

Talarico, H. R., G. J. Jordan, E. J. Woehler, R. Hamer, and J. L. Lavers. 2025. The efficacy of acoustic indices in detecting the post-migration return of Short-tailed Shearwaters (*Ardenna tenuirostris*) to their colonies. *Avian Conservation and Ecology* 20(1):26. <https://doi.org/10.5751/ACE-02894-200126>

Copyright © 2025 by the author(s). Published here under license by the Resilience Alliance. Open Access. CC-BY 4.0

Research Paper

## The efficacy of acoustic indices in detecting the post-migration return of Short-tailed Shearwaters (*Ardenna tenuirostris*) to their colonies

Harrison R. Talarico<sup>1,2</sup> , Gregory J. Jordan<sup>1</sup> , Eric J. Woehler<sup>3</sup> , Rowena Hamer<sup>1,4</sup>  and Jennifer L. Lavers<sup>2,5,6</sup> 

<sup>1</sup>Biological Sciences, University of Tasmania, Hobart, Tasmania, Australia, <sup>2</sup>Adrift Lab, Underwood, Tasmania, Australia,

<sup>3</sup>Australasian Seabird Group, BirdLife Australia, <sup>4</sup>Tasmanian Land Conservancy, <sup>5</sup>Gulbali Institute, Charles Sturt University, Wagga Wagga, New South Wales, Australia, <sup>6</sup>Bird Group, The Natural History Museum, Tring, Hertfordshire, United Kingdom

**ABSTRACT.** The migratory timings of many seabirds are shifting because of climate change, and viable methods to quantify these patterns are required. Using acoustic recording units (ARUs) to detect vocalizations can be effective in monitoring seabird migration. Specifically, ARUs can help detect the arrival dates of seabirds and assess the impact of environmental changes. However, processing large volumes of acoustic data can be challenging, particularly in seabird colonies where overlapping vocalizations and non-target noise are prevalent. Acoustic indices may mitigate some methodological challenges associated with processing acoustic data from colonies as they are fully automated and easy to use for researchers with limited experience in bioacoustics. However, little is known about the efficacy of acoustic indices in quantifying the arrival dates of seabirds. Short-tailed Shearwaters (*Ardenna tenuirostris*) are ecologically and culturally important migratory birds. Their late arrival to breeding areas in Tasmania in recent years created a need for efficient monitoring. This study aimed to explore the use of acoustic indices in identifying the timing of post-migratory return to their colonies in southeastern Tasmania. Five nightly subsamples were taken from recordings over three to four weeks at nine nesting sites, which spanned the period around the normal arrival time of the birds. The samples were manually tagged for vocalizations and other biotic and abiotic sounds. We used two acoustic indices (ACI and NDSI; applied to both raw data and data filtered using a novel wind-filter) to detect change points in number of vocalizations, as a proxy for arrival dates. By using a combination of the NDSI on raw data and ACI on wind-filtered data, we were able to detect the arrival of Short-tailed Shearwaters at seven of the nine study sites

### Efficacité des indices acoustiques à détecter le retour post-migratoire des Puffins à bec grêle (*Ardenna tenuirostris*) dans leurs colonies

**RÉSUMÉ.** Les périodes de migration de nombreux oiseaux marins varient en raison des changements climatiques, et des méthodes viables visant à quantifier ces tendances sont nécessaires. L'utilisation d'enregistreurs automatisés (EA) pour détecter les vocalisations peut s'avérer efficace dans le suivi de la migration des oiseaux marins. En particulier, les EA peuvent servir à détecter les dates d'arrivée des oiseaux marins et à évaluer les effets des changements environnementaux. Cependant, le traitement d'un grand volume de données acoustiques peut s'avérer difficile, en particulier dans les colonies d'oiseaux marins où les vocalisations qui se chevauchent et les bruits non ciblés sont fréquents. Les indices acoustiques peuvent atténuer certains défis méthodologiques associés au traitement de données acoustiques récoltées sur des colonies, car ils sont entièrement automatisés et faciles à utiliser pour les chercheurs ayant une expérience limitée en bioacoustique. Toutefois, on sait peu de choses sur l'efficacité de ces indices à quantifier les dates d'arrivée des oiseaux marins. Les Puffins à bec grêle (*Ardenna tenuirostris*) sont des oiseaux migrateurs importants d'un point de vue écologique et culturel. Leur arrivée tardive à leur site de nidification en Tasmanie ces dernières années a engendré un besoin de surveillance efficace. La présente étude visait à explorer l'utilisation d'indices acoustiques pour identifier le moment du retour post-migratoire des puffins vers leurs colonies dans le sud-est de la Tasmanie. Cinq sous-échantillons nocturnes ont été prélevés à partir d'enregistrements effectués pendant trois à quatre semaines sur neuf sites de nidification, couvrant la période habituelle du retour des oiseaux. Les échantillons ont été marqués manuellement pour les vocalisations et les autres sons biotiques et abiotiques. Nous avons utilisé deux indices acoustiques (l'Acoustic Complexity Index [ACI] et le Normalized Difference Soundscape Index [NDSI], appliqués à la fois aux données brutes et aux données filtrées à l'aide d'un nouveau filtre pour éliminer le vent) pour détecter les points d'inflexion du nombre de vocalisations, en tant qu'indicateurs des dates d'arrivée. En combinant l'utilisation de l'indice NDSI sur les données brutes et de l'ACI sur les données filtrées pour éliminer le vent, nous avons réussi à détecter l'arrivée des Puffins à bec grêle à sept des neuf sites d'étude.

**Key Words:** *acoustic indices; migration; monitoring methods; phenology; seabirds; soundscape; vocalization*

## INTRODUCTION

Seabirds respond to biological and physical oceanographic conditions at a broad range of temporal and spatial scales (Haury et al. 1978), and determining the demographic and phenological

characteristics of seabird species can provide a unique insight into the marine environment (Woehler 2012, Lavers and Bond 2023). However, capturing data on such characteristics can be labor and resource intensive, impractical for large-scale studies, and

impossible over the life-spans of individual seabirds whose life expectancies can exceed 60 years in the wild (e.g., Jiménez-Uzcátegui et al. 2016, Yap et al. 2021). Even single-season studies involving manual monitoring and data collection require substantial commitment of resources (Birkhead 2014).

Acoustic recording units (ARUs) can improve the efficiency of data collection and processing within seabird colonies on a variety of temporal scales (Oppel et al. 2014), both in single-species (e.g., Lavers et al. 2019) and mixed-species (Ratcliffe et al. 2015) colonies. Several studies have demonstrated that acoustic monitoring can be effective in measuring aspects of breeding season phenology and demography (e.g., Lynch et al. 2015, Hinke et al. 2018).

The arrival of colonial seabirds at their nesting colonies presents an opportunity to quantify the interannual variability in the timing of return, which can provide a basis for examining potential cause(s) of observed changes or trends. For example, local or regional prey resources and other environmental stressors can strongly influence the onset of breeding following return to nest sites in seabirds (Cappello and Boersma 2021, Fayet et al. 2021, Glencross et al. 2021). However, focusing solely on first arrival dates can be misleading, as early individuals may not represent broader population trends (Tryjanowski and Sparks 2001, Mizel et al. 2019). The arrival time(s) of nesting seabirds can be quantified using passive acoustic monitoring, and can be made more efficient when used in conjunction with automated techniques for quantifying changes in vocalizations of the targeted species (Sueur and Farina 2015, Sugai et al. 2019). Devising such methods can help researchers better understand how seabirds may be responding to environmental changes at a range of spatial scales (Woehler and Hobday 2024).

The two main ways of automating the analysis of acoustic recordings is (a) to use acoustic indices, and (b) to develop automated recognizers. Recognizers involve training a computer to detect the acoustic signature of a targeted species' vocalization, and are often created using machine learning techniques (Xie et al. 2023). However, the performance of an automated call recognizer may be impacted by the presence of non-target noise (Buxton and Jones 2012, Willacy et al. 2015), overlapping calls (Sugai et al. 2019), or variation in calls (including both dialects and variation within individuals) of a given species (Stowell et al. 2019). Moreover, building machine learning models can be challenging for those with little experience in the field, however this technique is becoming more accessible through the application of multispecies pre-trained models such as BirdNET and Perch (Dumoulin 2022, Cornell University 2025). Acoustic indices are mathematical summaries of the distribution of acoustic energy within a recording that focus on the characteristics of biotic and abiotic sound (Buxton et al. 2016, Brownlie et al. 2020). Although acoustic indices do not directly identify the calls of specific birds in the way that recognizers do, they are less impacted by some of the limitations of recognizers (Pérez-Granados and Traba 2021). However, acoustic indices can also be influenced by dominant sounds in the soundscape, which may mask vocalizations of a targeted species, especially if the targeted animal's vocalizations are quiet and infrequent (Buxton et al. 2013, Bradfer-Lawrence et al. 2020).

