail in Gillespie *et al.* (2013) and briefly outlined below. PAMGuard is designed for flexibility with core functionalities and a range of additional plugins that can be integrated within a single interface. In this study, we used two plugins tailored to tonal sounds: the WMD and the whistle classifier. The WMD automatically detects and extracts whistle contours by searching for spectral peaks within a user-specified frequency band. We ran the WMD on click and noise free FFT data (summarized in Table I), between 3 and 24 kHz with the amplitude threshold set to 8 dB (spectrogram settings FFT length 512 samples, FFT hop size 256 samples, Hanning window, time resolution 5.33, time step size 2.67 ms). The 3 kHz threshold was chosen after trials of the WMD demonstrated a high false

detection rate due to low frequency engine and water noise in the recordings in the 0–3 kHz bandwidth.

When conducting this analysis there exists a trade-off in the choice of fragment and section length. Short fragment (e.g., <50 ms) and section (e.g., 20 fragments) lengths can lead to unstable measurement of parameters and suboptimal classification success rate, whereas long section lengths requiring many fragments may need more whistles before classification can occur (Gillespie *et al.*, 2013). Choice of fragment length and section length was made by running a subset of data through the classifier for varying fragment length (from 5 to 70 bins, i.e., 16.1–191.6 ms) and section length (from 10 to 70 fragments). After following the advice set out in Gillespie *et al.* (2013), Caillat (2013), and a comprehensive testing period, we selected the following settings to conduct the analysis: the whistle classifier was run using a fragment length of 40 bins (equivalent to 110.6 ms) and section length of 60 fragments, over 100 bootstraps for the different datasets. The bootstrapping process allows the classification output to be presented with a standard deviation (SD) for each classification score. These parameters were similar to those used in Gillespie *et al.* (2013) in terms of the number of bins generated from the spectrogram with a FFT length of 512 (Gillespie *et al.*, 2013, used a fragment length of 30 bins and section length of 60 fragments). However, the time resolution of the spectrograms used in the two studies differs due to the different sampling rates: in this study a 96 kHz sampling rate was used resulting in a time resolution of 5.33 ms; in Gillespie *et al.* (2013) a 48 kHz sampling rate was used, resulting in a time resolution of 10.67 ms. Classification in this study is therefore based on shorter whistle fragments compared to Gillespie *et al.* (2013).

Files containing at least one whistle were considered suitable for analysis in PAMGuard. We sub-sampled the data from some regions to minimize differences in sample sizes and prevent over sampling of individuals or behavioral contexts. The small population of *T. truncatus* in Walvis Bay (∼100 individuals) was substantially recorded, therefore, we sub-sampled this dataset to one recording per encounter. Although the number of encounters were fewer in Plettenberg Bay, the group sizes were large (mean 129 individuals; Gridley, 2011) and whistle rates were very high compared to the other populations. Several short recordings were usually made per encounter. Therefore we sub-sampled these data so that 60%–80% of the recordings from each encounter were analyzed. Except for analysis III, we included all recordings with whistles of *T. aduncus* from Zanzibar, as recording effort was distributed over many encounters and individuals. In analysis III, only data from Zanzibar North and South were included and we randomly sub-sampled the acoustic data from Zanzibar North to run the analysis with similar sample sizes between the two locations. We included all available acoustic data from *D. delphis*, which was the smallest dataset in terms of recording occasions. However, group sizes and whistle rates were high.

In undertaking this analysis, we identified that the detection of BP sounds could possibly be a source of error affecting classification accuracy. BP sounds are broadband click-based sounds that appear as tonal sounds when displayed in the frequency domain on a spectrogram (Watkins, 1968),

and could be detected by the WMD as whistle-like frequency-modulated contours with sidebands and overtones (Fig. 2). To investigate the influence of false positive BP detection on the whistle classification, we conducted an analysis regarding BP sound detection by the WMD. For the *Tursiops* populations, a 30 s sub-sample of acoustic recording was taken from each file, representing a total of 15% of the duration of the combined dataset. As fewer recordings were available for *D. delphis*, we took additional 30 s sub-samples from within the available files ($n = 19$ sections from 11 files). All sub-sampled file sections were visually reviewed in Adobe Audition (FFT 512, 50% overlap, Hanning window) to manually count BP events, and files were run in the WMD. Detection output was manually reviewed in the PAMGuard "Viewer" mode to identify the rate of false detections attributed to BP sound detection.

III. RESULTS

Acoustic recordings of three species of delphinid from four populations were analyzed during this study. The number of encounters varied from 2 (*D. delphis*) to 46 (*T. truncatus*) with the most comprehensive sampling conducted in Walvis Bay, Namibia. However, due to variations in whistling rate between these populations, differences in the overall number of segments included in the analysis were not great. Between 14 777 (*D. delphis*) and 38 230 (*T. truncatus*) segments were included from each population (Table I).

A. Descriptive statistics and pair-wise comparison of whistle parameters

There were clear differences in the whistle characteristics of the four dolphin populations (three species) when
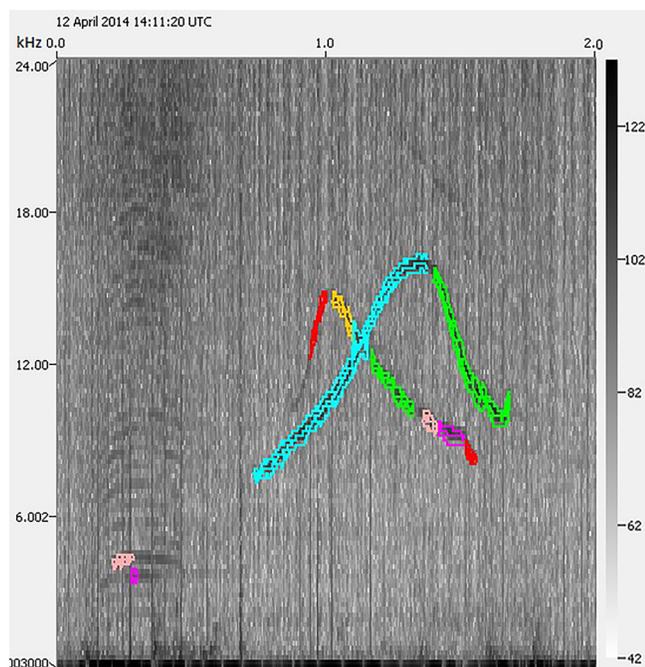
FIG. 2. (Color online) A spectrogram of 2 s of common dolphin recording showing the detected whistles segments and two detected BP segments using the PAMGuard WMD (FFT 512, 50% overlap, Hanning window). (Spectrogram segment is from recordings of *Delphinus delphis* from Plettenberg Bay on 12 April 2014.)

