Oregon State University

Detecting Ultrasonic Harbor Porpoise Clicks
with Hetrodyning & Teager-Kaiser Energy Operator

Patrick Donnelly, David Mellinger, and Daniel De Leon

soundbendor lab

College of Engineering | Computer Science



INTRODUCTION

Bay of Fundy

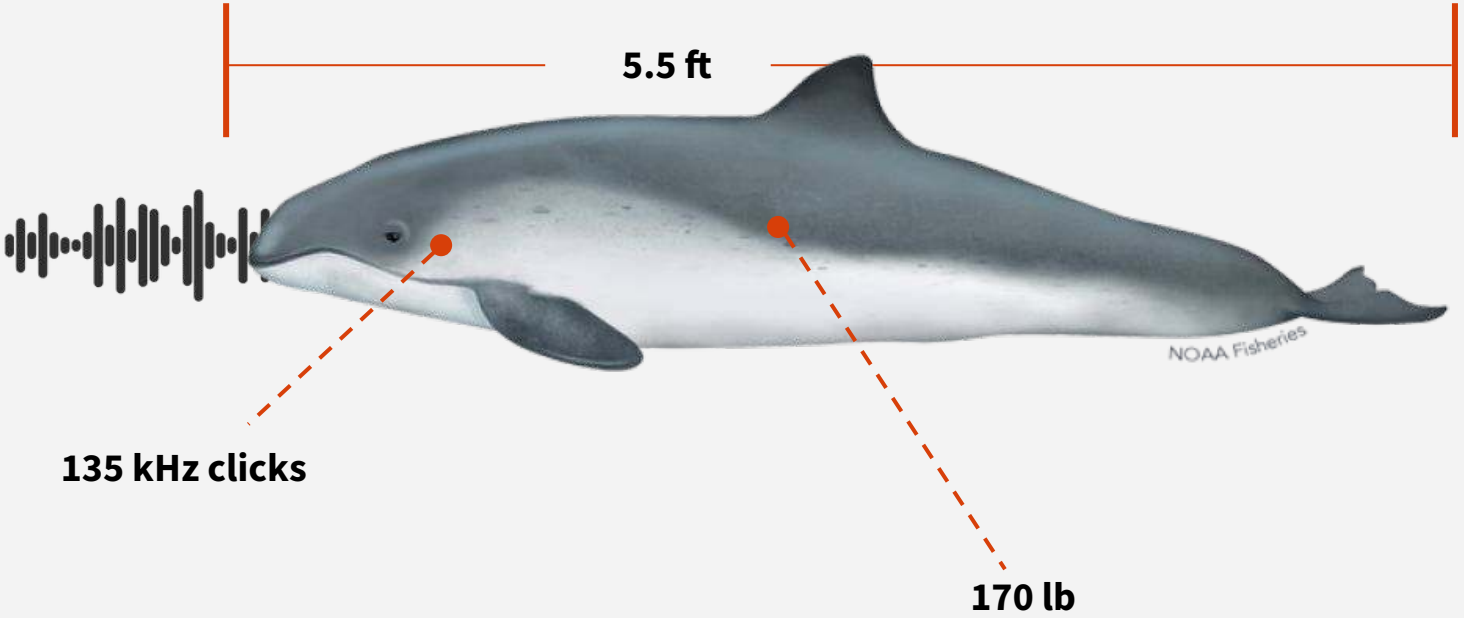
- tides up to 38 ft





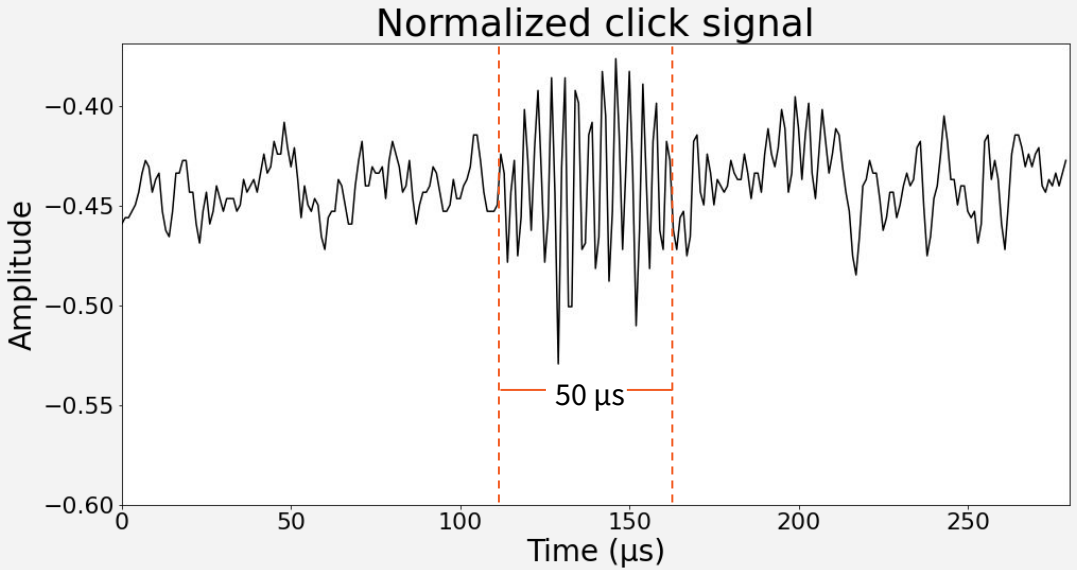
Harbor Porpoise

Oregon State University



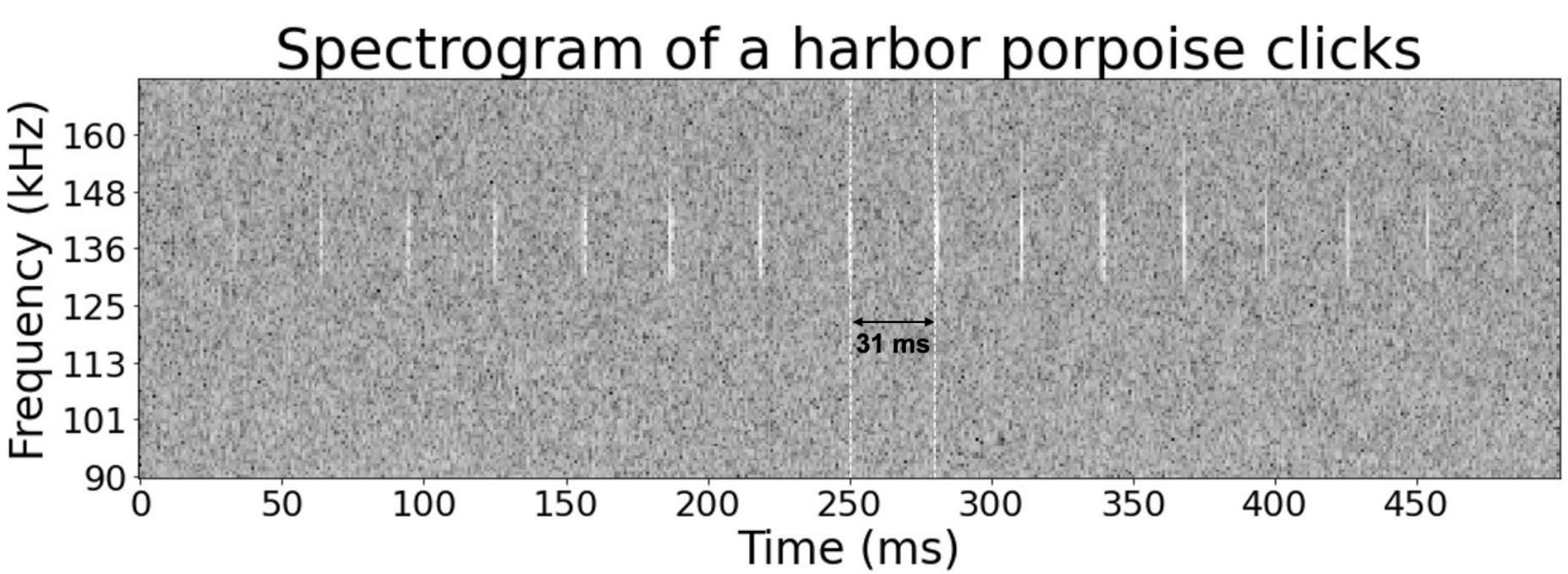
Harbor Porpoise Clicks

- Narrow-bandwidth
  - 6 to 26 kHz
- High-frequency
  - 130 to 140 kHz
- Duration
  - 44 and 115 µs
- Interclick interval
  - 30 to 100 ms



Harbor Porpoise Clicks

Oregon State University



Problem Statement

- Machine listening agent that can automatically predict the presence of a harbor porpoise click.
  - signal automatic detections to the turbine
  - capturing harbor porpoise clicks requires a **lot of data**

