

RESEARCH ARTICLE

Assemblage of Focal Species Recognizers—AFSR: A technique for decreasing false indications of presence from acoustic automatic identification in a multiple species context

Ivan Braga Campos^{1,2*}, Todd J. Landers^{1,3}, Kate D. Lee¹, William George Lee^{1,4}, Megan R. Friesen¹, Anne C. Gaskett¹, Louis Ranjard⁵

1 Centre for Biodiversity and Biosecurity, School of Biological Sciences, University of Auckland, Auckland, New Zealand, **2** Chico Mendes Institute for Biodiversity Conservation, Serra do Cipó National Park, Serra do Cipó/MG, Brasil, **3** Research and Evaluation Unit, Auckland Council, Auckland, New Zealand, **4** Landcare Research, Dunedin, New Zealand, **5** Research School of Biology, ANU College of Medicine, Biology and Environment, The Australian National University, Canberra, ACT, Australia

* icam765@aucklanduni.ac.nz



OPEN ACCESS

Citation: Campos IB, Landers TJ, Lee KD, Lee WG, Friesen MR, Gaskett AC, et al. (2019) Assemblage of Focal Species Recognizers—AFSR: A technique for decreasing false indications of presence from acoustic automatic identification in a multiple species context. *PLoS ONE* 14(12): e0212727. <https://doi.org/10.1371/journal.pone.0212727>

Editor: Dennis M. Higgs, University of Windsor, CANADA

Received: February 7, 2019

Accepted: November 18, 2019

Published: December 5, 2019

Copyright: © 2019 Campos et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The MatlabHTK package can be downloaded from: <https://github.com/LouisRanjard/matlabHTK> The use of MatlabHTK requires the beforehand installation of Octave (Eaton et al., 2014) and HTK (HTK, Young et al., 2016). The AFSR_summarizing script can be downloaded from: <https://github.com/klee8/AFSR>. Data (PAM sampled sound files) deposited in the Dryad Repository: <http://dx.doi.org/10.5061/dryad.6kv00>.

Abstract

Passive acoustic monitoring (PAM) coupled with automated species identification is a promising tool for species monitoring and conservation worldwide. However, high false indications of presence are still an important limitation and a crucial factor for acceptance of these techniques in wildlife surveys. Here we present the Assemblage of Focal Species Recognizers—AFSR, a novel approach for decreasing false positives and increasing models' precision in multispecies contexts. AFSR focusses on decreasing false positives by excluding unreliable sound file segments that are prone to misidentification. We used MatlabHTK, a hidden Markov models interface for bioacoustics analyses, for illustrating AFSR technique by comparing two approaches, 1) a multispecies recognizer where all species are identified simultaneously, and 2) an assemblage of focal species recognizers (AFSR), where several recognizers that each prioritise a single focal species are then summarised into a single output, according to a set of rules designed to exclude unreliable segments. Both approaches (the multispecies recognizer and AFSR) used the same sound files training dataset, but different processing workflow. We applied these recognisers to PAM recordings from a remote island colony with five seabird species and compared their outputs with manual species identifications. False positives and precision improved for all the five species when using AFSR, achieving remarkable 0% false positives and 100% precision for three of five seabird species, and < 6% false positives, and >90% precision for the other two species. AFSR' output was also used to generate daily calling activity patterns for each species. Instead of attempting to withdraw useful information from every fragment in a sound recording, AFSR prioritises more trustworthy information from sections with better quality data. AFSR can be applied to automated species identification from multispecies PAM recordings worldwide.

Funding: IBC received a Science Without Borders PhD scholarship funded by CNPq-Brazil. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Recent technical advances in sound-recording technologies and analyses have considerably enlarged the potential application of bioacoustics in conservation studies. Acoustic automated identification has been applied to numerous taxa including insects [1,2], anurans [3,4], bats [5,6], canids [7,8], birds [9–12], marine mammals [13] and elephants [14]. These automated techniques facilitate analysis of thousands of hours of sound files generated by a passive acoustic monitoring (PAM) approach, which could not realistically be done manually by a researcher. The integration of PAM recording with automated identification is a considerable advance that can be applied in biodiversity assessments in diverse environmental conditions and ecosystems. Although these methods show great potential for wildlife monitoring, limitations and uncertainties remain that discourage incorporation of these analyses into management programs. Several researchers have identified high number or proportions of false positives [6,7,15–17]. These are typically associated with recordings involving multiple species [18], different environmental sound background [15], recording quality and overlapping calls [19]. Attempts to improve detection rates can also lead to more misidentifications [5,17]. High number of false positives can be particularly problematic when acoustic sampling is the only method used to generate indications of presence for species focus of conservation actions. This is the case for many seabird species in remote offshore islands. These species, which usually return to their breeding colonies only during nighttime and can nest in island with difficult and intermittent access, are a major conservation focus internationally [20]. In this context, PAM studies aiming to indicate which species are present in each island of an offshore archipelago need to ensure that their indications of presence for each species have high probability to be correct. For that, automated identification techniques with low proportion of false positives for each species is crucial.

Automated identification studies commonly focus on increasing detection rates in order to maximise the number of identified target calls in a sound file. However, in PAM, the amount of sound being recorded can easily reach terabytes of data. Potamitis et al. [21], for example, rejected 90–95% of their recordings because they did not meet their target specifications in the signal pre-processing stage. It is probably impossible to extract useful information from every single sound segment when monitoring long term. Instead, we advocate in some cases focus should be given to extracting more precise and trustworthy information from the segments with the highest probabilities of providing a true indication of species presence. In other words, automated identification models should prioritise low false positives instead of high detection rates and accuracy. Otherwise, even with adequate similarity and accuracy, a high false positive rate renders automated identification of low utility for conservation managers in a PAM context. Techniques for decreasing false positives are especially important in circumstances known to generate high number of misidentifications, such as studies involving multiple species, which is a common situation in natural environments. It is noteworthy that the prioritization between aiming for low false positives or low false negatives (increasing the detection rate) may vary with the taxon and monitoring questions. Acoustic inventories aiming to detect rare (or acoustically rare) species may prioritize a high detection rate. However, it does not seem to be the proper choice when working with seabirds that establish their breeding colonies in sites where other sampling protocols are logistically challenging. Since these seabirds commonly have intense acoustic activity during the night, even with a lower detection rate, PAM recordings will still provide plenty of sound passages from which will be possible to generate correct indications of presence. In this context, it can be of more utility to be more confident in the correct species identification of each indication of presence. This way, conservation managers can be more confident in assuming the presence of a given species in a

specific offshore island based only in the indication of presence generated by acoustic automatic identification.

