Whistle while you work
Automating the detection of dolphin whistles in complex auditory scenes

Marie A. Roch

Computer Science

San Diego State University

Scripps Institution of Oceanography UCSD

Partners in crime

- Echolocation clicks
  H. Klinck, S. Baumann-Pickering, D.K. Mellinger, S. Qui, M.S. Soldevilla, and J.A. Hildebrand
- Whistle detection
  T.S. Brandes, B. Patel, Y. Barkley, S. Baumann-Pickering, and M.S. Soldevilla
- Whistle classification
  H. Bassett, K. Frasier*, and E. E. Henderson

* Classification slides from Kait Frasier's DCLMMPA 2011 presentation

Passive acoustic monitoring

Motivation

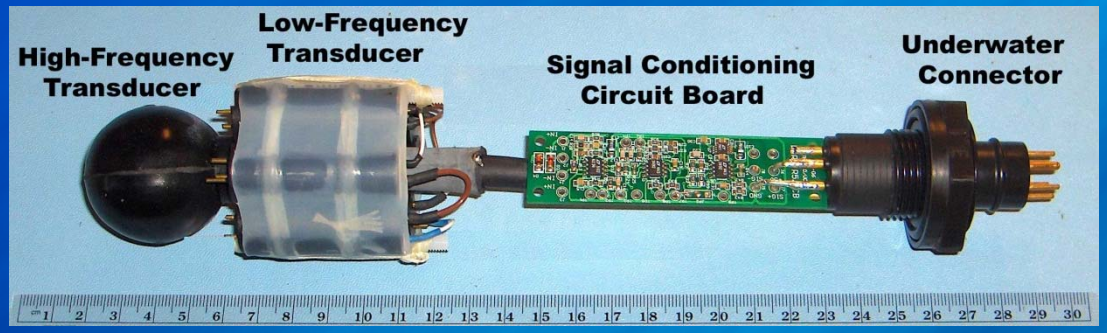
- Visible observation is difficult & expensive
- Most cetaceans are vocal

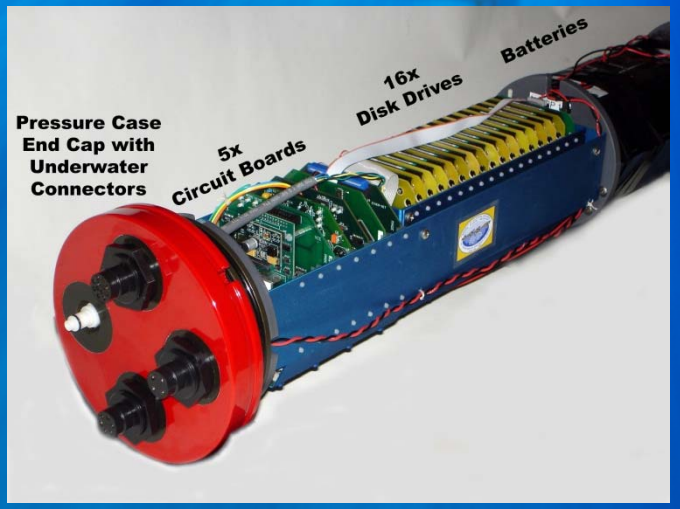
Goals

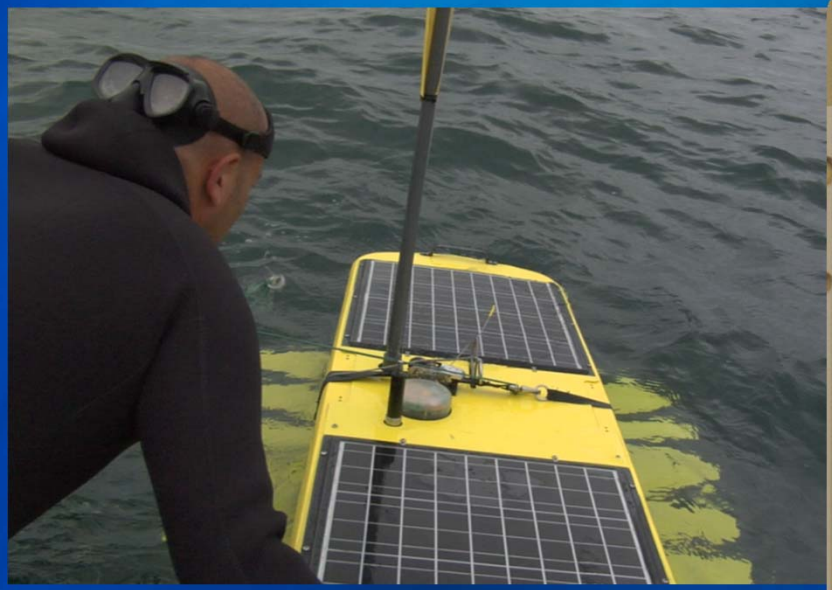
- Understanding eco-systems
- Science-based policy decisions
- Mitigation

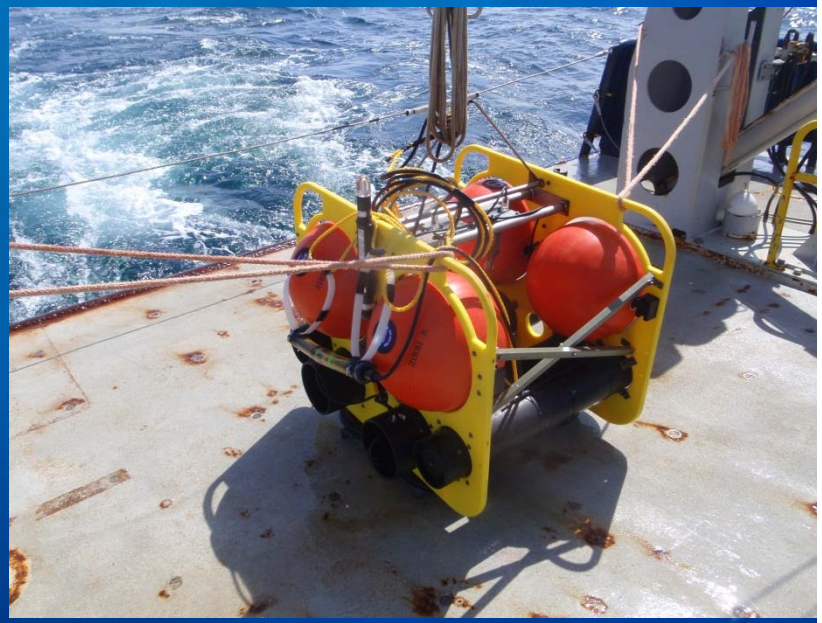
photo: Amanda Cummins

Passive acoustic monitoring









photos: SIO (Sean Wiggins, Josh Jones)


Data, data, everywhere...
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Image IBCAO
Image © 2011 TerraMetrics
Image © 2011 DigitalGlobe
photo: Amanda Cummins

Vocal production of odontocetes

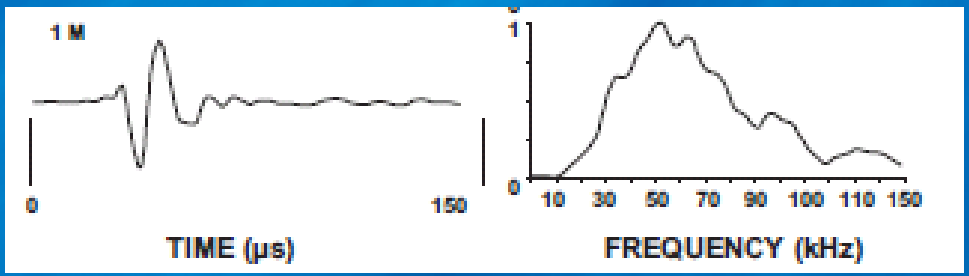
- Echolocation clicks
  - Navigation
  - Foraging
  - Communication

- Whistles
  - Communication

photo: Amanda Cummins

Echolocation clicks

- Short, broad band signals ≈ 100-200 µs



- Production
  - Highly specialized anatomy
  - Good candidate for deriving information about the animal or species

Echolocation click challenges

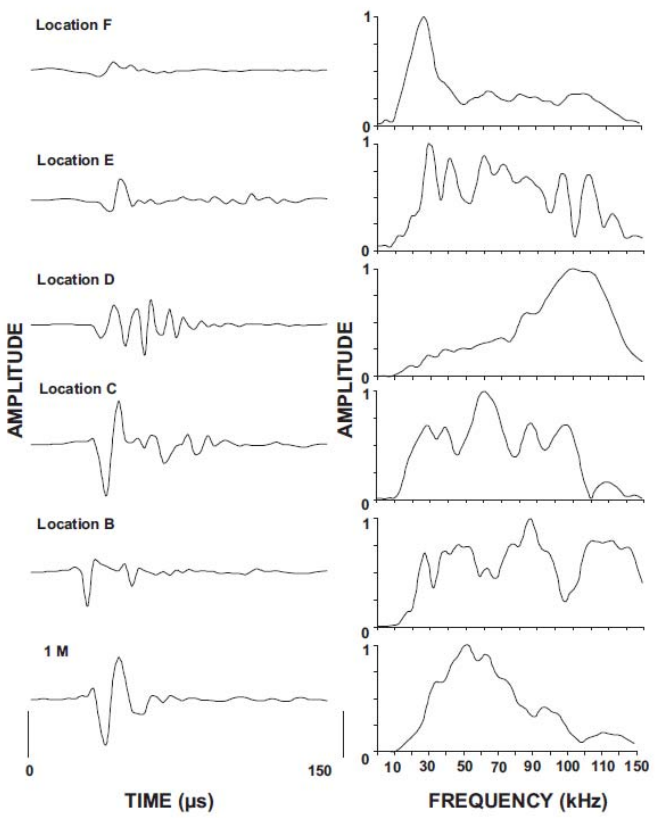
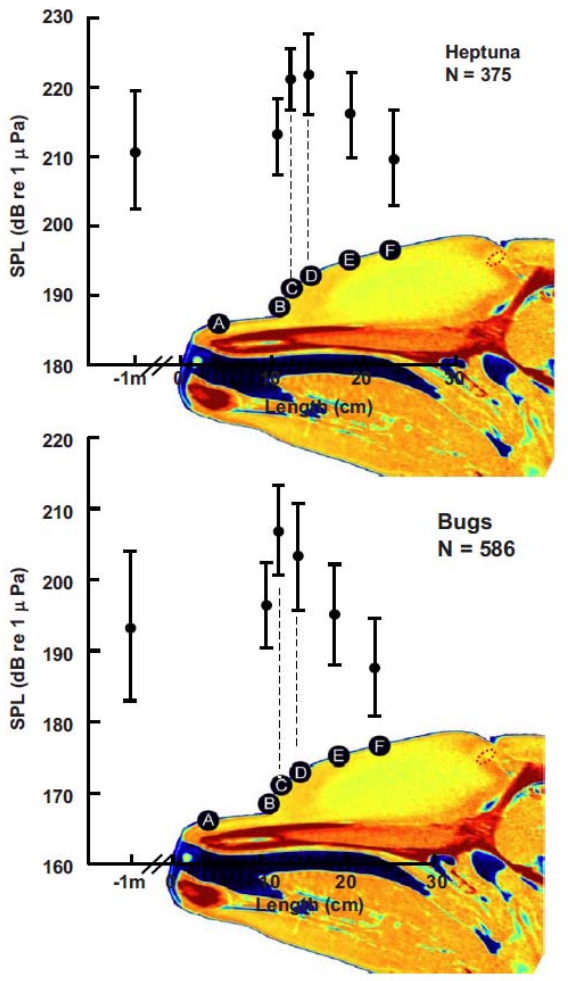


FIG. 3. Examples of a single echolocation click as measured by the hydrophones in the linear-array geometry; the waveforms are shown on the left and spectra are shown on the right.



Au, W.W.L., D.S. Houser, J.J. Finneran, W.-J. Lee, L.A. Talmadge, P.W. Moore (**2010) The acoustic field on the forehead of echolocating Atlantic bottlenose dolphins (*Tursiops truncatus*). J. Acoust. Soc. Am., 128(3), 1426-1434, doi: 10.1121/1.3372643.**

Species identification

6-species classification task

Short-beaked common

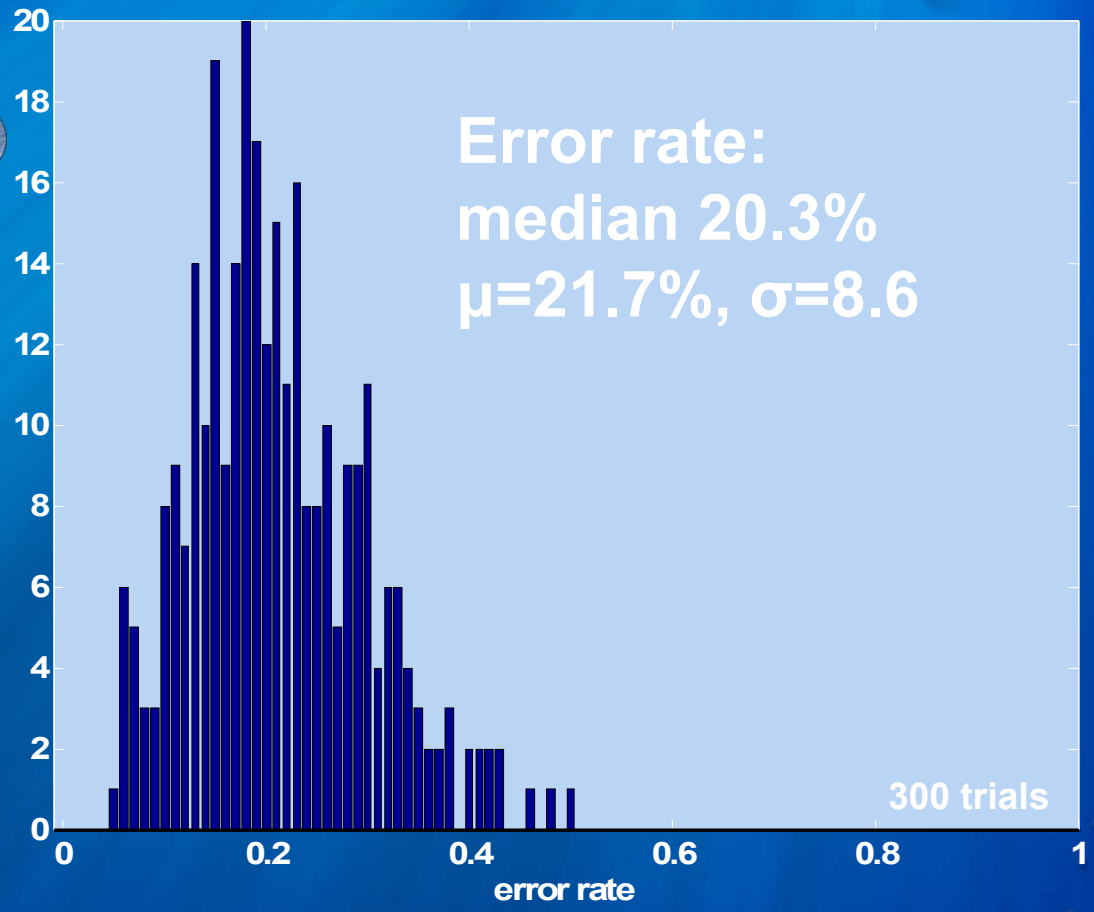
Long-beaked common

Bottlenose

Pacific white-sided

Cuvier's beaked

Risso's



Roch, et al. (2011) *J. Acous. Soc. Am., 129(1), 467-475.*

photos: Amanda Cummins (Dc, Dd), Dave Weller (Zc) and unattributed

Whistles

- Call description
- Individual identification
- Species identification
- Quantification of social calls
- Localization

