Intra-specific variability in delphinid whistle structure: implications for acoustic species identification

Julie Oswald[1], Kerry Dunleavy[1], Cory Hom-Weaver[1], Tina Yack[1], Shannon Rankin[2], Lynne Hodge[3], Danielle Cholewiak[2], Melissa Soldevilla[2], Anthony Martinez[2], Andrew Read[3]

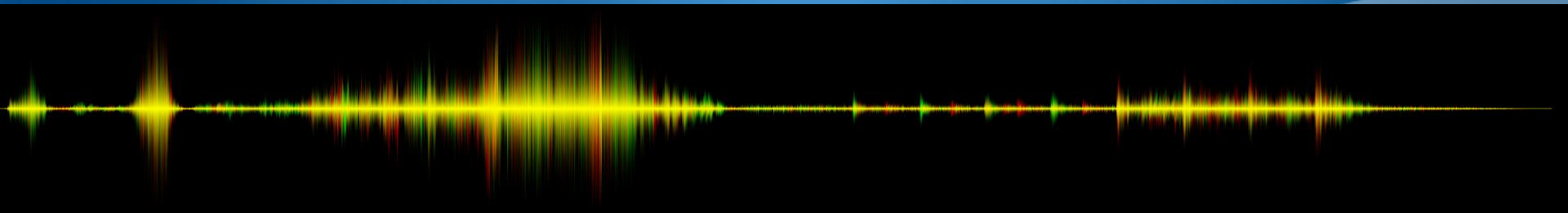
1 Bio·Waves Incorporated

2 NOAA National Oceanic and Atmospheric Administration U.S. Department of Commerce

3 Nicholas School of the Environment and Earth Sciences Duke University

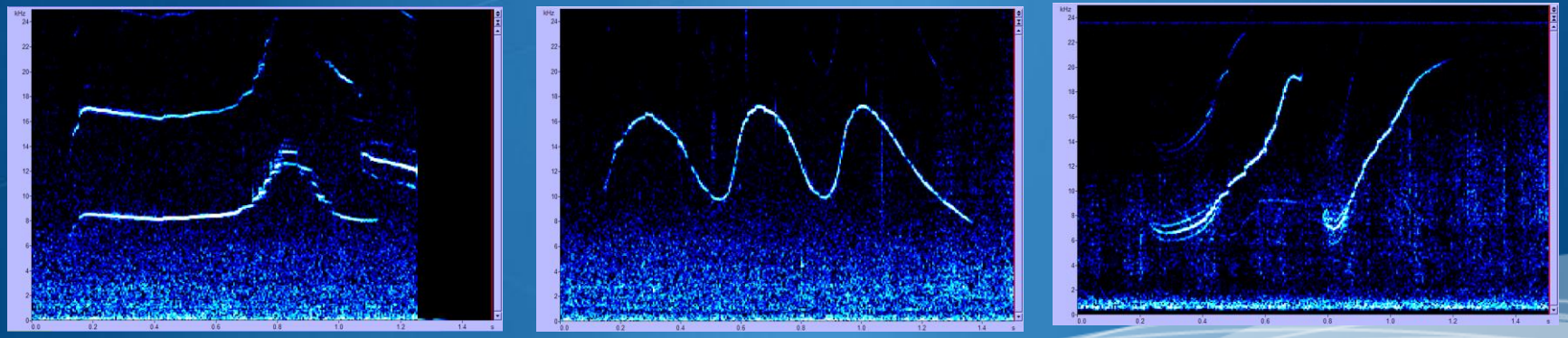
Introduction

- Sound is a primary modality for communication for all cetacean species
- We can eavesdrop and use passive acoustic methods to learn about these species
- A challenging first step in analyzing passive acoustic data is identifying species present in recordings



Introduction

- Marine mammal sounds:
  - Variable within species
  - Time-frequency characteristics often overlap among species
- Statistical classifiers are necessary for species identification from acoustic recordings



Striped dolphin

Bio·Waves Incorporated

Statistical Classifiers

- Require large amounts of acoustic data
- Recordings must
  - Have visual confirmation of species identity
  - Contain a single species
- Time consuming, expensive and difficult
- Combining datasets from different locations could increase sample sizes
- Complicated by geographic variation in signal structure

Bio·Waves Incorporated

Geographic Variation

- Geographic variation in whistle structure has been shown for many species
- If intra-species variability is large enough, it could affect classifier performance
- Does geographic variation exist in whistle variables used for classification?
- Does geographic variation affect classifier performance?