Another important factor for the utility of automatic recognizers is “precision”, b. Precision can also be described as the probability that a given positive indication of presence is correct. Low false positives can be achieved by decreasing detection rates, but decreasing detection can also result in decreased true positives. When the true positives and false positives are similar, even if both are low, any indication of a species presence would have similar chances of being correct or incorrect, so any technique that decreases false positives must also be reliable in terms of precision.

Here we present the Assemblage of Focal Species Recognizers—AFSR, a technique for decreasing false positives of automated identification models for acoustic recordings made in a multiple species context. It uses MatlabHTK, a hidden Markov models interface for bioacoustics analyses [22]. AFSR prioritises the extraction of information from more trustworthy sections of the recordings that can provide better quality data. It is a novel approach for increasing the precision of acoustic recognizers, enhancing the applicability of automated identification techniques for wildlife surveys.

Materials and methods

Study site and species

Pokohinu/Burgess Island ($35^{\circ} 54' S$, $175^{\circ} 07' E$) is part of the Mokohinau archipelago, located ca. 90km northeast of Auckland’s east coast, in Aotearoa/New Zealand (Fig 1). The island became pest free after the eradication of the kiore or pacific rat (*Rattus exulans*) in 1990, and has since seen an increase in bird fauna [23]. We focus on the development of an automatic identification model for the calls of five Procellariiformes seabird species: *Pelecanoides urinatrix* (Kūaka, Common diving petrel), *Puffinus gavia* (Pakahā, Fluttering shearwater), *Pterodroma*

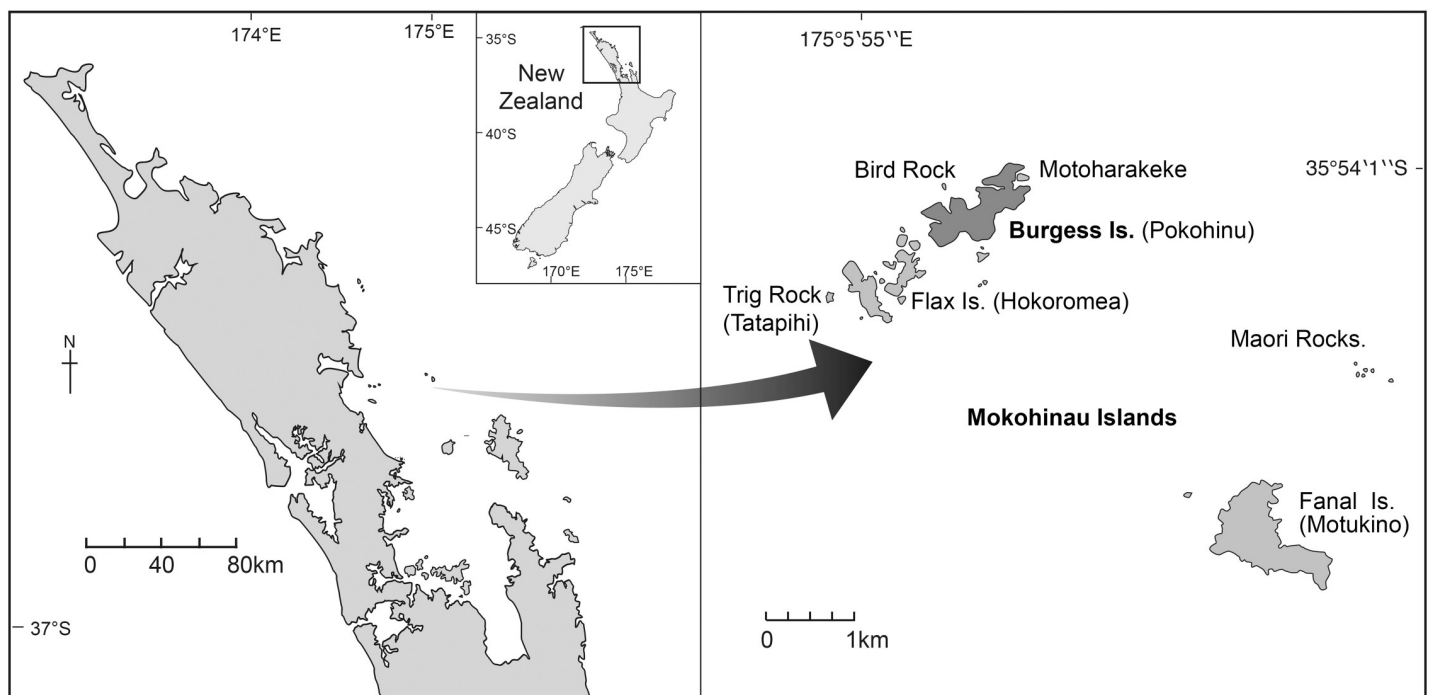


Fig 1. Pokohinu/Burgess Island, in the Mokohinau archipelago, Aotearoa/New Zealand.

<https://doi.org/10.1371/journal.pone.0212727.g001>

gouldi (Ōi, Grey-faced petrel), *Puffinus assimilis* (Little shearwater), and *Pelagodroma marina* (Takahikare-moana, White-faced storm petrel). These species are known to have breeding colonies at Burgess Island, although they do not co-occur at all locations on the island [23].

Acoustic recordings

We used manual recordings made by researchers in the field, and automatic PAM recordings made without researchers' presence using a pre-programmable acoustic sensor. The PAM recordings were made from 25th - 30th of September 2014, on Burgess Island (35.9057S, 175.1140E) using a sound recorder Song Meter SM2 (Wildlife Acoustics); at the sampling frequency of 44.1kHz; 32 bit resolution; from 6pm to 6am; totalling 60h of recordings. The manual recordings were performed by seabird experts to obtain good examples of the birds' typical calls, with confident species identification. Manual recordings of Grey-faced petrels were made from April to late May 2015 at Bethells Beach, New Zealand, using a 722 Digital Audio Recorder (Sound Devices, LLC) with a Sennheiser highly directional microphone (model K6 ME 66; Wennebostel, Wedemark, Germany). Manual recordings of the other four species were made on Burgess Island in September 2014 using a FR-2 Field Recorder (Foster Electric Co., Ltd.) with an Audio Technica shotgun microphone (model AT835b) housed in a Rycote wind-kit. All files were converted to 44.1kHz and 32 bit format.