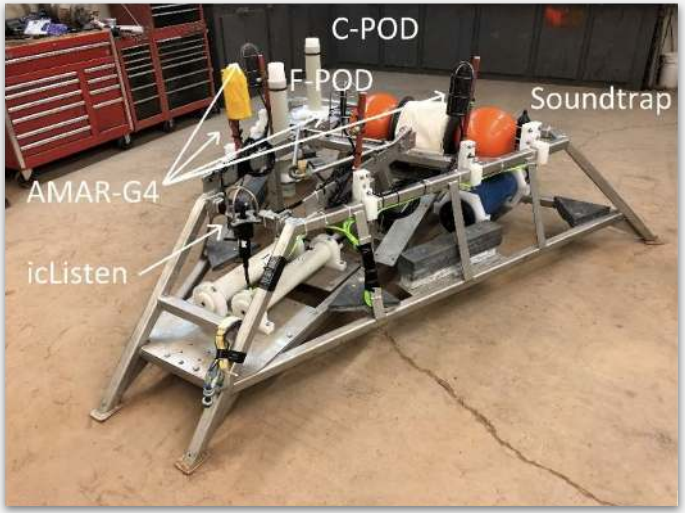


1 s = 512,000
1 hr = 1.8 B

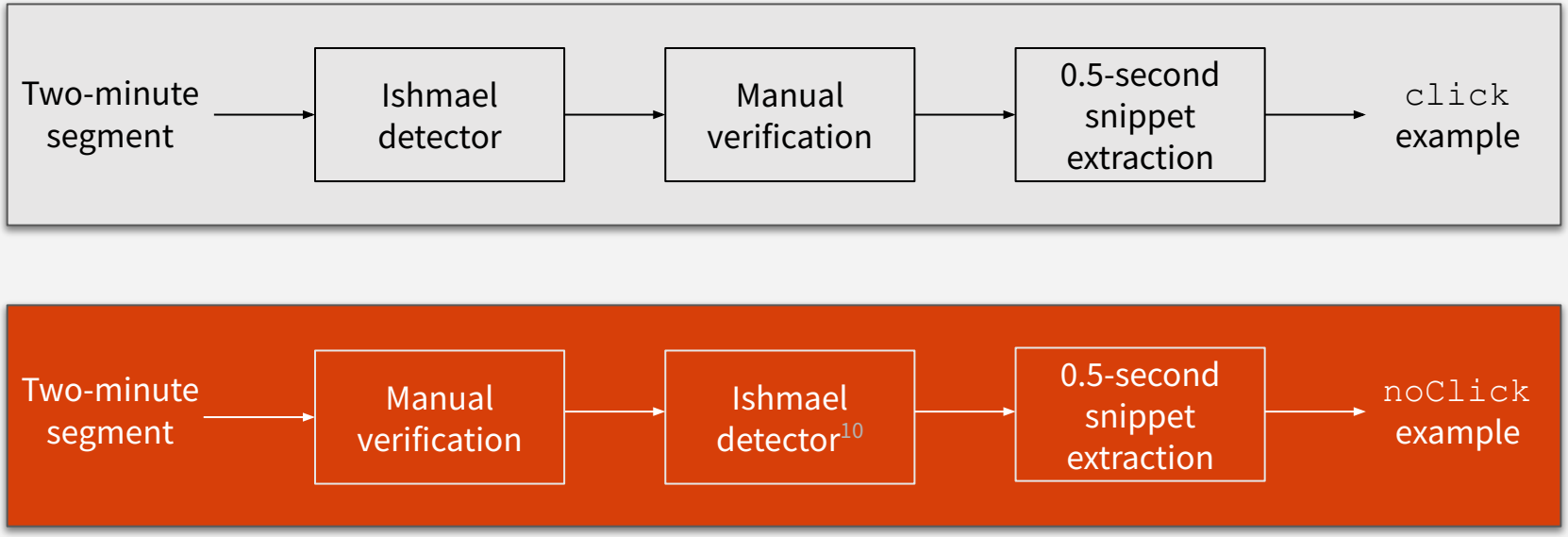

DATA

Hydrophone

- Fundy Ocean Research Center for Energy (FORCE)
- Autonomous Multi-channel Acoustic Recorder (AMAR)
- two deployments in 2019
  - 6/4 – 7/29
  - 9/5 – 9/13
- 512 kHz sample rate



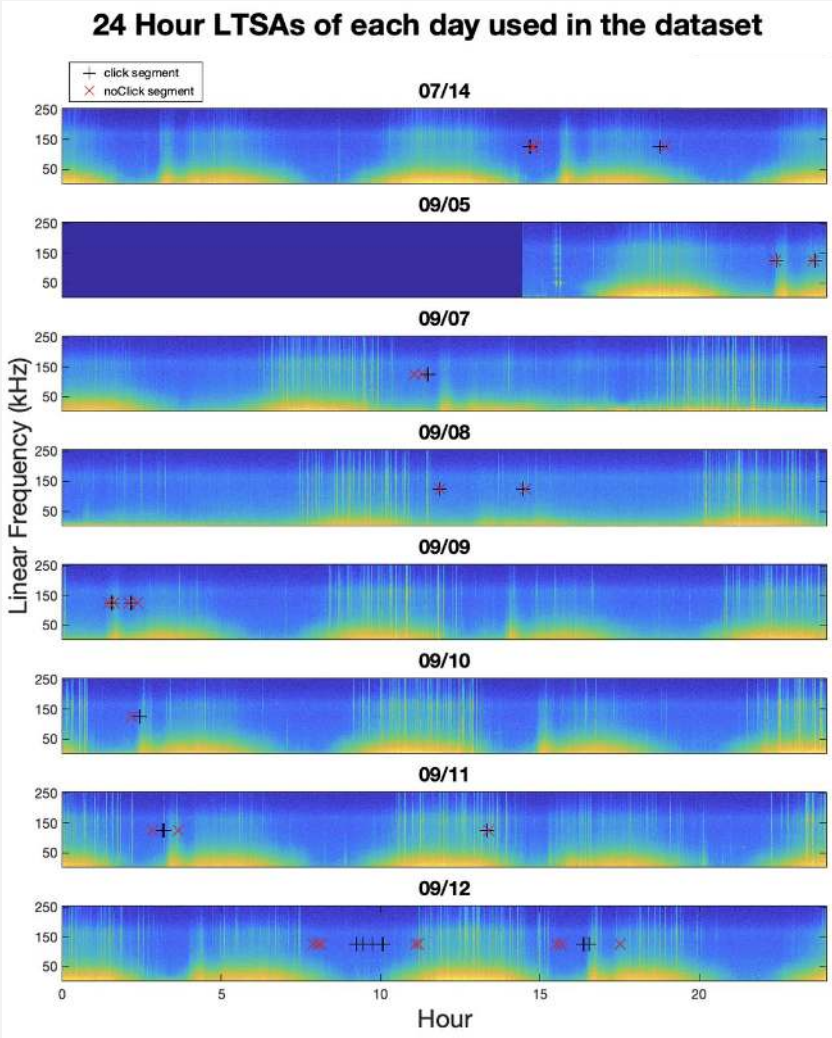
Data Collection



Data Collection

day	example count	average RMS_s	RMS_s std
07/14	1,848	0.555	0.163
09/05	674	0.320	0.047
09/07	1,538	0.480	0.103
09/08	668	0.340	0.058
09/09	1,662	0.347	0.055
09/10	426	0.452	0.124
09/11	814	0.430	0.143
09/12	656	0.410	0.133