Bio·Waves Incorporated

Whistle Data

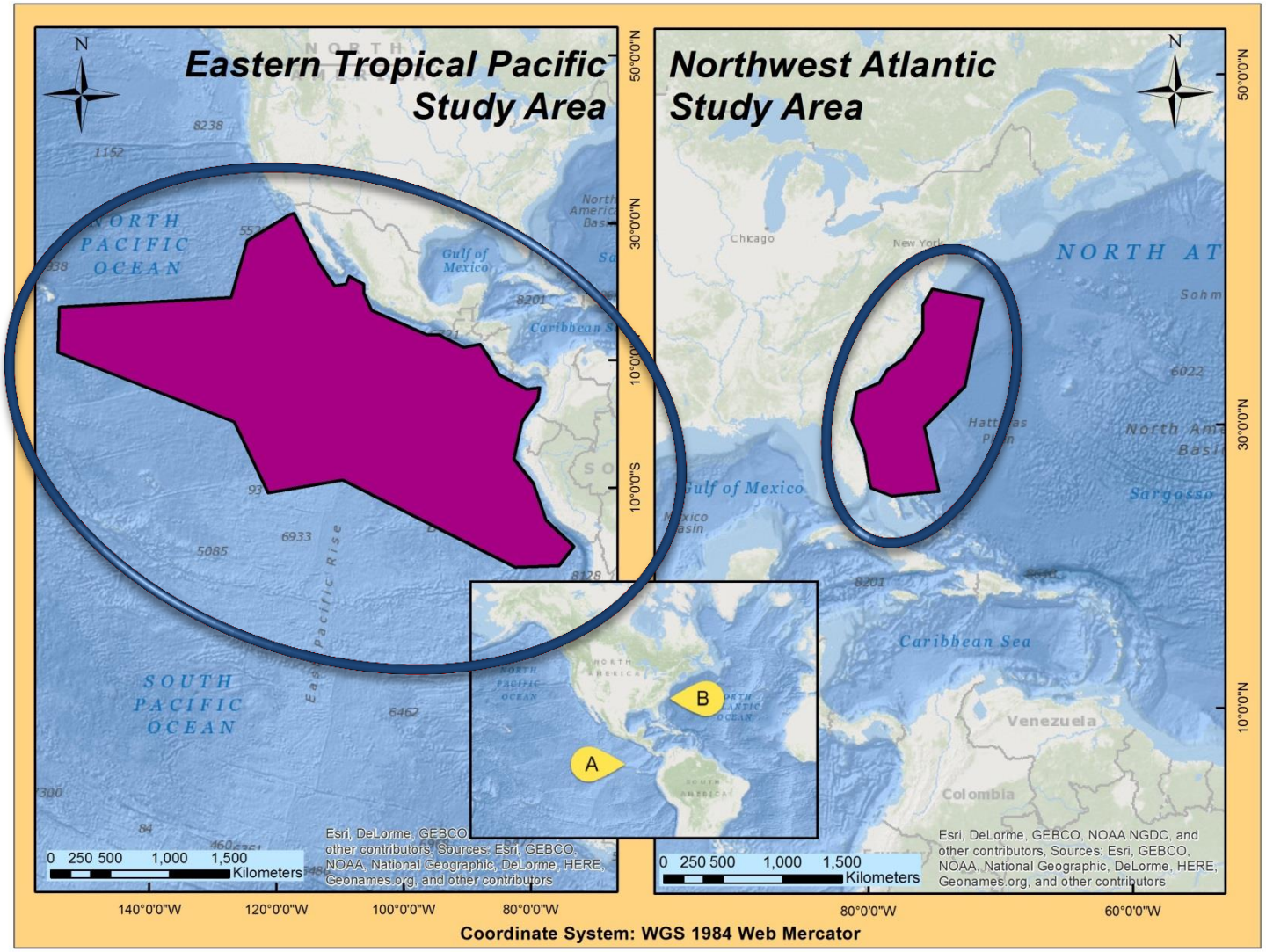
- Recorded during visual and acoustic marine mammal surveys
  - Duke University, Southeast Fisheries Science Center, Northeast Fisheries Science Center, Southwest Fisheries Science Center
- Towed hydrophone arrays, DTAGs
- Sample rates: 48 kHz – 192 kHz