Building the training sound files data set

We built a data set of training sound files (total 179 MB) of good examples of the different species calls and their environmental background sound from sections of both the manual and automatic PAM recordings. From the manual recordings, we selected and annotated calls that could be identified to species by seabird experts, and had good sound quality. We used these selected calls to create a preliminary species recogniser (S1 Supporting Information), that we then ran over the PAM recordings to fast track finding more examples of calls. All these calls extracted from the manual and PAM recordings were combined into a single set of training sound files, which were then used for both our Multispecies Recogniser and our AFSR (assemblage of focal species recognizers). Even though the Multispecies Recogniser and AFSR were built from the same sound files, these files were associated with different sets of annotation text files in each approach and were processed through different workflows. An overall modelling workflow diagram is shown in the Fig 2.

1) Multispecies Recognizer. Annotation files ("label" format) were manually created using the software AUDACITY[®]. These annotation files assigned a category name to each sound within the data set of training sound files described above. The categories used in the Multispecies Recogniser were: *Background*; *Noise*; *Diving Petrel*; *Grey-faced Petrel*; *Little Shearwater*; *Fluttering Shearwater*; and *White-faced Storm Petrel*. The '*Background*' category was assigned to sound fragments where there were no existing petrel calls. This recognizer followed the same processing workflow as described by Ranjard et al. [22].

2) Assemblage of Focal Species Recognizers—AFSR. Five independent species-specific recognisers were built using exactly the same data set of sound files data set previously described. In each case the sound files were associated with different annotation text files. For example, in the Little shearwater independent recogniser, all of this species' calls were assigned in the annotation files as '*Little Shearwater*', while all the other four species' calls were assigned as '*Other Species*'. The categories *Background* and *Noise* remained unchanged. This framework was applied to all the five independent recognisers, one for each of our five seabird species. According to Potamitis et al. [21], a detection model needs to differentiate a target call from all non-target sounds, including other species' calls and background sounds. In our system the

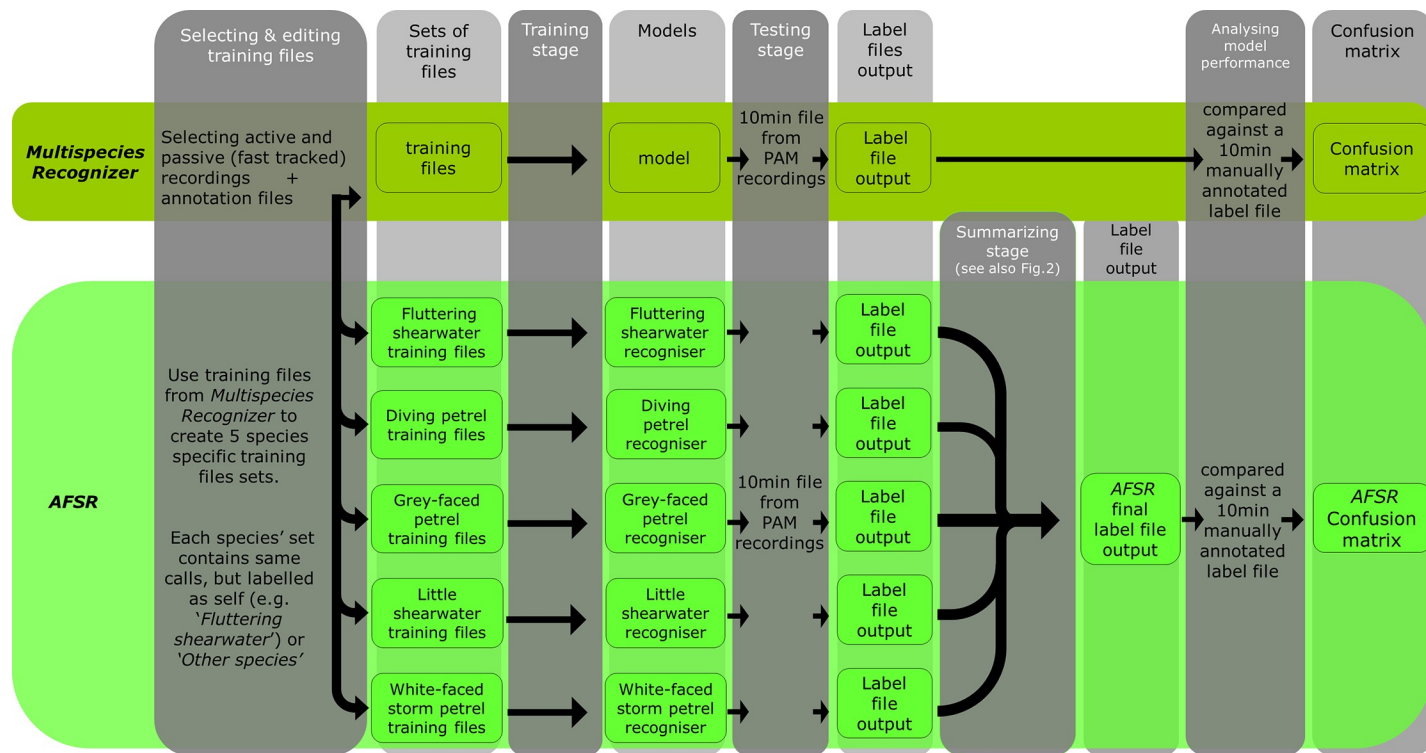


Fig 2. Modelling Workflow diagram. Each modelling approach is represented in horizontal lines. The workflow within each approach runs from left to right (columns) and the workflow from one modelling approach to the next runs from top to bottom.

<https://doi.org/10.1371/journal.pone.0212727.g002>

category ‘Other Species’ was used in each one of the independent recognisers to help discriminate the target call from other species’ calls.

In this way, we obtained five species-specific outputs, which we then compared to detect and remove unreliable sections of the recordings prone to misidentification. We did this by creating a script named “AFSR_summarizing” that applies a set of rules to summarize the independent outputs into one final annotation text file. Whenever the five recognisers disagreed about the species identification of any segment of the sound recording, the section was then labelled as ‘Unidentified’. Only the recording segments that showed consistent species identifications by all five independent recognisers were considered a valid indicator of species presence. The data accessibility information containing the for “AFSR_summarizing” script and the link for the MatlabHTK package are presented on the [S2 Supporting Information](#).

To this processing approach which consists in, from a single sound files data set, to create and run independent recognizers and then summarize their results into a single output following a specific set of rules, we named AFSR (Assemblage of Focal Species Recognizers). The set of rules used to summarize the five annotation text files (label format) into one is presented in the [S3 Supporting Information](#). An illustration of the summarizing process is presented in the [Fig 3](#).

Training and testing stages

The analyses were performed using the package MatlabHTK [22]. MatlabHTK processes the outputs of the Hidden Markov Toolkit (HTK; [24]), which was first designed for human speech analysis. In HTK, sound signal is typically represented as a sequence of parameter vectors, e.g. cepstrum coefficients, calculated for consecutive overlapping analysis windows.

