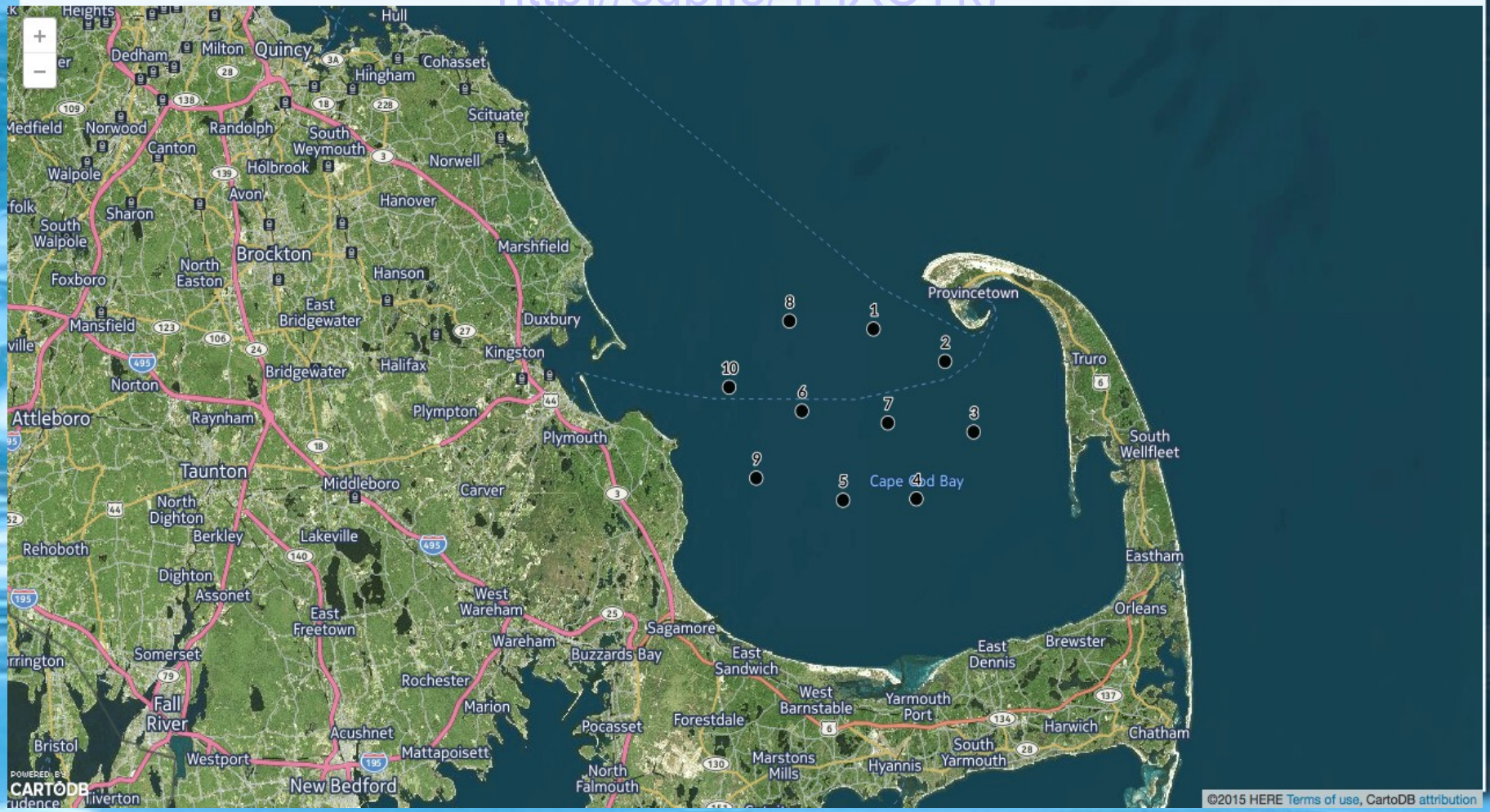
Near-Fully-Automatic Large-Scale Detection and Localization of Baleen Whale Vocalization in Fixed-Hydrophone Array