analyzed individually with a non-parametric Kruskal-Wallis one-way ANOVA. For every acoustic parameter tested, there was at least one population significantly different from the others (X^2_3 ranges from 124.45 to 1200.50, $p < 0.0001$). The frequency parameters (start, end, minimum, and maximum) of *D. delphis* were higher than both the *Tursiops* species, ranging from mean minimum values of 8.81 kHz ($\pm$SD 2.00) to mean maximum values of 15.68 kHz ($\pm$SD 4.02). Within *Tursiops*, the whistles of *T. truncatus* were consistently higher than for *T. aduncus* for all frequency parameters apart from end frequency. *Tursiops truncatus* had the longest duration whistles of those measured (1.08 s $\pm$ SD 0.65). The relatively low frequency whistles of *T. aduncus* ranged from 4.37 kHz ($\pm$ SD 1.35) to 11.44 kHz ($\pm$ SD 4.30) in Zanzibar and from 3.92 kHz ($\pm$ SD 1.64) to 9.62 kHz ($\pm$ SD 3.21) in Plettenberg Bay. The whistles from these populations were short in duration (0.37 s $\pm$ SD 0.22 and 0.44 s $\pm$ SD 0.28, respectively) and contained fewer inflection points (0.42 $\pm$ SD 0.77 and 0.66 $\pm$ SD 0.88, respectively) compared to *T. truncatus* or *D. delphis*. Overall, the Zanzibar population had higher frequencies than the Plettenberg Bay population for each of the four frequency parameters (minimum, maximum, start, and end frequency; Table II).

Dunn's multiple pair-wise comparison showed that the whistles from each of the four populations were statistically different ($p < 0.0001$) from the others for all apart from two acoustic parameters measured—no significant difference was identified between the end frequency of *T. truncatus* and *T. aduncus* from Plettenberg Bay ($p = 0.0510$) or between *T. truncatus* and *T. aduncus* from Zanzibar for frequency range ($p = 0.0826$; Table III). Although statistical differences in the whistle parameters between *T. aduncus* from North and South Zanzibar were identified, these differences were slight and reflected in four tested parameters: start frequency, end frequency, maximum frequency, and frequency range (Mann-Whitney U test, $p < 0.05$).

As this is the first description of whistles recorded from *D. delphis* in southern Africa, we compared the frequency parameters of whistle contours from this study with those from other *Delphinus* populations (Table IV). Comparisons demonstrate variability mainly in start and end frequency parameters, whereas minimum and maximum frequency and the number of inflections points are generally similar. The parameters of *D. delphis* from South Africa (this study) most closely match those reported in *D. delphis* from the Celtic Sea and English Channel (Ansmann *et al.*, 2007).

B. Analysis I: Automated classification to species

Species classification in PAMGuard using contours detected with the integrated WMD was highly accurate with an overall mean correct classification score of 87.3% for the three species (Table V). Classification scores were significantly different from that expected by chance (*t*-test $p < 0.001$ in all cases) and SDs around the classification scores were relatively low (range of SD 0.7–7.7). The whistles of *T. aduncus* were classified with the greatest success with 96.3% ($\pm$ SD 2.6) correct classification. Misclassification between *T.*

J. Acoust. Soc. Am. **141** (4), April 2017

Erbs *et al.*   2493

TABLE II. Summary of key acoustic characteristics of whistles produced by 3 species of delphinid from populations in the southern African subregion, including Mean, Standard Deviation (SD), and Coefficient of Variation (CV) values. Species names abbreviated to first letters of genus and species only: *Dephinus delphis* (*Dd*), *Tursiops truncatus* (*Tt*), *Tursiops aduncus* (*Ta*). Populations abbreviated as follows: Plettenberg Bay, South Africa (PB, SA), Walvis Bay, Namibia (WB, NAM) and Zanzibar, Tanzania (ZB, TZ). The values presented for *Delphinus delphis* and *Tursiops truncatus* are from this study, the values for *Tursiops aduncus* are from Gridley *et al.* (2012) and this study.

Species (Population)	N	Statistics	Start frequency (kHz)	End frequency (kHz)	Min frequency (kHz)	Max frequency (kHz)	Frequency range (kHz)	Inflection points (n)	Duration (s)
Dd	409	Mean (± SD)	12.6 (4.6)	12.5 (4.7)	8.8 (2.0)	15.7 (4.0)	6.9 (3.6)	1.2 (1.3)	0.7 (0.4)
(PB,SA)		CV	37	37	23	26	53	113	65
Tt	801	Mean (± SD)	8.7 (3.5)	7.4 (3.1)	5.8 (1.9)	13.2 (3.0)	7.4 (3.4)	2.00 (2.2)	1.1 (0.6)
(WB, NAM)		CV	40	42	33	23	46	109	60
Ta	1677	Mean (± SD)	5.8 (3.0)	10.1 (4.7)	4.2 (1.5)	11.3 (3.9)	7.1 (4.0)	0.5 (0.8)	0.40 (0.2)
(All)		CV	52	47	34	34	56	165	61
Ta	1166	Mean (± SD)	5.9 (3.1)	11.4 (4.3)	4.4 (1.3)	12.1 (3.9)	7.7 (4.0)	0.4 (0.8)	0.4 (0.2)
(ZB, TZ)		CV	52	38	31	35	52	183	59
Ta	511	Mean (± SD)	5.53 (2.9)	6.98 (4.3)	3.92 (1.6)	9.62 (3.2)	5.70 (3.4)	0.66 (0.9)	0.44 (0.3)
(PB,SA)		CV	52	61	42	33	60	134	62

aduncus and *T. truncatus* was rare (3.3%) and virtually absent between *T. aduncus* and *D. delphis* (0.04%). *Delphinus delphis* was correctly classified in 87.6% (± SD 7.7) of cases, although whistles were occasionally misclassified as *T. truncatus* (12.9%). While still well classified, the whistles of *T. truncatus* had the lowest correct classification score 79% (± SD 6.7). Misclassification with *D. delphis* was relatively common (19.4%), but the whistles of *T. truncatus* were rarely confused with those of *T. aduncus* (1.6%). In summary, most error in classification could be attributed to confusion between *T. truncatus* and *D. delphis*, whereas the whistles of *T. aduncus* were unlikely to be confused with either of these two species.

C. Analysis II: Investigating the impact of macro-geographical variation within species

We tested the PAMGuard whistle detection and species classification as in analysis I but with data from the two *T. aduncus* populations from Plettenberg Bay and Zanzibar Island included separately. The overall correct classification rate in this analysis was high (78.4%) and above that expected by chance (*t*-test $p < 0.001$, Table VI). Within this analysis, correct classification to species remained almost unchanged: 86.7% (± SD 7.6) for *D. delphis* and 79.3% (±SD 5.9) for *T. truncatus*. High classification scores for *T. aduncus* from Plettenberg Bay and Zanzibar populations (70.7% ± SD 8.6 and 77.1% ± SD 10.2, respectively) showed it was possible to discriminate between these two

geographically separate populations of the same species. However, the correct classification rate fell by 25.6% and 19.2% for Plettenberg Bay and Zanzibar Island, respectively, compared to the combined *T. aduncus* rate found in analysis I and the level of uncertainty in this classification test increased from a coefficient of variation [CV = (SD/mean) × 100] of 7% for the three-species classifier to 10.2% for the four-population classifier. These results demonstrate that although correct classification rates may fall, classification to population is possible in PAMGuard and can take place without affecting the classification scores of other species simultaneously tested.