Summarizing Independent Recognizers: Assemblage of Focal Species Recognizers - AFSR

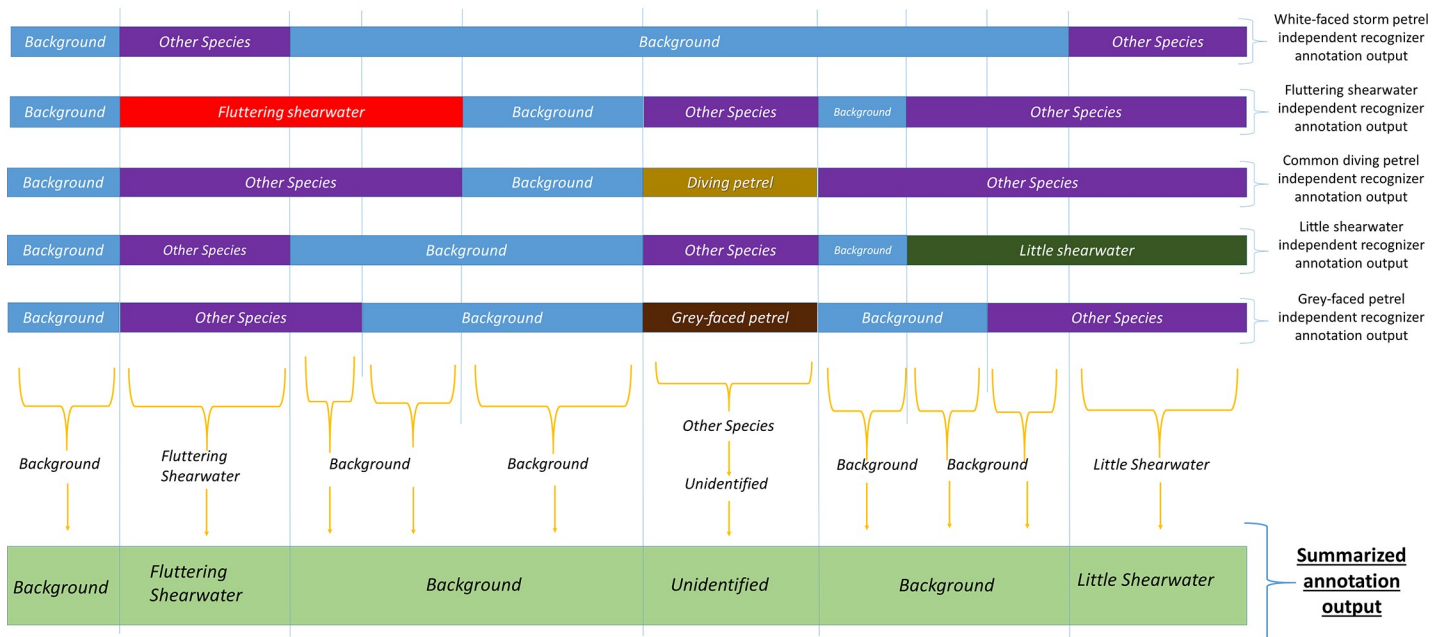


Fig 3. Illustration of the summarizing process. In this example of how AFSR converts category labels on a sound file fragment for five independent focal species recognisers into one summarized annotation based on specific rules, each of the upper horizontal lines represent one independent focal species recogniser’s annotation output and the lowest horizontal line represents the final summarized annotation output.

<https://doi.org/10.1371/journal.pone.0212727.g003>

Hidden Markov Models (HMMs) statistically represent transitions between these windows. Therefore, HMMs are flexible and capable of dealing with variation of a given signal such as differing intensity, variations within species repertoire, diverse sound background, etc. MatlabHTK allows users to easily build and runs HTK acoustic models and to process the results. In MatlabHTK, data analysis is typically performed in two stages; first, in the training stage the HMM probabilities are estimated from training data (annotation text files + “.wav” sound files) and, second, recognition is performed on the data during the testing stage. The testing stage outputs annotation files that indicate which categories have been identified throughout the data sound files. Only the category (among the ones used in the training) with higher probability of matching to a window in the testing files is assigned to each sound passage in the annotation outputs. A detailed description of MatlabHTK’s features including the training and the recognition stages is presented by Ranjard et al. [22]. Putland et al. [13] give an example of the MatlabHTK use for Bryde’s whale acoustics behaviour. The signal processing and statistical analyses were performed through an Octave environment [25].

Analysing the recognizers

To assess the performance of the two modelling approaches, we created a 10 minute sound file containing calls of all the five species as well as background, which were extracted from Burgess Island PAM recordings. The seabird experts in our team listened to the 10 minute file to manually assign the correct species to every bird call in the sound file. The annotation of the 10 minute was made using 0.1 second time window, resulting in 6000 windows that were manually checked for species identification. The time window of 0.1 second corresponds to the minimal interval between any two consecutive bird calls in the training sound files (for White-

faced storm petrel calls). It means that calls from the same species were separately annotated if the interval between them was longer than or equal to 0.1 second. This avoided inclusion of unnecessary background sound within sound fragments used to create the recogniser statistical profile for each category during the training stage (S1 Supporting Information). Species identifications were verified by different team members, based on their extensive experience and knowledge of wild seabird calls. Additionally to the bird species, ‘Background’ was also used as a category assigned to background sound.

We then ran the Multispecies Recognizer and the AFSR over this 10 min file and compared their outputs with the manual identification made by researchers. A 0.4 second time window was selected for comparisons because this was the minimum length of any bird call in the training sound files (for a White-faced storm petrel call). To compare the two annotations, a similarity measure was defined: for each time window and for the two files, vectors indicating the percentage of the window’s duration assigned to each category were constructed. Then, the percentage similarity score “S” was defined as one minus the average Euclidean distance between vectors over all time windows [22]. Let w be the window size, a sound file containing n_w windows, the similarity score between two annotation files a and b with n_a possible annotation categories is

$$S = 1 - \frac{\sum_1^{n_w} \sqrt{(\sum_1^{n_a} (d_a - d_b)^2) / 2w}}{n_w},$$

where d_a and d_b are the duration of annotation n_a in the current time window. We generated a confusion matrix tables (“.csv” format) to compare the manually annotated file with the automatically annotated files (outputs) generated by the Multispecies Recognizer and AFSR. All the similarity scores and confusion matrices were calculated using the same 0.4 second time window. Precision scores for each target species was calculated from the confusion matrix values as follows:

$$P = \frac{\text{true positives}}{(\text{true positives} + \text{false positives})}$$

In a final step, to benchmark the biological relevance of the AFSR, we ran it over the five nights of PAM recordings and used the output to generate activity pattern data showing the time periods in which our five target species were actively calling at their shared colony on Burgess Island.