Whistle challenges

- Do we need to isolate individual whistles?
- Less stereotyped
- Animals tend to whistle less often
- Some whistles may be associated with:
  - individuals
  - behavioral contexts

photo Amanda Cummins

freq (kHz)

time (s)

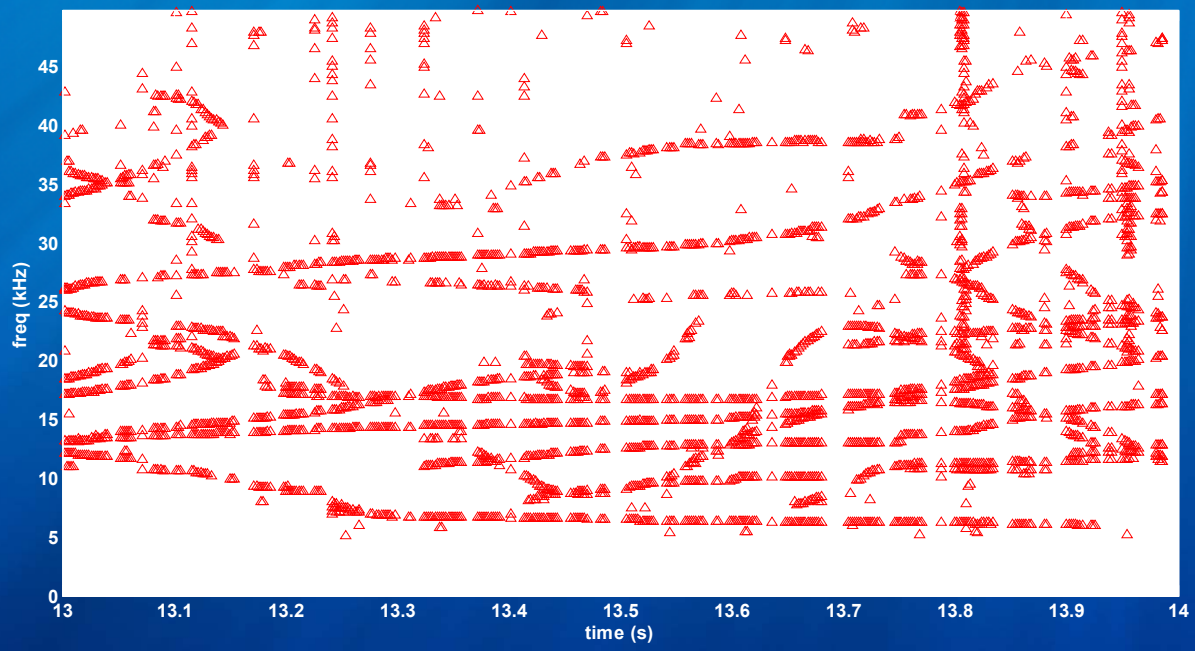
Spectrum Level [dB re counts ^{2}/Hz]

advance 2 ms, length 8 ms (125 Hz)

Graph search: Connect the dots

Edges

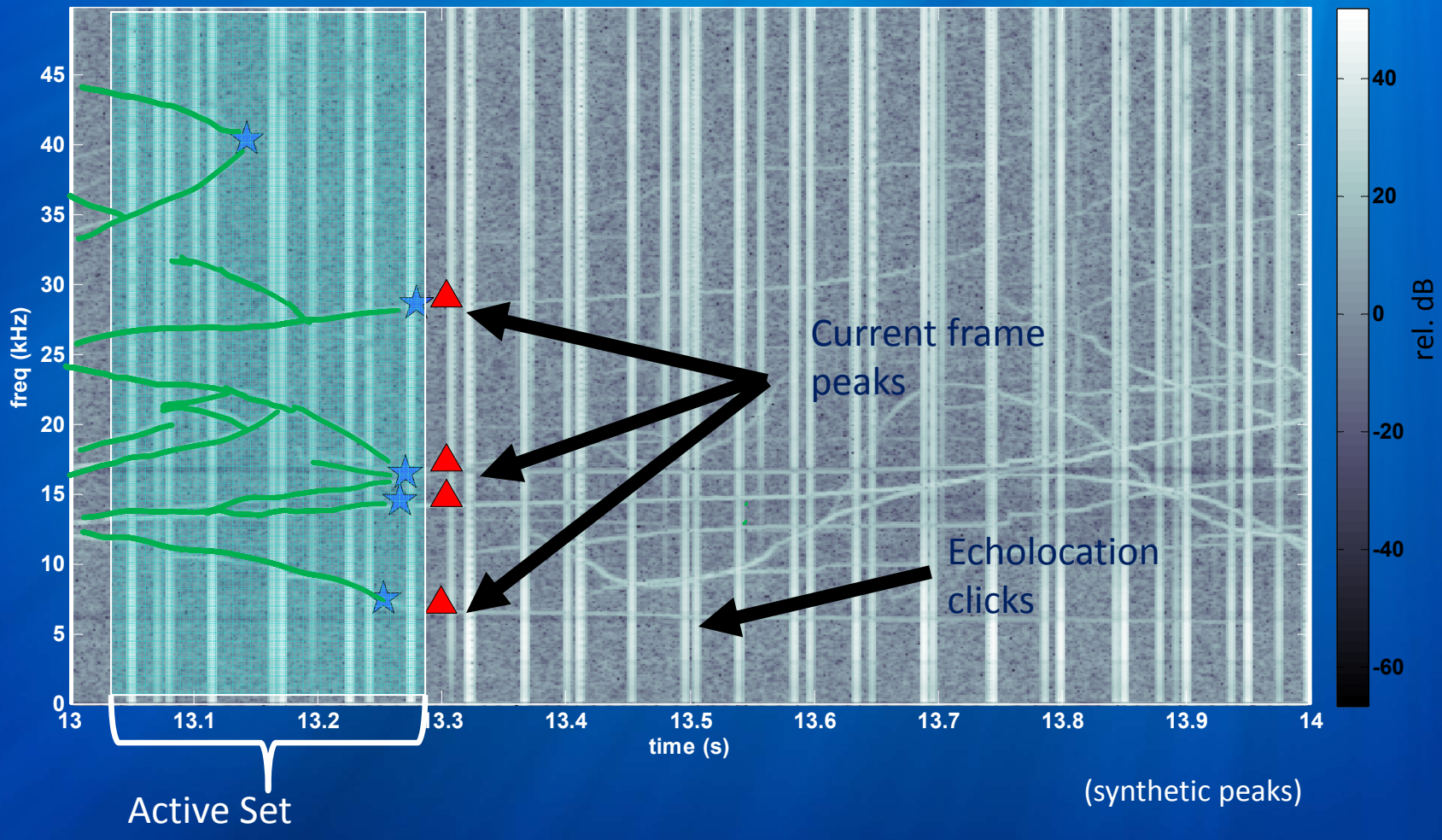
- define paths between peaks
- tonal calls comprised of coherent paths



Coherent connections

- Slope/Shape
  - polynomial fit [Mellinger - Ishmael]
  - adaptive polynomial fit
- Morphological operators (Mallawaarachchi et al. 2008)
- Momentum (Mallawaarachchi et al. 2008)
- Phase