NOAA

NICHOLAS SCHOOL OF THE ENVIRONMENT AND EARTH SCIENCES
DUKE UNIVERSITY

Bio·Waves Incorporated

Study Areas



BioWaves Incorporated

Whistle analysis

- Single species recordings
  - Visual confirmation of species identity
- Randomly selected maximum of 50 whistles per encounter
- Whistles traced manually using ROCCA software
  - PAMGuard module
- ROCCA automatically measured 53 variables
  - Duration, frequencies, slopes, shape variables

Bio·Waves Incorporated

Whistle Analysis cont'd

Species	# Whistles Northwest Atlantic	# Whistles Tropical Pacific
Short-beaked common dolphins	308	226
Pilot whales	250	109
Striped dolphins	250	109
Rough-toothed dolphins	225	145
Bottlenose dolphins	250	109

- Whistle variables compared using Mann-Whitney U tests

Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

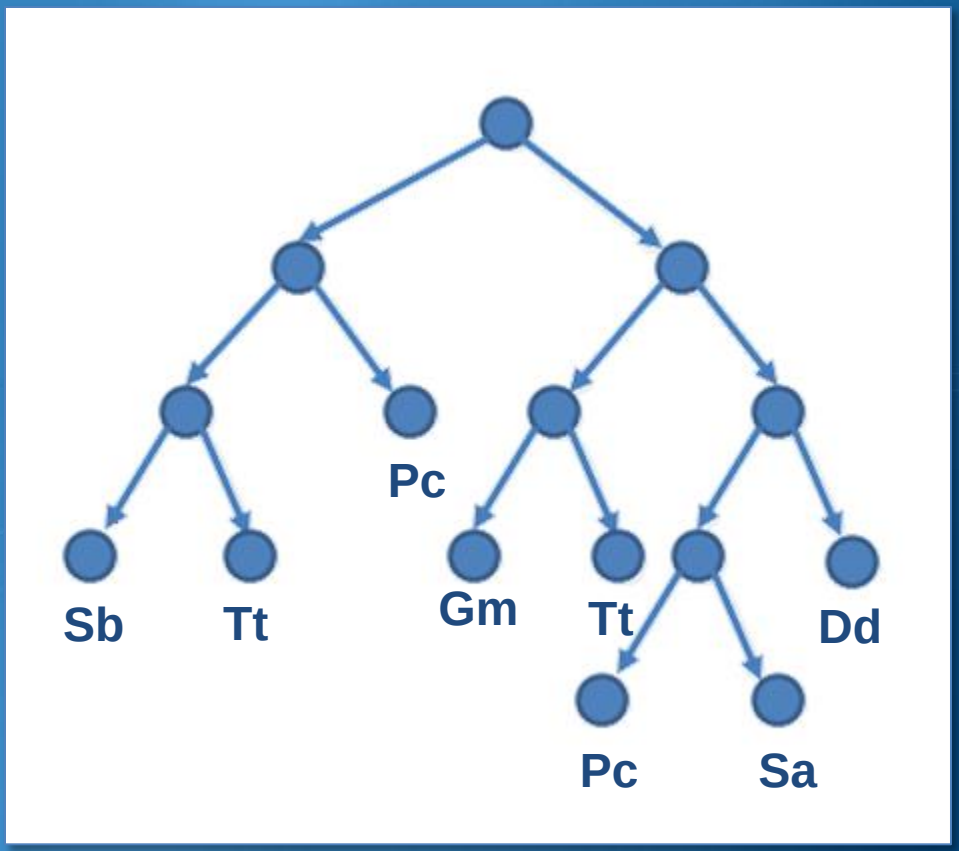
Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Do differences in whistle variables between study areas affect classifier performance?