Results

The “Noise” category, which corresponded in the training files to sounds typically produced when something touches the microphone, was not present in any of the recognizers’ outputs. For this reason, this category is not further mentioned in the results.

Since the analysis of performance here are based on the time length assigned to each category, the false positives are given in proportion of time. This way, the false positives presented here are the proportion of the time assigned to each species in which the species identification was incorrect.

1) Multispecies Recognizer

This recognizer achieved an overall similarity of 82% when compared with the 10 minute manually annotated file. The proportion of false positives for four petrel species were lower than 10% (Little shearwater 5%; White-faced storm petrel 2%, Fluttering shearwater 1%, Common

Table 1. Total false positive rate and Precision per species achieved by Multispecies Recognizer and AFSR.

Species	Multispecies Recognizer		AFSR	
	Proportion of false positive	Precision	Proportion of false positive	Precision
Common diving petrel	<u>0.09</u>	<i>0.88</i>	<u>0.01</u>	<i>0.97</i>
Grey-faced petrel	<u>0.1</u>	<i>0.88</i>	<u>0.05</u>	<i>0.92</i>
Little shearwater	<u>0.05</u>	<i>0.93</i>	<u>0</u>	<i>1</i>
Fluttering shearwater	<u>0.01</u>	<i>0.99</i>	<u>0</u>	<i>1</i>
White-faced storm petrel	<u>0.02</u>	<i>0.98</i>	<u>0</u>	<i>1</i>
Background	<u>0.28</u>	<i>0.72</i>	<u>0.17</u>	<i>0.81</i>

Proportion of false positives (underlined) and precision (*italic*) for each species were calculated from the values generated by confusion matrices for each model and are presented here in a scale from 0 to 1, being 1 equals to 100%.

<https://doi.org/10.1371/journal.pone.0212727.t001>

diving petrel 9%). However, the Grey-faced petrel had a higher false positive (10%). The precision was higher than 90% for three species: Little shearwater (93%), Fluttering shearwater (99%) and White-faced storm petrel (98%). However, the probabilities of false positives were > 10% for the Common diving petrel and Grey-faced petrel (precision was 88% for both species). The confusion matrix comparing the manual identification with the Multispecies Recognizer output is presented in the [S4 Supporting Information](#).

2) Assemblage of Focal Species Recognizers—AFSR

The overall similarity score achieved by AFSR was lower than the Multispecies Recognizer (74% vs. 82%) but it considerably reduced the false positives and increased precision scores for all five species. For White-Faced storm petrels, Fluttering shearwaters and Little shearwaters, the proportion of false positives achieved was 0%—all the indications of presence for these three species were correctly assigned. The proportion of false positives were also much lower for Common diving petrels (1%) and Grey-faced petrels (5%). For all species, precision values were >90%, with a remarkable 100% for Little shearwaters, Fluttering shearwaters, and White-faced storm petrels. [Table 1](#) presents the proportion of false positives and precision achieved by the AFSR and the Multispecies Recognizer. [Table 2](#) presents the confusion matrix comparing the manual identification with the AFSR output.

Table 2. Normalized Confusion Matrix comparing manual species identification Versus AFSR summarized output for a 10 minute long sound file.

Manually Annotated	AFSR							
	Unidentified	Background	Common diving petrel	Grey-faced petrel	Little shearwater	Fluttering shearwater	White-faced storm petrel	Proportion of false positives
Background	<i>0.08</i>	<i>0.75</i>	<i>0.04</i>	<i>0.01</i>	<i>0.03</i>	<i>0.01</i>	<i>0.08</i>	<u><i>0.17</i></u>
Common diving petrel	<i>0.33</i>	<i>0.29</i>	<u><i>0.37</i></u>	<i>0</i>	<i>0</i>	<i>0.01</i>	<i>0</i>	<u><i>0.01</i></u>
Grey-faced petrel	<i>0.19</i>	<i>0.21</i>	<i>0.04</i>	<u><i>0.55</i></u>	<i>0</i>	<i>0.01</i>	<i>0</i>	<u><i>0.05</i></u>
Little shearwater	<i>0.26</i>	<i>0.37</i>	<i>0</i>	<i>0</i>	<u><i>0.37</i></u>	<i>0</i>	<i>0</i>	<u><i>0</i></u>
Fluttering shearwater	<i>0.22</i>	<i>0.22</i>	<i>0</i>	<i>0</i>	<i>0</i>	<u><i>0.56</i></u>	<i>0</i>	<u><i>0</i></u>
White-faced storm petrel	<i>0.25</i>	<i>0.2</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>0</i>	<u><i>0.55</i></u>	<u><i>0</i></u>

The proportion of the time in which each category indicated at the manually annotated text file is assigned to each one of the categories at the AFSR's output text file is presented in a scale from 0 to 1 (being 1 equals to 100%) as follows: cells with *values in italic*: negative indications of presence; underlined values: true positive indication of presence, values with no special formatting: false positive indication of presence; values underlined and italic: normalized false positive for each one of the categories (sum of the cells with no special formatting in each line).

<https://doi.org/10.1371/journal.pone.0212727.t002>

AFSR results provided comprehensive daily activity patterns for individual species and the colony. Grey-faced petrels are the first species to vocalize at the colony after dusk (from ~19:20 hrs), shortly followed by Common diving petrels, Fluttering shearwaters, and White-faced storm petrels (all vocalizing from around ~19:30 hrs; Fig 4). The White-faced storm petrels are the first to depart before dawn around 04:20 hrs, followed by Grey-faced petrels and Fluttering shearwaters (about 5:00 hrs), then Common diving petrels (around 5:10 hrs). The low acoustic activity of Little shearwaters prevented a specific daily activity pattern from being produced.

Fig 5 shows the average percentage of identification from 18:00 to 6:00 hrs for all categories combined (the five seabird species, as well as the *Unidentified* category). For the colony, the first peak of activity happens after the sunset when the birds arrive, from 19:20 to 20:30 hrs. The activity then declines but is persistent at some level through the night, with exception to the period between 23:00 and 00:00 hrs in which a moderate peak of activity occurs. The calling activity raises to the most intense level from 3:30 to 5:00 hrs, which occurs during the birds' departure. After 5:00 hrs the activity reduces, and remains low until more birds arrive following the sunset. The overall acoustic activity pattern presented at Fig 5 is similar to the daily pattern for calls assigned to the *Unidentified* category (Fig 4F).