D. Analysis III: Investigating micro-geographic differences

In analysis III, we tested the PAMGuard whistle detection and classifier on only *T. aduncus* whistles from North and South Zanzibar (respectively, 4952 and 4743 segments contours detected) to investigate whether PAMGuard could discriminate between the whistles from these two areas. Although some statistical differences did exist in the standard whistle parameters such as start frequency, end frequency, maximum frequency, and frequency range, demonstrating micro-geographic variation between these areas, the PAMGuard classification results were poor. The overall correct classification score for this analysis did not differ from chance (mean correct classification of 52.7%

TABLE III. Results of multiple pair-wise comparisons using the Dunn test with Holm correction on acoustic characteristics between populations. Significant differences are noted by "*." Species names abbreviated to first letters of genus and species only: *Dephinus delphis* (*Dd*), *Tursiops truncatus* (*Tt*), *Tursiops aduncus* (*Ta*). Populations abbreviated as follows: Plettenberg Bay (PB) and Zanzibar (ZB).

Pair-wise comparisons		Statistics	Start frequency	End frequency	Minimum frequency	Maximum frequency	Frequency range	Inflection points	Duration
Dd	*Tt*	*P* (adjust)	<0.0001*	<0.0001*	<0.0001*	<0.0001*	0.0024*	<0.0001*	<0.0001*
	Ta (ZB)	*P* (adjust)	<0.0001*	0.0105*	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*
	Ta (PB)	*P* (adjust)	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*
Tt	*Ta* (ZB)	*P* (adjust)	<0.0001*	<0.0001*	<0.0001*	<0.0001*	0.0826	<0.0001*	<0.0001*
	Ta (PB)	*P* (adjust)	<0.0001*	0.0510	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*
Ta (ZB)	*Ta* (PB)	*P* (adjust)	0.0173*	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*

TABLE IV. Summary of key acoustic parameters (mean ± SD) of whistle contours from short-beaked common dolphins *Delphinus delphis* and long-beaked common dolphins *Delphinus delphis* (*previously *capensis*) from available published studies.

Species	Location	N	Start frequency (kHz)	End frequency (kHz)	Min frequency (kHz)	Max frequency (kHz)	Frequency range (kHz)	Inflection points (n)	Duration (s)	Source
*Delphinus delphis**	South Africa	409	12.60 (4.62)	12.52 (4.71)	8.81 (2.00)	15.68 (4.02)	6.87 (3.63)	0.66 (0.43)	1.16 (1.31)	This study
	Eastern tropical Pacific	174	10.87 (4.89)	14.46 (5.12)	8.48 (2.70)	16.21 (4.94)	-	1.59 (3.29)	0.62 (0.34)	Oswald *et al.* (2007)
Delphinus delphis	Eastern tropical Pacific	314	11.63 (4.84)	12.18 (4.38)	8.30 (2.69)	15.04 (4.39)	-	1.64 (1.87)	0.70 (0.39)	Oswald *et al.* (2007)
	Celtic Sea	1835	12.02	11.97	9.45	14.68	5.24	0.64	0.65	Ansmann *et al.* (2007)
	English Channel	435	12.64	12.48	9.80	15.83	6.03	0.56	0.64	Ansmann *et al.* (2007)
	Mediterranean Sea	120	10.91 (3.53)	11.92 (2.90)	8.53 (1.94)	13.15 (2.70)	4.62 (2.74)	1.1 (1.2)	0.47 (0.29)	Gannier *et al.* (2010)
	Atlantic Ocean	514	13.02 (4.98)	11.77 (3.98)	8.14 (1.77)	16.74 (3.55)	8.60 (3.43)	1.06 (1.24)	0.95 (0.38)	Papale *et al.* (2014)
	Mediterranean Sea	188	11.90	12.17	8.31	16.08	7.76	2.03	0.92	Azzolin *et al.* (2014)

with a CV of 38%). The correct classification scores were slightly greater for Zanzibar North (59.3% ± SD 18.0) compared to Zanzibar South (46.1% ± SD 21.5).

E. Analysis IV: Assessing the effect of regional differences in training and testing data

We re-ran the species classifier, observing the influence of using training and testing data from different *T. aduncus* populations. The findings from this analysis were contradictory. In a species classification where the classifier was trained to discriminate whistles of *T. aduncus* using data from Zanzibar but tested with *T. aduncus* data from Plettenberg Bay, the correct classification score for *T. aduncus* decreased by 14.4% compared to if it was trained and tested with Zanzibar data. This confirmed our hypothesis that classification success would decrease if there was geographic variation in training and testing datasets (Table VII). However, the classifier trained with recordings from *T. aduncus* from Plettenberg Bay and tested with Zanzibar data exhibited a small increase in classification success (∼3%) compared to if the classifier was trained and tested with data from Plettenberg Bay, resulting in a perfect correct classification score of 100% for *T. aduncus* under this scenario (Table VII).

F. Analysis V: Influence of BP sounds on classification

A total of 161 BP sounds were visually detected from the sub-sampled files. Among those visual detections, 16 were falsely detected in the PAMGuard WMD. Therefore 10% of BP sounds recorded were falsely identified as segments in the WMD (Table VIII). There were slight differences in the BP sound detection per species, with the BP sounds from *D. delphis* most likely to be falsely detected by the WMD. However, as the number of BP sounds recorded was much lower than whistle sounds, the overall false positive detections were very low, representing just 0.3% of the total number of detections by the WMD (range 0%–0.7% depending on species). Consequently, false detection of BP sounds is considered unlikely to influence whistle classification accuracy.

IV. DISCUSSION

This study was motivated by the increasing interest in applied acoustic research in Africa, and particularly the growing field of PAM of cetaceans in the southern African subregion. A key goal was to determine whether commonly encountered near-coast delphinid species (*D. delphis*, *T. truncatus*, and *T. aduncus*) from this region could be readily discriminated based on both standard whistle parameters and using PAMGuard, a freely available and widely utilized automated acoustic detection and classification system (Yack *et al.*, 2009; Oswald *et al.*, 2011; Bittle and Duncan, 2013; Gillespie *et al.*, 2013; Keating *et al.*, 2015). An

TABLE V. PAMGuard classification results for three species of delphinids found in coastal areas of the southern African subregion. Numbers in parentheses are the SD of the estimates of correct classification rates. The final column shows the results of a one-tailed *t*-test to test the null hypothesis that the results are due to chance. Species names abbreviated to first letters of genus and species only: *Delphinus delphis* (*Dd*), *Tursiops truncates* (*Tt*), *Tursiops aduncus* (*Ta*).