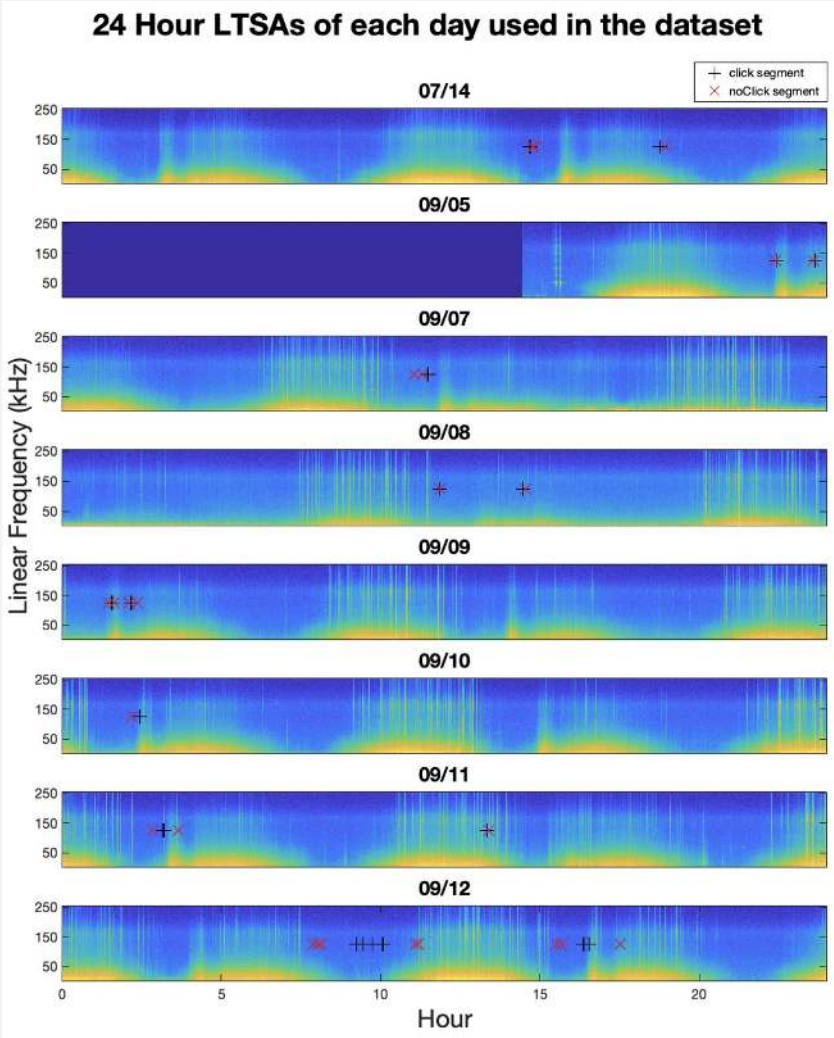
Total examples: 8,286



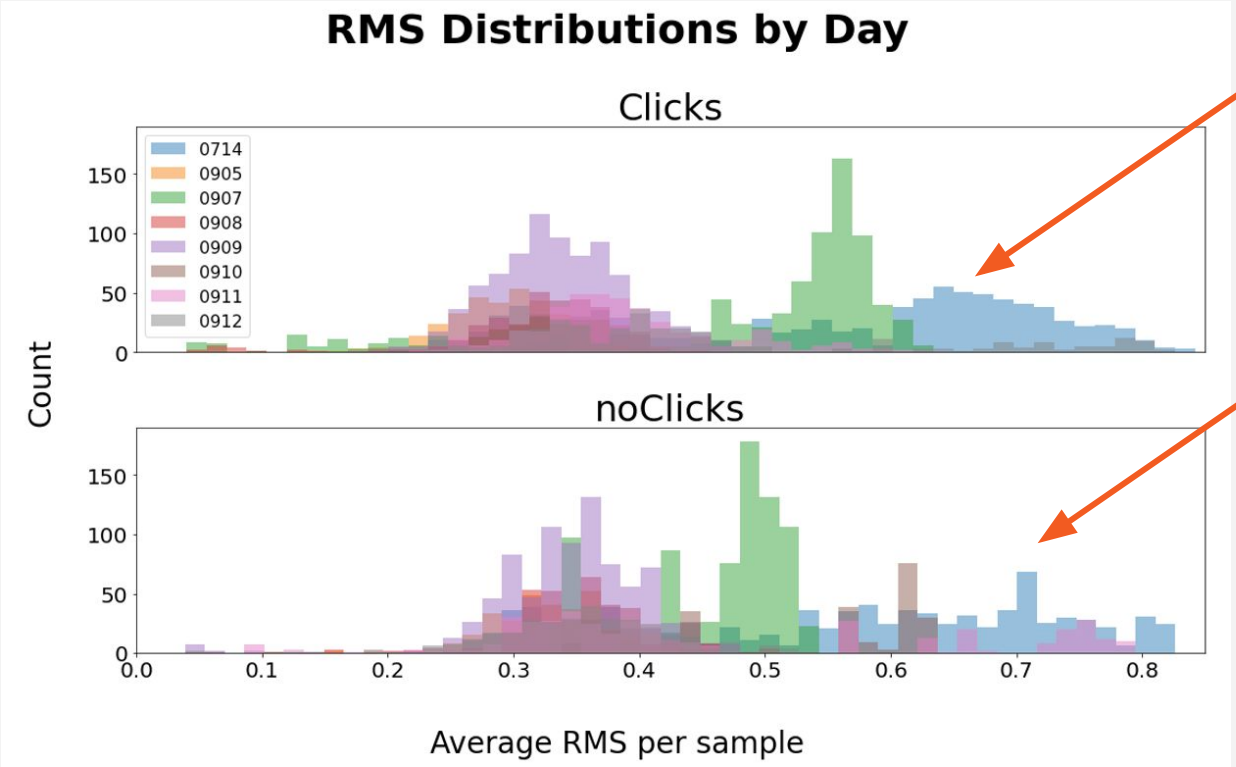
Data Collection

day	example count	average RMS_s	RMS_s std
07/14	1,848	0.555	0.163
09/05	674	0.320	0.047
09/07	1,538	0.480	0.103
09/08	668	0.340	0.058
09/09	1,662	0.347	0.055
09/10	426	0.452	0.124
09/11	814	0.430	0.143
09/12	656	0.410	0.133