Short-tailed Shearwaters (*Ardenna tenuirostris*) are a trans-hemispheric seabird species that breeds in southeastern Australia then migrates to the North Pacific Ocean for the non-breeding

season (Price et al. 2020a). They are culturally important for Tasmanian Aboriginal communities (Lavers and Bond 2013). Short-tailed shearwaters typically return to their breeding colonies in the last week of September (Carey et al. 2014, Price et al. 2020b, Glencross et al. 2021), arriving in high numbers over five nights, but sometimes over a longer period of up to 10–12 nights (Naarding 1981). These birds were once considered to be highly reliable in their migratory phenology (Serventy 1966, Carey et al. 2014, Cleeland et al. 2014), however observations have suggested an apparent increase in the variability and extent of the return to colonies in recent years (Glencross et al. 2021). These recent observations are largely subjective, and need to be examined to determine if there has been a phenological change in the species.

Short-tailed Shearwaters are an ideal avian species for studying migratory and breeding behaviors because of their reliable vocalization patterns at their breeding colonies. Upon returning from their migration, Short-tailed Shearwaters are vocally active after sunset, as they engage in activities such as burrow preparation, pair bond reestablishment, and defense against conspecifics. Our study investigated the performance of different acoustic indices in assessing the return of Short-tailed Shearwater to their breeding colonies in southeast Tasmania. Specifically, we used change point analyses to identify changes in the values of acoustic indices and numbers of vocalizations present on recordings obtained within nesting colonies. Although there were other sources of biotic sound at the study sites, it was hypothesized that Short-tailed Shearwater vocalizations would be highly correlated with acoustic index values as their arrival would result in a detectable shift in the soundscape. The arrival of returning Short-tailed Shearwater could then be inferred from the changes in acoustic index values.

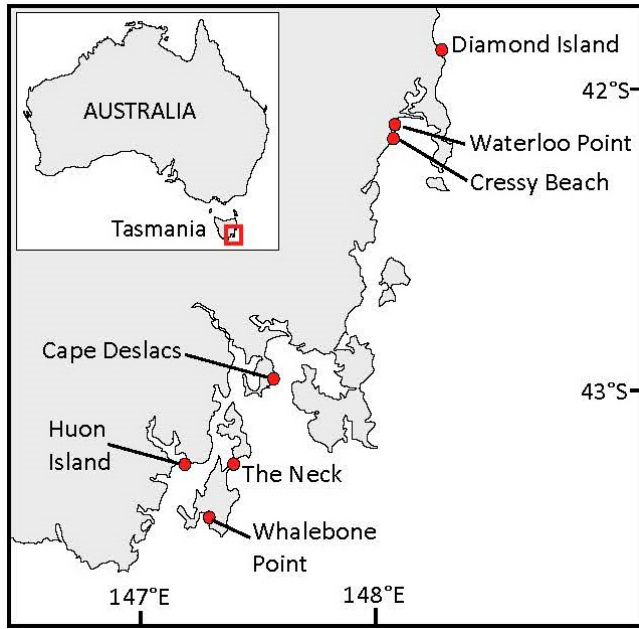
## MATERIALS AND METHODS

### Field data collection

Seven Short-tailed Shearwater breeding colonies in southeast Tasmania were selected based on breeding population size, logistic access for deployment and recovery of the equipment, and relative freedom from human disturbance and interference (Fig. 1). Within each colony, Song Meter SM4 acoustic recorders were placed in areas where there was evidence of Short-tailed Shearwater activity (e.g., fresh digging, footprints), but avoiding areas with high potential exposure to wind or that were visible from walking tracks, roads, or other areas with human activity. A single recorder was deployed at five of the colonies, but two recorders were deployed in two of the largest colonies used in the study (Table 1). The recorders in these latter colonies were placed in microsites that differed in wind exposure and were widely spaced from each other.

The recorders were deployed between mid-September and mid-October 2020, spanning the expected arrival time of the birds (Table 1). Each recorder was scheduled to record continuously from one hour before sunset to two hours after sunset each night because this is when nocturnal shearwaters are active at the colony (Warham 1960). The acoustic recorders were deployed with the manufacturer's default settings of stereo mode with equal 16dB gain from each microphone and a sampling rate of 24 kHz, and built-in microphones. Recordings were saved in uncompressed WAV format, and no high-pass filter was applied.

**Fig. 1.** Location of the Short-tailed Shearwater (*Ardenna tenuirostris*) colonies monitored in this study in southeast Tasmania. Two acoustic recorders were deployed at each of the Diamond Island and Whalebone Point colonies, and one recorder was deployed at the remaining colonies.



**Table 1.** Short-tailed Shearwater (*Ardenna tenuirostris*) colonies monitored in this study. Sites are listed from south to north (see Fig. 1). Population data (estimated burrows) from Skira et al. (1996).

Colony	Date deployed	Date retrieved	Approx. colony size (ha)	Estimated population (burrows)
Whalebone Point 1	19/09/2020	16/10/2020	2.1	7830–10,530
Whalebone Point 2	19/09/2020	16/10/2020		
The Neck	19/09/2020	16/10/2020	3.4	15,000–23,000
Huon Island	21/09/2020	17/10/2020	1.0	300–500
Cape Deslacs	20/09/2020	17/10/2020	3.0	17,820–19,800
Cressy Beach	20/09/2020	11/10/2020	0.1	2000–3000
Waterloo Point	20/09/2020	11/10/2020	0.2	500–1000
Diamond Island 1	20/09/2020	11/10/2020	3.0	No estimate
Diamond Island 2	20/09/2020	11/10/2020		

## Subsampling the recordings

Overall, the recorders produced 228 raw data files, each three hours in length. To reduce the substantial computational cost involved in using acoustic indices to analyze recordings, and to provide relative homogeneity within samples, one 15s subsample was taken every 30 minutes from the raw recordings, starting at the 50th minute (i.e., 10 minutes before sunset). Fifteen-second subsamples were chosen because manual identification and tagging of shearwater vocalizations was time-consuming because of the large volume of calls in each file. In total, 1140 fifteen-second subsamples were created (five per recording). Each subsample was reviewed manually using Song Scope 4.1.5A

(Wildlife Acoustics 2011). Both the spectrogram and waveform were used during review. Spectrogram settings were set to a 1-second background filter and a FFT window size of 512, with display settings adjusted to brightness -9, contrast 3, and luminosity 36.

## Manual review process

The subsamples were manually reviewed to extract relevant biological information for analyses. At their colonies, Short-tailed Shearwater have two distinct vocalizations in their repertoire, a guttural “kooka-rooka” call and a rising “koo-roo-rah” call (Szabo 2013). The “koo-roo-rah” calls (Fig. 2) occur mainly within the 2 to 10 kHz range, with harmonics extending across multiple frequency bands. As is common for many biological sounds, including seabird vocalizations, most of the energy of these calls fall between 2000 Hz and 8000 Hz (Kasten et al. 2012). Conversely, the “kooka-rooka” call (Fig. 2) occurs at very low frequencies, with most of the energy falling between 100 Hz and 2000 Hz. Abiotic noise often occurs in these lower frequencies (as apparent in Fig. 2), so we restricted our analyses to the “koo-roo-rah” calls. Each subsample typically took between 15 seconds and 1 minute to tag, depending on the density of Short-tailed Shearwater vocalizations in the recording.

Qualitative estimates of wind noise and other sources of interference were also scored on a six-step scale (from very low to extreme) based on how frequently their signals appeared in the recording, and in the case of wind, the presence of waveform distortion (clipping). Wind was scored on a scale of 1 to 6: (1) very low presence and no waveform clipping, (2) low presence and no waveform clipping, (3) medium presence and/or 1–5% of the sample displayed waveform clipping due to wind, (4) 5–20% of the sample displayed waveform clipping, (5) 20–50% of the sample displayed waveform clipping, (6) > 50% of the sample displayed waveform clipping.