		Percent classified as			
		Dd	*Tt*	*Ta*	P
Actual species	*Dd*	86.7 (7.7)	12.9 (7.6)	0.3 (1.0)	<0.001
	Tt	19.4 (7.2)	79.0 (6.7)	1.6 (2.4)	<0.001
	Ta	0.4 (0.7)	3.3 (2.6)	96.3 (2.6)	<0.001

TABLE VI. PAMGuard classification results for four populations of delphinids with *Tursiops aduncus* from Plettenberg Bay (PB) and Zanzibar (ZB) tested separately. Numbers in parentheses are the SDs of the estimates of correct classification rates. The final column shows the results of a one-tailed *t*-test to test the null hypothesis that the results are due to chance. Overall classification success was 78.4%. Species names abbreviated to first letters of genus and species only: *Delphinus delphis* (*Dd*), *Tursiops truncatus* (*Tt*), *Tursiops aduncus* (*Ta*).

		Percent classified as				
		Dd	*Tt*	*Ta* (PB)	*Ta* (ZB)	P
Actual species	*Dd*	86.7 (7.6)	12.6 (7.5)	0.1 (0.6)	0.6 (1.3)	<0.001
	Tt	18.3 (6.2)	79.3 (5.9)	2.4 (2.7)	0.0 (0.0)	<0.001
	Ta (PB)	0.5 (0.9)	3.8 (3.3)	70.7 (8.6)	25.0 (9.7)	<0.001
	Ta (ZB)	0.1 (0.7)	0.1 (0.5)	22.8 (10.1)	77.1 (10.2)	<0.001

TABLE VII. PAMGuard classification results for the three-species classifiers trained with *T. aduncus* data from one location, respectively, Zanzibar (normal) and Plettenberg Bay (bold). Numbers in parentheses are the SDs of the estimates of correct classification rates. The final column shows the results of a one-tailed *t*-test to test the null hypothesis that the results are due to chance. Final lines show classification rate when the trained classifier is tested with *T. aduncus* data from the other location, respectively, Plettenberg Bay and Zanzibar. Species names abbreviated to first letters of genus and species only: *Delphinus delphis* (*Dd*), *Tursiops truncatus* (*Tt*), *Tursiops aduncus* (*Ta*).

| | | Percent classified as | | | |
		Dc	*Tt*	*Ta*	P
Training					
Actual species	*Dc*	87.0 (8.1)	12.9 (8.1)	0.1 (0.6)	<0.001
		87.4 (6.7)	**12.6 (6.6)**	**0.1 (0.4)**	**<0.001**
	Tt	16.2 (6.2)	83.5 (6.0)	0.3 (1.0)	<0.001
		17.0 (6.2)	**81.0 (5.1)**	**2.1 (2.4)**	**<0.001**
	Ta	0.0 (0.0)	1.2 (2.7)	98.8 (2.7)	<0.001
		0.1 (0.3)	**2.4 (2.5)**	**97.6 (2.6)**	**<0.001**
Testing					
	Ta (PB)	0.52	15.1	84.4	
	Ta (ZB)	**0.0**	**0.0**	**100.0**	

additional objective was to investigate the impact of documented geographical variation in whistle parameters on the automated classification accuracy.

There was a distinct pattern in the whistle frequency parameters ranging from higher to lower frequency whistling species (*D. delphis*, *T. truncatus*, *T. aduncus*), with the end frequency of *T. aduncus* from Zanzibar the exception. Although there was overlap in the whistle parameters, pair-wise comparisons demonstrated statistical differences between them. We found that the PAMGuard classification outcomes reflected the patterns observed in standard whistle parameters. The overall correct classification score of 87.3% to species on a high quality and representative dataset, combined with the reasonable variability (SD ≤ 7.7) demonstrates the potential of automated detection and classification in acoustic species recognition for southern African species. Although other methods and programs may have delivered similar results, we found the automated contour detection feature of the PAMGuard WMD particularly useful and time efficient for large acoustic datasets. Furthermore, we found high correct classification scores to population,

demonstrating that PAMGuard is sensitive to intra-specific geographic variation in whistle parameters and supporting previous work using standard whistle parameters (Gridley *et al.*, 2012). Our investigation of PAMGuard and the potential sources of bias in its use provide a baseline on which future studies can be developed.

The high classification scores in our three-species (four-population) classifier can be explained in several ways. First, this classifier only contained three to four groups, thus, limiting the probability of misclassification (Gillespie *et al.*, 2013). Gillespie *et al.* (2013) report correct classification scores ranging from 94.5% for a 4-species classifier to 58.6% for a 12-species classifier. Likewise, our three-species classifier outperformed the four-population classifier. A review of correct classification scores from multiple sources demonstrate the negative linear relationship between the number of species or populations involved in a classification task and classification success (Fig. 3). As the acoustic space occupied by whistles from different groups (species or populations) overlaps, classifiers involving more groups are expected to have greater misclassification rates and lower classification success (Gillespie *et al.*, 2013). In the southern African subregion, such species could include humpback dolphins (*Sousa plumbea*) or pantropical spotted dolphin (*Stenella attenuata*), although acoustic data from these species in African waters are currently lacking.

Our study found excellent discrimination scores between *T. aduncus* and *T. truncatus* (96%) with almost no misclassification. Differences in frequency parameters in *Tursiops* species have been reported previously (Hawkins, 2010; Gridley *et al.*, 2012, this study) with *T. aduncus* generally having lower frequency whistles for most parameters considered. However, this is the first study demonstrating that species within the *Tursiops* genus can be reliably distinguished using automated contour classification techniques. The accurate discrimination of *T. truncatus* and *T. aduncus* from their whistle vocalizations is of particular importance as these species are morphologically similar and occur in sympatry and parapatry within the Indo-Pacific region. Our correct classification scores (near certainty) using whistle vocalizations for *T. aduncus* highlight the potential of this method to distinguish between *T. aduncus* and *T. truncatus*. However, our dataset used recordings from coastal *T. truncatus* from Namibia and *T. aduncus* from South Africa. As we have demonstrated, geographic variation in whistle

TABLE VIII. Summary of contour detection in analysis (V), including false detection of BP sounds by the PAMGuard WMD.