Bio·Waves Incorporated

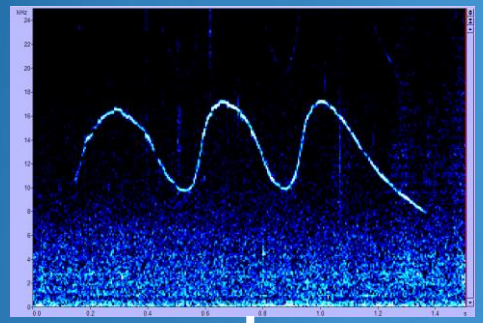
Random Forest

- Collection of decision trees
- Binary partitioning of data
- Each split based on a single variable
- Splitting variable chosen randomly at each node



Bio·Waves Incorporated

Random Forest



Striped | Striped | Common | Bottlenose | Striped | Common | Striped | Striped | Common | Bottlenose

Bio·Waves Incorporated

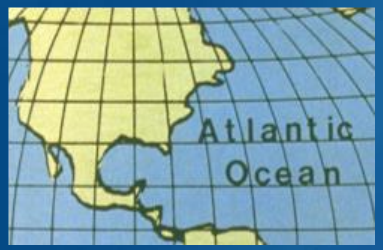
Training dataset

Testing dataset

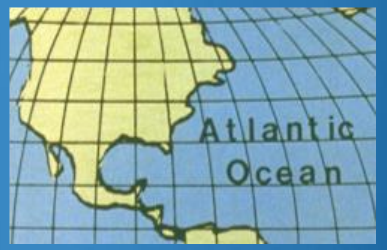
- Equal sample sizes per species
- Divided each dataset into 4 subsets
  - 3 used to train, 1 used to test
  - Entire process repeated 50 times

