Assessing the cross platform performance of marine mammal indicators between two collocated acoustic recorders

Citation	Denes, S. L., Miksis-Olds, J. L., Mellinger, D. K., & Nystuen, J. A. (2014). Assessing the cross platform performance of marine mammal indicators between two collocated acoustic recorders. Ecological Informatics, 21, 74-80. doi:10.1016/j.ecoinf.2013.10.005
DOI	10.1016/j.ecoinf.2013.10.005
Publisher	Elsevier
Version	Version of Record
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Ecological Informatics 21 (2014) 74–80

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Assessing the cross platform performance of marine mammal indicators between two collocated acoustic recorders

Samuel L. Denes [a,*], Jennifer L. Miksis-Olds [a], David K. Mellinger [b], Jeffrey A. Nystuen [c]

[a] Applied Research Laboratory, The Pennsylvania State University, PO Box 30, University Park, PA 16804-0030, USA
[b] Cooperative Institute for Marine Resources Studies, Oregon State University and NOAA Pacific Marine Environmental Laboratory, Hatfield Marine Science Center, 2115 S.E. OSU Drive, Newport, OR 97365, USA
[c] Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, WA 98105-6698, USA

A R T I C L E I N F O

Article history:
Received 21 June 2013
Received in revised form 5 October 2013
Accepted 29 October 2013
Available online 6 November 2013

Keywords:
Acoustic diversity
Ocean acoustics
Spectral detection
Acoustic entropy
Marine mammal classification

A B S T R A C T

Equipment and deployment strategies for remote passive acoustic sensing of marine environments must balance memory capacity, power requirements, sampling rate, duty-cycle, deployment duration, instrument size, and environmental concerns. The impact of different parameters on the data and applicability of the data to the specific questions being asked should be considered before deployment. Here we explore the effect of recording and detection parameters on marine mammal acoustic data across two platforms. Daily classifications of marine mammal vocalizations from two passive acoustic monitors with different subsampling parameters, an AURAL and a Passive Aquatic Listener (PAL), collocated in the Bering Sea were compared. The AURAL subsampled on a preset schedule, whereas the PAL sampled via an adaptive protocol. Detected signals of interest were manually classified in each dataset independently. The daily classification rates of vocalizations were similar. Detections from the higher duty-cycle but lower sample rate AURAL were limited to species and vocalizations with energy below 4 kHz precluding detection of echolocation signals. Temporal coverage from the PAL audio files was limited by the adaptive sub-sampling protocol. A method for classifying ribbon (*Histriophoca fasciata*) and bearded seal (*Erignathus barbatus*) vocalizations from the sparse spectral time histories of the PAL was developed. Although application of the acoustic entropy as a rapid assessment of biodiversity was not reflective of the number of species detected, acoustic entropy was robust to changes in sample rate and window length.

1. Introduction

Determining habitat usage by marine mammals contributes to the overall understanding of the ecology of these animals. Conducting visual surveys of marine mammals is expensive and difficult. Some marine mammal species are at the surface for only a short period of time making visual detection even more difficult. Acoustic monitoring permits the surveying of vocalizing animals without relying on visual detection. Remotely deployed autonomous passive acoustic monitoring enables persistent monitoring of a region for vocalizing marine mammals over long periods of time without requiring the presence of human observers. Improvements in hardware technology now permit the collection of enormous passive acoustic data sets from remotely deployed recorders (Van Parijs et al., 2009; Wiggins and Hildebrand, 2007). The use of autonomous passive acoustic monitoring (PAM) for studies of marine mammals greatly increases our capacity for collecting information about vocalizing animals in the absence of concurrent visual observations, which is critical for acquiring information in remote, inaccessible, or hazardous areas (Mellinger et al., 2007). Increasing the amount of data collected leads to an increase in the required storage space and post-processing demands. Analysis requirements have traditionally been met by long and often tedious hours from a human classifier listening to and looking through spectrograms of the recordings. Automated detection and classification algorithms are now replacing the previously required man hours with computer hours (Mellinger et al., 2011; Roch et al., 2008). Many different recording systems are being designed and employed for various studies around the world (Moore et al., 2012; Sousa-Lima et al., 2013). The results stemming from these efforts provide information on animal distribution, behavior, and reactions to environmental change, all of which have the potential to inform resource management, research efforts, and industry. As not all recordings are the same, understanding the relative strengths, weaknesses, and impacts of sampling strategies on data interpretation and results becomes increasingly important. Comparisons of detections from different species and across different recording systems will greatly increase the inferential power from the results of the analyses of the individual units. As deployment durations are increased and PAM recorders are deployed in new and increasingly remote locations, methods are being developed to handle the collection and processing load to yield results for interpretation regardless of recording strategy

* Corresponding author. Tel.: +1 814 863 5313.
E-mail addresses: sld980@psu.edu (S.L. Denes), jlm91@psu.edu (J.L. Miksis-Olds), david.mellinger@oregonstate.edu (D.K. Mellinger), nystuen@apl.washington.edu (J.A. Nystuen).

1574-9541/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecoinf.2013.10.005

(Wiggins and Hildebrand, 2007). This will make all data more valuable due to the wider use and applicability for short and long term studies.