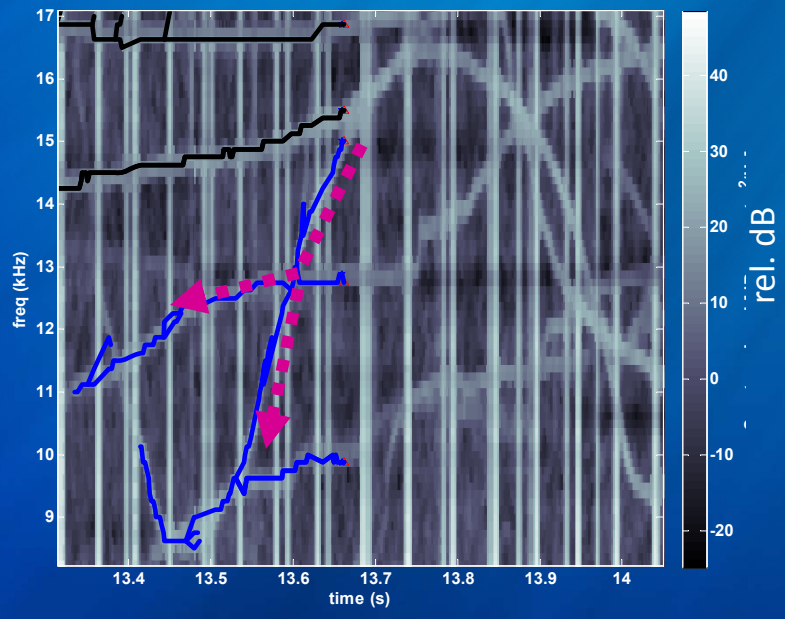
Silbido – contour detector



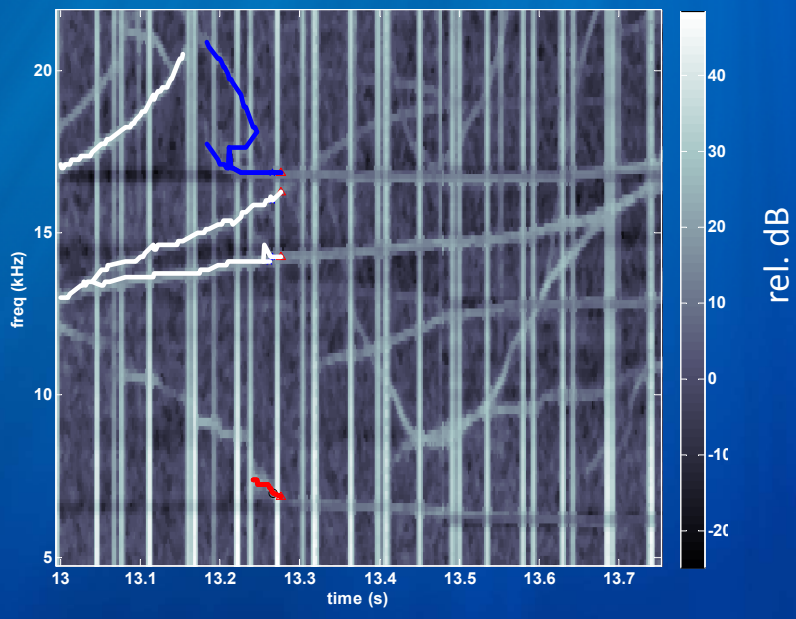
Graph extension

Connecting to valid candidate peaks

- Multipath adaptive polynomial prediction



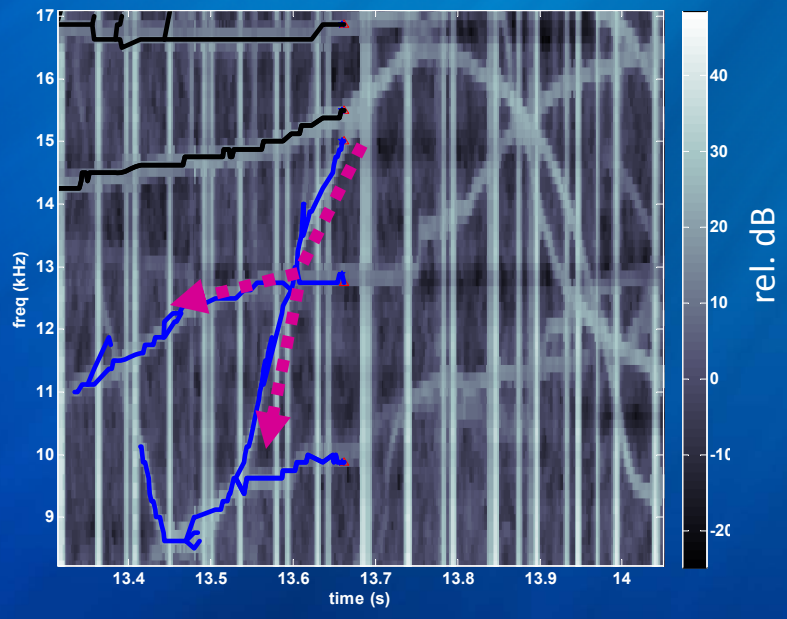
- Track/merge graphs



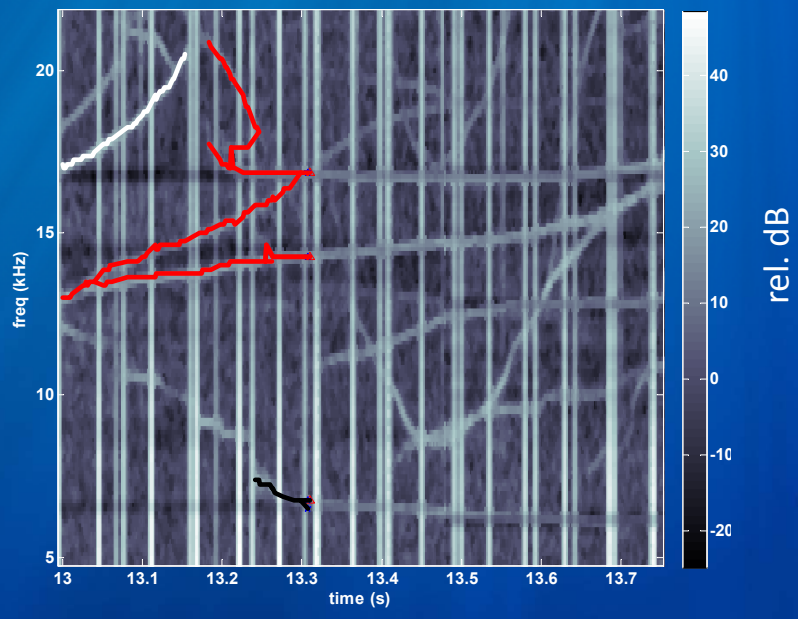
Graph extension

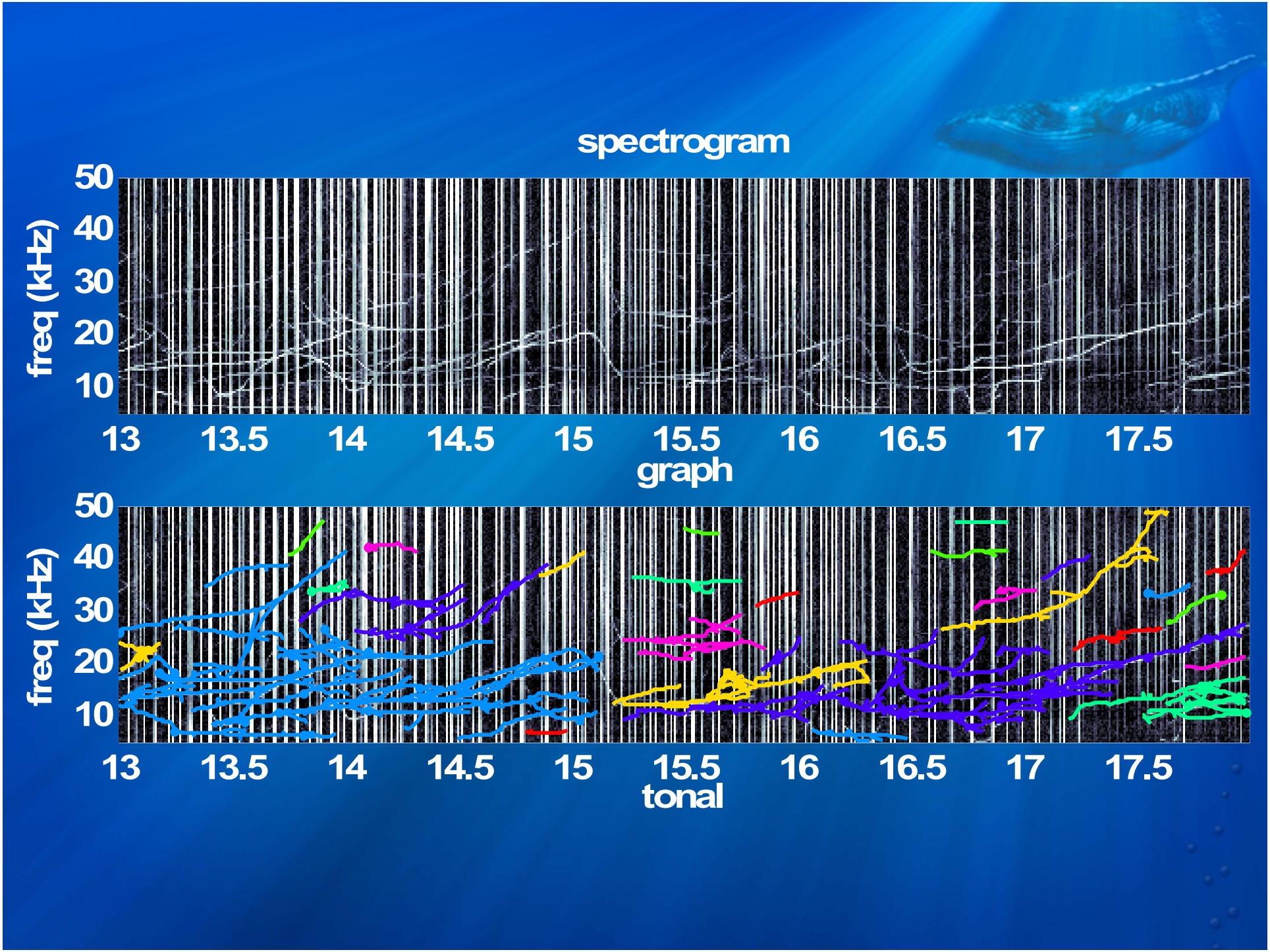
Connecting to valid candidate peaks

- Multipath adaptive polynomial prediction

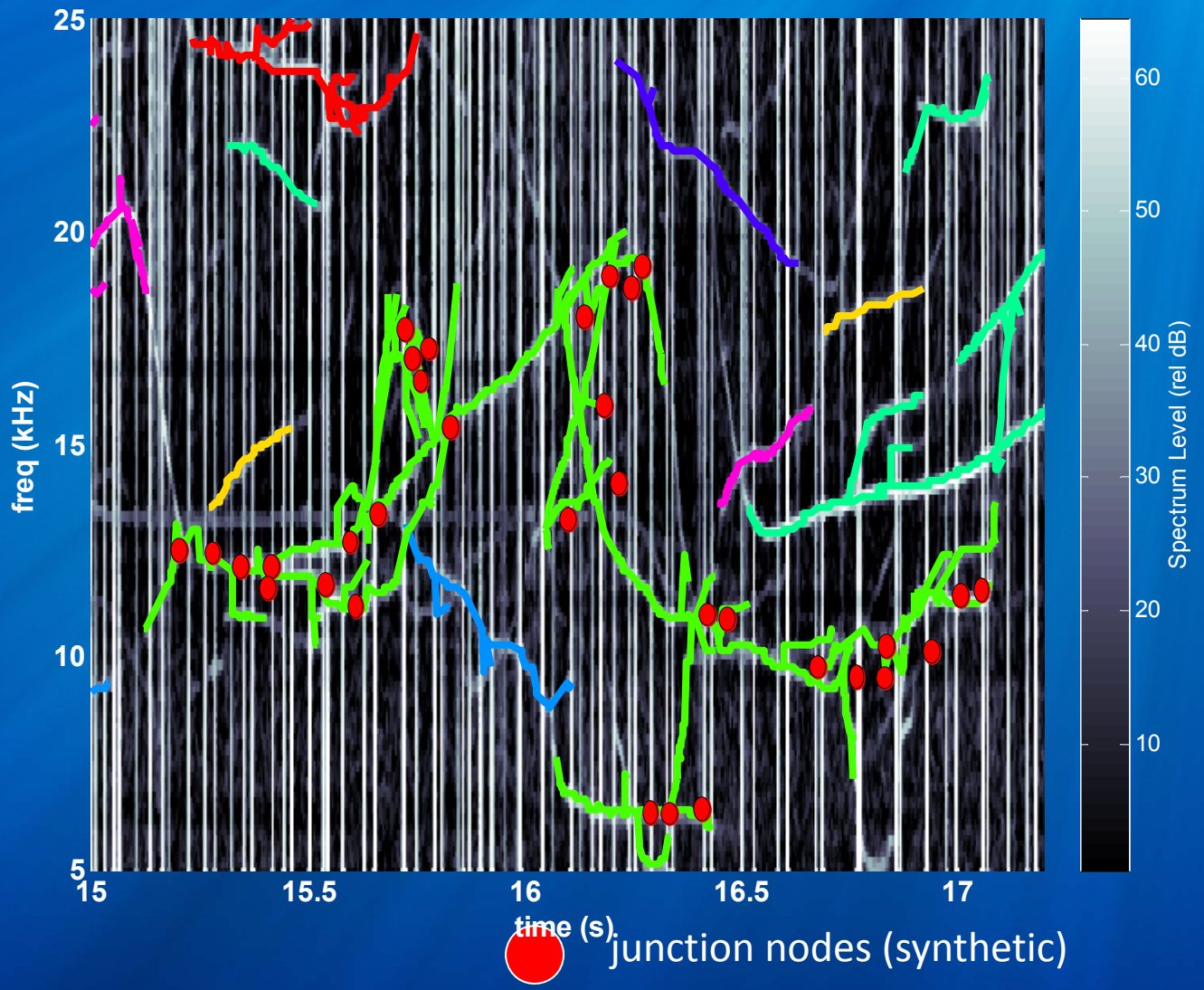


- Track/merge graphs

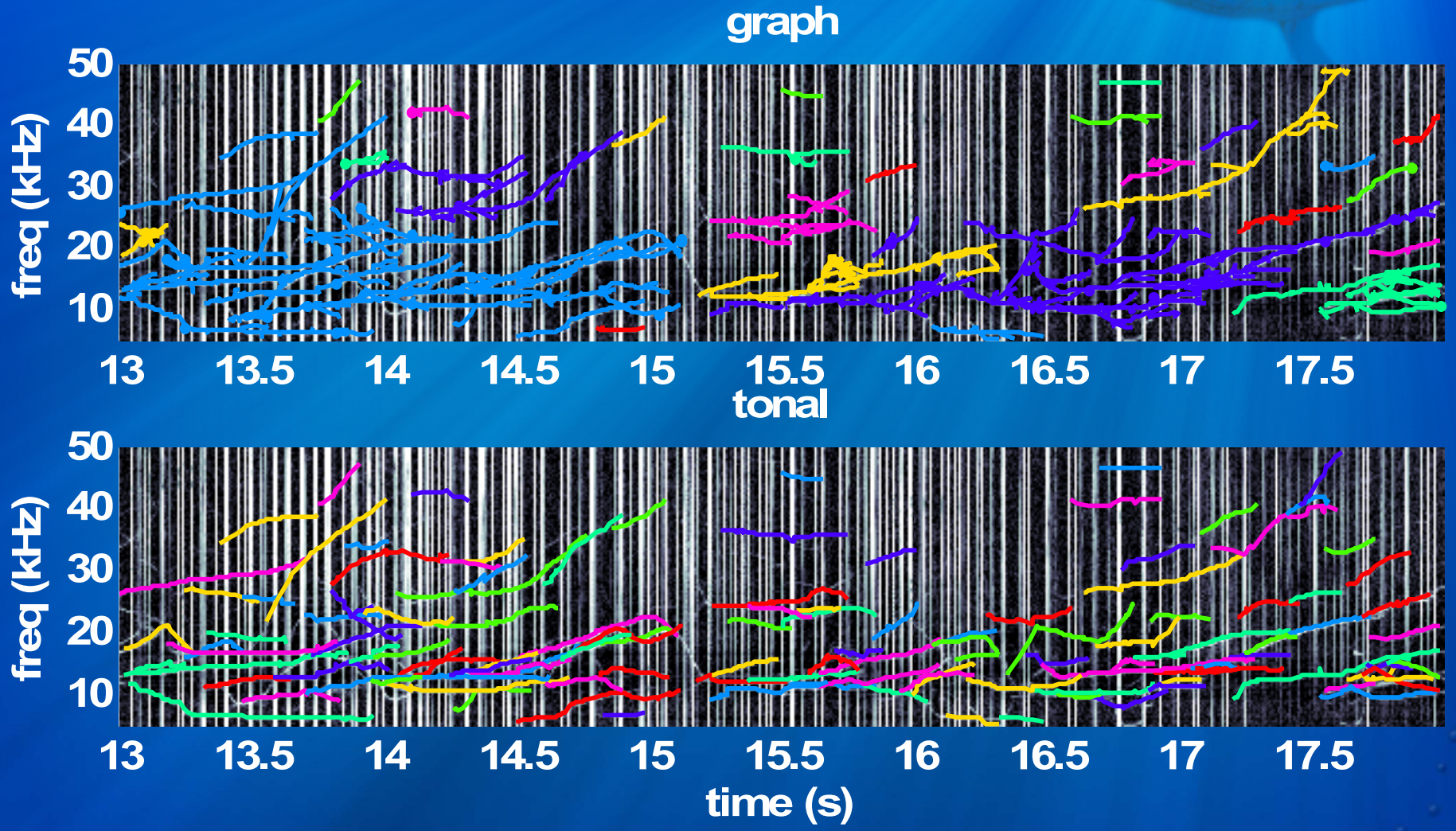



spectrogram
freq (kHz)
50
40
30
20
10
13
13.5
14
14.5
15
15.5
16
16.5
17
17.5
graph
freq (kHz)
50
40
30
20
10
13
13.5
14
14.5
15
15.5
16
16.5
17
17.5
tonal