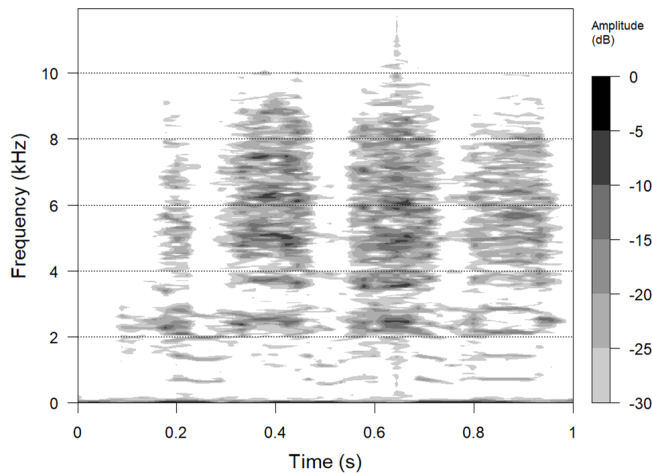
## Acoustic indices selection

Two acoustic indices were chosen to analyze the subsamples from the two datasets: the Acoustic Complexity Index (ACI) and the Normalized Difference Soundscape Index (NDSI) using R (R Core Team 2022) and the packages *soundecology* (Villanueva-Rivera and Pijanowski 2022) and *tuneR* (Ligges et al. 2023). For ACI, we used a cluster length of 1 second, a FFT window size of 512, and limited the frequency range to 2000–8000 Hz to match the dominant energy of Short-tailed Shearwater calls. For NDSI, the biotic band was set to 2000–8000 Hz and the abiotic band to 1000–2000 Hz. Each index calculation returned one value each for the left and right microphone channels, and these were averaged to provide a single value per sample.

Preliminary analyses indicated that these two indices were more powerful in recognizing Short-tailed Shearwater vocalizations than other available indices (i.e., the Acoustic Evenness Index, the Acoustic Diversity Index, and the Bioacoustic Index; Appendix 1; Farina 2025). The ACI relies on the assumption that biotic sounds are characterized by an intrinsic variability of intensities. It is calculated as the difference in spectral amplitude within a frequency band from one time sample to the next, and averaged over all frequency bands for the duration of the recording (Pieretti et al. 2011). The frequency scale of 2000–8000 Hz was selected for the ACI analysis, as this interval corresponds to the predominant acoustic energy of Short-tailed Shearwater vocalizations.



**Fig. 2.** Spectrogram of the higher frequency “koo-roo-rah” vocalization of Short-tailed Shearwater (*Ardena tenuirostris*). The images were generated using the seewave package in R v4.1.0 with a Hanning window (FFT size - 512, 75% overlap, sample rate = 24kHz).



The NDSI relies on a theoretical frequency split between non-biotic noise and biotic noise. It is calculated by computing the ratio of non-biotic noise and biotic noise in a recording on a range of -1 to +1 (with +1 indicating more biotic noise in the soundscape; Kasten et al. 2012). Because non-biotic noise typically occurs at lower frequencies than biotic sounds, the non-biotic frequency range was set between 1000 and 2000 Hz and the biotic frequency range was set between 2000 and 8000 Hz in this study. Notably, the “kooka-rooka” vocalizations of Short-tailed Shearwater fall within the NDSI’s non-biotic frequency range. Despite this potential limitation, the NDSI is still considered a suitable index. This is based on the consistent observation that when calls were detected in the lower frequency band (1000–2000 Hz), the higher frequency “koo-roo-rah” calls tended to co-occur. Additionally, when both types of calls were present, most of the acoustic energy within the soundscape was concentrated within the 2000–8000 Hz frequency band. These observations indicate that when Short-tailed Shearwater vocalizations overlapped into the non-biotic range, the ratio of biotic to non-biotic noise remained positive.

### Wind-filtered dataset

Because many of the recordings were affected by wind, we compiled two data sets (see Appendix 1 and Fig. S1 and S2 for details). The first data set included all samples except those noted as having extreme levels of wind (which were deemed unusable), leaving a total of 1005 samples (88% of the 1140 samples). We also compiled a wind-filtered dataset by removing samples with an automated wind filter. Wind filters are usually applied at lower frequency ranges (e.g., 0–1000 Hz; Brown et al. 2018a, 2018b). However, we chose a higher frequency range to prevent filtering out of recordings that contained Short-tailed Shearwater guttural calls because these recordings also often included the targeted “koo-roo-rah” calls. Thus, we plotted ACI values for frequency bands 11000–12000 Hz against manually scored “wind” values

(on a six-point scale) to examine the correlation strength. Subsequently, the wind-filtered dataset was created by removing samples with ACI scores in this frequency band greater than 45. The wind filter removed 100% of recordings in the extreme wind category, 96.2% in the very high category, 67.8% in the high category, and only 5.6% in the medium, low, and very low categories. The wind filter also removed 45% of recordings that contained rain in the soundscape. The resulting “wind-filtered” dataset consisted of 722 samples (63% of the full data set), and included samples from every night at every site, except for one day at one site well before arrival of the birds.

### Statistical analyses

To determine whether changes in index values corresponded with the migratory arrival of Short-tailed Shearwaters, we performed change point analyses with the *cpt.mean* function from the *changept* package in R (Killick and Eckley 2014). The change point analyses infer transitions between different states within the time series data by identifying abrupt changes in the mean values. Thus, at each site, we applied the At Most One Change (AMOC) algorithm on the time series of observations to detect a single change point in values of ACI, NDSI, and vocalizations, using the CUSUM test statistic and a penalty of 0.5 for the number of vocalizations and ACI, and the normal statistic and a penalty of 1 for NDSI. The penalty values for each index were chosen at a threshold that would maximize the likelihood of detecting one change point per site (although the penalty values remained consistent between sites). We accepted changes that were significant (i.e.  $p < 0.05$ ) for all analyses as change points.

We used the timing of a significant change in the average number of Short-tailed Shearwater vocalizations detected per night in the unfiltered data as a proxy for the true value of the main arrival of Short-tailed Shearwater in the colony. We then anecdotally compared these against change points in the nightly average of ACI and NDSI values on both the unfiltered and wind-filtered datasets.

## RESULTS

Vocalization change points were detected at eight of the nine study sites (Table 2), representing 89% of the colonies. Across both unfiltered and wind-filtered datasets, the detection abilities of the NDSI and ACI varied.

In the unfiltered dataset, NDSI successfully detected vocalization change points exactly on the day (+0 days) at 50% of the sites where a vocalization change point was detected (Fig. 3, Tables 2 and 3). ACI identified the vocalization change points at 12.5% of the sites. When considering a +1 day margin, NDSI’s detection rate increased to 62.5%, while ACI’s improved to 37.5%. Extending to a +2 day margin, NDSI maintained a detection rate of 62.5%, and ACI improved to 50%. At a +3 day margin, the detection rates for both NDSI and ACI remained the same, at 62.5% and 50% respectively.

In the wind-filtered dataset, both NDSI and ACI detected vocalization change points at +0 days at 25% of the sites (Fig. 3). With a +1 day margin, NDSI’s detection rate increased to 50% whereas ACI’s jumped to 75%. Extending further to a +2 day margin, NDSI’s rate remained at 50%, while ACI’s detection rate climbed to 87.5%. Both indices maintained these rates at a +3 day

**Table 2.** The day of the first significant change in index values at each site, calculated for Short-tailed Shearwater (*Ardenna tenuirostris*) vocalizations in Tasmania during the Austral spring of 2020. NDSI = Normalized Difference Soundscape Index; ACI = Acoustic Complexity Index; (unfiltered) refers to data analyzed without wind filtering; (wf) refers to data processed using a wind filter. The table shows the differences in days between the index change point and the vocalization change point. Positive values indicate the index detected a change after the vocalization change point, while negative values indicate the index detected a change earlier. “No change” indicates no significant change point was detected for that index at the respective site. “No. of vocalizations” is the total number of vocalizations recorded per site across all the recordings with 15 seconds reviewed per recording per day, and up to 6 recordings per day were examined.

Colony	No. of vocalizations	Day of first vocalization	Change point (vocalizations)	Δ NDSI (unfiltered)	Δ ACI (unfiltered)	Δ NDSI (wf)	Δ ACI (wf)
Diamond Island 1	724	277	278	0 days	-1 day	0 days	-1 day
Diamond Island 2	701	277	278	0 days	0 days	0 days	0 days
Whalebone Point 1	928	274	277	0 days	-6 days	-1 day	0 days
Whalebone Point 2	611	277	277	0 days	-2 days	+3 days	+4 days
The Neck	286	277	278	No change	+10 days	No change	-1 day
Huon Island	544	280	281	No change	-12 days	No change	+1 day
Waterloo Point	250	277	280	+1 day	+1 day	+1 day	+1 day
Cape Deslacs	174	278	282	No change	-3 days	No change	-2 days
Cressy Beach	12	281	No change	No change	279	No change	No change

margin. Notably, ACI detected the vocalization change points at two more sites than the NDSI across both datasets when considering detections within a +2 day margin.

The mean absolute difference between the vocalization change point and the detection by NDSI in the unfiltered dataset was  $0.2 \pm 0.45$  days SD, indicating a close alignment with actual vocalization change points (Table 3). For ACI, the mean absolute difference was higher at  $4.38 \pm 6.26$  days SD. In the wind-filtered dataset, the mean absolute difference for NDSI worsened to  $1.0 \pm 1.52$  days SD, while ACI showed a significant improvement to  $1.25 \pm 2.19$  days SD.