There are tradeoffs in remote passive acoustic sensing between memory capacity, power requirements, sampling rate, duty-cycle, deployment duration, and instrument size. For any signal of interest, the Nyquist theorem requires that the minimum sample rate must be at least twice the highest frequency in the signal or aliasing may occur and spectral data will be compromised (Nyquist, 1932). This places a minimum constraint on the sample rate for a signal with known frequency content. For low frequency vocalizations, like those of blue whales (*Balaenoptera musculus*), the required sample rate can be below 1 kHz because the highest frequency in known blue whale vocalizations does not exceed 400 Hz (Berchok et al., 2006). With these low sample rates and modern data storage capabilities, deployments are often limited by battery power. However, for signals containing high frequency components found in many odontocete vocalizations and echolocation signals, the required bandwidth of the recording introduces data storage constraints along with limitations from battery power. Reducing the duty-cycle of the recorder, so that it is recording only part of the time, can extend the battery and memory capacity. With reduced duty-cycle, the ability of a system to capture a particular sound depends on the probability of the signal being present with substantial signal to noise ratio and the probability of the system recording at that moment (Miksis-Olds et al., 2010; Richardson et al., 1995). For species with seasonally ubiquitous vocalizations, like bearded seals (*Erignathus barbatus*) in the Bering Sea during late winter and early spring, this is not an issue because vocalizations from multiple individuals are almost continuous (Miksis-Olds et al., 2010). If the received level of the vocalizations is consistently above the background noise, the signal will be recorded when the system is active. A caveat exists for those species where vocalizations are rare, either because the species vocalizes infrequently, or the number of vocalizing individuals is low e.g. North Pacific Right whales (*Eubalaena japonica*) (Mellinger et al., 2004). A sub-sampling protocol operates under the assumption that rare vocalizations may be missed yet permits the collection of long term data sets from recorders deployed in remote locations for long durations.

Sub-sampling methods can be adequate to address many research questions such as those pertaining to the presence and absence of marine mammals in a region over time and acoustic biodiversity (Lammers et al., 2008; Sueur et al., 2012). For binary presence/absence research questions, long recordings dominated by the repeated vocalizations of a single species provide the same result as a recording of a single vocalization from that species. Research questions addressing topics such as the vocalization behavior and population density may benefit from long recordings of a single species' vocalizations. Adaptive recorders with on-board decision making algorithms permit the collection of a limited amount of data with feature triggers focusing the effort on periods containing signals of interest; thus reducing battery, memory and post-processing requirements (Miksis-Olds et al., 2010).

Describing the biodiversity of the environment is often limited to the species richness or the number of different species present (Sueur et al., 2012). Marine mammal classifications from acoustic recordings obtained using autonomous recorders provide a measure of biodiversity. An automated assessment of biodiversity comparing the temporal and spectral entropies of the acoustic signals in the terrestrial environment was presented by Sueur et al. (2008). Biodiversity assessments amongst and between varying passive acoustic monitors have not been examined. Utilizing instruments with different duty cycles and sampling rates may not provide comparable results for detection, classification and relative vocal activity for different species.

Expanded effort to monitor the marine environment with sub-sampling acoustic recorders equipped with increasingly complex on-board processing raises the question of how to integrate data across acoustic monitoring systems. The species level classifications or acoustic biological diversity, and a statistically based acoustic biodiversity index can be generated from each dataset. Understanding the relative performance of systems with different sub-sampling recording paradigms is useful for comparisons between the systems. Two systems currently deployed on a single mooring (subsurface buoy) in the Bering Sea implement different strategies to achieve year-long deployments: 1) semi-continuous sampling; and 2) an adaptive sub-sampling paradigm with an on-board event detector for initial processing and adaptive control. The concurrent deployment of these two systems enables a comparison between detection and classification of marine mammal vocalizations necessary to identify the strengths and weaknesses of the two different sampling methods. The daily species level classifications of the two recording systems were examined. The acoustic biodiversity index calculated for each system was compared to species detected within a single recorder and across recording platforms.

2. Materials and methods

Two autonomous passive acoustic recorders with different sampling strategies were collocated on an oceanographic mooring maintained by NOAA's Pacific Marine Environmental Laboratory as part of the Fisheries Oceanography Coordinated Investigations (Eco-FOCI) Program (Stabeno et al., 2008). The Passive Aquatic Listener (PAL) is an adaptively sub-sampling recorder developed by Jeffrey Nystuen at the University of Washington (Nystuen, 1998) and the AURAL-M2 (Multi-Electronique Inc, Quebec) is a commercially available, programmable passive acoustic recorder. The mooring was located on the 70 m isobath southeast of St. Matthew Island in the eastern Bering Sea (59° 54.285′ N, 171° 42.285′ W)(Stabeno et al., 2008). The PAL and AURAL were deployed serially in the mooring line at depths of 65 m and 67 m, respectively. The data for this work are from a single deployment, September 2008 through May 2009, as part of a multiyear study.

2.1. Passive Aquatic Listener (PAL)

The PAL consists of a wide-band (0–50 kHz), low-noise hydrophone (HTI-96-MIN), pre-amplifier, and recording computer. An internal battery pack provided power for instrument operation. On-board memory consisted of a 2 GB compact flash card. The PAL was programmed to record a 4.5 s audio clip at a sampling rate of 100 kHz every 10 min. Eight spectra were created from 10.24 ms subsamples spaced equally throughout the 4.5 s clip. The spectral values were compressed by integrating the frequency bins over 200 Hz bandwidths from 100 to 3000 Hz and 1 kHz bandwidths from 3 to 50 kHz (Nystuen, 1998). The eight individual compressed spectra, or spectra cluster, were analyzed against predetermined detection thresholds. A signal of interest was detected in the sample if one of three criteria were met: 1) matching predefined spectral patterns for rain, 2) a 12 dB amplitude threshold difference for sequential samples, or 3) peaks in frequency bins indicating tonal signals (Miksis-Olds et al., 2010). Exceeding any of the criteria resulted in a detection and the implementation of the adapted sampling protocol. If the recording was determined to contain a signal of interest, the audio clip and the spectra cluster were saved to memory, otherwise the audio clip was cleared, the spectra were averaged, and only the average spectrum was saved. Additionally, if a signal of interest was present, the PAL reduced the sampling interval from 10 to 2 min until a signal of interest was no longer detected (Fig. 1a). These two sampling intervals resulted in duty cycles of 0.75% and 3.75%, respectively. A daily quota limiting the number of saved audio clips was selected based upon the expected deployment duration to ensure adequate disk space. For this deployment, the daily quota was six audio clips. If the number of audio clips saved for any day was less than the quota, the excess allocation was made available to subsequent days up to a maximum of 21 total audio clips per day. If the daily limits were exceeded, the PAL continued to operate with the same adaptive sub-sampling protocol with the exception that no further audio clips were saved; spectra clusters continued to be saved throughout the deployment regardless of whether the daily audio clip limit was exceeded (Miksis-Olds et al., 2010). This programming paradigm