Discussion

AFSR successfully decreased false positives and increased precision for all five target species in this comparative study of seabirds on an offshore island in New Zealand. Passive acoustic monitoring (PAM) recordings made in the natural settings can register complex soundscapes containing sound passages susceptible to generate misidentification. These sound passages can be a result of different individuals calling at the same time, call variations among individuals, different sound background, non-target species, among other reasons. When using the MatlabHTK interface [22], for every segment of sound in a testing file, only the category which best matches with the sound segment is indicated in the annotation output. The use of MatlabHTK combined with AFSR helps to significantly reduce false positives. The false positive and precision improvements achieved by AFSR are a result of the identification of the problematic sound fragments and their assignment to a new category ("*Unidentified*"), avoiding potentially incorrect indications of presence.

The three main sources of sound in a soundscape are biophony (sound produced by biological organisms); geophony (non-biological sound); and anthrophony (human-induced noise) [26]. While geophony and anthrophony are irrelevant for biodiversity monitoring, a portion of the biophony present in passive acoustic recordings is also not useful due to calls overlapping, signal masking, and other phenomena, which can produce poor quality sound segments. In general, when applying automatic identification to PAM recordings we should target sound segments with reliable and relevant information and leave poor quality remaining passages out of the analyses. In order to prioritize reduced false positives and increased precision we accepted an overall decrease in similarity. Consequently, we assumed we would have an increased level of false negatives. An increase in false negative is not necessarily a problem for PAM if this increase is related to sound passages with poor quality and low chances of generating trustable indications of presence.

It is important to highlight that the activity pattern of calls assigned to the *Unidentified* category (Fig 4[f]) follows the overall acoustic activity pattern (Fig 5). This confirms our assumption that when more birds are calling, there is a higher chance of recording overlapping calls, and hence more misidentifications and false positives. These coincident patterns show our AFSR approach was effective in categorising calls that are problematic to identify and thus they were assigned to the *Unidentified* category. In this way, it reduces the false positives.

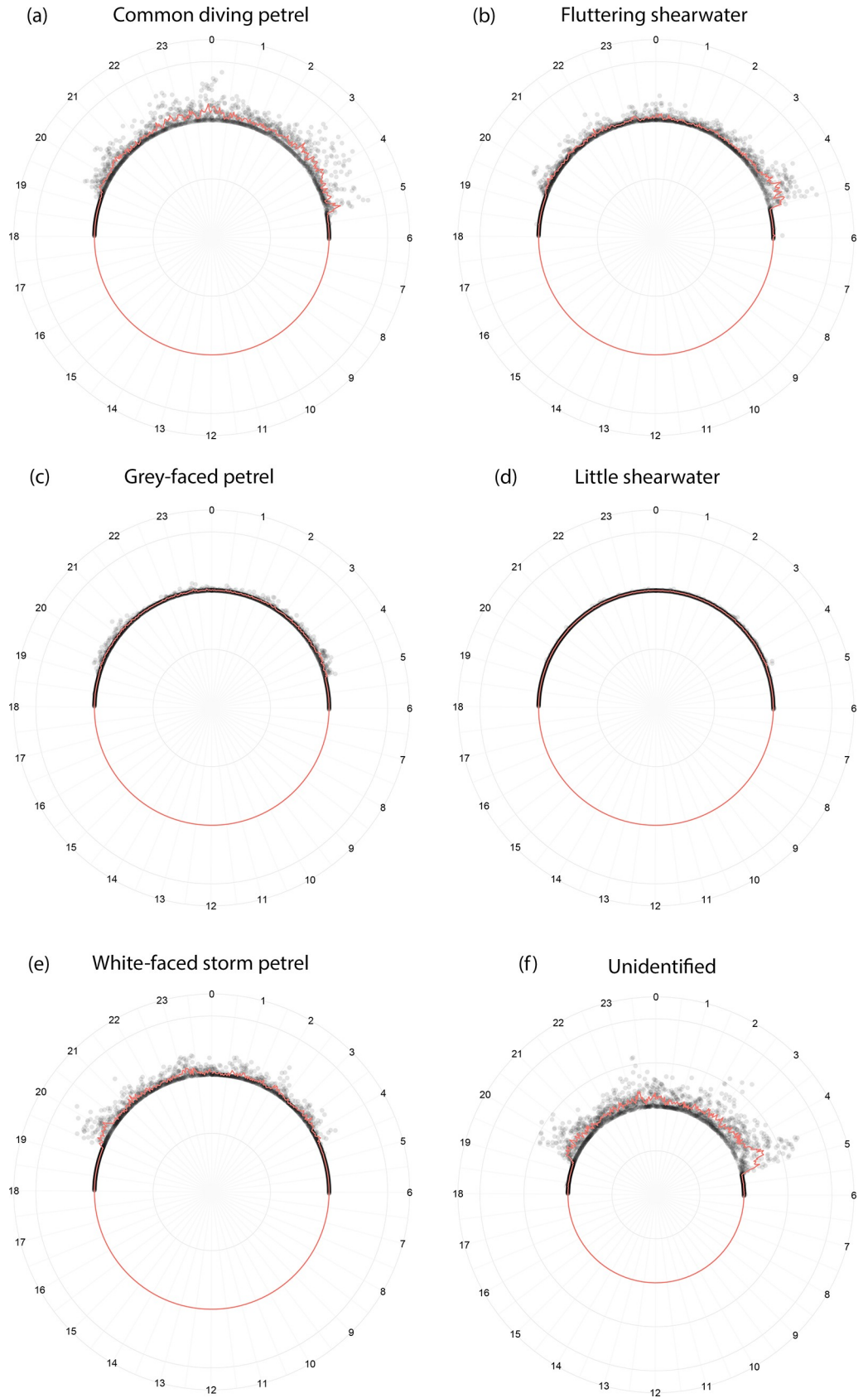


Fig 4. Daily pattern of acoustic activity identified by AFSR for the five seabird species. Common diving petrel [a]; Fluttering shearwater [b]; Grey-faced petrel [c]; Little shearwater [d]; White-faced storm petrel [e]; and the category Unidentified [f]; all identified from PAM recordings made for 5 consecutive days. Each grey circle is one indication of presence using 2 minute long analysis windows. The red line represents the average.

<https://doi.org/10.1371/journal.pone.0212727.g004>

Nevertheless, all the seabird species were detected by AFSR in this study, emphasising the success of this approach and its utility for multispecies monitoring.