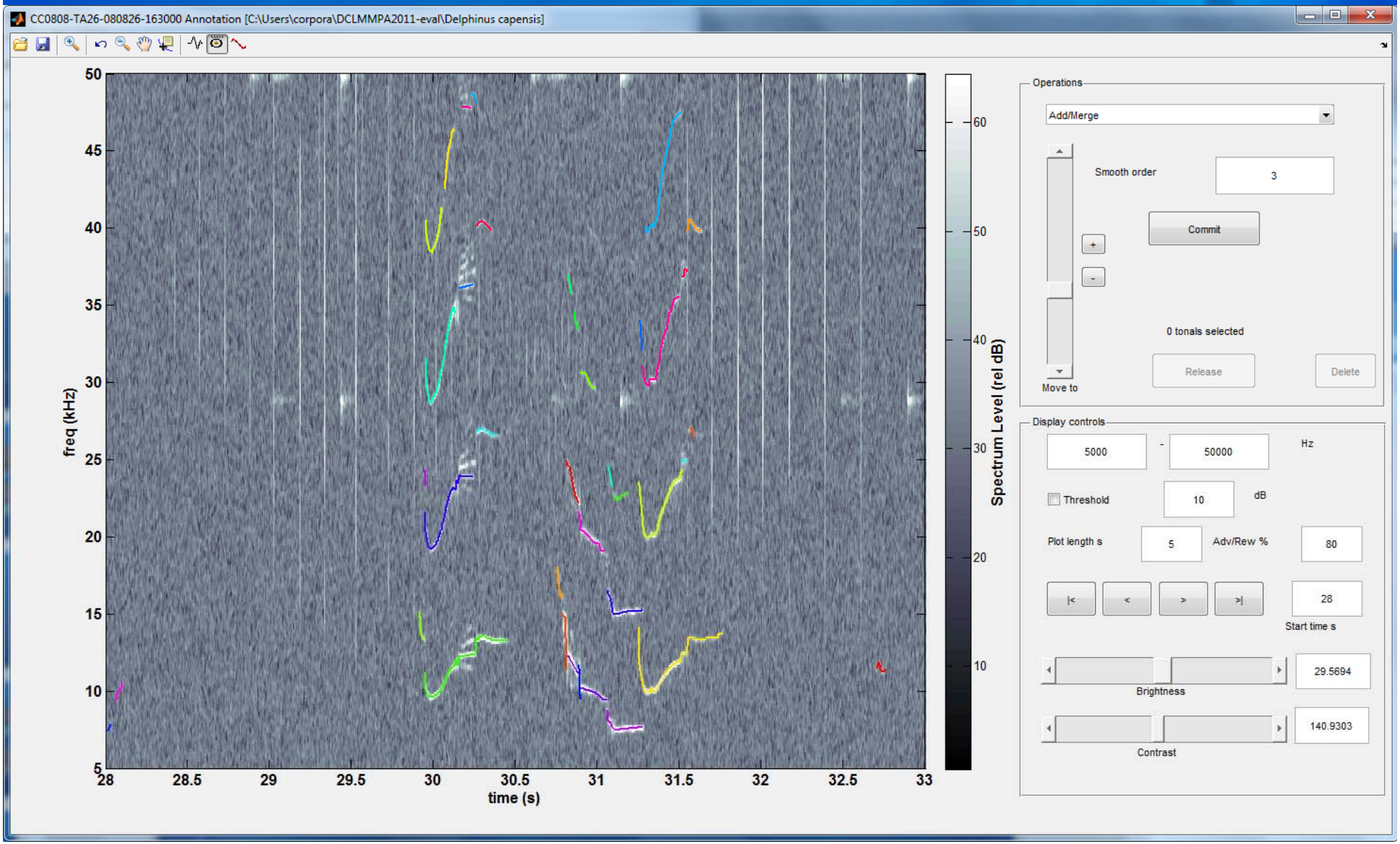
Disambiguation



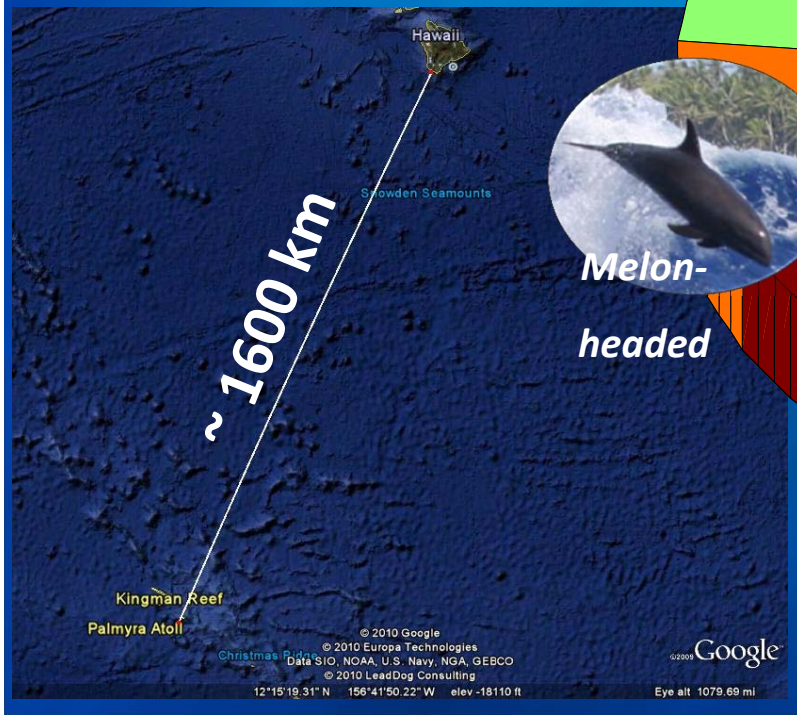
Whistles

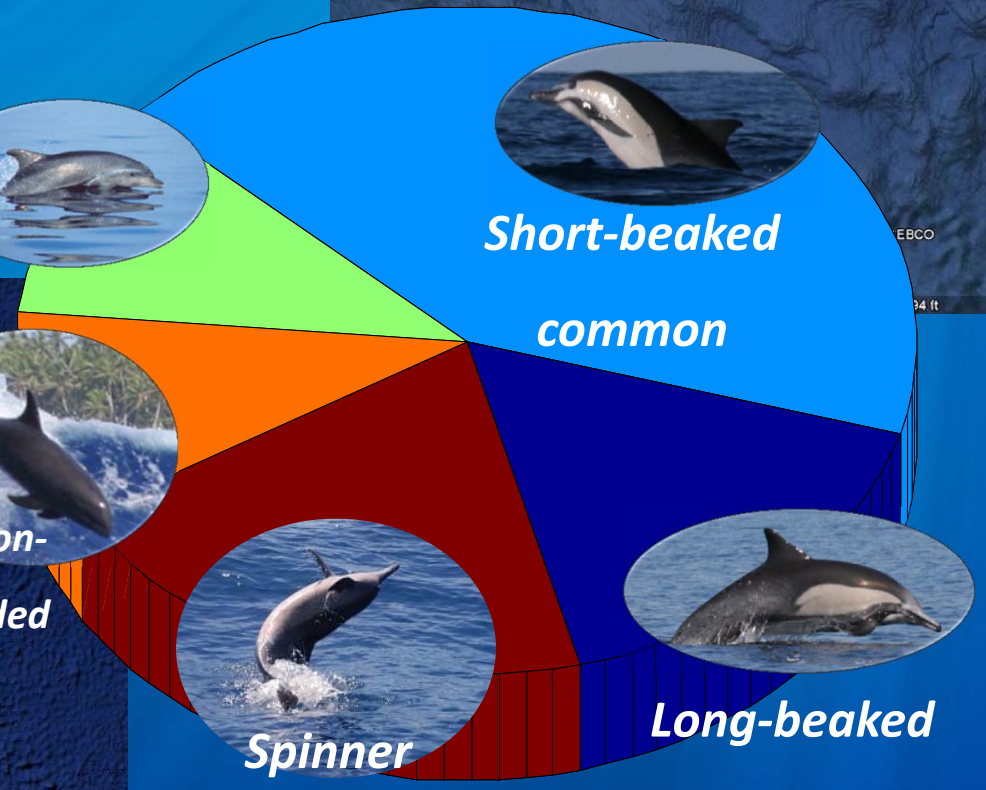

graph
freq (kHz)
50
40
30
20
10
13
13.5
14
14.5
15
15.5
16
16.5
17
17.5
tonal
freq (kHz)
50
40
30
20
10
13
13.5
14
14.5
15
15.5
16
16.5
17
17.5
time (s)