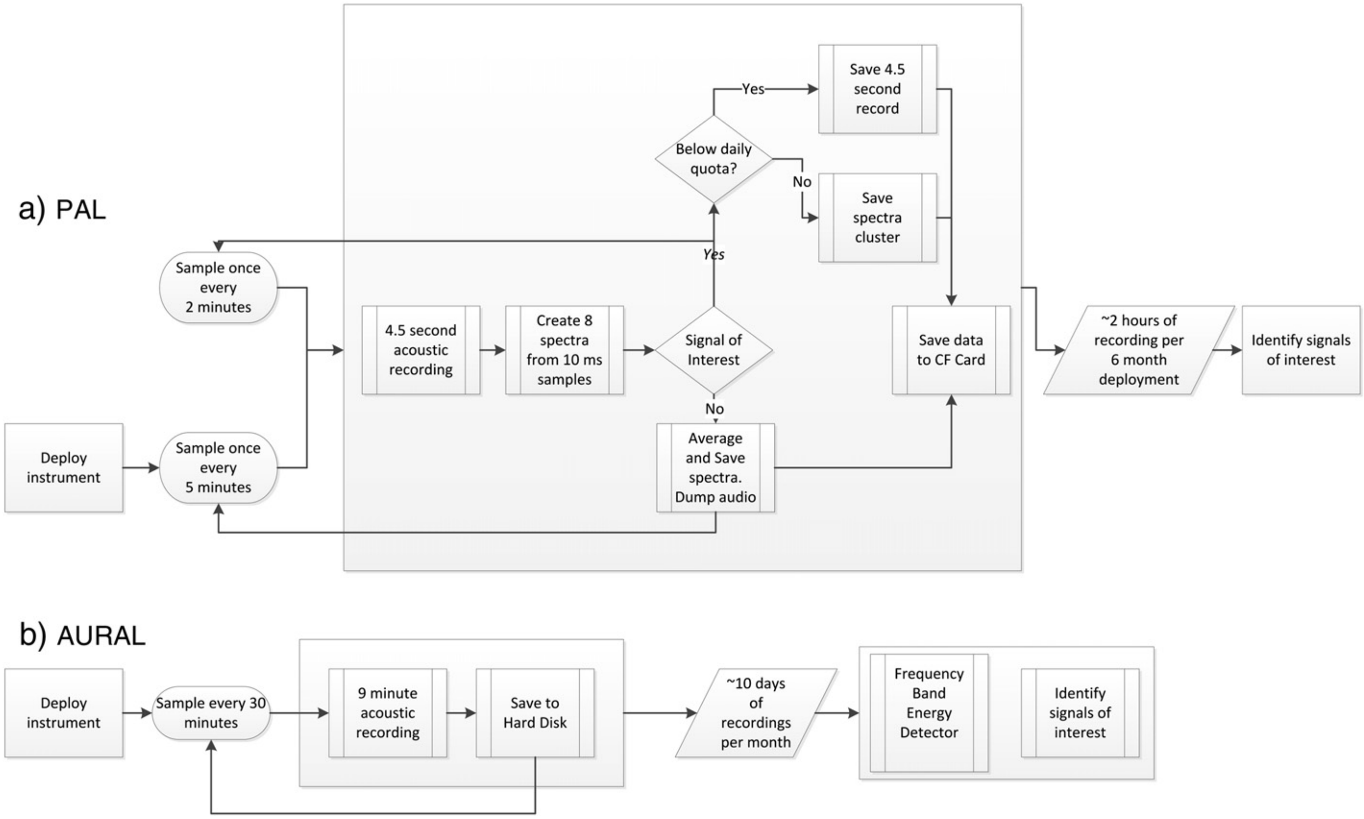

Fig. 1. Flow chart for the operation of acoustic recorders at Bering Sea mooring M5 - (a) PAL, a 50 kHz sample rate adaptive monitor & (b) AURAL, a 30% duty cycle recorder with an 8192 Hz sample rate.

provided three subsets of data: a quasi-continuous spectral record of the background acoustic environment, a limited number of high sample rate audio clips triggered by on-board signal detection analysis, and a series of spectral clusters triggered by on-board signal detection analysis throughout the collection period.

2.2. AURAL

The AURAL-M2 consists of a wide-band (2 Hz–30 kHz), low-noise hydrophone (HTI-96-MIN), preamplifier and recording computer. Data were saved to a 320 GB internal hard drive. The system has a user selectable sample rate, gain, and duty-cycle. The AURAL was programmed to record semi-continuously for nine minutes every half hour, a 30% duty cycle, at a sampling rate of 8192 Hz. These parameters were selected to ensure battery life and memory capacity was sufficient for a yearlong deployment. Recordings from the AURAL were initially processed by two frequency band energy detectors. Using Ishmael software, audio clips of signals containing elevated energy in the 0.1–1 kHz and 0.9–4 kHz frequency bands were created (Fig. 1b)(Mellinger, 2001). Detection occurred if the average energy over either band exceeded a threshold of 0.025 normalized units, and a 10 s sound clip was extracted. If successive detections resulted in overlapping sound clips, a single non-duplicated sound clip was created. The low threshold was selected to ensure that almost all transients would be detected. For many of the days, detections from elevated energy in the frequency bands resulted in near continuous detections from the nine minute sampling period.