Yu Shiu & Christopher W. Clark
Bioacoustics Research Program
Cornell University

Fixed-Hydrophone-Array in Cape Cod Bay, MA

http://cdb.io/1HXCYk7

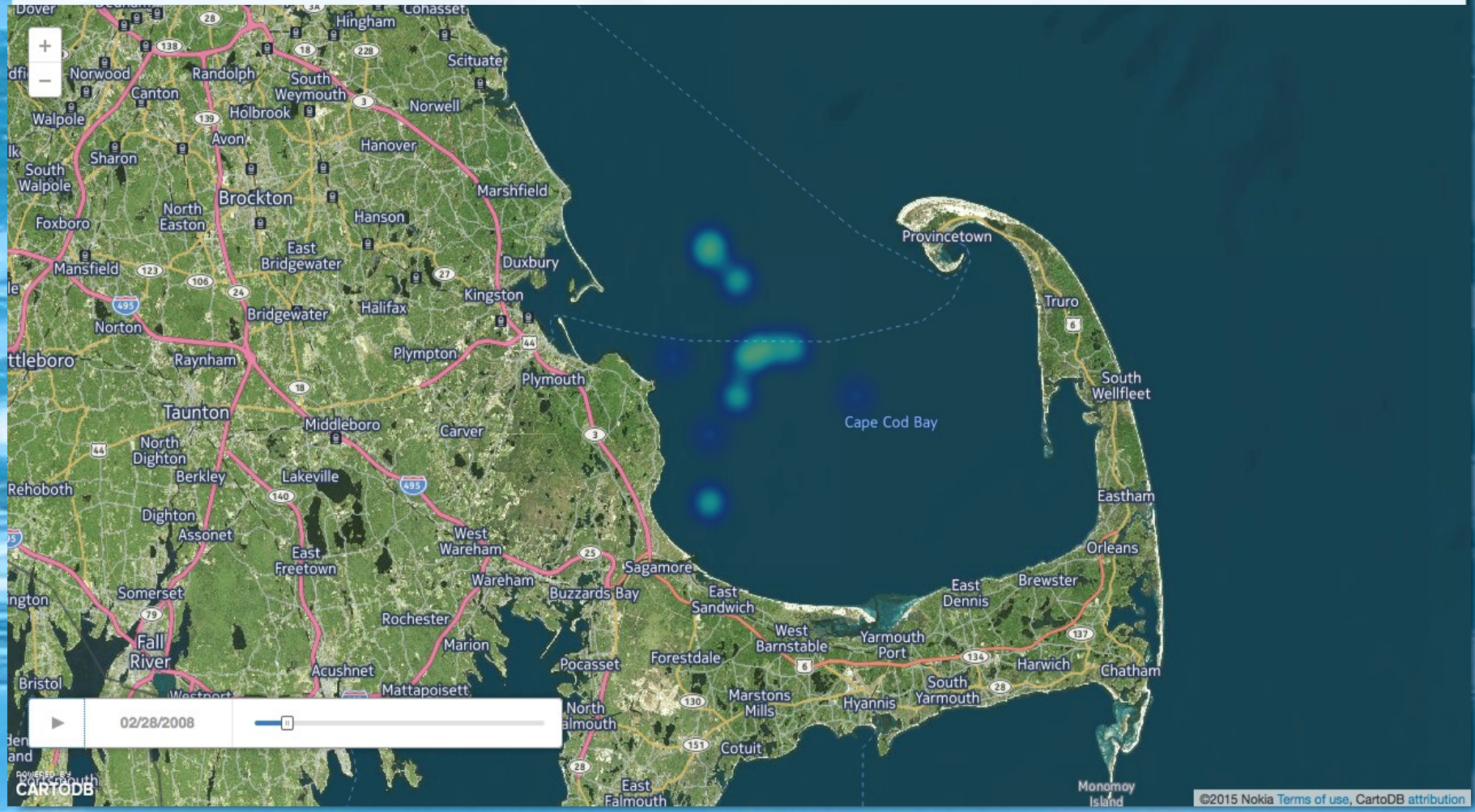


What Information on NARW can we get from an hydrophone-array?