Ground truth tools



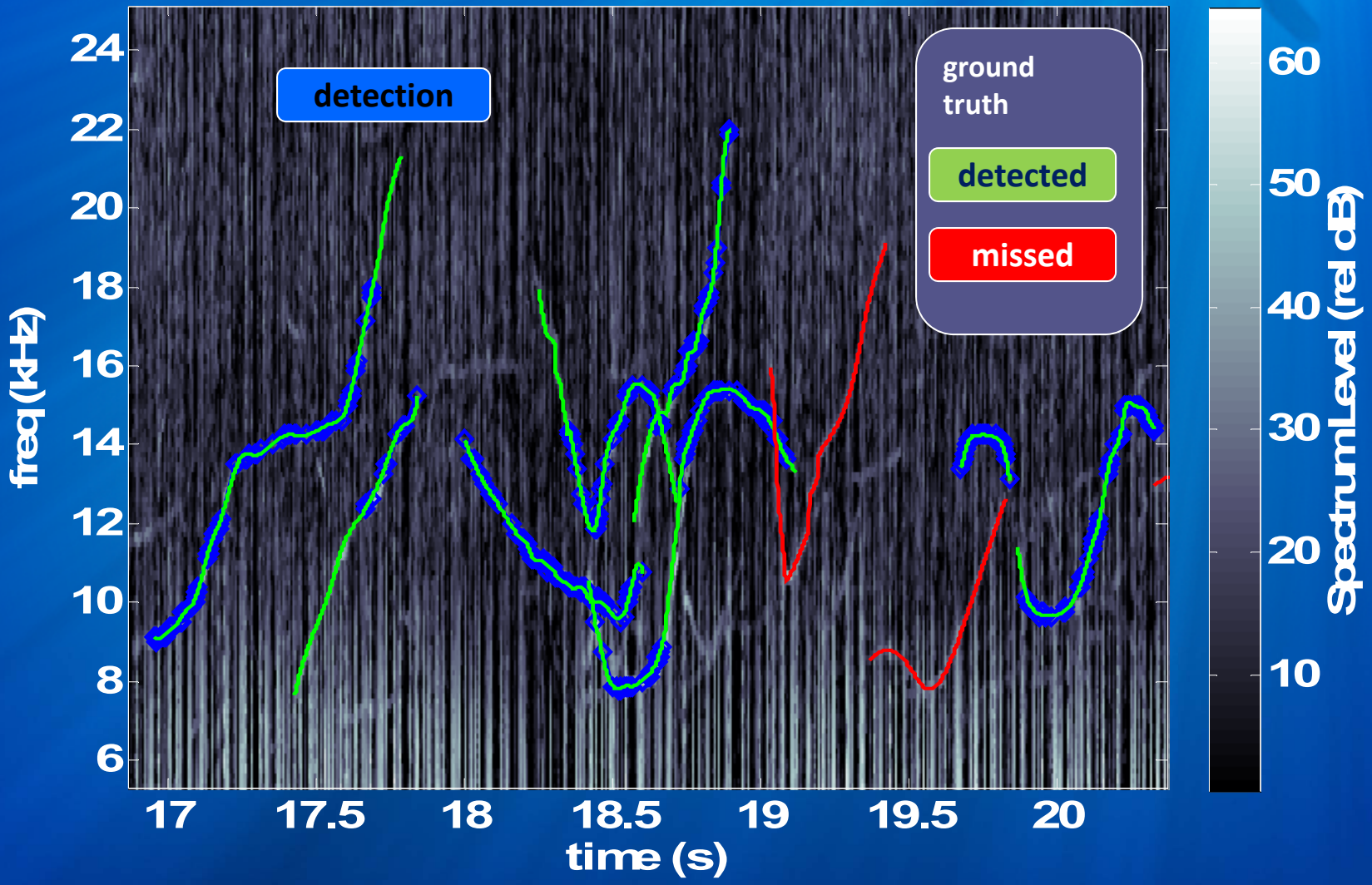
Corpus



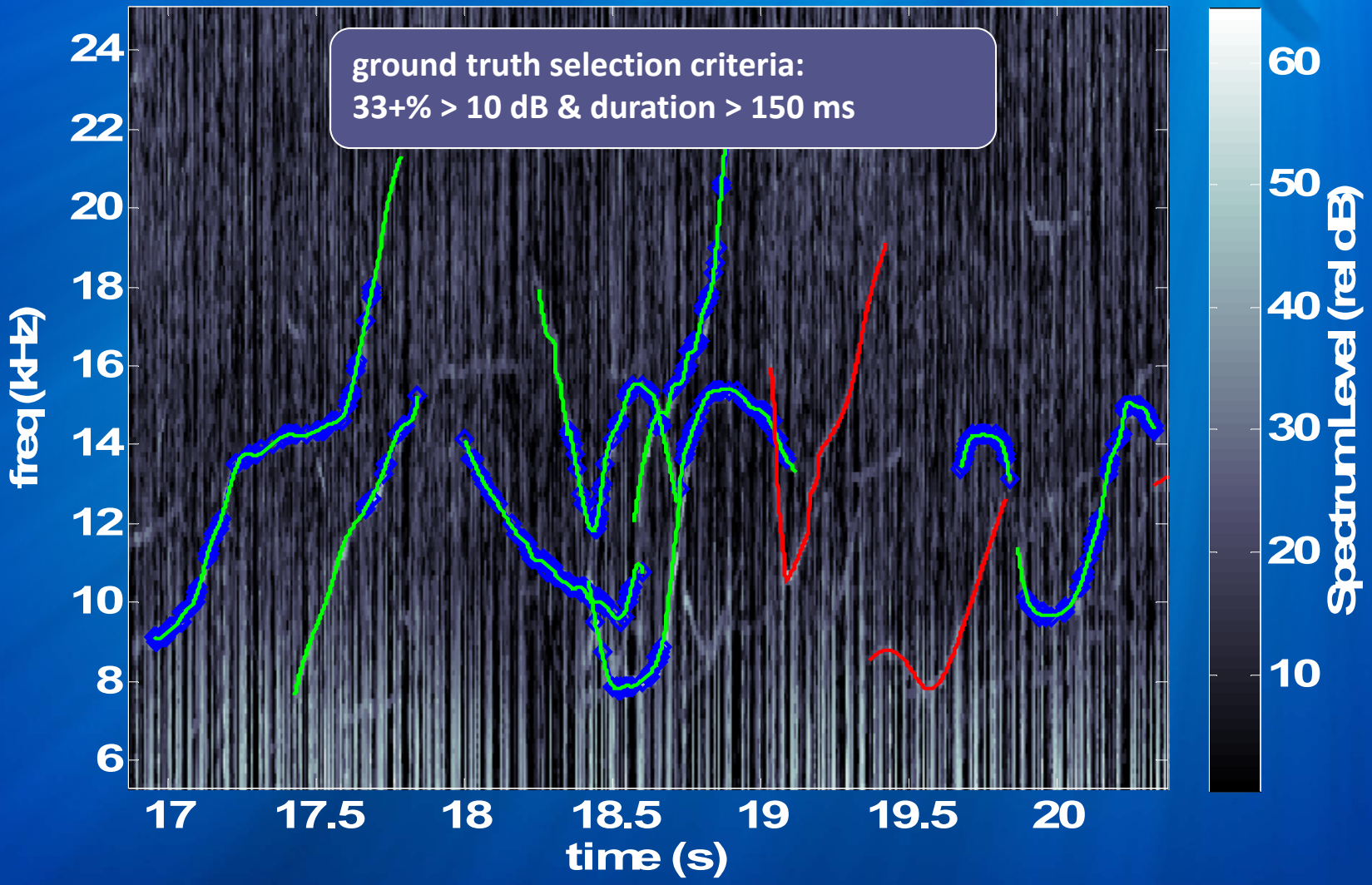


images: Baumann-Pickering, Cummins, unattributed

Short-beaked common

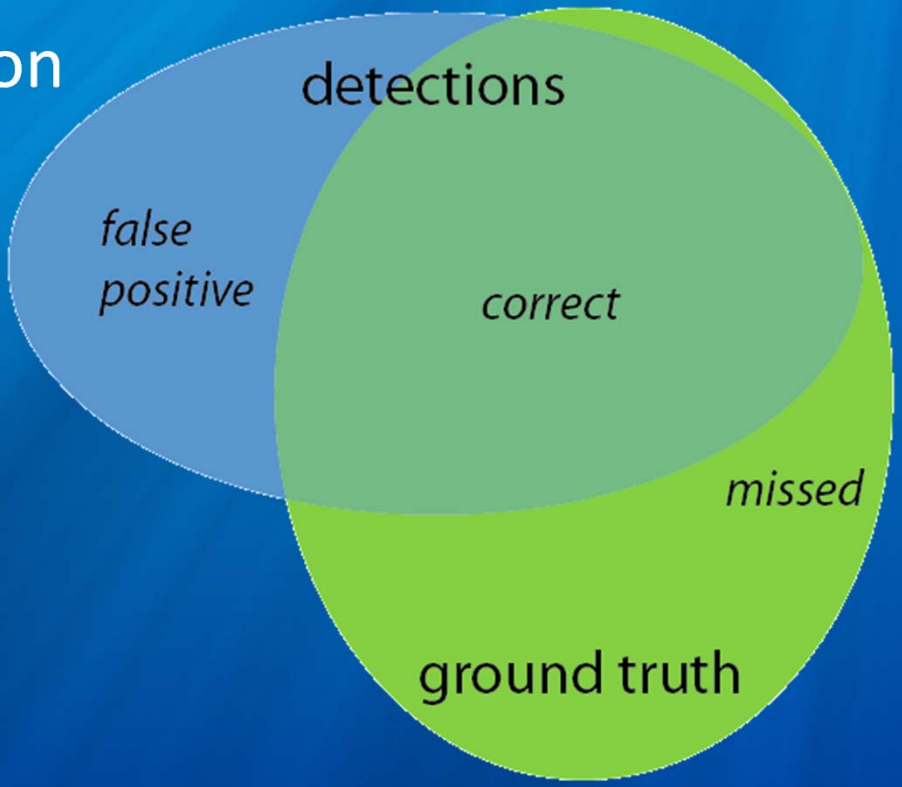


Short-beaked common



Recall and Precision

- Compared to analyst detected whistles
  - 80% precision
  - 77% recall



Long-beaked common

87%

94%

20%

8%

80%

100%

Short-beaked common

84%

67%

91%

97%

90%

79%

3,372 tonals meeting duration & SNR criteria

Recall

Precision

Bottlenose

84%

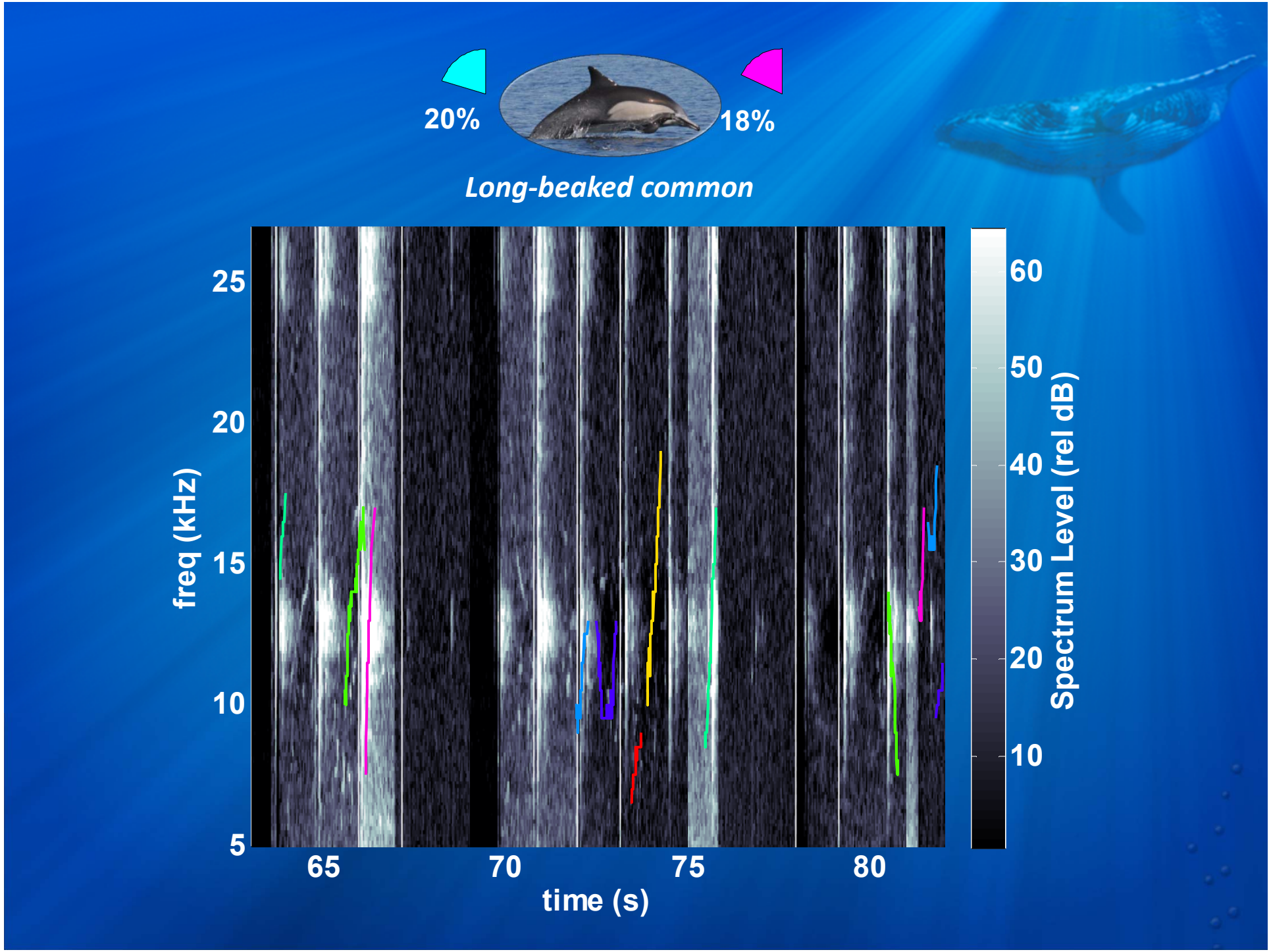
68%

85%

100%

80%

90%


20%
18%
Long-beaked common
freq (kHz)
25
20
15
10
5
65
70
75
80
time (s)
Spectrum Level (rel dB)
60
50
40
30
20
10