Species	Number of files	Total length (min) of subsampled data	Total number of contour detections by WMD	Overall contour detection rate (number detection/ min)	Number of BP sounds manually detected	Number of BP detections by WMD	Proportion of BP detections
Delphinus delphis	11	9.5	1687	178	50	12	0.007
Tursiops truncatus	46	23	405	18	47	2	0.005
Tursiops aduncus Zanzibar	150	75	845	11	10	0	0.000
Tursiops aduncus Plettenberg Bay	47	23.5	2318	99	54	2	0.001
Total	254	131	5255	—	161	16	0.003
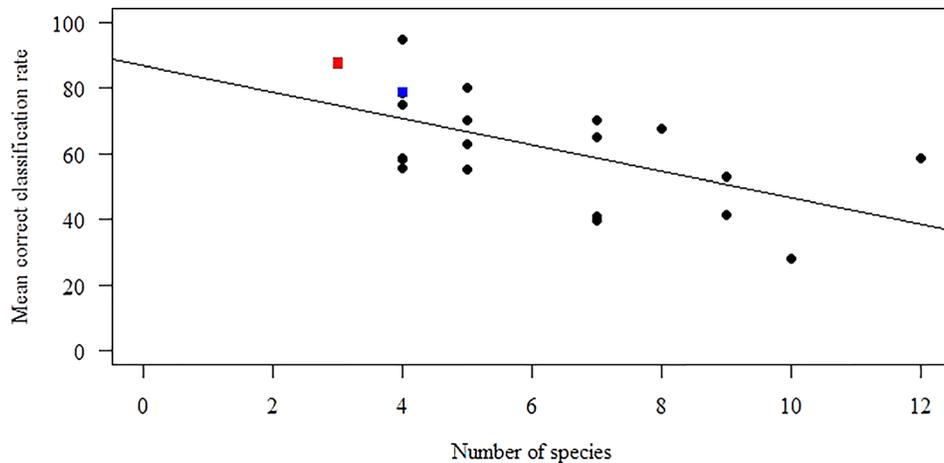
FIG. 3. (Color online) The success rates from published whistle classification tests run with different numbers of species (red and blue squares show, respectively, the results of this study for the three species-classifier and the four-population classifier; Steiner, 1981; Wang et al., 1995; Matthews et al., 1999; Rendell et al., 1999; Oswald et al., 2003; Oswald et al., 2007; Roch et al., 2007; Gannier et al., 2010; Caillat, 2013; Gillespie et al., 2013; Azzolin et al., 2014; Parada and Cardenal-López, 2014; Lin and Chou, 2015).

characteristics can interfere with classification success and future classification trials should be conducted with recordings from offshore *T. truncatus* from South Africa to better understand how accurately these can be distinguished from the adjacent coastal South African populations of *T. aduncus*. Our results agreed with other studies demonstrating that the whistles of *Delphinus and T. truncatus* can be misclassified at relatively high rates (e.g., 22% in Oswald et al., 2007). However, these are species that due to morphological and coloration differences are unlikely to be confused when observed in the field.

Geographic variation in delphinid whistle characteristics is well documented (e.g., Wang et al., 1995; Baron et al., 2008; Hawkins, 2010) and may result from ecological differences, morphological or cultural variability between populations, as well as differences in ambient noise levels (Morisaka et al., 2005; May-Collado and Wartzok, 2008). We tested the influence of intra-specific differences in whistle characteristics on classification success. The overall correct classification score in the four-population test fell relative to the three-species classifier (and the level of classification uncertainty increased), as misclassification resulted from confusion between the two populations of *T. aduncus*. However, the species level classification remained largely unchanged. Intra-specific variations in whistles characteristics of *T. aduncus* between Zanzibar and South Africa have been previously reported for all key frequency characteristics, however, such differences are less than those observed between species (Gridley et al., 2012) and the PAMGuard output largely reflects this. Our results confirm that the whistle classification method embedded in PAMGuard may be useful in population discrimination and can take place without affecting classification at the species level for the other groups included in the classification trial. This could assist in the PAM of both species and population units in areas where the range of ecologically distinct units overlap, such as where transient and resident killer whale ecotypes are found in sympatry (Foote and Nystuen, 2008; Riesch and Deecke, 2011; Oswald et al., 2015).

Previous studies reported significant population structuring in *T. aduncus* inhabiting Zanzibar North and Zanzibar South based on mtDNA molecular analysis (Särnblad et al., 2011), however, there is movement of individuals between

these two localities (Gridley, 2011). The two-region classifier performed poorly when tested using whistles from *T. aduncus* recorded from North and South Zanzibar. Subtle differences in several standard whistle parameters were observed between these two regions and the results therefore indicate that the PAMGuard whistle classification method is not as sensitive as some traditional statistical methods for discriminating between acoustically similar datasets. Oswald et al. (2007) experienced a similar problem when trying to differentiate between common dolphin species (*D. delphinus* and *D. capensis*), and pooled their data in subsequent analyses.

Our study represents a rare attempt to quantify the impact of intra-specific geographic variation in whistle characteristics on species classification results. The correct classification score of *T. aduncus* varied when the classifiers were trained with one population of *T. aduncus* and tested with the other population (trained with Zanzibar population and tested with South Africa population and vice versa). As predicted, the classifier trained with Zanzibar data and tested with South Africa data performed worse than if trained and tested with Zanzibar data, and the resulting drop in classification success was relatively large (14%). However, an unexpected slight increase (3%) in classification success was found for the classifier trained with Plettenberg Bay data and tested with Zanzibar data. Examination of the misclassification rates in the four-population classifier provides some insight into this result. Whereas the Zanzibar population exhibits almost no misclassification for *T. truncatus* or *D. delphis*, the South Africa *T. aduncus* population exhibits a 3.8% misclassification rate with *T. truncatus*. Following on, it seems logical to expect that when trained with data from Zanzibar (a population so acoustically distinct from the two other species that it displays almost no misclassification in the classifier), and tested with South Africa (a population showing more misclassification), the confusion rates increase and correct identification scores reduce. Conversely, classification rates increase when trained with data less distinct from other species in the classifier and tested with data from a more distinct population—driving the results to zero misclassification for this population. The results support the recommendation by Gillespie et al. (2013) to train and test the species classifier with data from the region of interest in order to obtain the most reliable results.

J. Acoust. Soc. Am. **141** (4), April 2017

Erbs *et al.* 2497

The field of PAM is becoming increasingly important in conservation management (Laiolo, 2010) and accurate species classification is fundamental to this. Many terrestrial and marine species can be recognized from the acoustic features of their commonly produced call types (Parsons and Jones, 2000; Fagerlund, 2007; Acevedo *et al.*, 2009; Heinicke *et al.*, 2015; Noda *et al.*, 2016). Geographic variation in the characteristics of species-specific sounds is widely reported (Krebs and Kroodsma, 1980; Mitani *et al.*, 1999; McDonald *et al.*, 2006; Delarue *et al.*, 2009). Therefore our results, demonstrating and measuring the impact of geographical variation on acoustic classification success and confirming the importance of using data from the region of interest, may be applicable to other scientists aiming to optimize classification accuracy using acoustic data.

Most delphinids produce BP vocalizations (Popper, 1980; Herzing, 2000) and during analysis we observed the harmonic structure of BP vocalizations being detected by the WMD. On-axis energy content and off-axis distortion of pulsed signals may lead to spectrogram detection of BP harmonic content in the mid-frequency range, i.e., that which overlaps with whistles. However, compared to whistles, pulsed vocalizations are highly directional (Au, 1993) and when recorded off axis can have lower received levels (Branstetter *et al.*, 2012). Therefore, although BP harmonics can be detected by the WMD, our investigation demonstrated a very low BP detection rate of 0.3%, unlikely to add notable confusion to classification success at the species-group level. However, as there is species-specific information in the clicks of several dolphin species (e.g., Soldevilla *et al.*, 2008; Roch *et al.*, 2011), whether BP sounds contain species specific information that could increase classification success warrants further investigation (Rankin *et al.*, 2016).