2.3. Species classifications

2.3.1. Classification from audio

Marine mammal vocalizations were manually classified from audio clips from both instruments. Classifications from the PAL were made from the 4.5 s audio clips. Classifications from the AURAL dataset were made from the output of the two frequency band energy detectors. Four days from each month (1st, 8th, 15th and 23rd) from October 2008 through May 2009 were selected for comparison. Comparison of audio and spectrographic representations of the records were reviewed and assigned sources when possible by two independent blind human classifiers using custom Matlab scripts. Discrepancies in classifications were reviewed using Adobe Audition. The presence of species specific marine mammal vocalizations for each day was tabulated independently for the audio recordings from each system.

2.3.2. Classifications from spectra clusters

Spectra clusters were obtained from the PAL adaptive sampling protocol. Classification of stereotyped vocalizations from the PAL spectra clusters was conducted by human classifiers. Clusters associated with saved audio clips from the deployment were used to develop templates of stereotyped vocalizations. Species with stereotyped vocalizations considered were bearded seals, ribbon seals (*Histriophoca fasciata*), bowhead whales (*Balaena mysticetus*), and Pacific walrus (*Odobenus rosmarus divergens*) (Risch et al., 2007). It has been observed that 56% of recorded bearded seal vocalizations in Alaskan waters are descending trills. These bearded seal trills have an average duration exceeding the 4.5 s duration of the audio clips, with some longer than a minute (Risch et al., 2007). In PAL spectra clusters, bearded seal trills manifested as two to three bin wide peaks descending in frequency in all 8 spectral samples from a 4.5 s audio clip (Fig. 2). Ribbon seal downsweeps have energy in frequencies from 200 Hz to 5 kHz and average duration of 2 s (Miksis-Olds and Parks, 2011; Watkins and Ray, 1977). These downsweeps manifested as a series of up to three narrow peaks descending in frequency over a subset of the PAL spectral samples (Fig. 3). Large baleen whale moans and songs contain energy in frequencies below 5 kHz (Cummings and Holliday, 1987). During classification, spectra clusters were displayed in random order without date stamps to

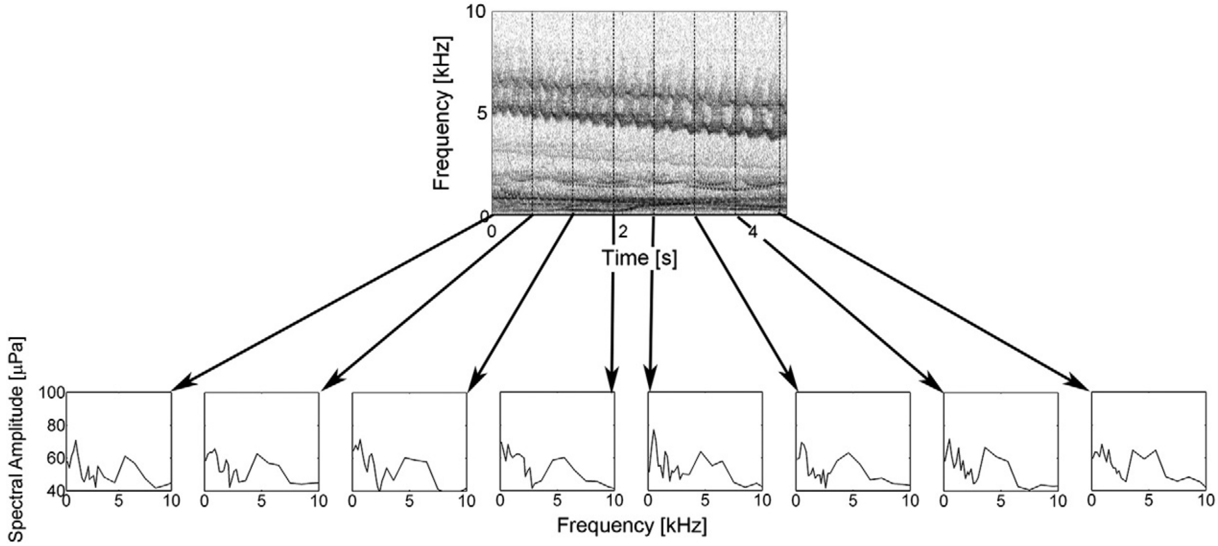

Fig. 2. Spectrogram and spectra from a bearded seal trill from a PAL audio clip. The vertical lines in the spectrogram indicate data from which spectra were calculated. This trill manifests as a 2 to 3 bin wide peak in the spectra from 3 to 7 kHz over all 8 spectra. The bandwidth has been limited to 10 kHz to highlight the signal.

prevent bias. Species classifications were recorded from user input in Matlab. If the human classifier observed a pattern matching the template for one of the species in a spectral cluster, the classification was recorded. Results for daily detection were tabulated by the software. Spectra clusters computed from saved audio clips were included in the classification process for groundtruthing purposes. Recall and false alarm rates were calculated to quantify the classification performance from spectra. Recall rate was the proportion of the clips previously identified in audio clips as containing a vocalization also classified using the spectral methods. False alarm rate was the proportion of the clips classified using the spectral methods in which the vocalization identified was not actually present.