Recall

Precision

88%

88%

71%

81%

18%

83%

87%

64%

Spinner

Melon-headed

82%

83%

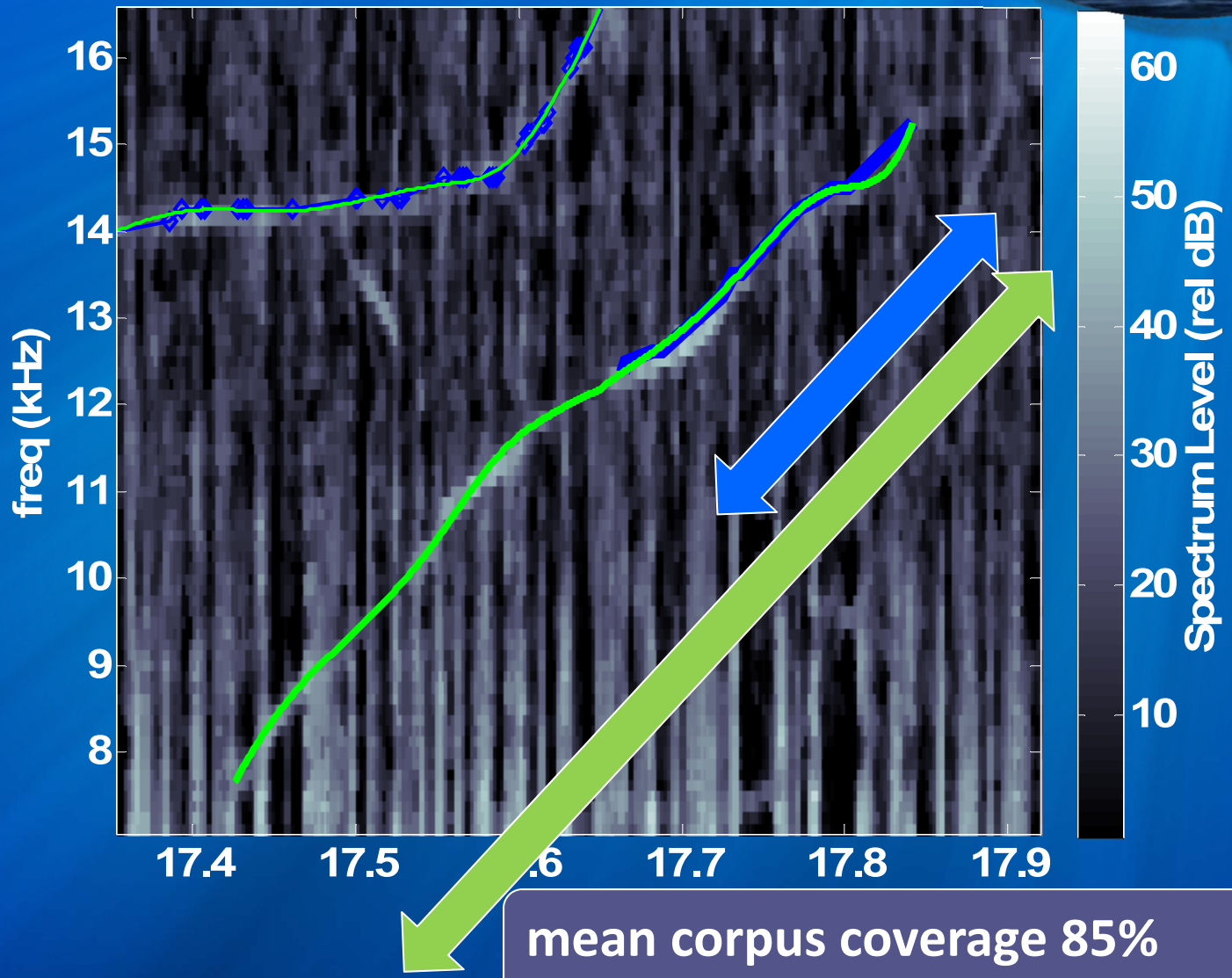
73%

88%

80%

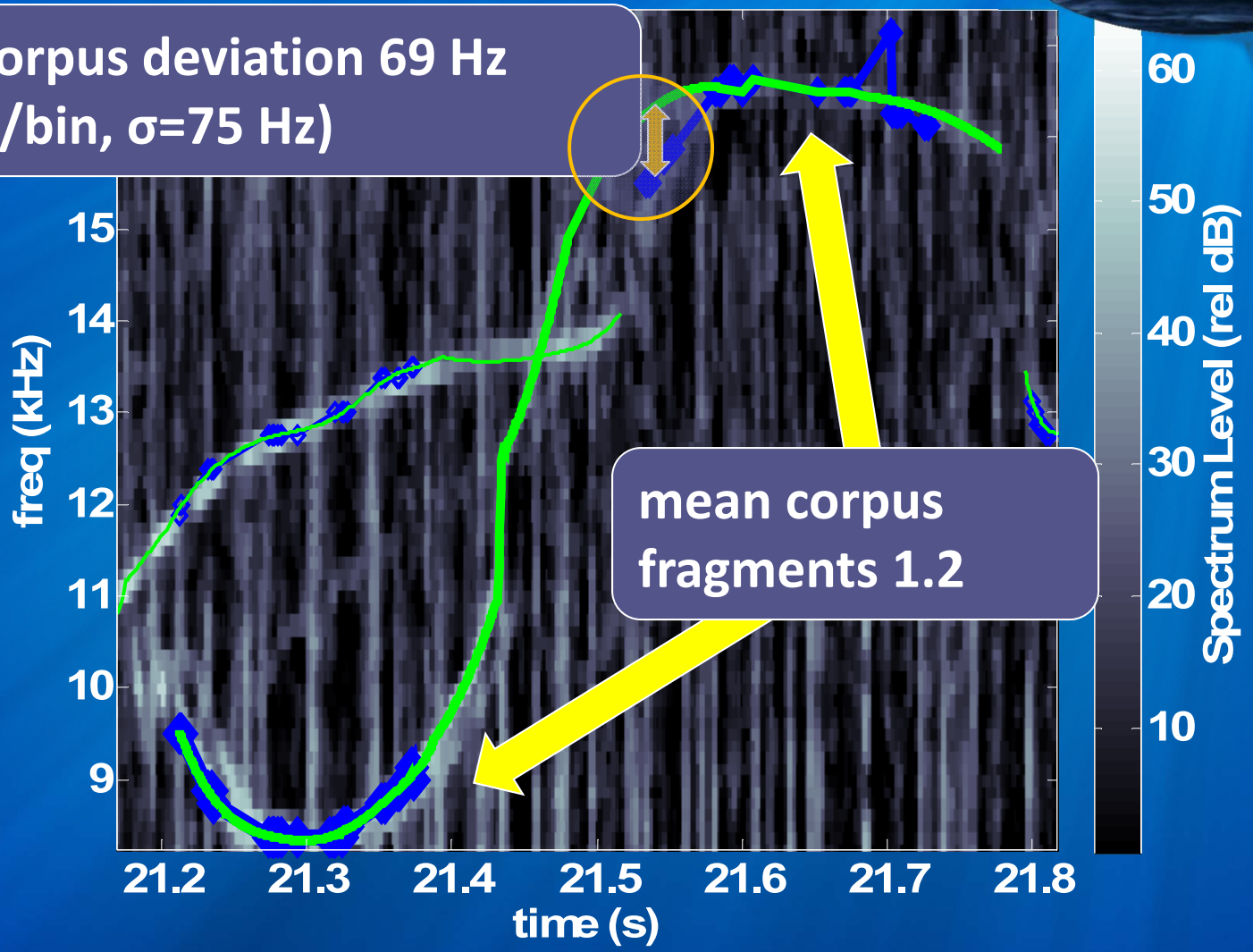
77%

Short-beaked common

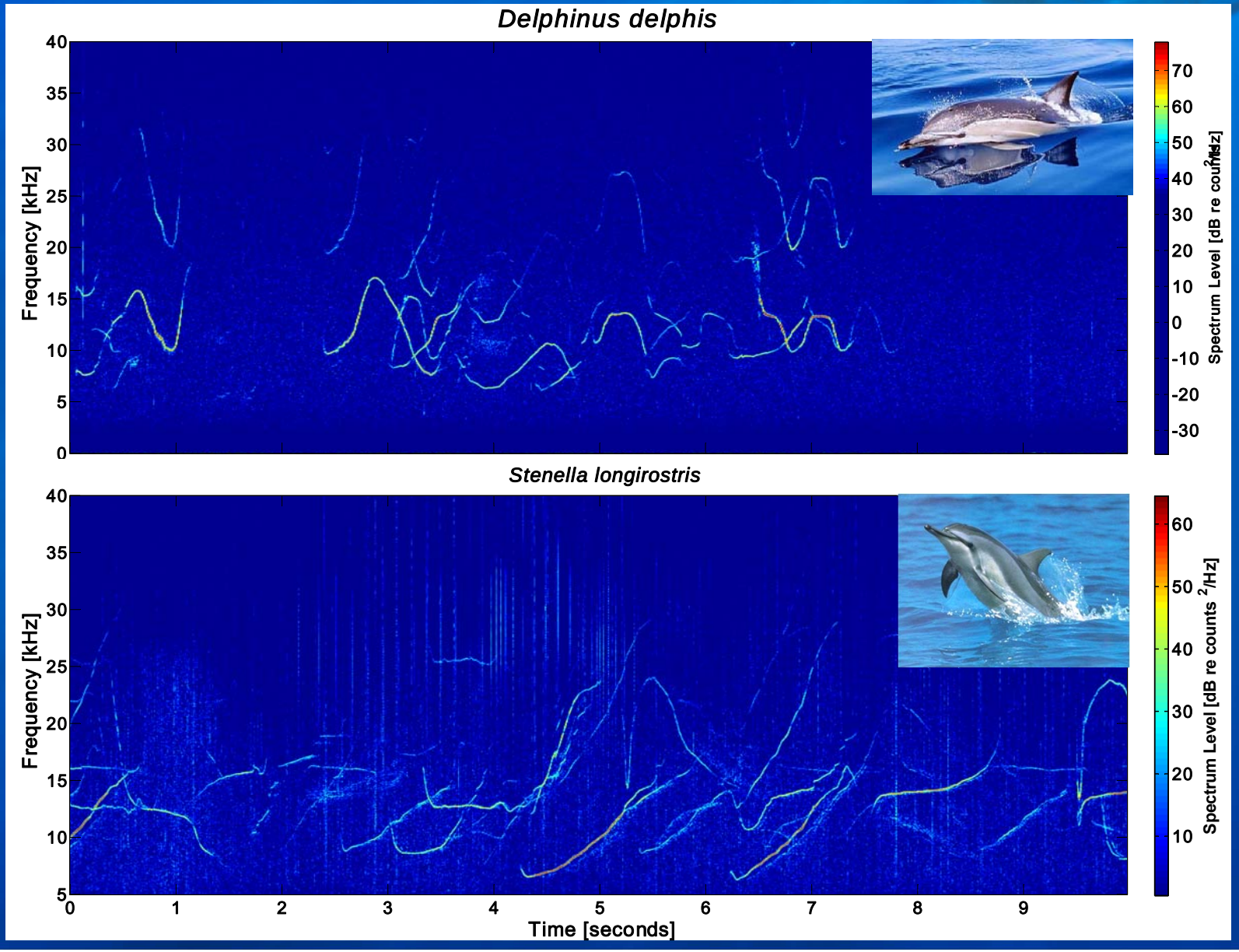


mean corpus coverage 85%

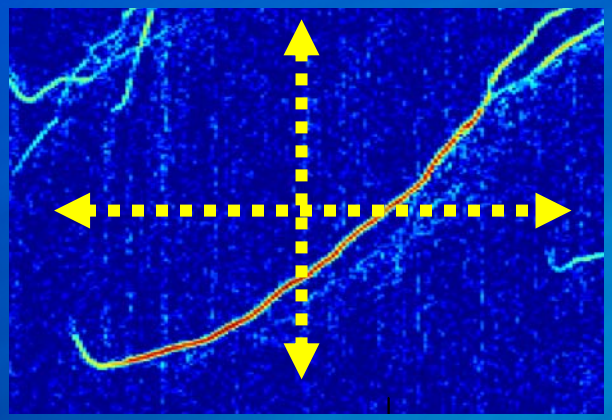
Short-beaked common

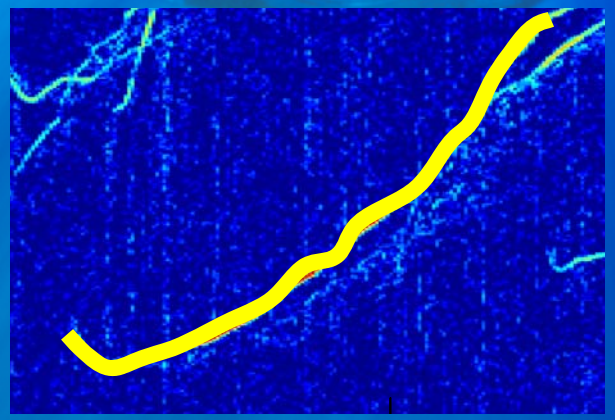


What's in a whistle?

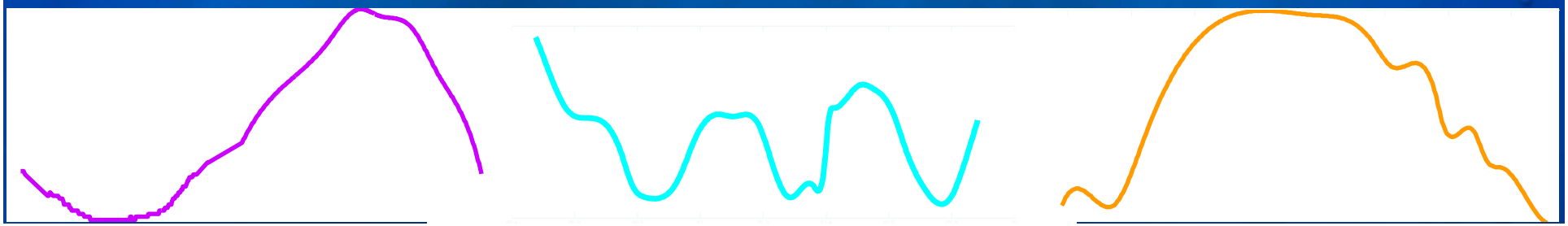


Mimicking trained analysts

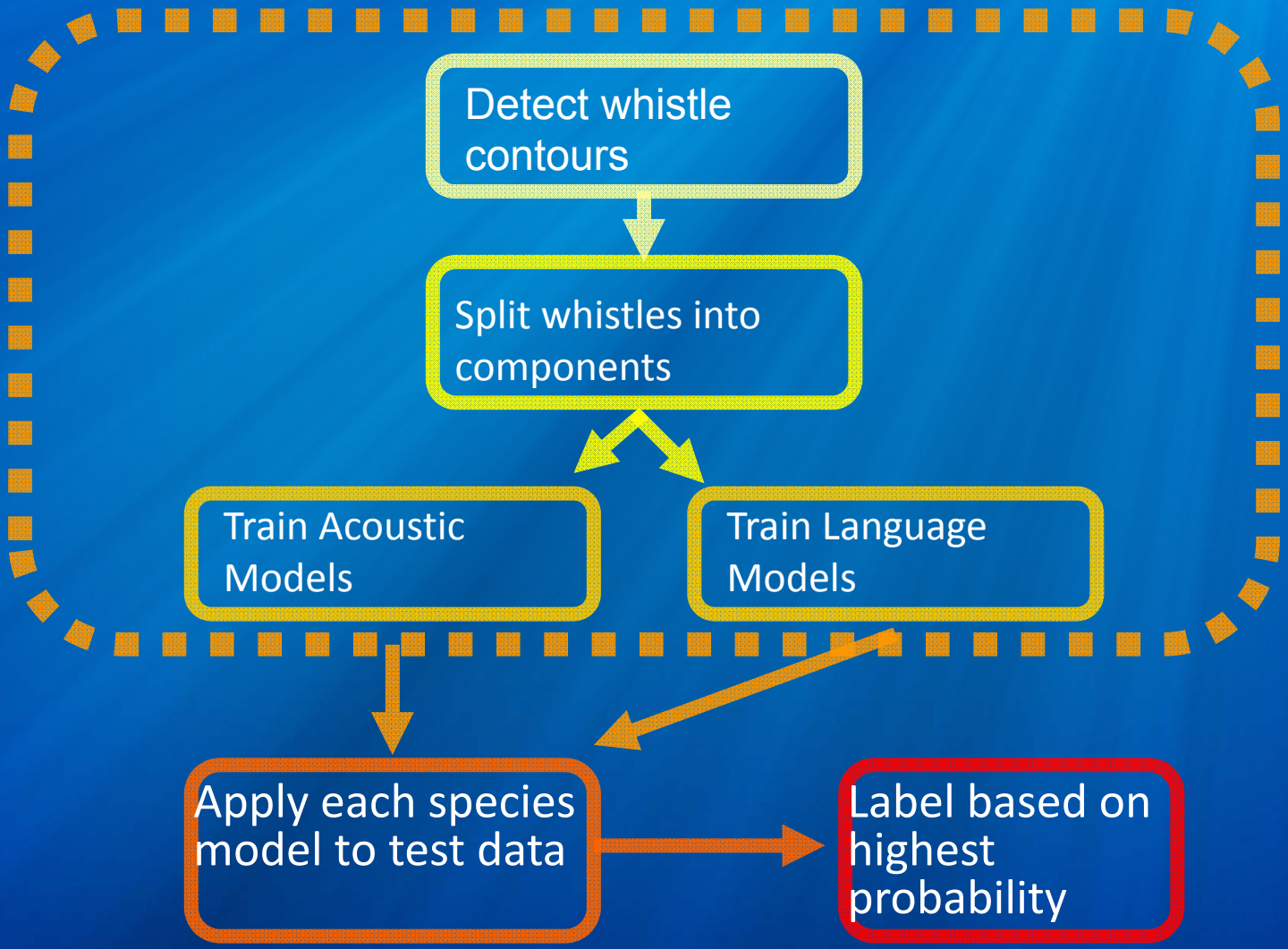




Hypothesis: Modeling shape might capture cues used by analysts to determine species.

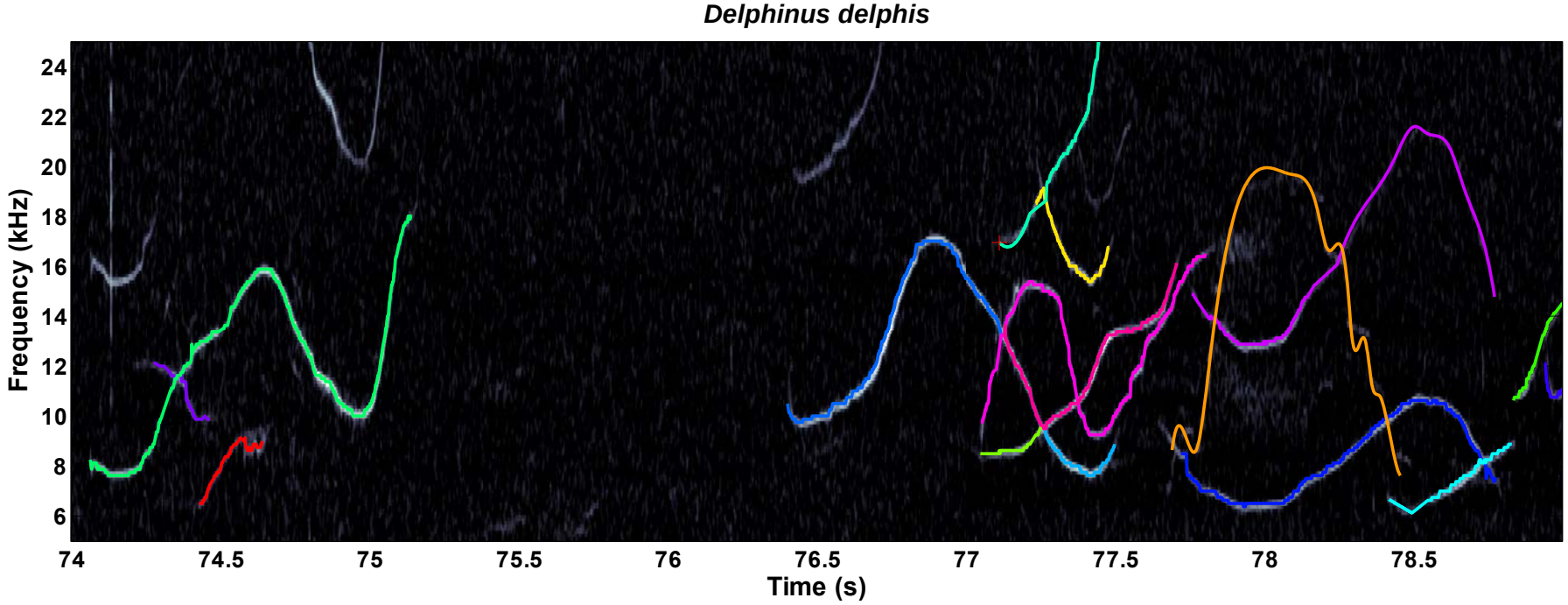


A model approach



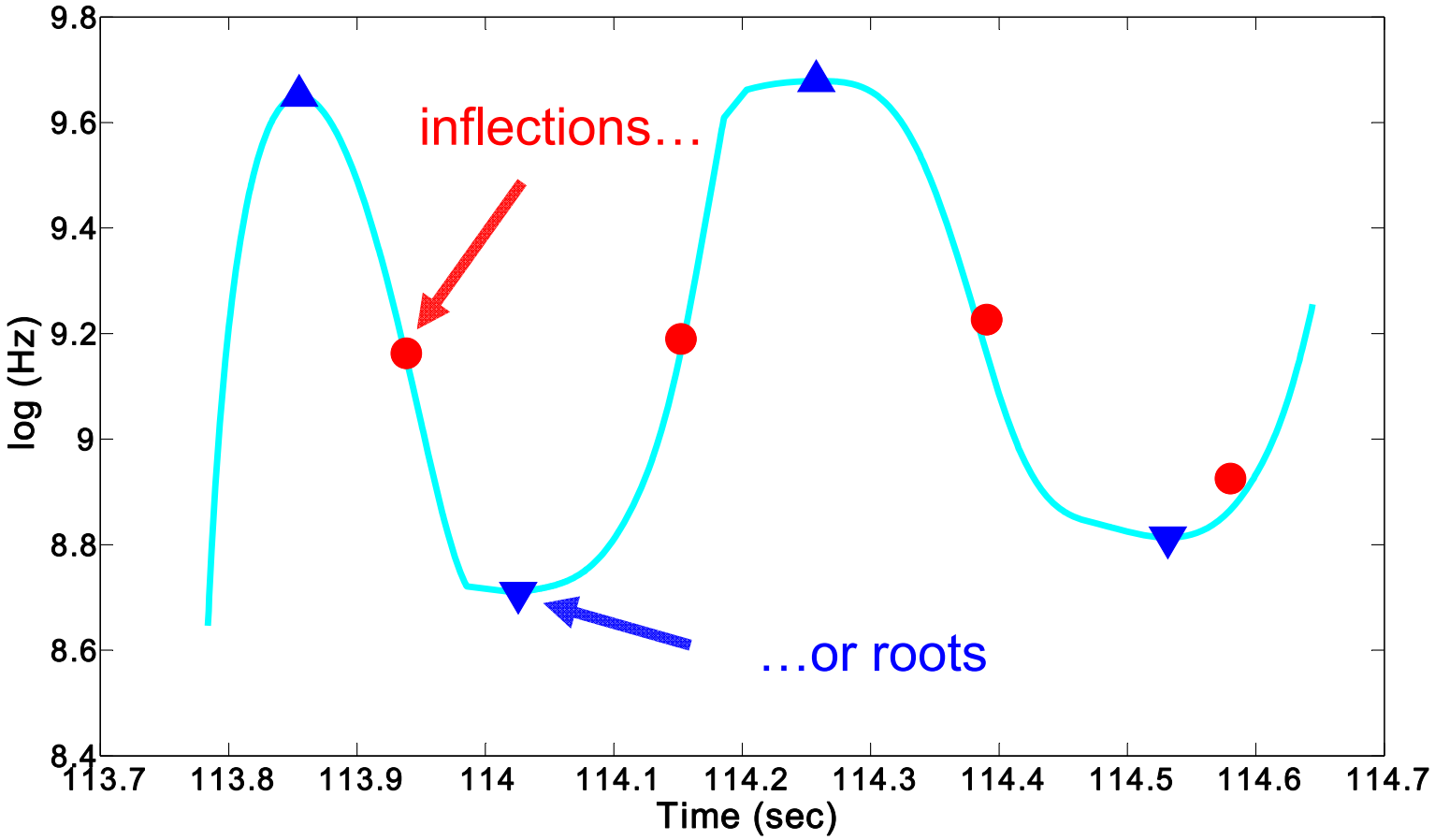
Detect whistle contours:

Silbido detector by Roch et al. (2011)



Roch, M.A., T.S. Brandes, B. Patel, Y. Barkley, S. Baumann-Pickering, M.S. Soldevilla (**2011**) Automated extraction of odontocete whistle contours. *J. Acous. Soc. Am.*, **130**(4), 2212-2223.