This is the first study in Africa to describe the acoustic features of whistles from *Delphinus*. The frequency of whistles from South African *Delphinus* fell into the range of other populations, although the duration of whistles reported in our study is the longest (1.16 s). The taxonomy of *Delphinus* is not well resolved globally and a recent study invalidated *D. capensis* from southern Africa (Cunha *et al.*, 2015). If whistle frequency is correlated to genetic similarity, our findings corroborate the lack of genetic differentiation in *Delphinus* as we found little difference in the key frequency parameters (such as minimum and maximum frequency) of South African *D. delphis* and other populations of *D. delphis* from the North Atlantic Ocean, Mediterranean Sea, and eastern Pacific Ocean.

This study represents a first step toward acoustic monitoring of coastal delphinids in southern Africa using automated acoustic techniques and will be particularly useful in future PAM studies on the distribution and habitat use of delphinids in this region. Other commonly encountered species are non-whistling (Heaviside's dolphin, *Cephalorhynchus heavisidii*, Morisaka *et al.*, 2011), functionally non-whistling (dusky dolphin, *Lagenorhynchus obscurus*, Vaughn-Hirshorn *et al.*, 2012, authors' unpublished data) or are known to whistle but there are no or very few recordings available from southern African waters (e.g., killer whales, *Orcinus orca*, Indian Ocean humpback dolphins, *Sousa plumbea*). Future data collection should focus on collecting data from those understudied species encountered in southern African waters. In addition, to fully realize the potential of PAM for cetaceans, careful data collection and analysis is necessary to obtain critical values required for acoustic density or abundance estimation such as whistle rates, detection ranges, and detection probability (Marques *et al.*, 2013).

V. CONCLUSIONS

Achieving a high correct classification rate with a low level of uncertainty is of crucial importance when monitoring species acoustically for studies of habitat use or abundance estimation (Caillat, 2013; Zimmer, 2011). PAMGuard was able to correctly classify three commonly encountered near shore species of delphinid with high classification rates and reasonable levels of uncertainty. High species classification scores and discrimination of macro-geographically separated populations of *T. aduncus* was a key outcome, especially as this species is regularly encountered in coastal southern Africa and categorized as "data deficient" by the International Union for Conservation of Nature (Hammond *et al.*, 2012). As geographic variation in whistle features reduces classification reliability, further dedicated acoustic data collection is required from delphinid populations within the southern African subregion. Such work will facilitate the monitoring of the species assemblage in near shore regions of the southern African subregion using passive acoustic methods.

Acevedo, M. A., Corrada-Bravo, C. J., Corrada-Bravo, H., Villanueva-Rivera, L. J., and Aide, T. M. (**2009**). "Automated classification of bird and amphibian calls using machine learning: A comparison of methods," Ecol. Inform. **4**, 206–214.

Aickin, M., and Gensler, H. (**1996**). "Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods," Am. J. Public Health **86**, 726–728.

André, M., Van Der Schaar, M., Zaugg, S., Houégnigan, L., Sánchez, A. M., and Castell, J. V. (**2011**). "Listening to the deep: Live monitoring of ocean noise and cetacean acoustic signals," Mar. Pollut. Bull. **63**, 18–26.

Ansmann, I. C., Goold, J. C., Evans, P. G., Simmonds, M., and Keith, S. G. (**2007**). "Variation in the whistle characteristics of short-beaked common dolphins, *Delphinus delphis*, at two locations around the British Isles," J. Mar. Biol. Assoc. U. K. **87**, 19–26.

Au, W. W. (**1993**). *The Sonar of Dolphins* (Springer, New York), pp. 1–277.

Azzolin, M., Gannier, A., Lammers, M. O., Oswald, J. N., Papale, E., Buscaino, G., Buffa, G., Mazzola, S., and Giacoma, C. (**2014**). "Combining whistle acoustic parameters to discriminate Mediterranean odontocetes during passive acoustic monitoring," J. Acoust. Soc. Am. **135**, 502–512.

Barlow, J., and Taylor, B. L. (**2005**). "Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey," Mar. Mammal Sci. **21**, 429–445.

Baron, S. C., Martinez, A., Garrison, L. P., and Keith, E. O. (**2008**). "Differences in acoustic signals from delphinids in the western North Atlantic and northern Gulf of Mexico," Mar. Mammal Sci. **24**, 42–56.

Best, P. B. (**2007**). *Whales and Dolphins of the Southern African Subregion* (Cambridge University Press, Cambridge, UK), pp. 1–338.

Bittle, M., and Duncan, A. (**2013**). "A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring," in *Proc. Acoust. Sci. Technol. Amenity Nov 17-20 2013*, edited by T. McMinn and V. Harbour (Australian Acoustical Society, South Australia), pp. 1–8.

Branstetter, B. K., Moore, P. W., Finneran, J. J., Tormey, M. N., and Aihara, H. (**2012**). "Directional properties of bottlenose dolphin (*Tursiops truncatus*) clicks, burst-pulse, and whistle sounds," J. Acoust. Soc. Am. **131**, 1613–1621.

Caillat, M. (**2013**). "Assessing and correcting for the effects of species misclassification during passive acoustic surveys of cetaceans," Ph.D. dissertation, University of St. Andrews, pp. 1–174.

Cornell Bioacoustics Research Program (**2011**). Raven Pro: Interactive Sound Analysis Software (Version 1.4) [Computer software]. (The Cornell Lab of Ornithology, Ithaca, NY). Available from http://www.birds.cornell.edu/raven.

Cunha, H. A., de Castro, R. L., Secchi, E. R., Crespo, E. A., Lailson-Brito, J., Azevedo, A. F., Lazoski, C., and Solé-Cava, A. M. (**2015**). "Molecular and morphological differentiation of common dolphins (*Delphinus* sp.) in the Southwestern Atlantic: Testing the two species hypothesis in sympatry," PLoS One **10**, e0140251.

Delarue, J., Todd, S. K., Van Parijs, S. M., and Di Iorio, L. (**2009**). "Geographic variation in Northwest Atlantic fin whale (*Balaenoptera physalus*) song: Implications for stock structure assessment," J. Acoust. Soc. Am. **125**, 1774–1782.

Dunn, R. A., and Hernandez, O. (**2009**). "Tracking blue whales in the eastern tropical Pacific with an ocean-bottom seismometer and hydrophone array," J. Acoust. Soc. Am. **126**, 1084–1094.

Elwen, S. H., Findlay, K. P., Kiszka, J., and Weir, C. R. (**2011**). "Cetacean research in the southern African subregion: A review of previous studies and current knowledge," Afr. J. Mar. Sci. **33**, 469–493.

Fagerlund, S. (**2007**). "Bird species recognition using support vector machines," EURASIP J. Appl. Signal Process **2007**, 38637.

Foote, A. D., and Nystuen, J. A. (**2008**). "Variation in call pitch among killer whale ecotypes," J. Acoust. Soc. Am. **123**, 1747–1752.