2.4. Acoustic entropy

Acoustic entropy has been shown to correlate to acoustic biodiversity (Sueur et al., 2008). The acoustic entropy was computed on all audio clips analyzed from both systems following the procedure described in Sueur et al. (2008). To compensate for biases that can result from the differing sampling parameters, the entropy was computed twice for each signal: 1) acoustic entropy was computed for each of the original signals from the PAL with 100 kHz sample rate and duration of 4.5 s and 2) for the PAL clips re-sampled to the sample rate of the AURAL recordings of 8192 Hz. There were a total of 192 PAL clips analyzed in each dataset. Acoustic entropy was computed twice for AURAL recordings: 1) at 8192 Hz with durations of 4.5 s to match those from the PAL, resulting in 117 529 clips and, 2) on samples with the same number of points in the waveform as found in the original PAL samples – 450 000 samples. This resulted in 9583 samples of just under 55 s.

2.5. Statistical analyses

The number of species per day identified manually from detections on each recording platform was compared. The adaptive sampling protocol of the PAL did not permit finer temporal comparisons. The classifications were reduced to the presence or absence of a vocalization from each species for each day. Daily classifications from the PAL audio and spectra clusters were considered equally. A Wilcoxon signed rank test was used to test the hypothesis that the number of daily species classifications was equal between the two systems.

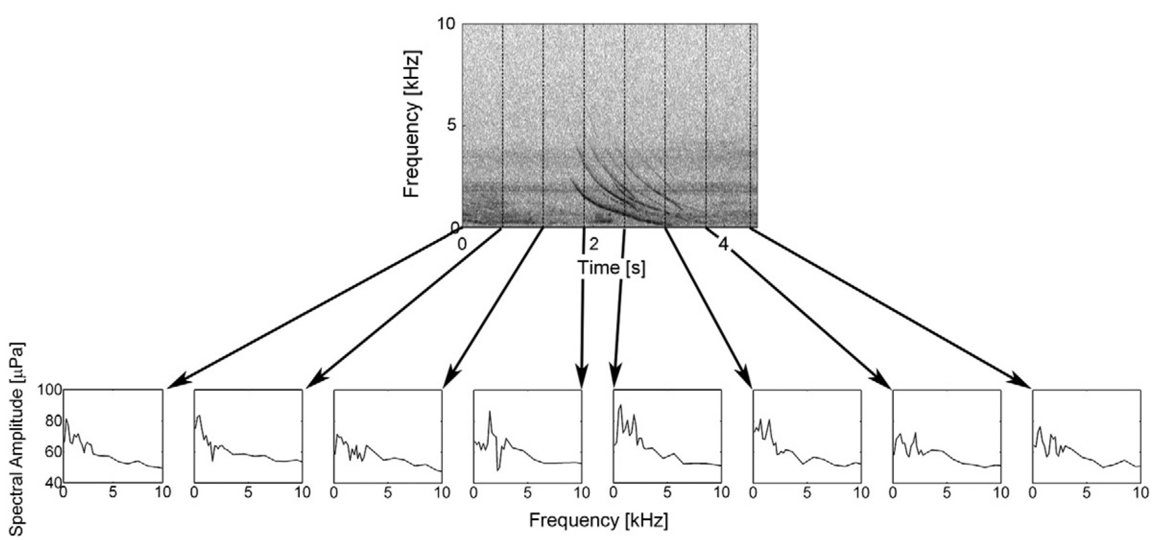

Fig. 3. Spectrogram and spectra from a ribbon seal downsweep from a PAL audio clip. The vertical lines in the spectrogram indicate data from which spectra were calculated. The 4th, 5th and 6th spectra contain peaks from the vocalization. The bandwidth has been limited to 10 kHz to highlight the signal.

To determine the robustness of the acoustic entropy for use as a biodiversity index across platforms, the acoustic entropy of each system was compared. The daily and monthly mean acoustic entropies were calculated. A linear regression was used to compare the daily mean acoustic entropies from each recorder. The daily and monthly mean entropy values were compared to the results from the marine mammal classification data over the same time periods for each recorder. Linear regression was used to test the relationship between the number of marine mammals classified and the acoustic entropy for daily and monthly time frames.

3. Results

3.1. Species classifications

Data from both recorders were analyzed on 32 days from an 8 month deployment. Table 1 summarizes the species detected and breakdown of detections by instrument. Totals for the PAL include audio and spectral detection. Bowhead whales were detected on 21 different days. Bowhead whale vocalizations were classified in the AURAL recordings on all 21 days, whereas the PAL recorded them on 18 days. Humpback whales (*Megaptera novaeangliae*) were recorded on 9 days: 2 days on the PAL only, 4 days on the AURAL only, 3 days on both. Non-echolocation odontocete vocalizations were classified on 7 days. These signals were classified only by the PAL on 2 days, only by the AURAL on 3 days, and on both instruments on 2 days. Bearded seals were classified on 17 days. Classifications were made from the AURAL on all of those days. The PAL data contained bearded seal vocalizations on 16 of those days. Walrus were detected 12 days, one day only on the PAL, 7 days only on the AURAL and 4 days on both instruments. Ribbon seals were classified 12 days as well. On 11 of the days, they were classified on both instruments. They were identified on the PAL uniquely once, with no unique detections on the AURAL. Due to the lower bandwidth of the AURAL system, the detection of echolocation clicks was made exclusively by the PAL. Echolocation clicks were detected on 18 of the 32 days examined.

Of the species identified from the PAL spectra clusters, only classifications for bearded and ribbon seals were included in the daily counts. Classification of bearded seal vocalizations from spectra clusters led to an additional three detection days for the PAL with a recall rate of 11% and a false alarm rate of 3%. Classification of ribbon seal vocalizations from the spectra led to an additional four detection days for the PAL with a recall rate of 7% and a false alarm rate of 2%. Despite repeated efforts at improvement in identification of walrus and bowhead whale spectral patterns, recall rates never exceeded false alarm rates for these classifications from spectra clusters. Therefore, no additional detection days for these species were included.

