True or False: Developing Methods for Pruning False and Inaccurate Detections from an Automated Whistle Detector

Robyn Walker[1], Julie Oswald[1], Cory Hom-Weaver[1], Kerry Dunleavy[1], Thomas Norris[1]

[1]Bio-Waves, Inc., 364 2nd Street, Suite #3, Encinitas, CA 92024 USA

Introduction

Automated methods for detecting and extracting delphinid whistles are becoming more popular for analyzing large datasets, however due to false detections, inaccurate contour extractions and whistle fragmentation, these methods may provide inaccurate results. In addition, detection and extraction errors are affected by noise in recordings, making it difficult to compare results across different methods. In this study, we develop methods for removing or pruning out false or 'Bad' extractions, reducing inaccurate results.

Objectives

- Determine which whistle variables can be used to identify and prune false or inaccurate whistle extractions in automated data.
- Determine values that can be used to prune the most 'Bad' extractions, while retaining the most "Good" extractions.

Methods

DATA

- Visually validated, single-species recordings from:
  - striped dolphin (*Stenella coeruleoalba*)
  - bottlenose dolphin (*Tursiops truncatus*)
  - short-beaked common dolphin (*Delphinus delphis*)
  - Atlantic spotted dolphin (*Stenella frontalis*)
  - pilot whale (*Globicephala spp.*)

Figure 1: Atlantic spotted dolphins (*Stenella frontalis*). Photo taken by Heather Foley, Duke University, under NOAA Permit No. 16185.

Methods (cont.)

MANUAL AND AUTOMATED WHISTLE EXTRACTIONS

- Whistles were detected, extracted and measured using the whistle and moan detector and ROCCA (Real-time Odontocete Call Classification Algorithm; Oswald et al. 2013) modules in PAMGuard (Gillespie et al. 2008).
- Fifty variables describing the whistle were measured for each extraction.
- Whistle extractions examined manually and labelled as 'Good' or 'Bad' (**Figure 2**) using custom MATLAB code.

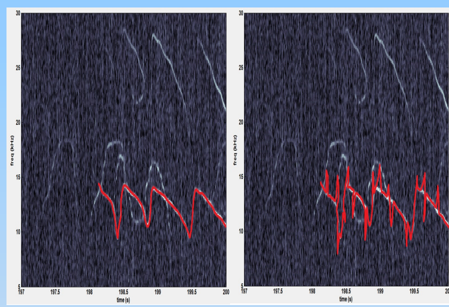

Figure 2: Example of 'Good' (left) and 'Bad' (right) whistle contour extractions.

FINDING PRUNING VALUES

- Pruning variables and values were chosen by comparing histograms of measurements taken from 'Good' and 'Bad' extractions (**Figure 3**).
- Variables that had different distributions for 'Good' and 'Bad' extractions were tested as pruning variables
- Variables were chosen that provided the best trade-off between pruning 'Bad' extractions and keeping 'Good' extractions
- Variables examined (singly and in combination) included:

▪Maximum Frequency	▪Absolute Slope
▪Minimum Frequency	▪Positive Slope
▪Frequency Range	▪Negative Slope
▪Beginning Frequency	▪Frequency Slope Ratio
▪End Frequency	▪Duration

Results

- **Maximum Frequency, Absolute Slope and Duration were used to concurrently prune the most 'Bad' extractions while retaining the most 'Good' extractions (Table 1, Figure 4).**

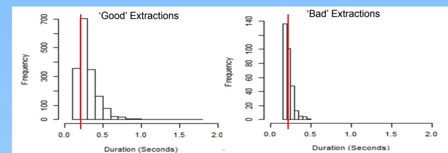

Figure 3: Duration of 'Good' vs. 'Bad' extractions for Atlantic spotted dolphins. Red line indicates lower pruning value for duration. Example of histogram used for choosing pruning values.

Table 1: Final values used for pruning 'Bad' whistle extractions out of an automated data set

Variable	Pruning Value (Less Than)	Pruning Value (Greater Than)
Maximum Frequency (Hz)	10600	27000
Absolute Slope Mean (Hz/s)	9100	82000
Duration (s)	0.15	2.5

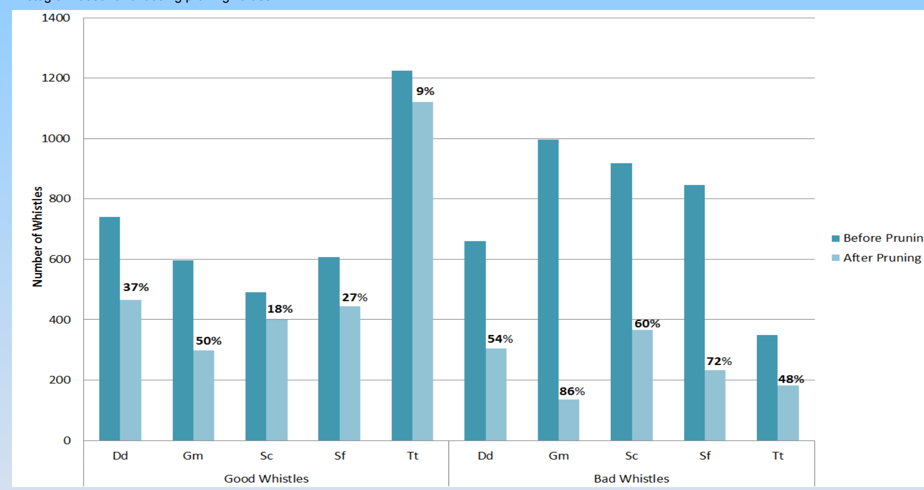

Figure 4: Number of 'Good' and 'Bad' whistles in the dataset before and after pruning, Percent of whistles removed during pruning is indicated.

Conclusions

Using the whistle variables frequency, slope and duration allows for a large proportion of 'Bad' extractions to be removed for each species, while retaining the majority of good extractions.

Future Work

In the future, these pruning variables and values will be used to attempt to increase the success of automated whistle classifiers. We will test this by comparing classification results from pruned and un-pruned datasets.

References

▪Gillespie, D., D.K. Mellinger, J. Gordon, D. Mclaren, P. Redmond, R. McHugh, P. W. Trinder, X. Y. Deng, and A. Thode. 2008. PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans. *The Journal of the Acoustical Society of America,* 30(5): 54-62.

▪Oswald, J.N., S. Rankin, J. Barlow, M. Oswald and M.O. Lammers. 2013. Real-time Call Classification Algorithm (ROCCA): software for species identification of delphinid whistles. In: Detection, Classification and Localization of Marine Mammals using Passive Acoustics, 2003-2013: 10 years of International Research, DIRAC NGO, Paris, France, pp. 245-266.

Acknowledgements

We would like to thank the Northeast Fisheries Science Center, the Southeast Fisheries Science Center, and Duke University who collected the visually confirmed single species recordings used in this analysis during combined visual and acoustic surveys of the northwest Atlantic coast. Funding for this project was provided by the Office of Naval Research (ONR), and the Living Marine Resources (LMR) Program.