Bio·Waves Incorporated

Training dataset



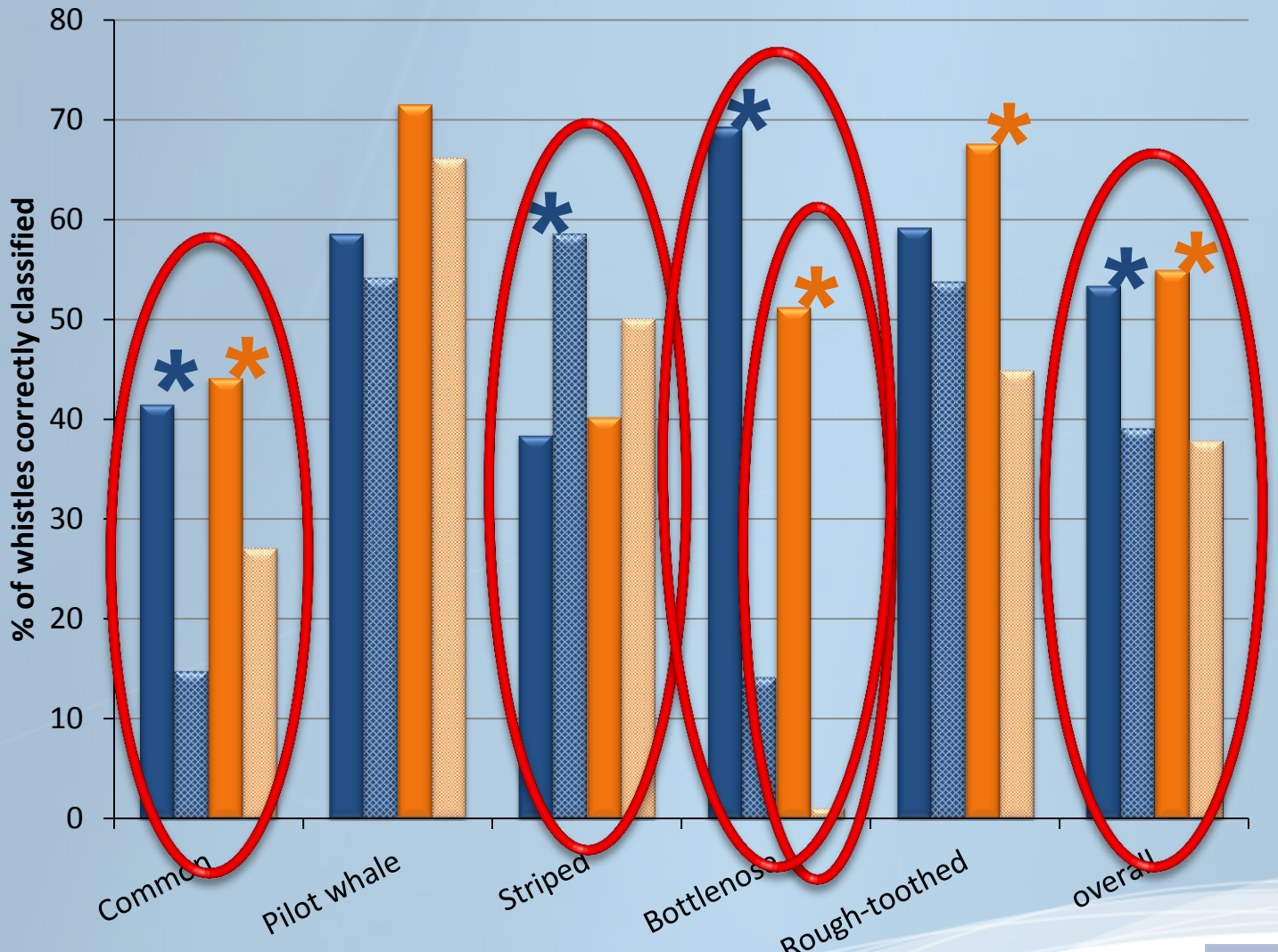
Testing dataset



- Equal sample sizes per species
  - Repeated 50 times

Bio·Waves Incorporated

Classifier results cont'd



Bio·Waves Incorporated

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Frequency Variables

Species	Min	Max	Beg	End	Mean	Median	Quarter	Half	Three quarter
Pilot whale		★	★		★	★	★	★	
Striped dolphin		★							★
Common dolphin	★	★	★	★				★	★
Bottlenose dolphin		★							
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★

Shape Variables

Species	Dura-tion	Mean slope	Mean abs slope	Mean pos slope	Mean neg slope	% up	% down	% flat	# steps	# inflect pts
Pilot whale				★	★		★		★	★
Striped dolphin			★	★				★	★	★
Common dolphin			★			★	★	★	★	★
Bottlenose dolphin	★		★	★	★	★	★	★	★	★
Rough-toothed dolphin	★	★	★	★	★	★	★	★	★	★

Classifier performance

- The presence of geographic variation does not always mean classifier performance will be negatively affected
- For some species and locations, classifiers trained using data from another study area will perform better
  - Ex. Striped dolphins

Bio·Waves Incorporated

What's going on with striped dolphins?

- Patterns of misclassification
  - Striped dolphins misclassified significantly less frequently as rough-toothed dolphins when the 'other' classifier was used
- Atlantic striped and rough-toothed whistles are more similar in the same study areas than in different study areas
  - So the 'other' study area classifier is more effective for classifying striped dolphins

Bio•Waves Incorporated

Summary

- Geographic variation evident for all five species
  - More geographic variation for some species than others
- Geographic variation affected classifier performance
  - Not always a good predictor of how classifier would be affected
- In general, classifier performed better when trained and tested with data from same location
  - But not always