The number of species detected on each day on each instrument was summed. The number of daily species detected from each instrument was compared using a paired two sided Wilcoxon signed rank test. The test statistic was computed using daily classifications for the PAL exclusively using audio clips and the combined classifications from the audio clips and spectra clusters. For both tests, the Wilcoxon test showed that there was no significant difference in daily classifications from each system: audio clips only ($p = 0.82$, $U = 80.5$, $n = 32$), audio clips and spectra clusters ($p = 0.29$, $U = 70.5$, $n = 32$).

Table 1
AURAL and PAL species identification by date.

Date	Species identified	
	PAL	AURAL
10/1/08	C	H
10/8/08	C	H
10/15/08	H	H
10/23/08		H
11/1/08		H
11/8/08		
11/15/08	H	H
11/23/08	H	H
12/1/08	B H	B
12/8/08	B	B
12/15/08	B H	B
12/23/08	C	B
1/1/09		B
1/8/09	B C (S)	B S W
1/15/09	B C S W	B S W
1/23/09	B C (R S) W	B R S W
2/1/09	B C (S)	B S W
2/8/09	B S	B O S W
2/15/09	B C O R S	B O R S W
2/23/09	B R W	B O R W
3/1/09	B R S W	B R S
3/8/09	B C	B W
3/15/09	B	B O S W
3/23/09	B C S W	B S W
4/1/09	B C (R) S	B R S
4/8/09	B C O R S	B R S R
4/15/09	B C (R) S	B R S
4/23/09	B C (R) S	B R S
5/1/09	C R S	B R S
5/8/09	C R S	R S
5/15/09	C R S	R S
5/23/09	C O R S	O S

3.2. Acoustic entropy

A linear regression was used to compare the daily mean acoustic entropies of the different recording systems. Four tests were conducted 1) default PAL recordings against down-sampled PAL recordings (Fig. 4a), 2) 4.5 s AURAL recordings against 55 s AURAL recordings (Fig. 4d), 3) default PAL recordings against 55 s AURAL recordings (Fig. 4b), and 4) down-sampled PAL recordings against 4.5 s AURAL recordings (Fig. 4c). This analysis excluded data from January 1st as there were no detections from the PAL and therefore no audio clips recorded. The comparison between the PAL recordings with different sample rates indicated a positive relationship ($R^2 = 0.74$, $F = 86.95$, $n = 31$, $p < 0.001$). The comparison between the daily mean acoustic entropy from the AURAL recordings with different lengths indicates a strong linear relationship ($R^2 = 0.86$, $F = 171.6$, $n = 32$, $p < 0.001$). The comparison between the daily average acoustic entropies from the default PAL recordings and the 55 s AURAL segments, showed a weakly linear relationship ($R^2 = 0.012$, $F = 0.325$, $n = 31$, $p = 0.57$). The comparison between the down-sampled PAL and the 4.5 s AURAL daily mean entropies similarly indicated a weakly linear relationship ($R^2 = 0.034$, $F = 1.01$, $n = 31$, $p = 0.323$).

The usefulness of the acoustic entropy as a proxy for biodiversity was examined. The number of marine mammal species classified daily and monthly was compared to the daily and monthly mean acoustic entropies. A linear regression between the daily species classified and mean daily entropy was not significant for either system (PAL: $R^2 = 0.005$, $F = 0.145$, $n = 31$, $p = 0.71$; AURAL: $R^2 = 0.061$, $F = 1.93$, $n = 32$, $p = 0.175$). Extending the time frame for detections and mean entropy to a month was not significant for either system (PAL: $R^2 = 0.18$, $F = 1.34$, $n = 8$, $p = 0.29$; AURAL: $R^2 = 0.14$, $F = 0.99$, $n = 8$, $p = 0.36$). Linear regressions were calculated between the count of daily species classified and the mean daily temporal and spectral entropies separately for each system. The linear regressions of the daily species counts with the spectral entropy indices were not significant for either system (PAL: $R^2 = 0.006032$, $F = 0.176$, $n = 31$, $p = 0.6779$; AURAL: $R^2 = 0.091$, $F = 3.01$, $n = 32$, $p = 0.093$). The linear regressions of species counts and temporal entropy indices for both systems were significant (PAL: $R^2 = 0.143$, $F = 4.829$, $n = 31$, $p = 0.036$; AURAL: $R^2 = 0.14$, $F = 5.069$, $n = 32$, $p = 0.0318$).

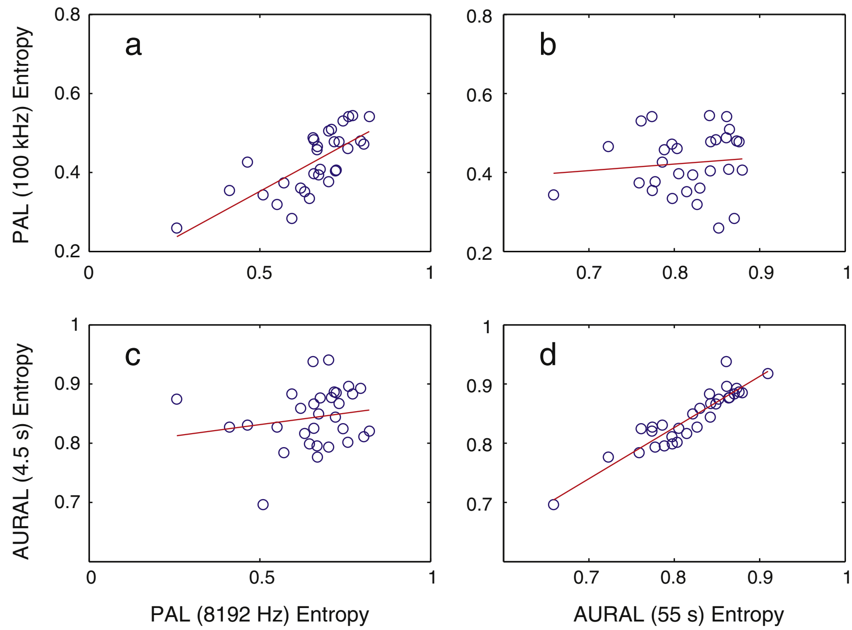

Fig. 4. Comparison of daily mean entropy within and between recorders. a) PAL entropy from 100 kHz as a function of down-sampled 8192 Hz. b) PAL entropy from 100 kHz recordings as a function of 55 s AURAL clips. c) Entropy from 4.5 s AURAL clips as a function of PAL clips at the same sample rate (8192 Hz). d) Entropy from the 4.5 s AURAL clips as a function of entropy from the 55 s AURAL clips.