The use of AFSR is especially appropriate for multispecies monitoring at roosting and breeding colonies where the acoustic activity can be very intensive in some hours of the night, but absent at other times. This intense acoustic activity produces overlapping calls which makes their correct identification very difficult for conventional recognizer approaches and increases the chances of misidentification.

The biological relevancy of AFSR results is supported by the successful indication of presence of all the five species' calls in our PAM recordings from Burgess Island. The activity patterns generated from the indications of presence confirm direct observations of these seabirds at natural colonies. The first and last peaks of activity (Fig 4) match the general pattern for petrels, with increased vocal activity just after sunset when birds arrive at the colony and before sunrise when they depart [27,28]. The first Grey-faced petrels' calls recorded from around one hour after sunset, agrees with Ross and Brunton's [29] data on arrival time at the colonies. The overall vocal activity pattern found for Common diving petrels is consistent with that reported by Ranjard et al. [22]. To our knowledge, there is no information about the daily activity pattern of the other studied species that we can use to compare to our automatically generated activity patterns. The low calling activity detected for Little shearwaters may be a consequence of a relatively smaller population size in comparison to the other four species at our fieldsite. Further studies are necessary to confirm this hypothesis.

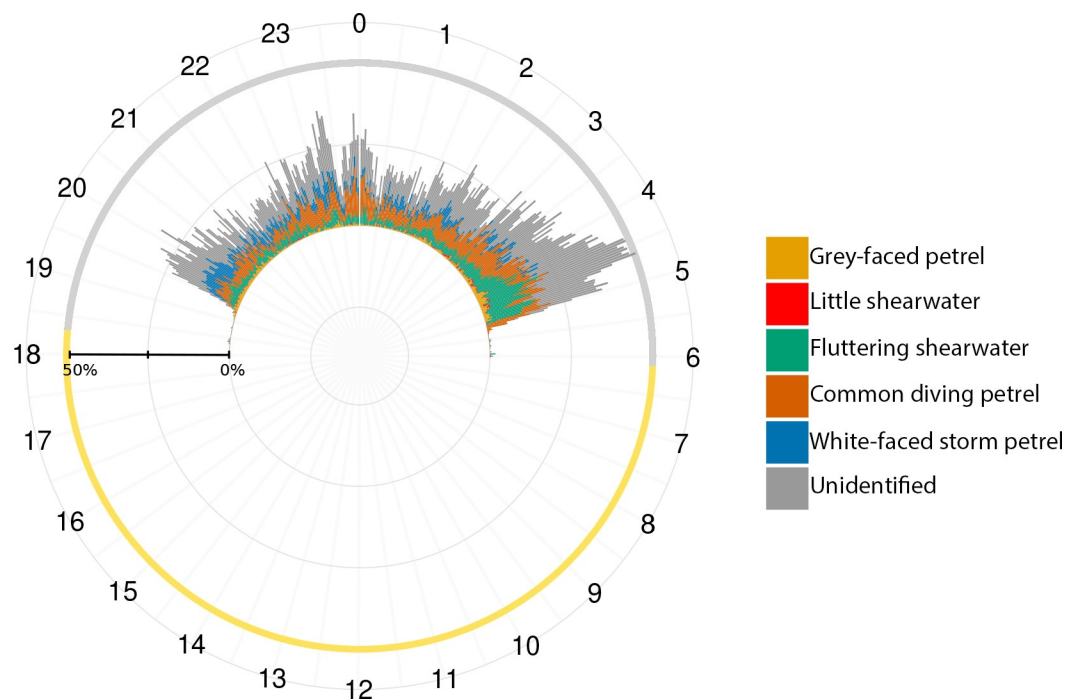


Fig 5. Colony mean daily acoustic activity for five seabird species, as well as sounds categorised as Unidentified by Assemblage of Focal Species Recognizers approach.

<https://doi.org/10.1371/journal.pone.0212727.g005>

Numerous islands in New Zealand have become pest free environments as a consequence of the elimination of invasive mammals which reduce or eliminate native bird's populations [30]. Following pest eradication, many seabird re-establish breeding colonies on islands [31]. PAM associated with the AFSR model used in this paper can be an important tool for detecting and monitoring the process of the re-establishment of seabird breeding colonies in New Zealand islands and South Pacific region. However, the methods described here are likely to have broader applicability to any situation where multiple species with acoustic activity are being monitored to determine occupancy.

The use of AFSR also allows the addition of more species into a multiple species analysis. The same set of rules used for summarizing the annotations from the five independent models can be extended depending on the number of focal taxa. It makes AFSR suitable for being applied to multiple species from different animal communities and diverse ecosystems monitored by a passive acoustic monitoring approach around the world.

Supporting information

S1 Supporting Information. Preliminary recognizers and overall modelling approach.
(PDF)

S2 Supporting Information. Data accessibility.
(PDF)

S3 Supporting Information. Summarizing rules.
(PDF)

S4 Supporting Information. Confusion matrix comparing manual species identification Versus the Multispecies Recognizer output for a 10 minute long sound file.
(PDF)

Acknowledgments

IBC thanks CNPq-Brazil for financial support through the Science Without Borders scholarship program.

Author Contributions

Conceptualization: Ivan Braga Campos, Todd J. Landers, William George Lee, Anne C. Gaskett, Louis Ranjard.

Data curation: Todd J. Landers, Megan R. Friesen, Louis Ranjard.

Formal analysis: Ivan Braga Campos, Anne C. Gaskett, Louis Ranjard.

Funding acquisition: Ivan Braga Campos.

Investigation: Ivan Braga Campos, Anne C. Gaskett, Louis Ranjard.

Methodology: Ivan Braga Campos, Kate D. Lee, Louis Ranjard.

Project administration: Ivan Braga Campos, Anne C. Gaskett.

Software: Louis Ranjard.

Supervision: William George Lee, Anne C. Gaskett, Louis Ranjard.

Validation: Ivan Braga Campos.

Visualization: Ivan Braga Campos, Anne C. Gaskett, Louis Ranjard.

Writing – original draft: Ivan Braga Campos.

Writing – review & editing: Ivan Braga Campos, William George Lee, Anne C. Gaskett, Louis Ranjard.