Total examples: 8,286



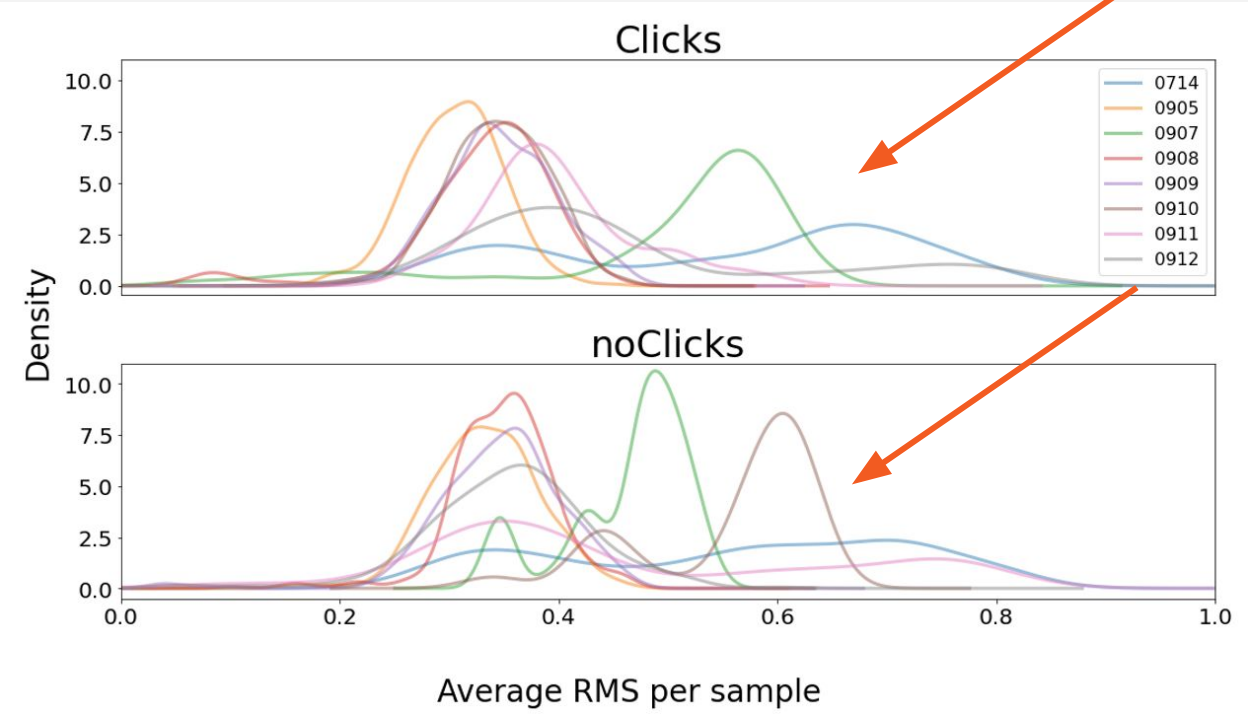
Data Collection



Data Collection

Oregon State University

RMS kernel density estimation by day



RELATED WORK



Related Work

- Acoustic feature extraction
  - Mel frequency cepstral coefficients[1]
  - Short-time Fourier Transform
- Algorithms
  - Decision Trees[2]
  - Deep Learning[3-4]
- TKEO for sperm whales[5]
- Heterodyne for bats[6-7]

[1]

Table 7

Species	Median Frequency Mean Hz	Total Spectrum Median Hz	Modal Spectrum Median Hz
B. mysticetus	$698\,\frac{1723}{200}$	$642\,\frac{1586}{187}$	$634\,\frac{1591}{188}$
C. marginata	$83\,\frac{86}{81}$	$82\,\frac{84}{80}$	$82\,\frac{86}{80}$
E. glacialis	$345\,\frac{593}{138}$	$317\,\frac{622}{139}$	$226\,\frac{330}{28}$
E. australis	$157\,\frac{235}{39}$	$145\,\frac{218}{27}$	$141\,\frac{219}{25}$
B. acutorostrata	$78\,\frac{110}{28}$	$75\,\frac{103}{28}$	$73\,\frac{100}{30}$
B. physalus	$66\,\frac{129}{25}$	$66\,\frac{129}{25}$	$67\,\frac{130}{25}$
P. catodon	$4906\,\frac{11283}{525}$	$4340\,\frac{11578}{234}$	$2606\,\frac{10089}{162}$
D. leucas	$2058\,\frac{2058}{2058}$	$2279\,\frac{2279}{2279}$	$2312\,\frac{2312}{2312}$
S. longirostris	$9451\,\frac{26124}{349}$	$9702\,\frac{26160}{320}$	$9664\,\frac{26018}{321}$
S. long. + P. cat.	$16107\,\frac{22928}{10992}$	$16848\,\frac{26772}{10116}$	$10874\,\frac{12547}{10013}$
S. bredanensis	$11038\,\frac{16942}{7180}$	$8837\,\frac{17320}{3849}$	$7230\,\frac{9013}{2832}$
C. commersonii	$3236\,\frac{3664}{2809}$	$3473\,\frac{3683}{3262}$	$3503\,\frac{3685}{3322}$
D. delphis	$13538\,\frac{26735}{3285}$	$13571\,\frac{30097}{1753}$	$11917\,\frac{19455}{2286}$
G. griseus	$6518\,\frac{11305}{939}$	$6763\,\frac{12346}{869}$	$6408\,\frac{12397}{286}$
G. macrorhynchus	$4339\,\frac{9384}{2630}$	$3859\,\frac{8766}{1573}$	$3398\,\frac{8568}{198}$
G. melaena	$5195\,\frac{7203}{2530}$	$5198\,\frac{8451}{2534}$	$5136\,\frac{8442}{2518}$
O. orca	$1850\,\frac{3357}{938}$	$1480\,\frac{2064}{622}$	$1290\,\frac{2062}{583}$
P. crassidens	$6558\,\frac{9636}{2091}$	$5623\,\frac{7964}{165}$	$5459\,\frac{7955}{148}$
P. phocoena	$1423\,\frac{2537}{536}$	$1346\,\frac{3017}{14}$	$625\,\frac{2520}{9}$
I. geoffrensis	$3622\,\frac{4250}{2994}$	$3387\,\frac{4072}{2702}$	$2840\,\frac{3129}{2551}$
A. phillipi	$263\,\frac{393}{185}$	$149\,\frac{161}{141}$	$125\,\frac{135}{118}$
Hammer on metal	$2718\,\frac{2718}{2718}$	$1217\,\frac{1217}{1217}$	$840\,\frac{840}{840}$