4. Discussion

4.1. Species classifications

For the comparison conducted here, the number of daily classification of marine mammal vocalizations from each dataset was not significantly different. Only walrus and echolocation clicks were identified in one of the datasets, and not the other, a majority of the days classified. Walrus vocalizations were classified on the AURAL a total of 11 days and only 5 days from the PAL. Echolocation signals were only classified in the PAL data. The difference in echolocation classification can be attributed to the lower usable bandwidth of the AURAL that was inadequate for recording these signals. The difference in classification counts for the walrus requires alternate theories. The short duration of these impulsive signals results in limited triggering of the PAL's onboard event detector. The current detection protocol is not well suited for impulsive signals. Including criteria designed specifically for these signals would improve performance. One such method would involve comparing the kurtosis of the waveform from a one second sample of the signal against a threshold (Mouy et al., 2012).

The spectra clusters saved by the PAL provide additional data from which marine mammal classifications are possible. During periods of high acoustic activity, such as when bearded or ribbon seals were calling nearby in late winter and early spring, the quota of PAL audio files was exceeded early in the day. This limited the temporal coverage of the recordings and may have biased classifications, especially for species that exhibit diel vocal behavior or with less frequent vocalizations. The recognition of patterns associated with stereotyped vocalizations in the spectra clusters was limited by the reduced temporal and frequency resolution of the saved spectra. Classification with this method was made more difficult by the varying noise conditions and by the presence of more than one vocalizing marine mammal within the sample. Classification of bearded and ribbon seal vocalizations from PAL spectra was possible when salient spectral features associated with these stereotyped vocalizations separated them from background noise and overlapping sounds from other vocalizing species. For more targeted species classification, adjustments to the protocol for sub-sampling of the spectra may improve our ability to classify from this sparse dataset. However, this would come at a cost of reduced resolution in other frequency ranges, reduced duty-cycle, or increased memory consumption. This method of classification requires that the spectral history have limited energy contribution from other signals. That is, there should be only one transient signal in the frequencies of interest for the sample to be classified confidently. It is not a very robust process, as evidenced by the low recall rates, but can serve as a highly conservative indicator of presence as the false alarm rates were even lower.

Comparing daily classifications from these two systems with recording parameters chosen without this comparison in mind exposed different limitations for each system. The data from the PAL had variable temporal coverage depending on the event detector output. The shorter duration of the individual PAL recordings permitted recordings with a much larger usable bandwidth covering the frequency range of some echolocating odontocetes including beaked whales. The 4 kHz usable bandwidth of the AURAL precluded detection of high frequency signals of odontocetes. While the low duty cycle, and adaptive sub-sampling of the PAL limited the classification of less frequent vocalizations. The AURAL recorded with high temporal resolution resulting in a large dataset. The collection of the full recordings permits the use of different detection algorithms without the need to specify them before deployment. This provides more flexibility, with the drawback of greater processing time. Due to the disparate record durations of the two datasets, the difference in time required to process each dataset was stark. On average, a month's worth of PAL recordings could be identified in a typical work day (8 h) of devoted processing. For the majority of the days considered, less than one day of the detections from AURAL recordings could be processed in the same period of time. Despite the difference in effort and bandwidth the number of species detected each day was similar.

4.2. Acoustic entropy

The large bandwidth of the PAL recordings contributed to the lower acoustic entropy values. The maximum value of the acoustic entropies calculated for the unaltered PAL recordings was below 0.7. Down-sampling the PAL recordings resulted in higher entropy values. This is most likely due to the absence of high frequency signals in most of the recordings. The higher frequencies did not contribute to the overall acoustic entropy. Down-sampling resulted in recordings with energy in a greater proportion of frequencies. Down-sampling did not change the trend of the acoustic entropy. The entropy from the unaltered and down-sampled PAL recordings was shown to be linearly related (Fig. 4a). The entropy was robust to this change in sample frequency.

The linear relationship between the average daily acoustic entropy computed from the AURAL recordings of different length shows the robustness to sample length. Comparisons of acoustic entropies between two recording systems will need to compensate for biases resulting from different sampling frequencies. Down-sampling the high frequency data to match the sample rate of other recorders brings the values closer together at the cost of considering high frequency signals.

Monthly and daily comparisons of marine mammal classifications and acoustic entropy yielded no relationship. Examining the relationship between the spectral and temporal entropy indices separately yielded mixed results. The temporal entropy indices were related to the daily classifications at the $p < 0.05$ level for both instruments, but the spectral entropy indices were not significantly related to the daily classifications. The temporal entropy index can be used as a proxy for a biodiversity index. There were a number of factors that may have influenced these mixed results. The recordings analyzed were the result of event detections. All of the samples contained at least one and up to four transient signals. The presence of the transient signals meant that there was no baseline period. Marine mammal classifications were made from signals that often overlapped in time and frequency. Non-biological transient sources were classified in the samples and their presence can contribute to the acoustic entropy introducing a bias. The presence of noise is an issue in many facets of bioacoustics.