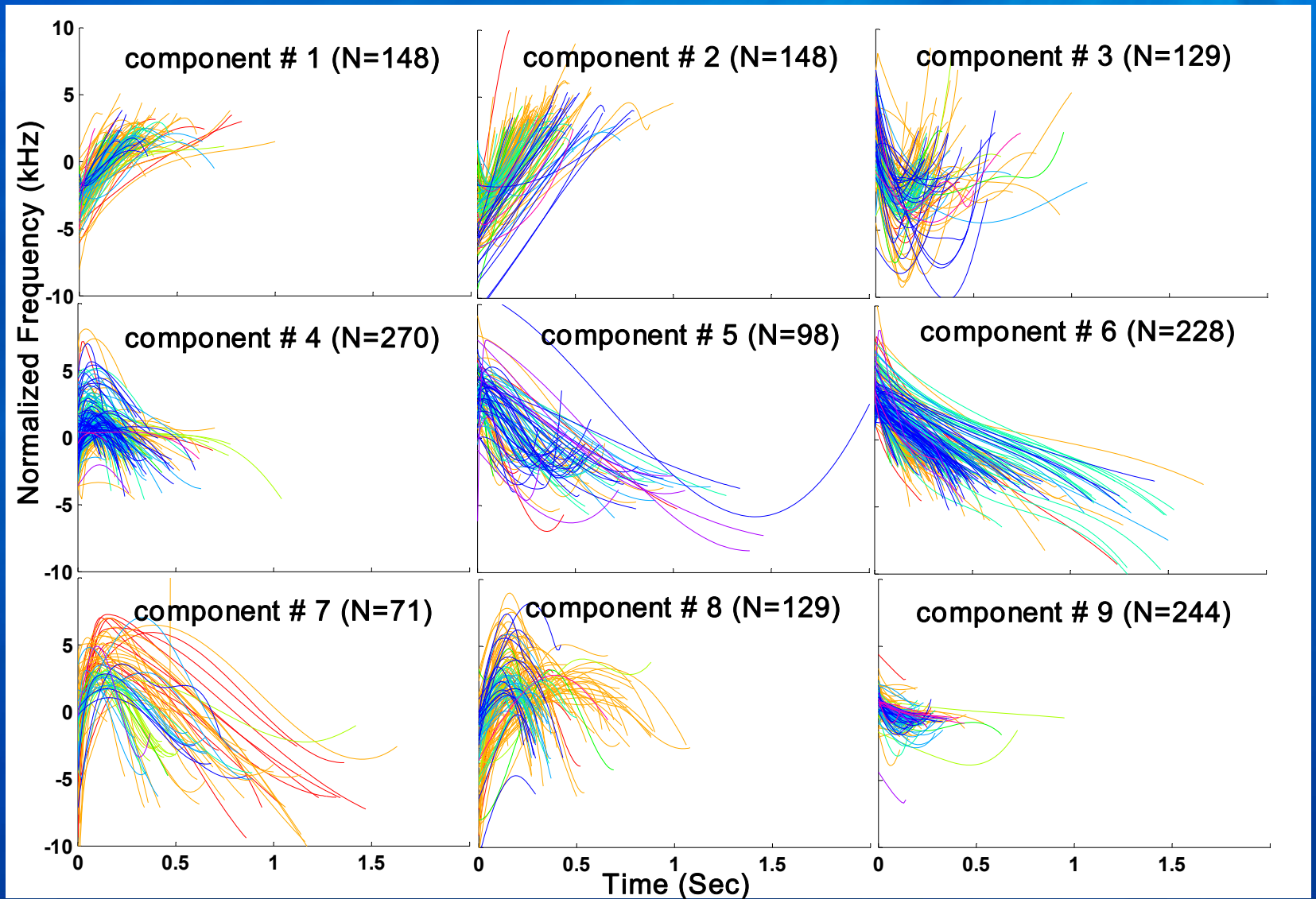
Split whistles into components



Split whistles into components

Component clusters

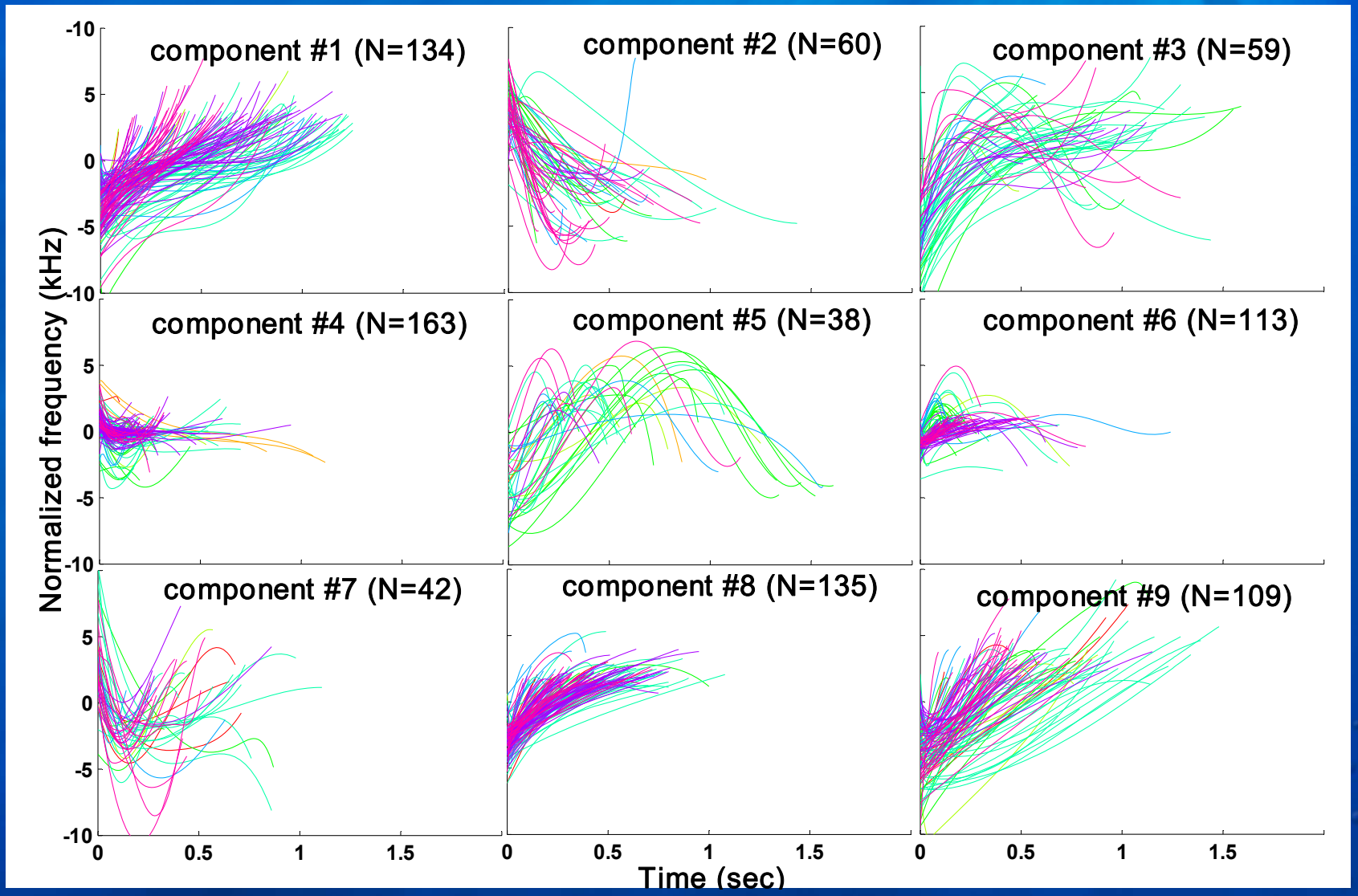
T. truncatus



Split whistles into components

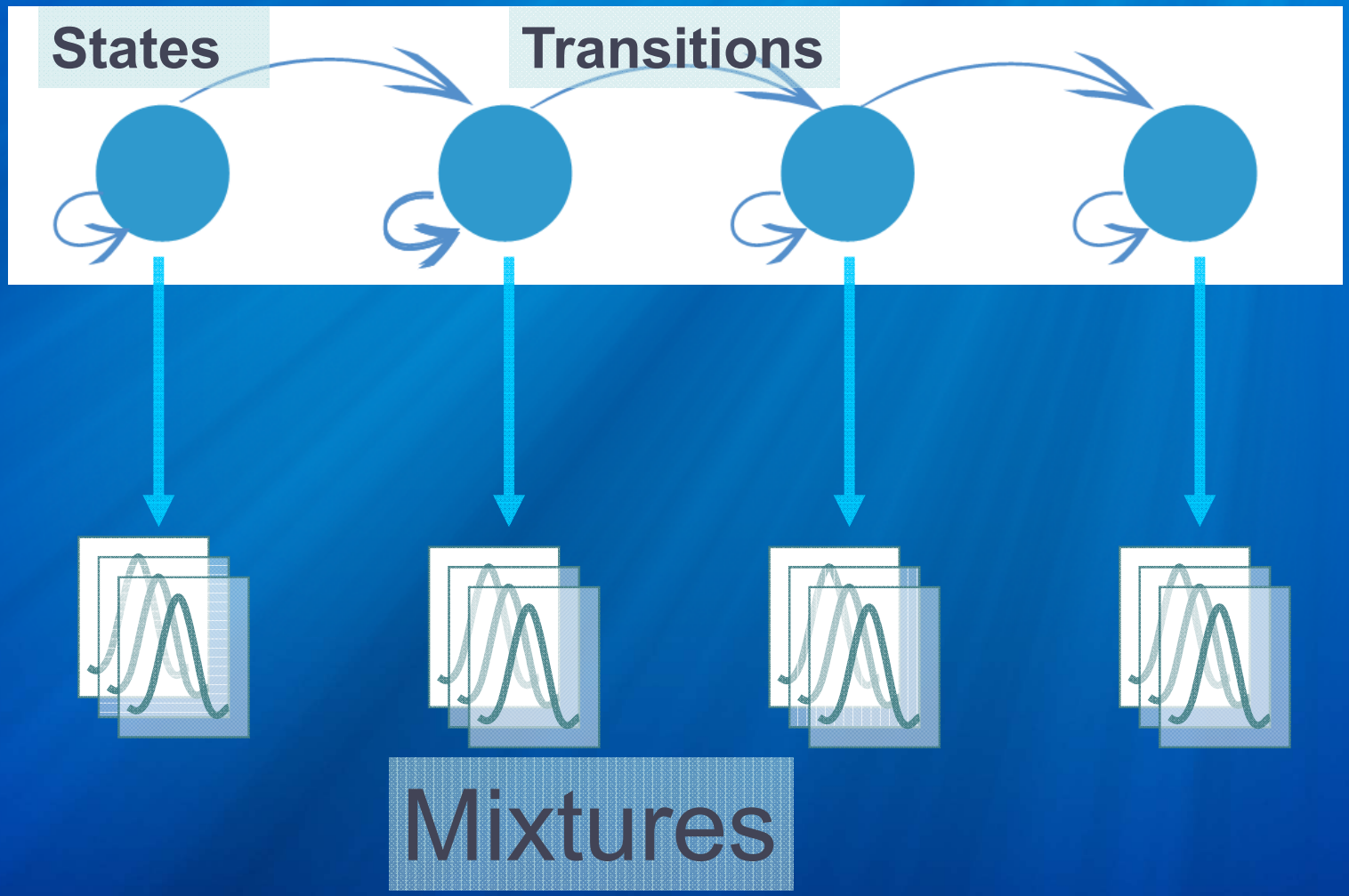
Component clusters

S. longirostris

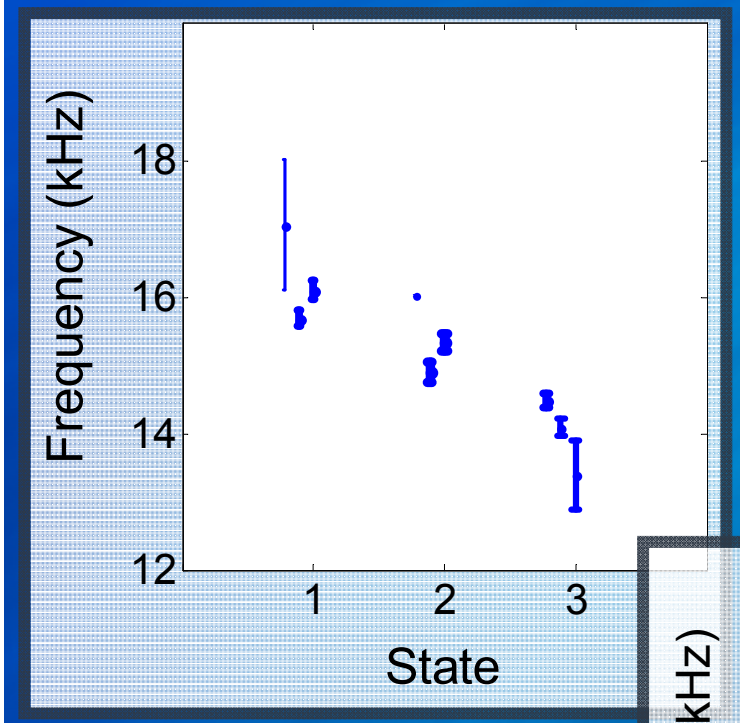


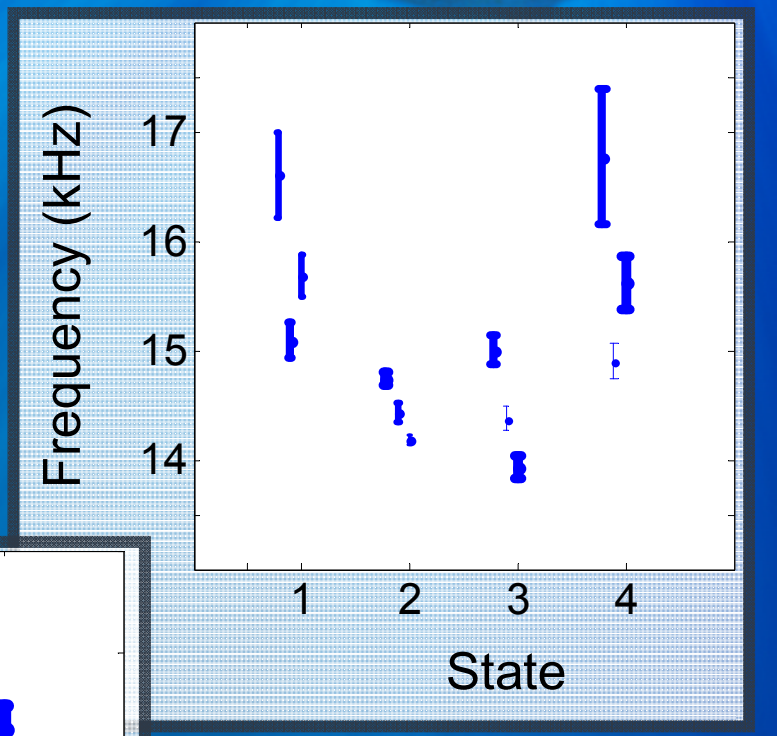
Train Acoustic Models

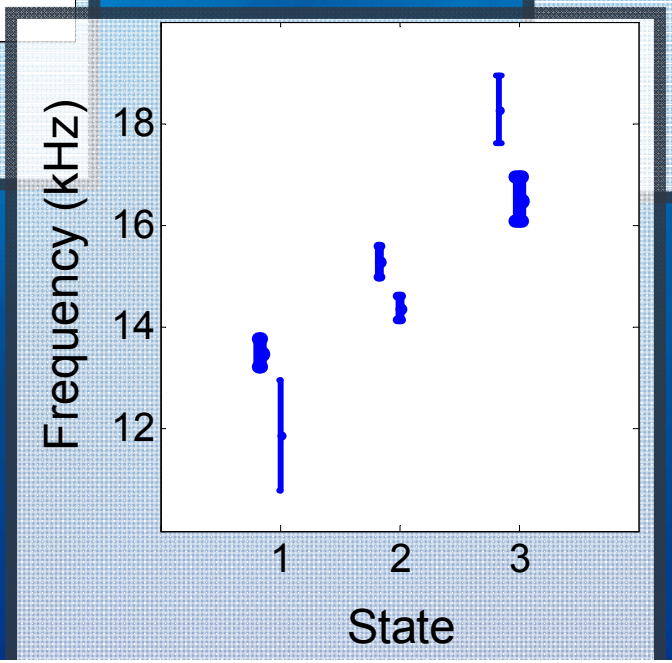
Hidden Markov model:



Training yields component models



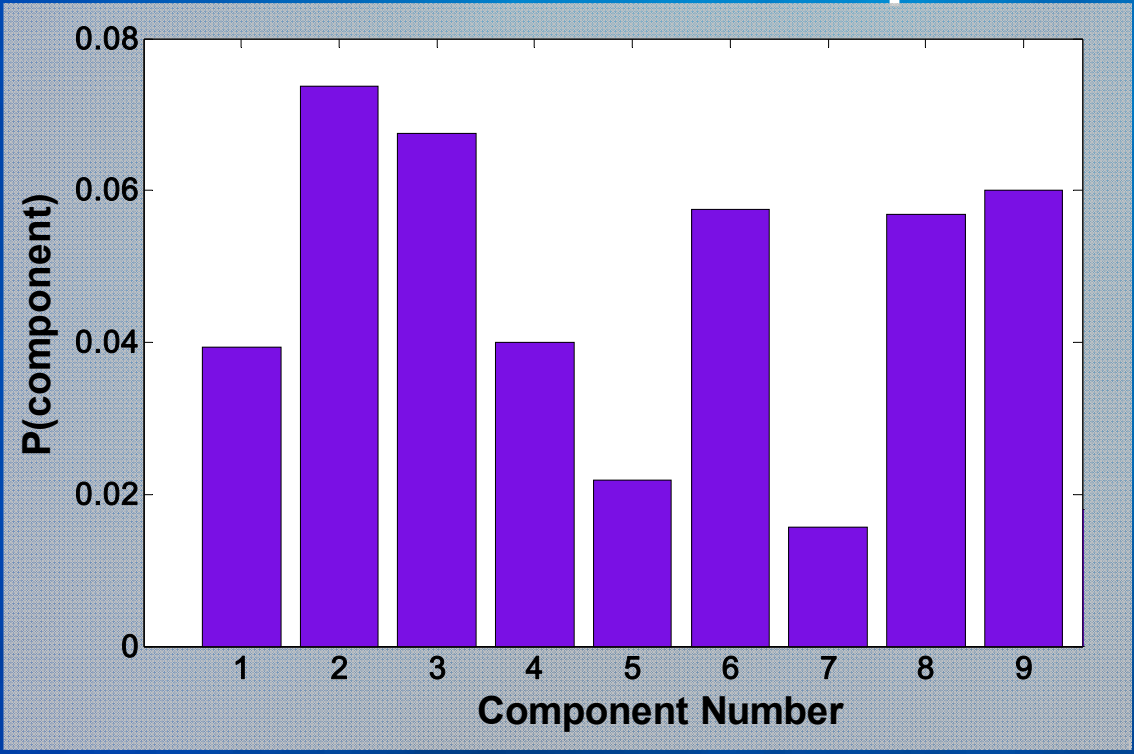


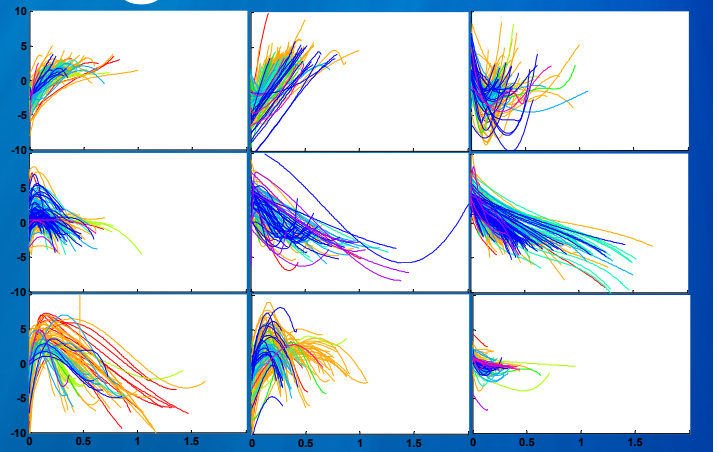


Train Language Models

Language Models

Bottenose dolphin unigram