Gannier, A., Fuchs, S., Quèbre, P., and Oswald, J. N. (**2010**). "Performance of a contour-based classification method for whistles of Mediterranean delphinids," Appl. Acoust. **71**, 1063–1069.

Gerrodette, T., Taylor, B. L., Swift, R., Rankin, S., Jaramillo-Legorreta, A. M., and Rojas-Bracho, L. (**2011**). "A combined visual and acoustic estimate of 2008 abundance, and change in abundance since 1997, for the vaquita, *Phocoena sinus*," Mar. Mammal Sci. **27**, E79–E100.

Gillespie, D., Caillat, M., Gordon, J., and White, P. (**2013**). "Automatic detection and classification of odontocete whistles," J. Acoust. Soc. Am. **134**, 2427–2437.

Gillespie, D., Mellinger, D. K., Gordon, J., McLaren, D., Redmond, P., McHugh, R., Trinder, P. W., Deng, X. Y., and Thode, A. (**2009**). "PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans," J. Acoust. Soc. Am. **125**, 2547.

Gridley, T. (**2011**). "Geographic and species variation in bottlenose dolphin (*Tursiops* spp.) signature whistle types," Ph.D. dissertation, University of St Andrews, pp. 1–212.

Gridley, T., Berggren, P., Cockcroft, V. G., and Janik, V. M. (**2012**). "Whistle vocalizations of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) inhabiting the south-west Indian Ocean," J. Acoust. Soc. Am. **132**, 4032–4040.

Gridley, T., Cockcroft, V. G., Hawkins, E. R., Blewitt, M. L., Morisaka, T., and Janik, V. M. (**2014**). "Signature whistles in free-ranging populations of Indo-Pacific bottlenose dolphins, *Tursiops aduncus*," Mar. Mammal Sci. **30**, 512–527.

Gridley, T., Nastasi, A., Kriesell, H. J., and Elwen, S. H. (**2015**). "The acoustic repertoire of wild common bottlenose dolphins (*Tursiops truncatus*) in Walvis Bay, Namibia," Bioacoustics **24**, 153–174.

Hammond, P. S., Bearzi, G., Bjørge, A., Forney, K. A., Karkzmarski, L., Kasuya, T., Perrin, W. F., Scott, M. D., Wang, J. Y., Wells, R. S., and Wilson, B. (**2012**). "Tursiops aduncus," in *The IUCN Red List of Threatened Species 2012: e.T41714A17600466*. Available at http://dx.doi.org/10.2305/IUCN.UK.2012.RLTS.T41714A17600466.en (Last viewed September 14, 2016).

Hawkins, E. R. (**2010**). "Geographic variations in the whistles of bottlenose dolphins (*Tursiops aduncus*) along the east and west coasts of Australia," J. Acoust. Soc. Am. **128**, 924–935.

Heinicke, S., Kalan, A. K., Wagner, O. J., Mundry, R., Lukashevich, H., and Kühl, H. S. (**2015**). "Assessing the performance of a semi-automated acoustic monitoring system for primates," Methods Ecol. Evol. **6**, 753–763.

Herzing, D. L. (**2000**). "Acoustics and social behavior of wild dolphins: Implications for a sound society," in *Hearing by Whales and Dolphins* (Springer, New York), pp. 225–272.

Hofmeyr-Juritz, L. H., and Best, P. B. (**2011**). "Acoustic behaviour of southern right whales in relation to numbers of whales present in Walker Bay, South Africa," Afr. J. Mar. Sci. **33**, 415–427.

Holm, S. (**1979**). "A simple sequentially rejective multiple test procedure," Scand. J. Stat. **6**, 65–70.

Keating, J. L., Oswald, J. N., Rankin, S., and Barlow, J. (**2015**). "Whistle classification in the California Current: A complete whistle classifier for a large geographic region with high species diversity." U.S. Dept. Commerce, NOAA Tech. Memo. NOAA-TM-NMFS-SWFSC-552, pp. 1–13.

Krebs, J. R., and Kroodsma, D. E. (**1980**). "Repertoires and geographical variation in bird song," Adv. Study Behav. **11**, 143–177.

Laiolo, P. (**2010**). "The emerging significance of bioacoustics in animal species conservation," Biol. Conserv. **143**, 1635–1645.

Lin, T.-H., Chan, H.-C., Chen, C.-F., Akamatsu, T., and Chou, L.-S. (**2012**). "Detection on the presence and frequency use pattern of cetacean tonal sound," J. Acoust. Soc. Am. **131**, 3361–3361.

Lin, T.-H., and Chou, L.-S. (**2015**). "Automatic classification of delphinids based on the representative frequencies of whistles," J. Acoust. Soc. Am. **138**, 1003–1011.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (**2013**). "Estimating animal population density using passive acoustics," Biol. Rev. **88**, 287–309.

Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc. Am. **125**, 1982–1994.

Matthews, J. N., Rendell, L. E., Gordon, J. C. D., and Macdonald, D. W. (**1999**). "A review of frequency and time parameters of cetacean tonal calls," Bioacoustics **10**(1), 47–71.

May-Collado, L. J., and Wartzok, D. (**2008**). "A comparison of bottlenose dolphin whistles in the Atlantic Ocean: Factors promoting whistle variation," J. Mammal. **89**, 1229–1240.

McDonald, M. A., Mesnick, S. L., and Hildebrand, J. A. (**2006**). "Biogeographic characterization of blue whale song worldwide: Using song to identify populations," J. Cetacean Res. Manag. **8**(1), 55–65.

Mellinger, D. K., Stafford, K. M., Moore, S., Dziak, R. P., and Matsumoto, H. (**2007**). "Fixed passive acoustic observation methods for cetaceans," Oceanography **20**, 36–45.

J. Acoust. Soc. Am. **141** (4), April 2017

Erbs *et al.* 2499

Merchant, N. D., Fristrup, K. M., Johnson, M. P., Tyack, P. L., Witt, M. J., Blondel, P., and Parks, S. E. (**2015**). "Measuring acoustic habitats," Methods Ecol. Evol. **6**, 257–265.

Mitani, J. C., Hunley, K. L., and Murdoch, M. E. (**1999**). "Geographic variation in the calls of wild chimpanzees: A reassessment," Am. J. Primatol. **47**, 133–151.

Morisaka, T., Karczmarski, L., Akamatsu, T., Sakai, M., Dawson, S., and Thornton, M. (**2011**). "Echolocation signals of Heaviside's dolphins (*Cephalorhynchus heavisidii*)," J. Acoust. Soc. Am. **129**, 449–457.

Morisaka, T., Shinohara, M., Nakahara, F., and Akamatsu, T. (**2005**). "Effects of ambient noise on the whistles of Indo-Pacific bottlenose dolphin populations," J. Mammal. **86**, 541–546.

Noda, J. J., Travieso, C. M., and Sánchez-Rodríguez, D. (**2016**). "Automatic taxonomic classification of fish based on their acoustic signals," Appl. Sci. **6**, 443.