## DISCUSSION

We observed that the temporal change points in acoustic indices largely occurred on the same day or within one day of the vocalization change points. This suggests that the change points represent the main arrival time of birds, indicating a significant increase in vocal activity and the mass arrival of individuals at the colony. In contrast, the vocalizations detected before the change points likely reflect the arrival of a few early individuals. This pattern is consistent with the staggered return observed in Short-tailed Shearwaters (Naarding 1981) and similar phenomena in other seabird species (Woehler 2012, Fayet et al. 2021). Although first arrival dates have previously formed the basis for inferences about migration timings (Tryjanowski and Sparks 2001, Mizel et al. 2019), these findings often do not reflect changes in the population (Miller-Rushing et al. 2008, Mizel et al. 2019). Indeed, a study on American passerines showed strikingly different patterns of change between earliest recorded arrival dates and mean arrival dates (Miller-Rushing et al. 2008). Therefore, identifying change points is crucial as they may provide a more accurate measure of population dynamics and the timing of key biological events.

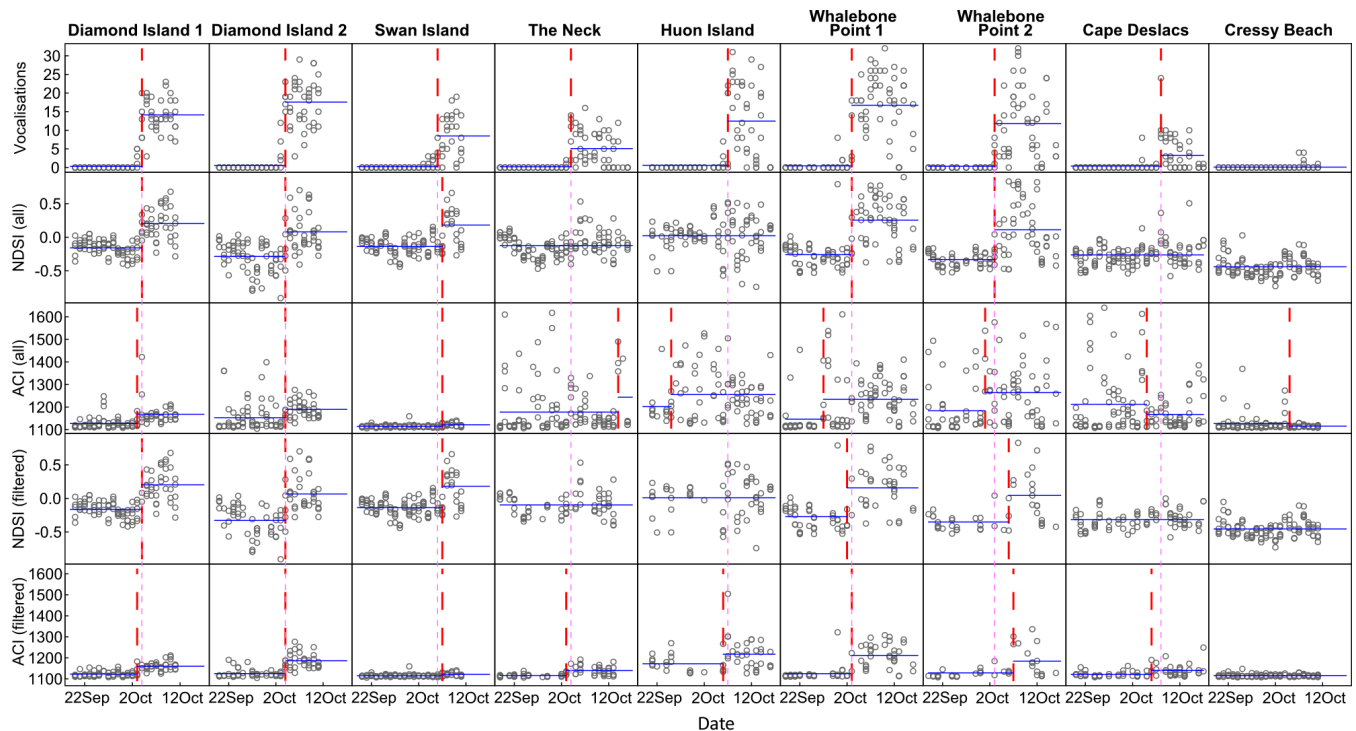
For the unfiltered dataset, the NDSI was more effective than the ACI at detecting the return of Short-tailed Shearwaters, although ACI outperformed NDSI once a wind filter was applied. This difference in performance stems from the way each index operates and their sensitivity to specific types of noise and frequency ranges. NDSI relies on the ratio of biotic to abiotic sounds, which

enables it to effectively capture shifts in the soundscape when both biotic and abiotic noise were present. The NDSI is more robust to abiotic noise because it measures the ratio between biotic (vocalizations) and abiotic sounds (such as wind). This allows the NDSI to factor in abiotic noise in lower frequency ranges without distorting the detection of biotic signals in higher frequency ranges. By focusing on the relative balance between the two, NDSI can effectively identify changes in biotic activity, making it more capable of detecting shifts in the soundscape, such as the arrival of birds, even in noisy environments.

Conversely, ACI measures the complexity of the soundscape, meaning it is more sensitive to all types of noise, including wind. This lack of clear distinction can make it harder to pinpoint significant shifts in biotic activity, as high ACI values may occur in both windy and vocalization-rich recordings. As a result, there may not be enough differentiation between these conditions to accurately capture change points related to the birds' arrival in unfiltered datasets. However, when a wind filter was applied, ACI's performance improved markedly. Whereas the NDSI exhibited a lower mean absolute difference ( $0.2 \pm 0.45$  days) in the unfiltered dataset, the wind-filtered ACI demonstrated greater predictive power. Despite having a slightly higher mean absolute difference of  $1.25 \pm 2.19$  days, the ACI successfully predicted vocalization change points within +1 day at 75% of sites where a change point was detected (8 out of 9 sites) and within +2 days at 87.5% of sites (compared to 50% at +0 days and 62.5% within +2 days for the NDSI in the unfiltered dataset). This level of consistency across multiple sites makes the wind-filtered ACI a more robust tool for monitoring migratory arrivals, especially in challenging field conditions where environmental noise may interfere with other indices.

In practical terms, the error margin of up to 1.25 days for the wind-filtered ACI falls within acceptable limits for management purposes. This level of error likely will not significantly affect the outcomes of monitoring programs that aim to detect shifts in arrival dates due to long-term environmental change. Although improvements in index performance could refine precision, the current level of error provides a reliable basis for detecting conspicuous shifts in phenology over time.

**Fig. 3.** Characteristics of Short-tailed Shearwater (*Ardenna tenuirostris*) recordings. Red dashed vertical lines represent change points. The fine pink dashed lines represent the change points in vocalizations extended to acoustic indices. Unbroken horizontal lines are estimates of the mean values before and after change points. First row: number of vocalizations; 2nd row: unfiltered Normalized Difference Soundscape Index (NDSI); 3rd row: unfiltered Acoustic Complexity Index (ACI); 4th row: wind-filtered NDSI; 5th row: wind-filtered ACI. Note that all indices detected the change point in vocalizations at Diamond Island 1, Diamond Island 2, and Swan Island; the wind-filtered ACI detected the change point in vocalizations at The Neck, Huon Island, and Whalebone Point 1; and unfiltered NDSI detected the change in vocalizations at Whalebone Point 1 and Whalebone Point 2. The change point in vocalizations at Cape Deslacs was not detected by the indices and no change point in vocalizations was detected at Cressy Beach.



The overall better performance of the ACI over NDSI may be due to flaws associated with the assumed frequency split between biotic and anthropogenic noise underpinning NDSI. For the biotic component at least, these limits seem artificial (Sueur et al. 2014). Many animals produce sounds below 2000 Hz, including some of the species detected at our study sites. For instance, although high-pitched Little Penguin (*Eudyptula minor*) vocalizations were also observed in the study (2000 Hz–8000 Hz), breeding Little Penguins also engage in low-frequency calls (0.2 Hz–2000 Hz), both before and after pairs engage in mutual displays (Colombelli-Négrel and Smale 2018). The guttural “kooka-rooka” Short-tailed Shearwater call also occurs below 2000 Hz. If a recording with targeted Short-tailed Shearwater vocalizations also included a considerable number of low-frequency calls (from either Little Penguins or Short-tailed Shearwaters), then this would result in comparatively lower NDSI values due to the greater energy in the anthropogenic frequency bands.