5. Conclusions

For acoustic biodiversity purposes, the use of these two systems provided similar results. The daily species counts from the two systems were not significantly different. This indicates that daily species counts, reflecting biodiversity, could be compared across systems deployed at locations with similar species makeup. However, the acoustic entropy was not related to the number of classifications from either dataset and was not appropriate as an acoustic biodiversity index.

In considering remotely deployed autonomous passive acoustic recorders, one must balance requirements for power, storage, deployment duration, duty cycle, usable bandwidth, and logistics. For the comparison in this work, the results from the systems were comparable. The main difference was in the time required for manual classification of detected signals. When considering man-hours required, the PAL was a superior instrument. However, the larger, more persistent data set from the 30% duty-cycled AURAL has greater applicability to a wider range of scientific questions related to or including rare vocalizations and vocalization rates.

Acknowledgments

The authors wish to thank the editors for organizing this special issue. Additionally, thanks are extended to the captains and crews of the R/V Miller Freeman and R/V Oscar Dyson for their invaluable help deploying and recovering the instruments. Thanks are also due to Phyllis Stabeno, Bill Floering and Carol Dewitt for the hard work incorporating passive acoustic recorders into their mooring program. Funding is provided by an Eric Walker Graduate Fellowship from ARL Penn State for SLD and Office of Naval Research Marine Mammal and Biology Program Award N000140810391to JLMO.

References

Berchok, C.L., Bradley, D.L., Gabrielson, T.B., 2006. St. Lawrence blue whale vocalizations revisited: characterization of calls detected from 1998 to 2001. J. Acoust. Soc. Am. 120, 2340–2354.

Cummings, W.C., Holliday, D.V., 1987. Sounds and source levels from bowhead whales of Pt. Barrow, Alaska. J. Acoust. Soc. Am. 82, 814–821.

Lammers, M.O., Brainard, R.E., Au, W.W.L., Mooney, T.A., Wong, K.B., 2008. An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. J. Acoust. Soc. Am. 123, 1720–1728.

Mellinger, David K., 2001. Ishmael 1.0 User's Guide. NOAA Technical Memorandum OAR PMEL-120, Pacific Marine Environmental Laboratory, Seattle, WA 30 pp.

Mellinger, D.K., Stafford, K.M., Moore, S.E., Munger, L., Fox, C.G., 2004. Detection of North Pacific right whale (*Eubalaena japonica*) calls in the Gulf of Alaska. Mar. Mamm. Sci. 20, 872–879.

Mellinger, D.K., Stafford, K.M., Moore, S.E., Dziak, R., Matsumoto, H., 2007. Fixed passive acoustic observation methods for cetaceans. Oceanography 20, 36–45.

Mellinger, D.K., Martin, S.W., Morrissey, R.P., Thomas, L., Yosco, J.J., 2011. A method for detecting whistles, moans, and other frequency contour sounds. J. Acoust. Soc. Am. 129, 4055–4061.

Miksis-Olds, J., Parks, S., 2011. Seasonal trends in acoustic detection of ribbon seal (*Histriophoca fasciata*) vocalizations in the Bering Sea. Aquat. Mamm. 37, 464–471.

Miksis-Olds, J., Nystuen, J., Parks, S., 2010. Detecting marine mammals with an adaptive sub-sampling recorder in the Bering Sea. Appl. Acoust. 71, 1087–1092.

Moore, S.E., Reeves, R.R., Southall, B.L., Ragen, T.J., Suydam, R.S., Clark, C.W., 2012. A new framework for assessing the effects of anthropogenic sound on marine mammals in a rapidly changing arctic. Bioscience 62, 289–295.

Mouy, X., Hannay, D., Zykov, M., Martin, B., 2012. Tracking of Pacific walruses in the Chukchi Sea using a single hydrophone. J. Acoust. Soc. Am. 131, 1349–1358.

Nyquist, H., 1932. Regeneration theory. Bell Syst. Tech. J. 11, 126–147.

Nystuen, J., 1998. Temporal sampling requirements for autonomous rain gauges. J. Atmos. Ocean. Technol. 15, 1254–1261.

Richardson, W.J., Greene Jr., C.R., Malme, C.I., Thomson, D.H., 1995. Marine Mammals and Noise. Academic Press, San Diego.

Risch, D., Clark, C., Corkeron, P., Elepfandt, A., Kovacs, K., Lydersen, C., Stirling, I., Van Parijs, S., 2007. Vocalizations of male bearded seals, *Erignathus barbatus*: classification and geographical variation. Anim. Behav. 73, 747–762.

Roch, M.A., Soldevilla, M.S., Hoenigman, R., Wiggins, S.M., Hildebrand, J.A., 2008. Comparison of machine learning techniques for the classification of echolocation clicks from three species of odontocetes. Can. Acoust. 36, 41–47.

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

Stabeno, P.J., Napp, J.M., Whitledge, T.E., 2008. Sentinels for ecosystem change: long-term biophysical moorings on the Bering Sea shelf (2007-2008). North Pacific Research Board Final Report RO701. 19 pp.

Sueur, J., Pavoine, S., Hamerlynck, O., Duvail, S.p., 2008. Rapid acoustic survey for biodiversity appraisal. PLoS One 3, e4065.

Sueur, J., Gasc, A., Grandcolas, P., Pavoine, S., 2012. Global estimation of animal diversity using automatic acoustic sensors. In: Le Galliard, J.F., Guarini, J.M., Gaill, F. (Eds.), Sensors for ecology: towards integrated knowledge of ecosystems. CNRS Editions, pp. 101–119.

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

Watkins, W., Ray, G., 1977. Underwater sounds from ribbon seal, *Phoca* (*Histriophoca*) *fasciata*. Fish. Bull. 75, 450–453.

Wiggins, S.M., Hildebrand, J.A., 2007. High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring, Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies. IEEE, Tokyo 7.