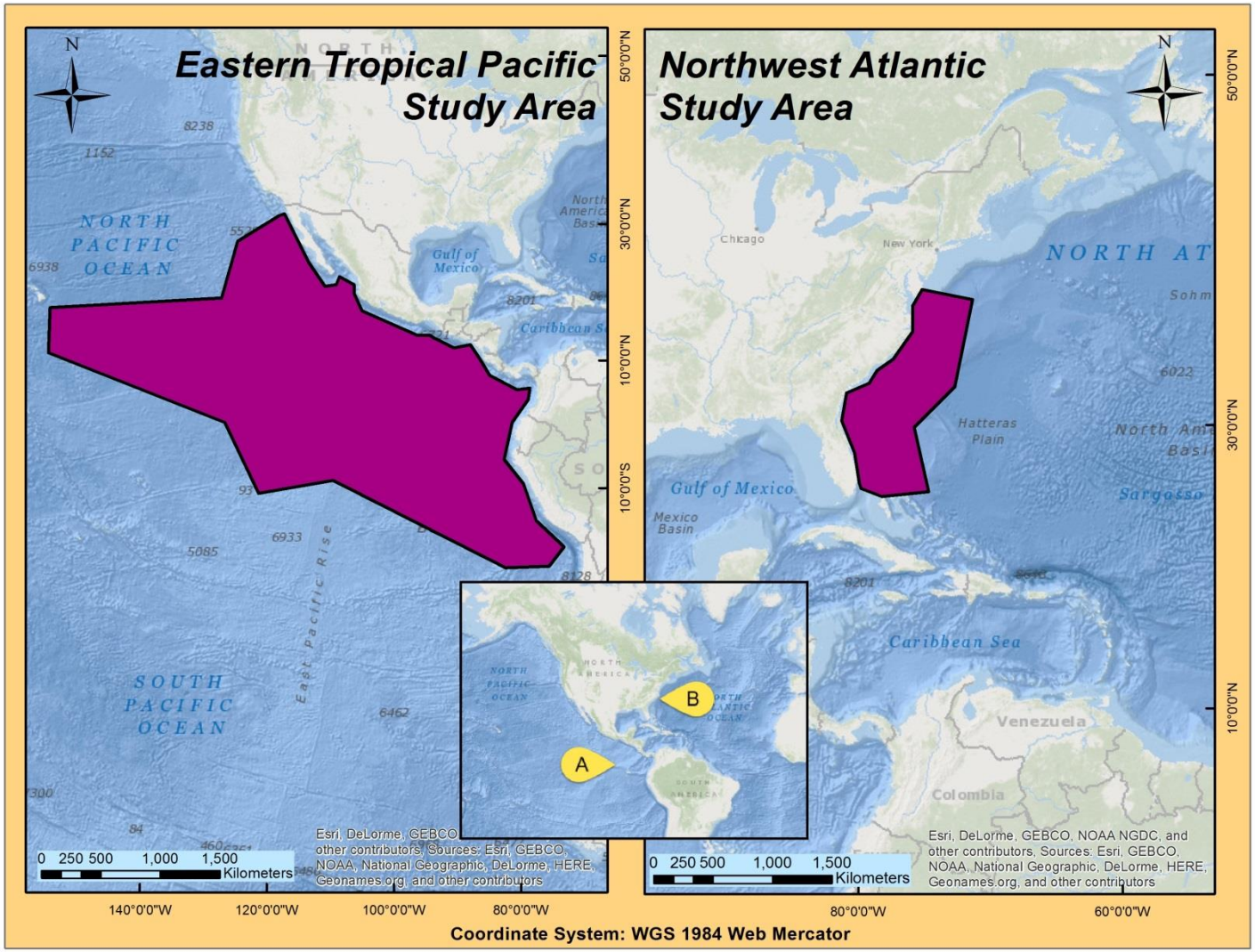
Bio·Waves Incorporated

Conclusions

- In general, classifiers should be trained using data collected in the study area where the classifiers will be used
- Different classifiers should be tested and training data chosen with study goals in mind
  - Ex. Striped dolphins

Bio·Waves Incorporated

Future Directions



BioWaves Incorporated

Future Directions

- On what geographic scale does variation in whistle structure occur?
- Over what geographic scale can classifiers be successfully used?
- How does geographic variation in the whistles of one species affect classification of another?

Bio•Waves Incorporated

Acknowledgements

- Acousticians, visual observers, officers and crew on all surveys
- Tom Norris, Liz Ferguson, Robyn Walker, Shannon Coates – Bio-Waves, Inc.
- Alexis Rudd, Theresa Weber, Stephanie Grassia
- Office of Naval Research
- Living Marine Resources Program
- NAVFAC Atlantic and Pacific


Thank you!