When attempting to detect the vocal activity of biota, it may seem biologically appropriate to calculate indices at frequency ranges that encompass the entire vocal repertoire of a targeted species. However, there is often a trade-off between the sensitivity of the index to targeted vocalizations and the extent of the frequency

range to which it can be applied. Indices calculated across broad frequency bands are less viable as ecological indicators, because broader frequency ranges are more likely to contain superfluous acoustic information that could mask correlations between vocal activity and index values (Metcalf et al. 2021). Although Metcalf et al. employed acoustic indices to characterize community-level metrics, calculating indices at narrow frequency bands is also appropriate in studies on single species. For instance, broadening the frequency range of the ACI to include Short-tailed Shearwater guttural calls (100 Hz–2000 Hz) would have increased the amount of non-target noise that was summarized by the index, even when the wind filter was applied. This is because most of the energy of abiotic noise falls in lower frequency bands, with even light wind and rainfall causing significant increases in energy between 0 Hz and 1000 Hz (Bedoya et al. 2017, Fairbrass et al. 2017).

Interestingly, an inverse change point was detected in the raw ACI at Cape Deslacs three days before the vocalization change point (Fig. 2, Table 2). ACI is a parabolic index representing both a quiet sound environment with only the occasional call interrupting silence, and the “saturated” sound environment with constant calls (Pieretti et al. 2011). It is possible that an extremely saturated recording containing only Short-tailed Shearwater vocalizations could result in a comparatively low ACI, however,

**Table 3.** Detection performance and mean absolute differences from vocalization change points for Normalized Difference Soundscape (NDSI) and Acoustic Complexity Index (ACI) across unfiltered and wind-filtered datasets. The percentages indicate the frequency with which the acoustic indices detected vocalization change points within 0 to 3 days. The mean absolute difference, along with the standard deviation ( $\pm$  SD), reflects the accuracy of each index in relation to the actual vocalization change points.

Dataset	Index	Index detection of vocalization change point (+0 days)	+1 Day of vocalization change point	+2 Days of vocalization change point	+3 Days of vocalization change point	Mean absolute difference from vocalization change point (days)
Unfiltered	NDSI	50%	62.5%	62.5%	62.5%	0.2 $\pm$ 0.45
	ACI	12.5%	37.5%	50%	62.5%	4.38 $\pm$ 6.26
Wind-Filtered	NDSI	25%	50%	50%	62.5%	1 $\pm$ 1.52 days
	ACI	25%	75%	87.5%	87.5%	1.25 $\pm$ 2.19

that is not what occurred here. Cape Deslacs had very few vocalizations compared to other sites, and only 10 vocalizations were recorded in the days around the change point. Therefore, the change point likely reflected a conspicuous shift in abiotic conditions instead. Although the ACI may still indicate colony arrival when no wind filter is applied, it is important to consider environmental factors that may affect its readings.

Despite its effectiveness, applying a wind-filter entails a certain amount of data loss, and this method may therefore be inappropriate at particularly windy sites (e.g., offshore islands). In contrast, the unfiltered dataset may pose practical problems because its preparation involved the manual removal of extreme wind days, requiring some potentially time-consuming pre-preparation of the recordings.

To further improve the performance of indices and reduce the need for filtering recordings, weather barriers could be constructed around the ARUs. However, depending on the wall construction, wind can sometimes interact with the weather barrier and lead to turbulence noise (Walker and Hedlin 2010). Targeted signals may also be attenuated when weather barriers are in place (Christie et al. 2007). Despite these challenges, weather barriers can be very effective in improving the signal-to-noise ratio in recordings (Fristrup and Mennitt 2012, Buxton et al. 2016).

Given their automated nature, the logistical costs of applying our approaches on a larger scale are relatively small. Our sampling design was chosen to test the accuracy of the methods. However, in a management scenario, finer resolution and greater information content may be achieved by deploying multiple ARUs within sites, increasing the number of samples taken per night, or possibly analyzing whole recordings (Bradfer-Lawrence et al. 2020). Although ground-truthing recordings through collecting visual observations of bird arrivals can be extremely valuable (e.g., Pérez-Granados et al. 2019), this is difficult for nocturnal seabirds that only attend the colony in darkness. Thus, improved capacity to characterize vocalizations through better sampling may help overcome a key limitation of studies of nocturnal seabird arrival time.

Because it is unlikely that a single index can describe the diverse soundscapes that exist within and between Short-tailed Shearwater colonies, a combination of these two indices may perhaps be used to monitor their return. Selection of which index to use should be based on an a priori knowledge of the study site and an understanding of the soundscape patterns that underlie index values.

## Conclusions

When attempting to quantify the migratory arrival of Short-tailed Shearwaters with acoustic indices, one should consider other factors that can influence vocal activity. Vocalizations of seabirds may vary not only because of the arrival of the birds, but also because of environmental and biological factors. For instance, studies have shown that lunar phases can affect vocal activity in various seabird species, with vocalizations shown to decrease with increasing lunar phase (Mougeot and Bretagnolle 2000). Therefore, in the presence of a full moon, Short-tailed Shearwater vocalizations may be reduced upon arrival, and migration timings may be misinterpreted. In a study that investigated the relationship between acoustic indices and Short-tailed Shearwater burrow density, Brownlie et al. accounted for variations in vocal activity by incorporating a new moon value into their models (Brownlie et al. 2020). Brownlie et al. also noted that Short-tailed Shearwaters are more vocally active before dawn, suggesting that sampling before sunset might miss peak activity. These methods may also be considered for studies on the migratory arrival of Short-tailed Shearwaters.

The ACI and NDSI both offer valuable tools for detecting the migratory arrival of Short-tailed Shearwaters, though their effectiveness depends on the characteristics of the soundscape. The wind-filtered ACI, with an error of  $\sim 1.25$  days, successfully detected change points at 7 of 9 sites, making it particularly accurate and the recommended index for managers in environments where abiotic noise is minimal or can be effectively filtered out. In these quieter conditions, ACI provides precise detection of arrival times and should be prioritized for monitoring. At sites with significant abiotic noise, such as particularly windy environments, NDSI may be a more suitable index because of its ability to handle noisy conditions. However, in our study, it was the less effective index overall, identifying Short-tailed Shearwater arrival at 5 of 9 sites.

Both indices should therefore be adopted when monitoring the migratory arrival of Short-tailed Shearwaters. Some adjustment of the methods here are likely to improve the capacity to both detect vocalizations and quantify arrival times, for instance, adjusting sampling times to suit biologically appropriate scales (Metcalf et al. 2021) and using weather barriers to reduce the effects of abiotic sound (Fristrup and Mennitt 2012, Buxton et al. 2016). Overall, the findings of this study may be used to quantify the migration timings of other seabirds, although the effectiveness of the indices will likely depend on the structure of the targeted vocalizations and the makeup of the soundscape in



which they arrive (for instance, acoustic indices may be a less reliable tool in complex soundscapes with overlapping vocalizations of different species). Notwithstanding, the ACI and NDSI will likely be effective in monitoring the migratory arrival of a species with a similar vocal repertoire, such as the threatened Sable Shearwater *Ardenna carneipes* (Bond and Lavers 2024).

---

#### Author Contributions:

*EJW and JLL formulated the initial research questions; HRT and JLL completed the field work; HRT, RH, and GJJ processed the data; EJW, GJJ, RH, and JLL supervised the research, HRT wrote the original draft of the paper to which all authors contributed through editing and revisions.*

#### Acknowledgments:

*First and foremost, the authors acknowledge the Traditional Owners of the land on which this research took place (lutruwita/Tasmania, with multiple sites adjacent to nipaluna/Hobart). As the First Scientists of Australia, we value and recognize the incredible knowledge possessed by those who've lived on this country for more than 50,000 years. We thank staff at the Tasmanian Department of Natural Resources and Environment (NRE; especially Dr Ross Monash) for guidance on access to the shearwater colonies, and the Tasmanian Land Conservancy for loan of acoustic recorders. The Pacific Seabird Group, The Seabird Group, and BirdLife Tasmania generously provided funding for field work. We are grateful for the remarkable generosity, kindness, and laughter provided by Michael Foley, the Friends of Huon Island, and Detached Cultural Organization. Peter and Caitlin Vertigan provided hands-on training and Lillian Stewart (Adrift Lab), Jacob Hull, and Steph Zeliadt greatly assisted in the field. Feedback and advice on earlier drafts of this manuscript, and project as a whole, were kindly provided by Drs Sue Baker, Erik Wapstra, and Barry Brook in addition to Mel Wells, Annie Foppert, Nathan Coffey, and Adam Reeves. Finally, we thank our fellow Adrift Lab team members for being such supportive and inspiring early-career scientists and five anonymous reviewers who volunteered their time to provide constructive feedback on this manuscript.*

#### Data Availability:

*The data and R scripts are available upon request.*

---

## LITERATURE CITED

Bedoya, C., C. Isaza, J. M. Daza, and J. D. López. 2017. Automatic identification of rainfall in acoustic recordings. *Ecological Indicators* 75:95-100. <https://doi.org/10.1016/j.ecolind.2016.12.018>