Table 1: Summary of Statistics

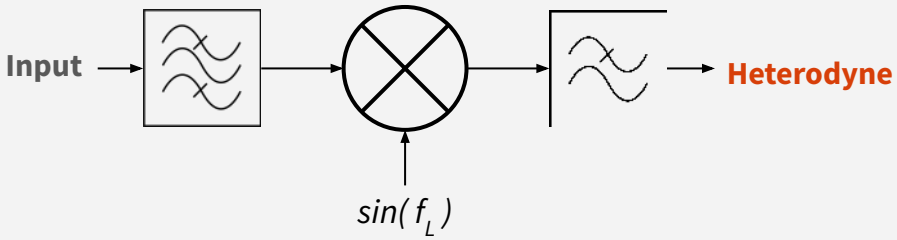
	Equation	Page
Duration: Results in Table 2		
Total Duration	2	8
Sound Concentration	3	8
Amplitude Modulation: Results in Table 3		
Amplitude Mean	4	9
Amplitude Standard Deviation	5	9
Attack Fraction	6	9
Attack Proportion	7	10
Amplitude Skewness	8	10
Frequency Modulation: Results in Table 4		
Upsweep Mean	9	10
Upsweep Fraction	10	11
Upsweep Proportion	11	11
Time Frequency Correlation	12	11
Time Upsweep Correlation	13	12
Short-term Bandwidth: Results in Table 5		
Short-term Bandwidth Mean	14	12
Short-term Spectral Concentration	15	13
Short-term Spectral Asymmetry	16	13
Aggregate Bandwidth: Results in Table 6		
Total Upper Frequency - F_{75}		5, 13
Total Lower Frequency - F_{25}		5, 13
Total Spectrum Concentration - N_{50}		5, 13
Modal Upper Frequency - F_{75}		5, 13
Modal Lower Frequency - F_{25}		5, 13
Modal Spectrum Concentration - N_{50}		5, 13
Center Frequency: Results in Table 7		
Median Frequency Mean	17	14
Total Spectrum Median Frequency		
Modal Spectrum Median Frequency		
Amplitude Frequency Interaction: Results in table 8		
Amplitude Frequency Correlation	18	14
Amplitude Upsweep Correlation	19	14

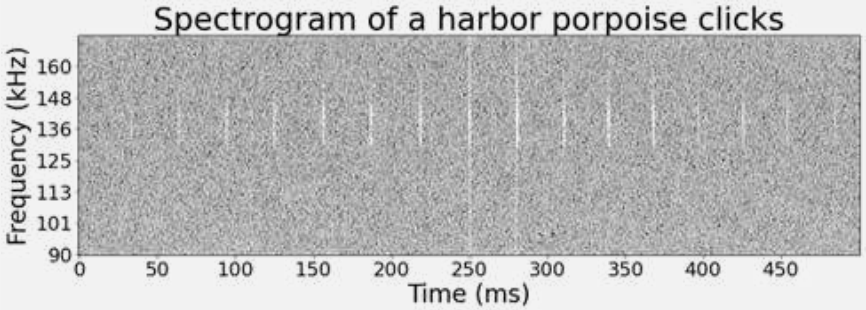


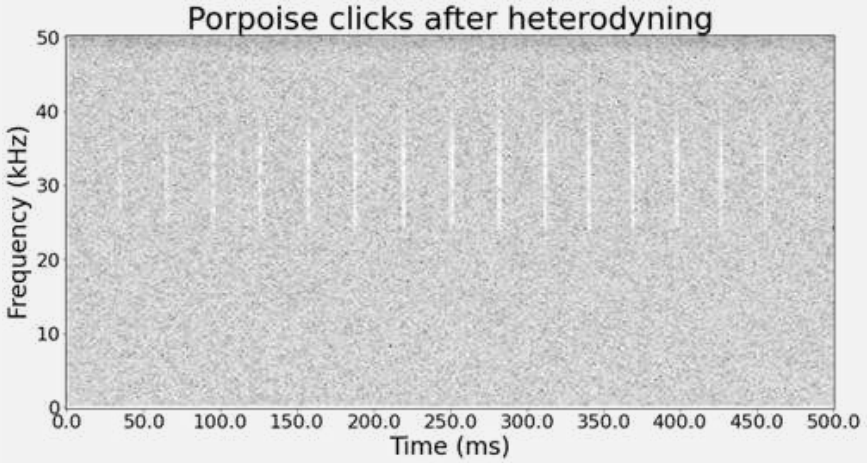
EXPERIMENTS

Heterodyne

- Heterodyne
  - band-pass filter
  - shift frequencies
  - low-pass filter
- Reduces sampling rate
  - from 512 kHz to 100 kHz