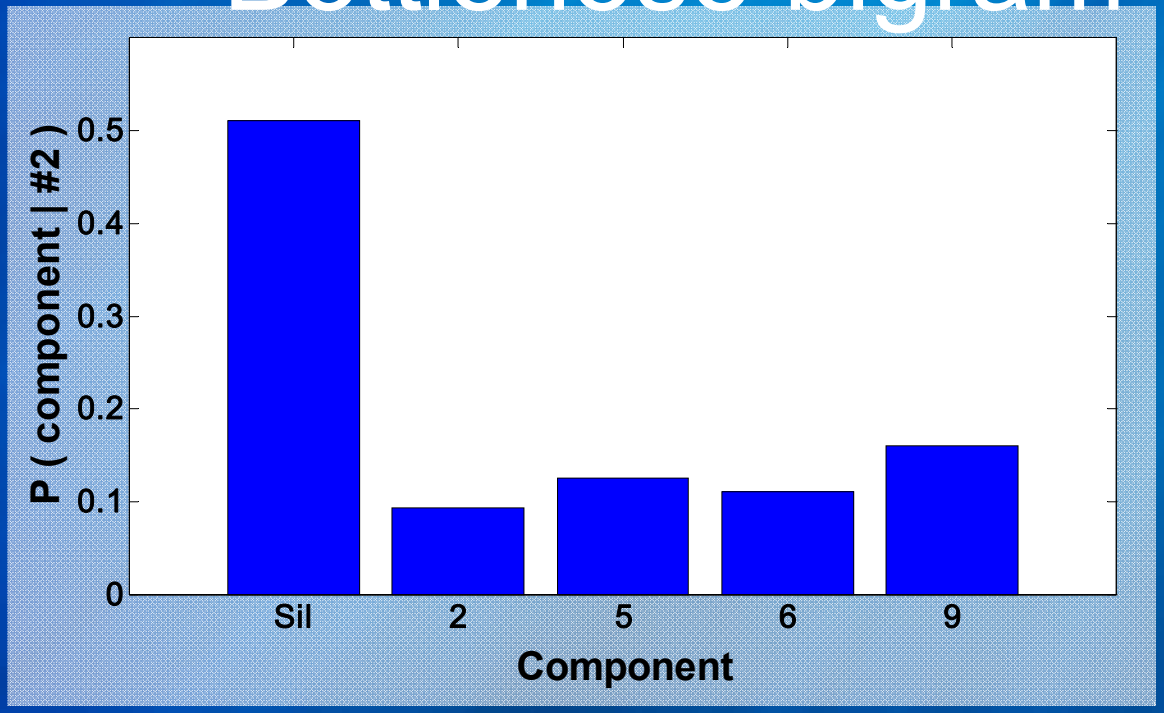


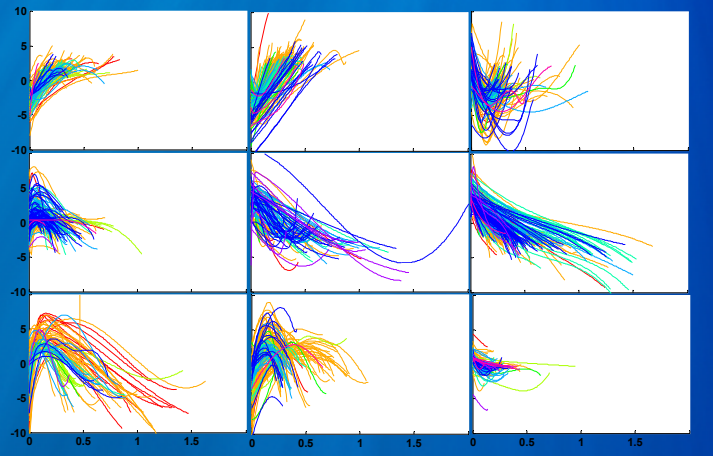


Train Language Models

Language Models

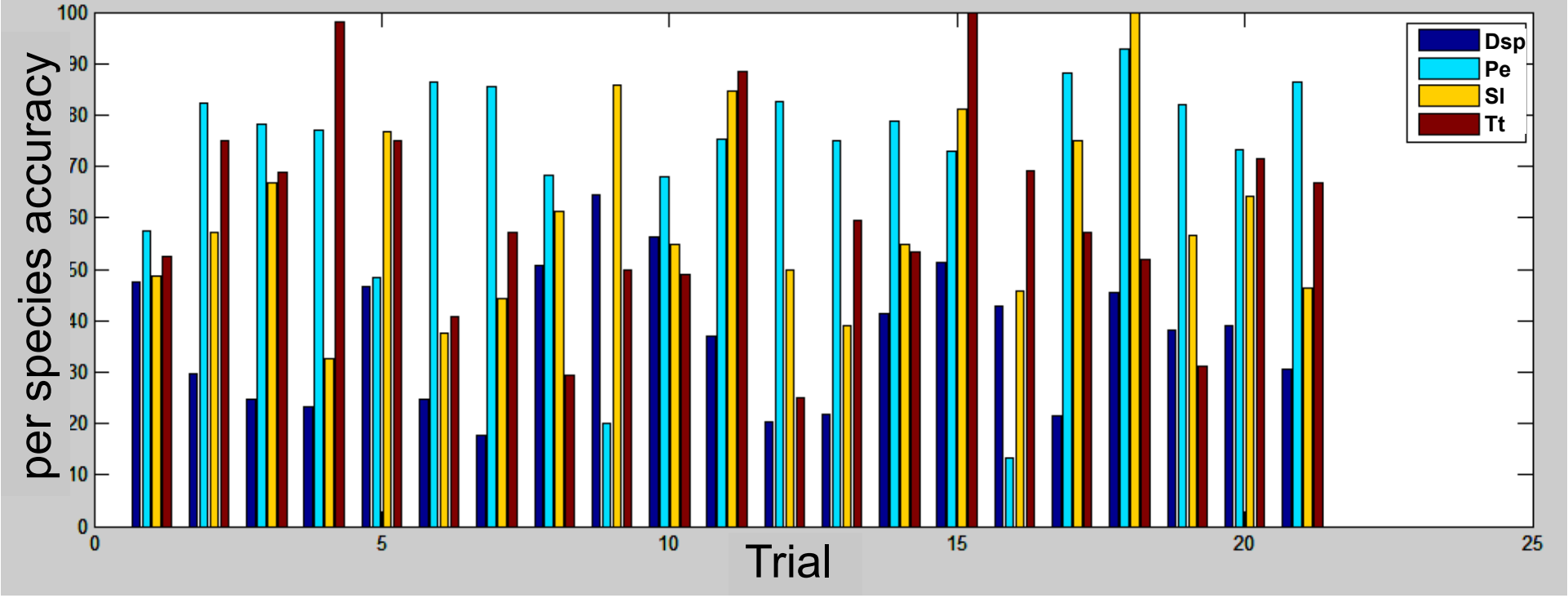
Bottlenose bigram





Preliminary Results

Accuracy %	*Common (Dsp)*	*Melon-headed (Pe)*	*Spinner (Sl)*	*Bottlenose (Tt)*
μ	36.9	71.1	60.1	60.5
σ	13.4	20.8	18.1	20.5



What next?

- Improvements to existing techniques in whistle and echolocation click domains
- Fusion of classification systems
- Behavior and localization…

Our thanks to…

Office of Naval Research – Michael Weise

US Navy CNO N45 – Frank Stone & Ernie Young

Cascadia Research Collective

folks at the SIO Whale Acoustics Lab


photo: Amanda Cummins
Questions?