Birkhead, T. 2014. Stormy outlook for long-term ecology studies. *Nature* 514:405. <https://doi.org/10.1038/514405a>

Bond, A. L., and J. L. Lavers. 2024. A feathered past: colonial influences on bird naming practices, and a new common name for *Ardenna carneipes* (Gould 1844). *Ibis* 166:1413-1418. <https://doi.org/10.1111/ibi.13356>

Bradfer-Lawrence, T., N. Bunnefeld, N. Gardner, S. G. Willis, and D. H. Dent. 2020. Rapid assessment of avian species richness and abundance using acoustic indices. *Ecological Indicators* 115:106400. <https://doi.org/10.1016/j.ecolind.2020.106400>

Brown, A., S. Garg, and J. Montgomery. 2018a. Automatic and efficient denoising of bioacoustics recordings using MMSE STSA. *IEEE Access* 6:5010-5022. <https://doi.org/10.1109/ACCESS.2017.2782778>

Brown, A., S. Garg, and J. Montgomery. 2018b. Scalable preprocessing of high volume environmental acoustic data for bioacoustic monitoring. *PLoS ONE* 13:e0201542. <https://doi.org/10.1371/journal.pone.0201542>

Brownlie, K. C., R. Monash, J. J. Geeson, J. Fort, P. Bustamante, and J. P. Y. Arnould. 2020. Developing a passive acoustic monitoring technique for Australia's most numerous seabird, the Short-tailed Shearwater (*Ardenna tenuirostris*). *Emu* 120:123-134. <https://doi.org/10.1080/01584197.2020.1732828>

Buxton, R. T., E. Brown, L. Sharman, C. M. Gabriele, and M. F. McKenna. 2016. Using bioacoustics to examine shifts in songbird phenology. *Ecology and Evolution* 6:4697-4710. <https://doi.org/10.1002/ece3.2242>

Buxton, R. T., and I. L. Jones. 2012. Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. *Journal of Field Ornithology* 83:47-60. <https://doi.org/10.1111/j.1557-9263.2011.00355.x>

Buxton, R. T., H. L. Major, I. L. Jones, and J. C. Williams. 2013. Examining patterns in nocturnal seabird activity and recovery across the Western Aleutian Islands, Alaska, using automated acoustic recording. *Auk* 130:331-341. <https://doi.org/10.1525/auk.2013.12134>

Cappello, C. D., and P. D. Boersma. 2021. Consequences of phenological shifts and a compressed breeding period in Magellanic Penguins. *Ecology* 102:e03443. <https://doi.org/10.1002/ecy.3443>

Carey, M. J., R. A. Phillips, J. R. D. Silk, and S. A. Shaffer. 2014. Trans-equatorial migration of Short-tailed Shearwaters revealed by geolocators. *Emu* 114:352-359. <https://doi.org/10.1071/MU13115>

Christie, D. R., B. L. N. Kennett, and C. Tarlowski. 2007. Advances in infrasound technology with application to nuclear explosion monitoring. Pages 825-835 in *Proceeding of the 29th monitoring research review on ground-based nuclear explosion monitoring technologies*. Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

Cleeland, J. B., M. A. Lea, and M. A. Hindell. 2014. Use of the Southern Ocean by breeding Short-tailed Shearwaters (*Puffinus tenuirostris*). *Journal of Experimental Marine Biology and Ecology* 450:109-117. <https://doi.org/10.1016/j.jembe.2013.10.012>

Colombelli-Négrel, D., and R. Smale. 2018. Habitat explained microgeographic variation in Little Penguin agonistic calls. *Auk* 135:44-59. <https://doi.org/10.1642/AUK-17-75.1>

Cornell University. 2025. BirdNET Sound ID. Cornell Lab of Ornithology, Ithaca, New York, USA. <https://birdnet.cornell.edu/>



- Dumoulin, V. 2022. Perch: A bioacoustic search platform for conservation. <https://github.com/google-research/perch>
- Fairbrass, A. J., P. Rennert, C. Williams, H. Titheridge, and K. E. Jones. 2017. Biases of acoustic indices measuring biodiversity in urban areas. *Ecological Indicators* 83:169-177. <https://doi.org/10.1016/j.ecolind.2017.07.064>
- Farina, A. 2025. The acoustic complexity index (ACI): theoretical foundations, applied perspectives and semantics. *Oikos* 2025: e10760. <https://doi.org/10.1111/oik.10760>
- Fayet, A. L., G. V. Clucas, T. Anker-Nilssen, M. Syposz, and E. S. Hansen. 2021. Local prey shortages drive foraging costs and breeding success in a declining seabird, the Atlantic Puffin. *Journal of Animal Ecology* 90:1152-1164. <https://doi.org/10.1111/1365-2656.13442>
- Fristrup, K., and D. Mennitt. 2012. Bioacoustical monitoring in terrestrial environments. *Acoustics Today* 8:16. <https://doi.org/10.1121/1.4753913>
- Glencross, J. S., J. L. Lavers, and E. Woehler. 2021. Breeding success of Short-tailed Shearwaters following extreme environmental conditions. *Marine Ecology Progress Series* 672:193-203. <https://doi.org/10.3354/meps13791>
- Haury, L., J. McGowan, and P. Wiebe. 1978. Patterns and processes in the time-space scales of plankton distributions. Pages 277-327 in J. Steele, editor. *Spatial pattern in plankton communities*. NATO Conference Series IV, Marine Science. Springer, New York, New York, USA. [https://doi.org/10.1007/978-1-4899-2195-6\\_12](https://doi.org/10.1007/978-1-4899-2195-6_12)
- Hinke, J. T., A. Barbosa, L. M. Emmerson, T. Hart, M. A. Juárez, M. Korczak-Abshire, G. Milinevsky, M. Santos, P. N. Trathan, G. M. Watters, and C. Southwell. 2018. Estimating nest-level phenology and reproductive success of colonial seabirds using time-lapse cameras. *Methods in Ecology and Evolution* 9:1853-1863. <https://doi.org/10.1111/2041-210X.13015>
- Jiménez-Uzcátegui, G., M. P. Harris, C. R. Sevilla, and K. P. Huyvaert. 2016. Longevity records for the Waved Albatross *Phoebastria irrorata*. *Marine Ornithology* 44:133-134. <https://doi.org/10.5038/2074-1235.44.2.1169>
- Kasten, E. P., S. H. Gage, J. Fox, and W. Joo. 2012. The remote environmental assessment laboratory's acoustic library: an archive for studying soundscape ecology. *Ecological Informatics* 12:50-67. <https://doi.org/10.1016/j.ecoinf.2012.08.001>
- Killick, R., and I. Eckley. 2014. changepoint: an R package for changepoint analysis. *Journal of Statistical Software* 58:1-19. <https://doi.org/10.18637/jss.v058.i03>
- Lavers, J. L., and A. L. Bond. 2013. Contaminants in indigenous harvests of apex predators: the Tasmanian Short-tailed Shearwater as a case study. *Ecotoxicology and Environmental Safety* 95:78-82. <https://doi.org/10.1016/j.ecoenv.2013.05.021>
- Lavers, J. L., and A. L. Bond. 2023. Long-term decline in fledging body condition in Flesh-footed Shearwaters (*Ardenna carneipes*). *ICES Journal of Marine Science* 80:1113-1119. <https://doi.org/10.1093/icesjms/fsad048>
- Lavers, J. L., I. Hutton, and A. L. Bond. 2019. Changes in technology and imperfect detection of nest contents impeded reliable estimates of population trends in burrowing seabirds. *Global Ecology and Conservation* 17:e00579. <https://doi.org/10.1016/j.gecco.2019.e00579>
- Ligges, U., S. Krey, O. Mersmann, and S. Schnackenberg. 2023. tuneR: Analysis of music and speech. Version 1.4.7. <https://CRAN.R-project.org/package=tuneR> <https://doi.org/10.32614/CRAN.package.tuneR>
- Lynch, T. P., R. Alderman, and A. J. Hobday. 2015. A high-resolution panorama camera system for monitoring colony-wide seabird nesting behaviour. *Methods in Ecology and Evolution* 6:491-499.
- Metcalf, O. C., J. Barlow, C. Devenish, S. Marsden, E. Berenguer, and A. C. Lees. 2021. Acoustic indices perform better when applied at ecologically meaningful time and frequency scales. *Methods in Ecology and Evolution* 12:421-431. <https://doi.org/10.1111/2041-210X.13521>
- Miller-Rushing, A. J., T. L. Lloyd-Evans, R. B. Primack, and P. Satzing. 2008. Bird migration times, climate change, and changing population sizes. *Global Change Biology* 14:1959-1972. <https://doi.org/10.1111/j.1365-2486.2008.01619.x>
- Mizel, J. D., C. Guglielmo, G. W. Mitchell, K. A. Hobson, and D. Norris. 2019. Migratory connectivity and population-specific migration timing in a long-distance migratory bird. *Ecological Monographs* 89:e01310.
- Mougeot, F., and V. Bretagnolle. 2000. Predation risk and moonlight in nocturnal seabirds. *Journal of Avian Biology* 31:376-386. <https://doi.org/10.1034/j.1600-048X.2000.310314.x>
- Naarding, J. A. 1981. Study of the Short-tailed Shearwater *Puffinus tenuirostris* in Tasmania. Wildlife Division Technical Report, Hobart, Tasmania.
- Oppel, S., S. Hervias, N. Oliveira, T. Pipa, C. Silva, P. Gerales, M. Goh, E. Immler, and M. McKown. 2014. Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. *Nature Conservation* 7:1-13. <https://doi.org/10.3897/natureconservation.7.6890>
- Pérez-Granados, C., J. Gómez-Catasús, D. Bustillo-de la Rosa, A. Barrero, M. Reverter, and J. Traba. 2019. Effort needed to accurately estimate Vocal Activity Rate index using acoustic monitoring: a case study with a dawn-time singing passerine. *Ecological Indicators* 107:105608. <https://doi.org/10.1016/j.ecolind.2019.105608>
- Pérez-Granados, C., and J. Traba. 2021. Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. *Ibis* 163:765-783. <https://doi.org/10.1111/ibi.12944>
- Pieretti, N., A. Farina, and D. Morri. 2011. A new methodology to infer the singing activity of an avian community: the Acoustic Complexity Index (ACI). *Ecological Indicators* 11:868-873. <https://doi.org/10.1016/j.ecolind.2010.11.005>
- Price, C. A., T. Emery, K. Hartmann, E. Woehler, R. Monash, and M. Hindell. 2020b. Inter-annual and inter-colony variability in breeding performance of four colonies of Short-tailed Shearwaters. *Journal of Experimental Marine Biology and Ecology* 537:151498. <https://doi.org/10.1016/j.jembe.2020.151498>