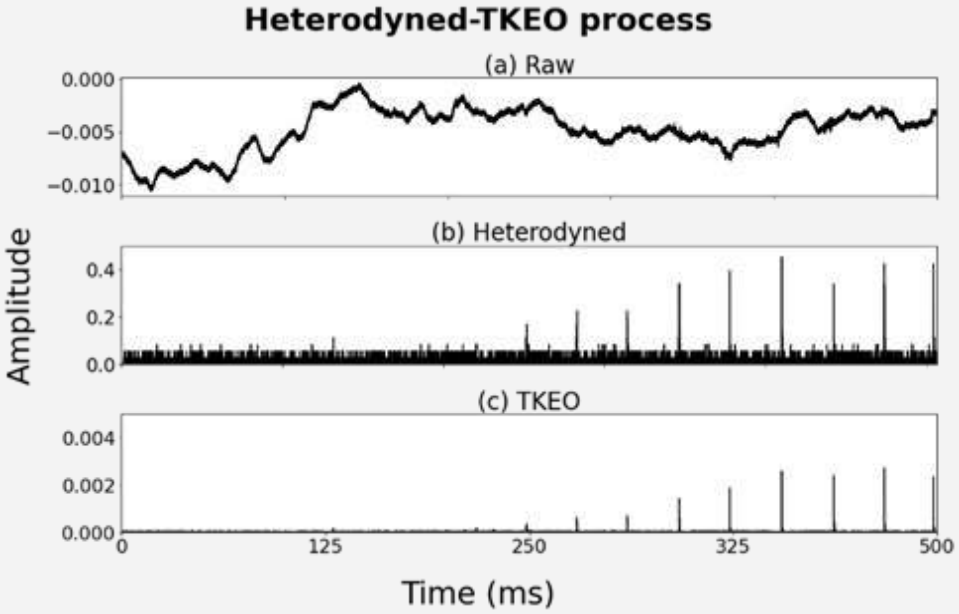




Teager-Kaiser Energy Operator

- measures change of energy
  - difference of the squared value
  - and product of the previous and following samples
- discrete operation

$$\psi[x(n)] = x^2(n) - x(n+1)x(n-1)$$



Experiment

- Convolutional neural network (CNN)
  - 2 × 2D convolutional layers with pooling
  - 2 × dense layers
- Long Short-Term Memory (LSTM)
  - 2 × LSTM layers with dropout
  - 2 × dense layers

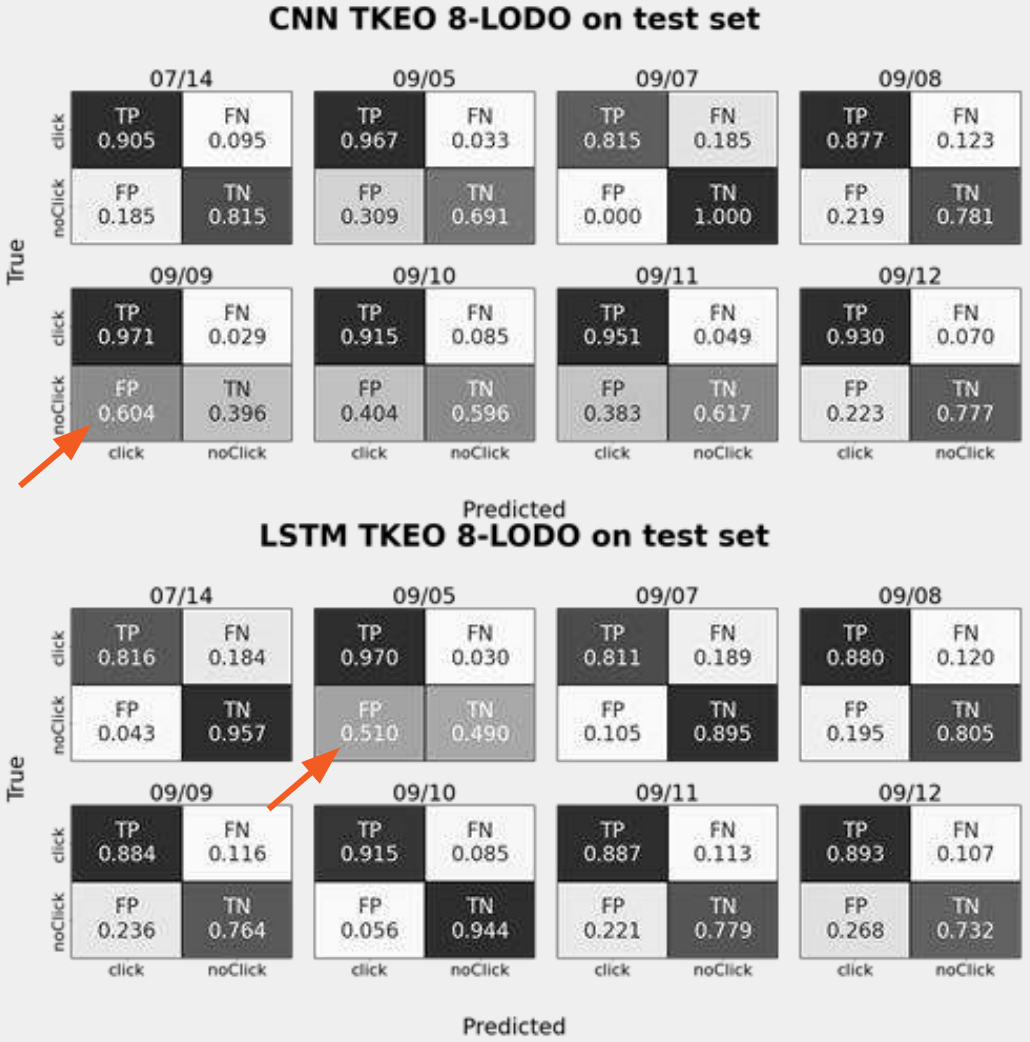
	CNN	LSTM
Precision	.78	.83
Recall	.92	.88
F_1 Score	.83	.85
Accuracy	.81	.84