References

1. Zilli D, Parson O, Merrett GV, Rogers A. A hidden Markov model-based acoustic cicada detector for crowdsourced smartphone biodiversity monitoring. *Journal of Artificial Intelligence Research* 2014; 51:805–827.
2. Aide TM, Corrada-Bravo C, Campos-Cerqueira M, Milan C, Vega G, Alvarez R. Real-time bioacoustics monitoring and automated species identification. *PeerJ* 2013; 1(e103):1–19.
3. Xie J, Towsey M, Zhang L, Yasumiba K, Schwarzkopf L, Zhang J, et al. Multiple-Instance Multiple-Label Learning for the Classification of Frog Calls with Acoustic Event Detection. In: Mansouri A, Nouboud F, Chalifour A, Mammass D, Meunier J, Elmoataz A, editors. *Image and Signal Processing: 7th International Conference, ICISP 2016, Trois-Rivières, QC, Canada, May 30–June 1, 2016, Proceedings* Cham: Springer International Publishing; 2016. p. 222–230.
4. Deichmann JL, Hernández-Serna A, Campos-Cerqueira M, Aide TM. Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest. *Ecol Ind* 2017; 74:39–48.
5. Andreassen T, Surlykke A, Hallam J. Semi-automatic long-term acoustic surveying: A case study with bats. *Ecological Informatics* 2014; 21:13–24.
6. Newson SE, Evans HE, Gillings S. A novel citizen science approach for large-scale standardised monitoring of bat activity and distribution, evaluated in eastern England. *Biol Conserv* 2015; 191:38–49.
7. Rocha LH, Ferreira LS, Paula BC, Rodrigues FH, Sousa-Lima RS. An evaluation of manual and automated methods for detecting sounds of maned wolves (*Chrysocyon brachyurus* Illiger 1815). *Bioacoustics* 2015; 24(2):185–198.
8. Palacios V, López-Bao JV, Llana L, Fernández C, Font E. Decoding Group Vocalizations: The Acoustic Energy Distribution of Chorus Howls Is Useful to Determine Wolf Reproduction. *PLoS one* 2016; 11(5):e0153858. <https://doi.org/10.1371/journal.pone.0153858> PMID: 27144887
9. Sanders CE, Mennill DJ. Acoustic monitoring of nocturnally migrating birds accurately assesses the timing and magnitude of migration through the Great Lakes. *The Condor* 2014; 116(3):371–383.
10. Stowell D, Benetos E, Gill LF. On-bird Sound Recordings: Automatic Acoustic Recognition of Activities and Contexts. *IEEE/ACM Transactions on Audio, Speech, And Language Processing* 2017; 25(6):1193–1206.
11. Ranjard L, Withers SJ, Brunton D, Ross HA, Parsons S. Integration over song classification replicates: evidence for song types and micro-geographic variation in the hihi. *Journal of the Acoustical Society of America* 2015; 137(5):2542–2551. <https://doi.org/10.1121/1.4919329> PMID: 25994687
12. Ranjard L, Withers SJ, Brunton DH, Parsons S, Ross HA. Geographic patterns of song variation reveal timing of song acquisition in a wild avian population. *Behavioral Ecology* 2017; 28(4):1085–1092.
13. Putland RL, Ranjard L, Constantine R, Radford CA. A hidden Markov model approach to indicate Bryde's whale acoustics. *Ecological Indicators* 2018; 84(Supplement C):479–487.
14. Wrege PH, Rowland ED, Keen S, Shiu Y. Acoustic monitoring for conservation in tropical forests: examples from forest elephants. *Methods in Ecology and Evolution* 2017; 8(10):1292–1301.
15. Cragg JL, Burger AE, Piatt JF. Testing the effectiveness of automated acoustic sensors for monitoring vocal activity of marbled murrelets *Brachyramphus marmoratus*. *Mar Ornithol* 2015; 43:151–160.
16. Zwart MC, Baker A, McGowan PJK, Whittingham MJ. The Use of Automated Bioacoustic Recorders to Replace Human Wildlife Surveys: An Example Using Nightjars. *PLOS ONE* 2014; 9(7):e0102770. <https://doi.org/10.1371/journal.pone.0102770> PMID: 25029035
17. Borker AL, McKown MW, Ackerman JT, EAGLES-SMITH CA, Tershy BR, Croll DA. Vocal activity as a low cost and scalable index of seabird colony size. *Conserv Biol* 2014; 28(4):1100–1108. <https://doi.org/10.1111/cobi.12264> PMID: 24628442
18. Jennings N, Parsons S, Pocock M. Human vs. machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. *Can J Zool* 2008; 86(5):371–377.
19. Buxton RT, Jones IL. Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. *J Field Ornithol* 2012; 83(1):47–60.
20. Mulder CPH, Anderson WB, Towns DR, Bellingham PJ editors. *Seabird islands: ecology, invasion, and restoration*. New York: Oxford University Press; 2011.

21. Potamitis I, Ntalampiras S, Jahn O, Riede K. Automatic bird sound detection in long real-field recordings: Applications and tools. *Appl Acoust* 2014; 80:1–9.
22. Ranjard L, Reed BS, Landers TJ, Rayner MJ, Friesen MR, Sagar RL, et al. MatlabHTK: a simple interface for bioacoustic analyses using hidden Markov models. *Methods in Ecology and Evolution* 2017; 8:615–621.
23. Ismar SMH, Baird KA, Gaskin CP, Taylor GA, Tennyson AJD, Rayner MJ, et al. A case of natural recovery after the removal of invasive predators—community assemblage changes in the avifauna of Burgess Island. *Notornis* 2014; 61:188–195.
24. Young SJ, Evermann G, Gales MJF, Hain T, Kershaw D, Moore G, et al. *The HTK Book*. Cambridge University Engineering Department, Cambridge 2006.
25. Eaton JW, Bateman D, Hauberg S, and Wehbring R. GNU Octave version 3.8.1 manual: a high-level interactive language for numerical computations. 2014.
26. Gage SH, Napoletano BM, Cooper MC. Assessment of ecosystem biodiversity by acoustic diversity indices. *J Acoust Soc Am* 2001; 109(5):2430.
27. Warham J. *The petrels: their ecology and breeding systems*. London: Academic Press; 1990.
28. Landers TJ, Bannock CA, Hauber ME. Dynamics of behavioural rhythms in a colonial, nocturnal, burrowing seabird: a comparison across different temporal scales. *Notornis* 2011; 58:81–89.
29. Ross EL, Brunton D. Seasonal trends and nightly variation in colony attendance of grey-faced petrels (*Pterodroma macroptera gouldi*). *Notornis* 2002; 49(3):153–157.
30. Howald G, Donlan C, Galván JP, Russell JC, Parkes J, Samaniego A, et al. Invasive rodent eradication on islands. *Conserv Biol* 2007; 21(5):1258–1268. <https://doi.org/10.1111/j.1523-1739.2007.00755.x> PMID: 17883491
31. Bellingham PJ, Towns DR, Cameron EK, Davis JJ, Wardle DA, Wilmshurst JM, et al. New Zealand island restoration: seabirds, predators, and the importance of history. *N Z J Ecol* 2010; 34(1):115.