- Price, C. A., K. Hartman, T. J. Emery, E. J. Woehler, C. R. McMahon, and M. A. Hindell. 2020a. Climate variability and breeding parameters of a transhemispheric migratory seabird over seven decades. *Marine Ecology Progress Series* 642:191-205. <https://doi.org/10.3354/meps13328>
- R Core Team. 2022. R: A language and environment for statistical computing. Version 4.2.1 [computer program].
- Ratcliffe, N., D. Guihen, J. Robst, S. Crofts, A. Stanworth, and P. Enderlein. 2015. A protocol for the aerial survey of penguin colonies using UAVs. *Journal of Unmanned Vehicle Systems* 3:95-101. <https://doi.org/10.1139/juvs-2015-0006>
- Serventy, D. 1966. Aspects of the population ecology of the Short-tailed Shearwater *Puffinus tenuirostris*. Pages 165-190. *Proceedings of the 14th International Ornithological Congress*, Oxford, UK.
- Skira, I., N. Brothers, and D. Pemberton. 1996. Distribution, abundance and conservation status of Short-tailed Shearwaters *Puffinus tenuirostris* in Tasmania, Australia. *Marine Ornithology* 24:1-14. <https://doi.org/10.5038/2074-1235.24.1.353>
- Stowell, D., T. Petrusková, M. Šálek, and P. Linhart. 2019. Automatic acoustic identification of individuals in multiple species: improving identification across recording conditions. *Journal of The Royal Society Interface* 16:20180940. <https://doi.org/10.1098/rsif.2018.0940>
- Sueur, J., and A. Farina. 2015. Ecoacoustics: the ecological investigation and interpretation of environmental sound. *Biosemiotics* 8:493-502. <https://doi.org/10.1007/s12304-015-9248-x>
- Sueur, J., A. Farina, A. Gasc, N. Pieretti, and S. Pavoine. 2014. Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica* 100:772-781. <https://doi.org/10.3813/AAA.918757>
- Sugai, L. S. M., T. S. F. Silva, J. W. Ribeiro Jr., and D. Llusia. 2019. Terrestrial passive acoustic monitoring: review and perspectives. *BioScience* 69:15-25. <https://doi.org/10.1093/biosci/biy147>
- Szabo, M. 2013. Short-tailed Shearwater. *New Zealand Birds Online*. <https://www.nzbirdsonline.org.nz/species/short-tailed-shearwater>
- Tryjanowski, P., and T. H. Sparks. 2001. Is the detection of the first arrival date of migrating birds influenced by population size? A case study of the Red-backed Shrike *Lanius collurio*. *International Journal of Biometeorology* 45:217-219. <https://doi.org/10.1007/s00484-001-0112-0>
- Villanueva-Rivera, L., and B. Pijanowski. 2022. Package 'soundecology'. Version 1.3.3. <http://lrvillanueva.github.io/soundecology/>
- Walker, K. T., and M. A. H. Hedlin. 2010. A review of wind-noise reduction methodologies. Pages 141-182 in A. Le Pichon, E. Blanc, and A. Hauchecorne, editors. *Infrasound monitoring for atmospheric studies*. Springer, Dordrecht, The Netherlands. [https://doi.org/10.1007/978-1-4020-9508-5\\_5](https://doi.org/10.1007/978-1-4020-9508-5_5)
- Warham, J. 1960. Some aspects of breeding behaviour in the Short-tailed Shearwater. *Emu - Austral Ornithology* 60:75-87. <https://doi.org/10.1071/MU960075>
- Wildlife Acoustics. 2011. Song Scope Bioacoustics Software. Version 4.1.5A. <https://www.wildlifeacoustics.com/images/documentation/Song-Scope-Users-Manual.pdf>
- Willacy, R. J., M. Mahony, and D. A. Newell. 2015. If a frog calls in the forest: bioacoustic monitoring reveals the breeding phenology of the endangered Richmond Range mountain frog (*Philoria richmondensis*). *Austral Ecology* 40:625-633. <https://doi.org/10.1111/aec.12228>
- Woehler, E. 2012. What do signals from seabirds tell us about the marine environments? Pages 218-225 in F. Huettmann, editor. *Protection of the three poles*. Springer, New York, New York, USA.
- Woehler, E. J., and A. J. Hobday. 2024. Impacts of marine heatwaves may be mediated by seabird life history strategies. *Marine Ecology Progress Series* 737:9-23. <https://doi.org/10.3354/meps14333>
- Xie, J., L. Zhang, J. Zhang, Y. Zhang, and B. W. Schuller. 2023. Cross-corpus open set bird species recognition by vocalization. *Ecological Indicators* 154:110826. <https://doi.org/10.1016/j.ecolind.2023.110826>
- Yap, V. H. S., L. G. Stewart, S. Stuckenbrock, A. L. Fidler, and J. L. Lavers. 2021. Longevity records for the Brown Booby *Sula leucogaster* and Flesh-footed Shearwater *Ardenna carneipes*. *Marine Ornithology* 49:167-170. <https://doi.org/10.5038/2074-1-235.49.1.1413>



## Appendix 1.

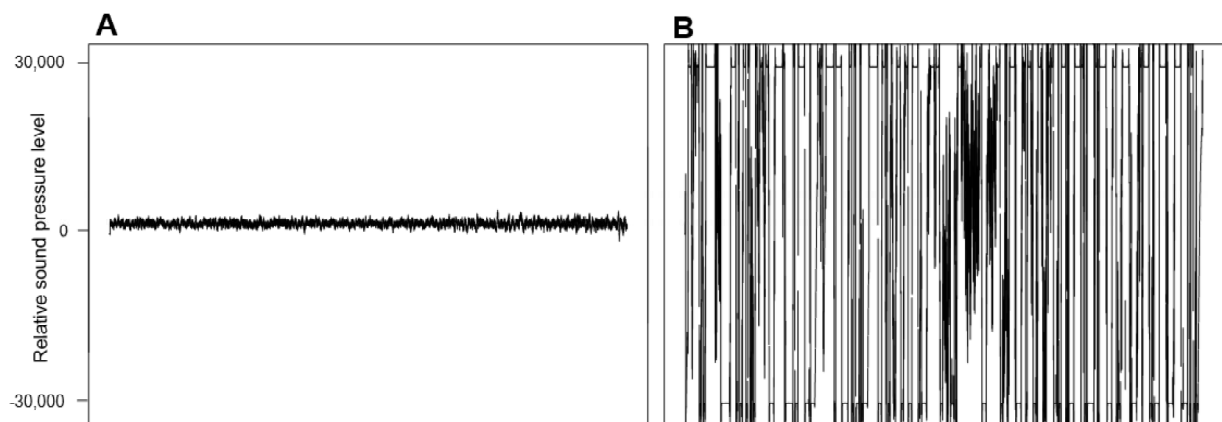
### Acoustic Index Selection

This supplementary material provides additional detail on the selection of acoustic indices based on their relationship with short-tailed shearwater (STSH) vocalisations under varying environmental conditions. These methods directly support the evaluation of how well each index captures STSH vocal activity, as presented in the main manuscript.