RESULTS & DISCUSSION

Results

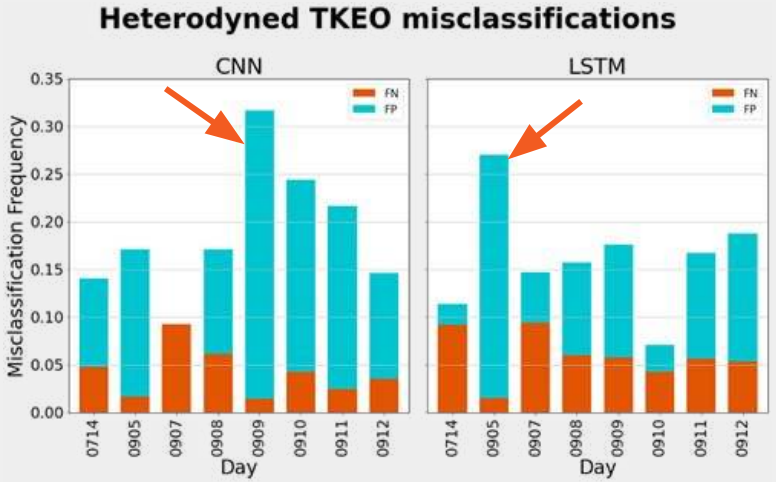
- CNN and LSTM
  - overall similar performance
- day to day variances
  - CNN for 09/09
    - 0.60 false positive rate
  - LSTM on 09/05
    - 0.51 false positive rate

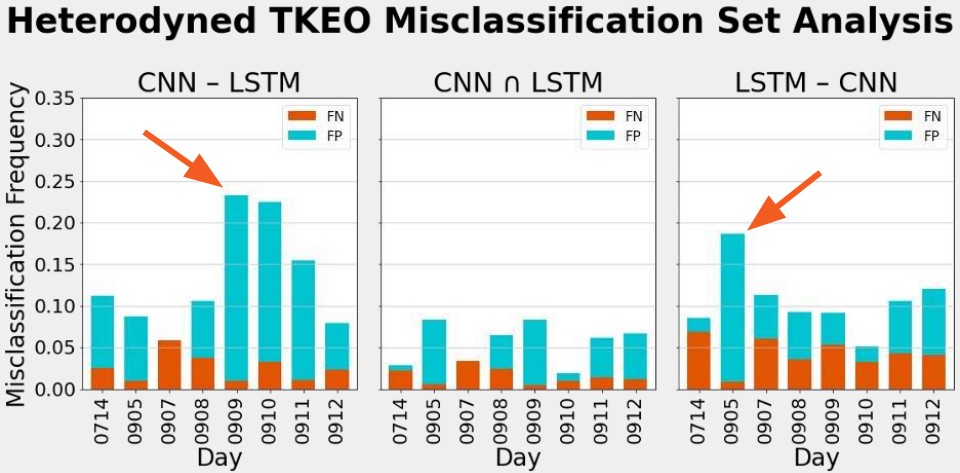


Misclassifications

- CNN struggled on 09/09
- LSTM struggled on 09/05

- intersection:
  - errors on different days
  - utility of both models
  - ensemble needed





CONCLUSION



Summary

- Heterodyne-TKEO
  - reduces data size
- CNN and LSTM
  - performance varied by day
- cross-validation by day
  - counter covariate shift
  - noise-varying environments

Future Work

- compare other feature sets
  - e.g., MFCCs, SFTF, wavelets
- explore ensemble approach
- collect more data
- investigate correlation
  - tide cycle & ambient noise

References

[1] K. Fristrup and W. Watkins, “Characterizing acoustic features of marine animal sounds,” Woods Hole Oceanographic Institution, Technical Report, Jan. 1992, doi: 10.1575/1912/3055.

[2] K. M. Fristrup and W. A. Watkins, “Marine animal sound classification,” Woods Hole Oceanographic Institution, Technical Report, Oct. 1993. doi: 10.1575/1912/546.

[3] P. C. Bermant, M. M. Bronstein, R. J. Wood, S. Gero, and D. F. Gruber, “Deep machine learning techniques for the detection and classification of sperm whale bioacoustics,” Scientific Reports, vol. 9, no. 1, Art. no. 1, Aug. 2019, doi: 10.1038/s41598-019-48909-4.

[4] D. Duan et al., “Real-time identification of marine mammal calls based on convolutional neural networks,” Applied Acoustics, vol. 192, p. 108755, Apr. 2022, doi: 10.1016/j.apacoust.2022.108755.

[5] V. Kandia and Y. Stylianou, “Detection of sperm whale clicks based on the Teager–Kaiser energy operator,” Applied Acoustics, vol. 67, no. 11, pp. 1144–1163, Nov. 2006, doi: 10.1016/j.apacoust.2006.05.007.

[6] G. Mirzaei et al., “The BIO-acoustic feature extraction and classification of bat echolocation calls,” IEEE International Conference Electro Information Technology., pp. 1–4, May 2012, doi: 10.1109/EIT.2012.6220700.

[7] Ingemar Ahlén, “Heterodyne and time-expansion methods for identification of bats in the field,” 2004. Department of Conservation Biology. July, 2004.