Oswald, J. N., Barlow, J., and Norris, T. F. (**2003**). "Acoustic identification of nine delphinid species in the eastern tropical Pacific Ocean," Mar. Mammal Sci. **19**, 20–37.

Oswald, M., Oswald, J. N., Lammers, M. O., Rankin, S., and Au, W. W. (**2011**). "Integration of real-time odontocete call classification algorithm into PAMGUARD signal processing software," J. Acoust. Soc. Am. **129**, 2639.

Oswald, J. N., Rankin, S., Barlow, J., and Lammers, M. O. (**2007**). "A tool for real-time acoustic species identification of delphinid whistles," J. Acoust. Soc. Am. **122**, 587–595.

Oswald, J. N., Yack, T. M., Dunleavy, K. D., and Zoidis, A. M. (**2015**). "Development of a tool for acoustic identification of killer whale communities in the Pacific Ocean," prepared for SeaWorld Busch Gardens 2014 Conservation Fund Grant award, Pacifica, CA.

Papale, E., Azzolin, M., Cascao, I., Gannier, A., Lammers, M. O., Martin, V. M., Oswald, J., Perez-Gil, M., Prieto, R., and Silva, M. A. (**2014**). "Macro- and micro-geographic variation of short-beaked common dolphin's whistles in the Mediterranean Sea and Atlantic Ocean," Ethol. Ecol. Evol. **26**, 392–404.

Parada, P. P., and Cardenal-López, A. (**2014**). "Using Gaussian mixture models to detect and classify dolphin whistles and pulses," J. Acoust. Soc. Am. **135**, 3371–3380.

Parsons, S., and Jones, G. (**2000**). "Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks," J. Exp. Biol. **203**, 2641–2656.

Popper, A. N. (**1980**). "Behavioral measures of odontocete hearing," in *Animal Sonar Systems* (Springer, New York), pp. 469–481.

Quick, N. J., and Janik, V. M. (**2008**). "Whistle rates of wild bottlenose dolphins (*Tursiops truncatus*): Influences of group size and behavior," J. Comp. Psychol. **122**, 305–311.

Rankin, S., Archer, F., Keating, J. L., Oswald, J. N., Oswald, M., Curtis, A., and Barlow, J. (**2016**). "Acoustic classification of dolphins in the California Current using whistles, echolocation clicks, and burst pulses," Mar. Mammal Sci. in press.

R Core Team (**2015**). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.

Rendell, L. E., Matthews, J. N., Gill, A., Gordon, J. C. D., and Macdonald, D. W. (**1999**). "Quantitative analysis of tonal calls from five odontocete species, examining interspecific and intraspecific variation," J. Zool. **249**, 403–410.

Richman, N. I., Gibbons, J. M., Turvey, S. T., Akamatsu, T., Ahmed, B., Mahabub, E., Smith, B. D., and Jones, J. P. (**2014**). "To see or not to see: Investigating detectability of Ganges River dolphins using a combined visual-acoustic survey," PLoS One **9**, e96811.

Riesch, R., and Deecke, V. B. (**2011**). "Whistle communication in mammal-eating killer whales (*Orcinus orca*): Further evidence for acoustic divergence between ecotypes," Behav. Ecol. Sociobiol. **65**, 1377–1387.

Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., and Hildebrand, J. A. (**2011**). "Classification of echolocation clicks from odontocetes in the southern California Bight," J. Acoust. Soc. Am. **129**, 467–475.

Roch, M. A., Soldevilla, M. S., Burtenshaw, J. C., Henderson, E. E., and Hildebrand, J. A. (**2007**). "Gaussian mixture model classification of odontocetes in the southern California Bight and the Gulf of California," J. Acoust. Soc. Am. **121**, 1737–1748.

Särnblad, A., Danbolt, M., Dalén, L., Amir, O. A., and Berggren, P. (**2011**). "Phylogenetic placement and population structure of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) off Zanzibar, Tanzania, based on mtDNA sequences," Mar. Mammal Sci. **27**, 431–448.

Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and Roch, M. A. (**2008**). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," J. Acoust. Soc. Am. **124**, 609–624.

Sousa-Lima, R. S., Norris, T. F., Oswald, J. N., and Fernandes, D. P. (**2013**). "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," Aquat. Mamm. **39**, 23–53.

Steiner, W. W. (**1981**). "Species-specific differences in pure tonal whistle vocalizations of five western North Atlantic dolphin species," Behav. Ecol. Sociobiol. **9**, 241–246.

Stevenson, B. C., Borchers, D. L., Altwegg, R., Swift, R. J., Gillespie, D. M., and Measey, G. J. (**2015**). "A general framework for animal density estimation from acoustic detections across a fixed microphone array," Methods Ecol. Evol. **6**, 38–48.

Swift, R., Rankin, S., Gerrodette, T., Taylor, B., Gillespie, D., Gordon, J., Rojas-Bracho, L., and Jaramillo-Legoretta, A. (**2011**). "Towed hydrophone surveys for monitoring trends in the abundance and distribution of the critically endangered Gulf of California porpoise, Vaquita," J. Acoust. Soc. Am. **130**, 2420.

Turvey, S. T., Pitman, R. L., Taylor, B. L., Barlow, J., Akamatsu, T., Barrett, L. A., Zhao, X., Reeves, R. R., Stewart, B. S., Wang, K., Wei, Z., Zhang, X., Pusser, L. T., Richlen, M., Brandon, J. R., and Wang, D. (**2007**). "First human-caused extinction of a cetacean species?," Biol. Lett. **3**, 537–540.

Van Parijs, S. M., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., and van Opzeeland, I. (**2009**). "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," Mar. Ecol. Prog. Ser. **395**, 21–36.

Vaughn-Hirshorn, R. L., Hodge, K. B., Würsig, B., Sappenfield, R. H., Lammers, M. O., and Dudzinski, K. M. (**2012**). "Characterizing dusky dolphin sounds from Argentina and New Zealand," J. Acoust. Soc. Am. **132**, 498–506.

Wang, D., Würsig, B., and Evans, W. (**1995**). "Comparisons of whistles among seven odontocete species," in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden, The Netherlands), pp. 299–323.

Watkins, W. A. (**1968**). "The harmonic interval: Fact or artifact in spectral analysis of pulse trains," in *Marine Bioacoustics 2*, edited by W. N. Tavogla (Pergamon, New York), pp. 15–43.

Weilgart, L. S. (**2007**). "The impacts of anthropogenic ocean noise on cetaceans and implications for management," Can. J. Zool. **85**, 1091–1116.

Yack, T. M., Barlow, J., Rankin, S., and Gillespie, D. (**2009**). "Integration of automated detection methods into NOAA Southwest Fisheries Science Center (SWFSC) acoustic marine mammal monitoring protocol," J. Acoust. Soc. Am. **125**, 2588.

Zar, J. H. (**2010**). *Biostatistical Analysis* (Prentice Hall, Englewood Cliffs, NJ), Chap. 11, pp. 226–244.

Zimmer, W. M. (**2011**). *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, Cambridge, UK), pp. 1–356.