### Environmental Covariate Scoring

Environmental factors such as wind noise and non-target biotic sounds were manually scored to assess their potential impact on acoustic recordings on a six-step scale. Wind, the most prevalent abiotic factor, was scored as follows:

1. **Very Low:** Minimal wind, no waveform clipping.
2. **Low:** Some wind noise, no clipping.
3. **Medium:** Moderate wind, 1-5% waveform clipping.
4. **High:** Significant wind, 5-20% waveform clipping.
5. **Very High:** Heavy wind, 20-50% waveform clipping.
6. **Extreme:** Severe wind, >50% waveform clipping.



**Fig. S1.** Two waveform plots that show varying levels of wind in the recording. Plot A shows a recording with limited wind noise, exhibited by an absence of waveform distortion (clipping). Plot B shows a recording with a very high wind noise, as >50% of the recording displays waveform clipping. For plot B, the sample would be given a ‘wind’ score of 6.

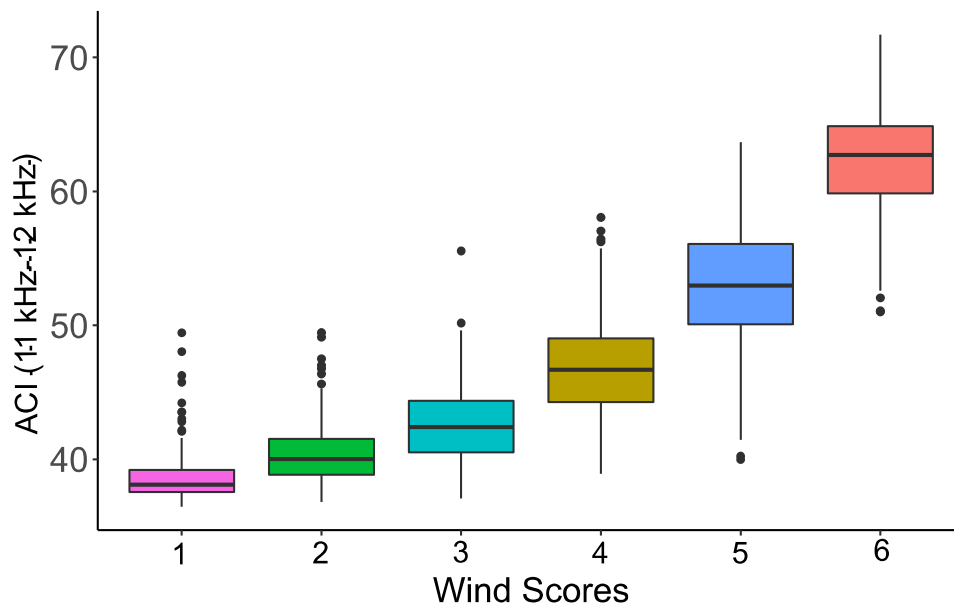


## Subsetting the Sample Data

We generated three subsets of the data to manage non-target noise:

1. **Raw Dataset:** Included 1203 samples, only excluding those where wind noise was extreme (score of 6 or more).
2. **Clean Dataset:** Contained 685 samples, excluding those with covariate scores above 3 for wind, rain, little penguin vocalisations, and other biotic noise.
3. **Wind-Filtered Dataset:** Contained 828 samples, using an automated wind filter to remove high-wind recordings.

ACI values between 11,000-12,000 Hz were used to detect wind noise, with samples exceeding a wind.ACI threshold of 45 being removed. This resulted in a wind-filtered dataset of 828 samples, representing approximately 68% of the total dataset. The wind filter also filtered out 45% of recordings that contained rain with a score of 4 or higher.



**Fig. S2.** Box plot of values in the wind-filter Acoustic Complexity Index (ACI) plotted against the manually logged wind scores for all samples. The ACI was calculated between 11,000–12,000 Hz. The wind-filter was applied by removing recordings from the dataset if the ACI value was greater than 45, which was the median of samples with a score of 4 for wind.

Model Selection for Acoustic Indices

Linear mixed models were fit to each dataset to assess the relationship between the acoustic indices (ACI, NDSI, ADI, AEI, BIO) and STSH vocalisations, along with environmental covariates like wind, rain, penguin vocalisations, and other biotic noise. Model selection was based on Akaike Information Criterion (AIC) and cross-validated R<sup>2</sup> values. The following tables provide ΔAIC and R<sup>2</sup> values for each index across the datasets:

**Table S1.** Factors affecting acoustic indices in short-tailed shearwater colonies in the raw dataset. Terms in the best fitting models were determined by Akaike information criterion (AIC) model selection. ΔAIC values indicating the relative importance of terms and cross-validated R2 values are also shown for each model.

	ACI		NDSI		ADI		AEI		BIO	
TERM (df)	ΔAIC	R <sup>2</sup>	ΔAIC	R <sup>2</sup>	ΔAIC	R <sup>2</sup>	ΔAIC	R <sup>2</sup>	ΔAIC	R <sup>2</sup>
Vocalisations (1)	498.7	<b>0.73</b>	602.1	<b>0.54</b>	354.2	<b>0.58</b>	341.1	<b>0.53</b>	49.7	<b>0.26</b>
Wind (4)	845.9		7.3		247.4		391.6			
Rain (5)	128.6		25.6		0.8		9.3		17.0	
Penguin (5)	36.3		44.4		9.7		3.15			
Other biotic (4)	113.4		24.0		10.3		19.8		8.6	

**Table S2.** Factors affecting acoustic indices in short-tailed shearwater colonies in the clean dataset. Terms in best fitting models were determined by Akaike information criterion (AIC) model selection.  $\Delta$ AIC values indicating the relative importance of terms and cross-validated R2 values are also shown for each model.

	ACI		NDSI		ADI		AEI		BIO	
TERM (df)	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>
Vocalisations (1)	608.1	<b>0.71</b>	388.7	<b>0.58</b>	176.1	<b>0.48</b>	211.6	<b>0.47</b>	56.8	<b>0.29</b>
Wind (4)	11.9				42.9		13.7			
Rain (5)	6.8				0.9		0.16		8.5	
Penguin (5)					2.8		6.0			
Other biotic (4)	36.5				26.5		20.0		25.7	



**Table S3.** Factors affecting acoustic indices in short-tailed shearwater colonies in the wind-filtered dataset. Terms in the best fitting models were determined by Akaike information criterion (AIC) model selection.  $\Delta$ AIC values indicating the relative importance of terms and cross-validated R2 values are also shown for each model.

TERM (df)	ACI		NDSI		ADI		AEI		BIO	
	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>	$\Delta$ AIC	R <sup>2</sup>
Vocalisations (1)	677.1	<b>0.71</b>	382.5	<b>0.55</b>	173.1	<b>0.45</b>	202.9	<b>0.43</b>	29.86	<b>0.30</b>
Wind (4)	16.3		8.82		54.0		32.8			
Rain (5)	120.3		3.5						28.7	
Penguin (5)	39.1		26.1		1.1		7.1			
Other biotic (4)	128.4		23.5		17.3		9.8		9.92	

### Results of Index Performance

The ACI and NDSI consistently outperformed other indices across all datasets:

- Raw Dataset:** ACI had the strongest fit ( $R^2 = 0.73$ ), but was heavily influenced by wind ( $\Delta$ AIC = 845.9). NDSI was less sensitive to wind but showed a coarser relationship with STSH vocalisations ( $R^2 = 0.54$ ).
- Clean Dataset:** ACI again showed the strongest relationship with STSH vocalisations ( $R^2 = 0.71$ ), with minimal influence from covariates. NDSI also performed well ( $R^2 = 0.58$ ), but BIO consistently performed poorly, with  $R^2$  values no higher than 0.30.
- Wind-Filtered Dataset:** ACI continued to detect STSH vocalisations effectively ( $R^2 = 0.71$ ), though it excluded some valid samples. NDSI, while slightly less sensitive ( $R^2 = 0.55$ ), performed better in noisy environments.

### Discussion on Index Selection and Site-Specific Conditions

The exploratory analysis demonstrates the importance of selecting acoustic indices based on site-specific environmental conditions. ACI showed the strongest relationship with STSH

vocalisations in clean environments but was sensitive to wind and biotic noise. NDSI, though coarser, was more robust in noisy environments. A combination of ACI and NDSI provided the most reliable detection of STSH arrival, particularly when working across varied environmental conditions.

## **Conclusion**

No single index was effective across all conditions. ACI worked best in clean and wind-filtered environments, while NDSI was more reliable in noisier settings. Using both indices together provides a complementary approach for monitoring STSH arrival, though careful consideration of site conditions is essential for accurate results.