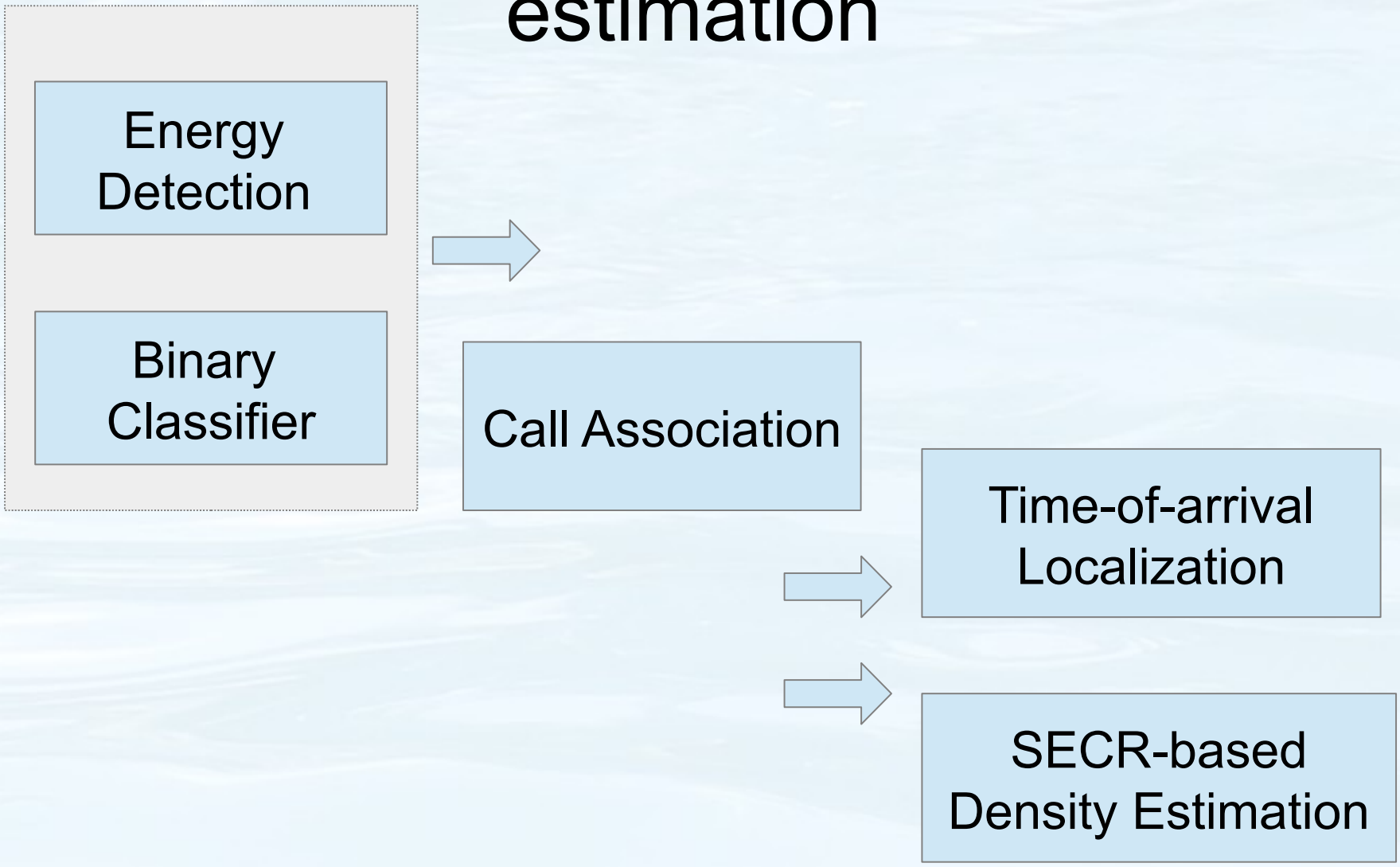
- Target species: North Atlantic Right Whale
- **D**C**LDE**
  - Detection: an important intermediate step
  - Localization
  - Density estimation
- Can they be done automatically?
  - Sort of.
- Without automation, how large scale can we go?

Case study: 45 Days in Cape Cod Bay

http://cdb.io/1HoWmFk

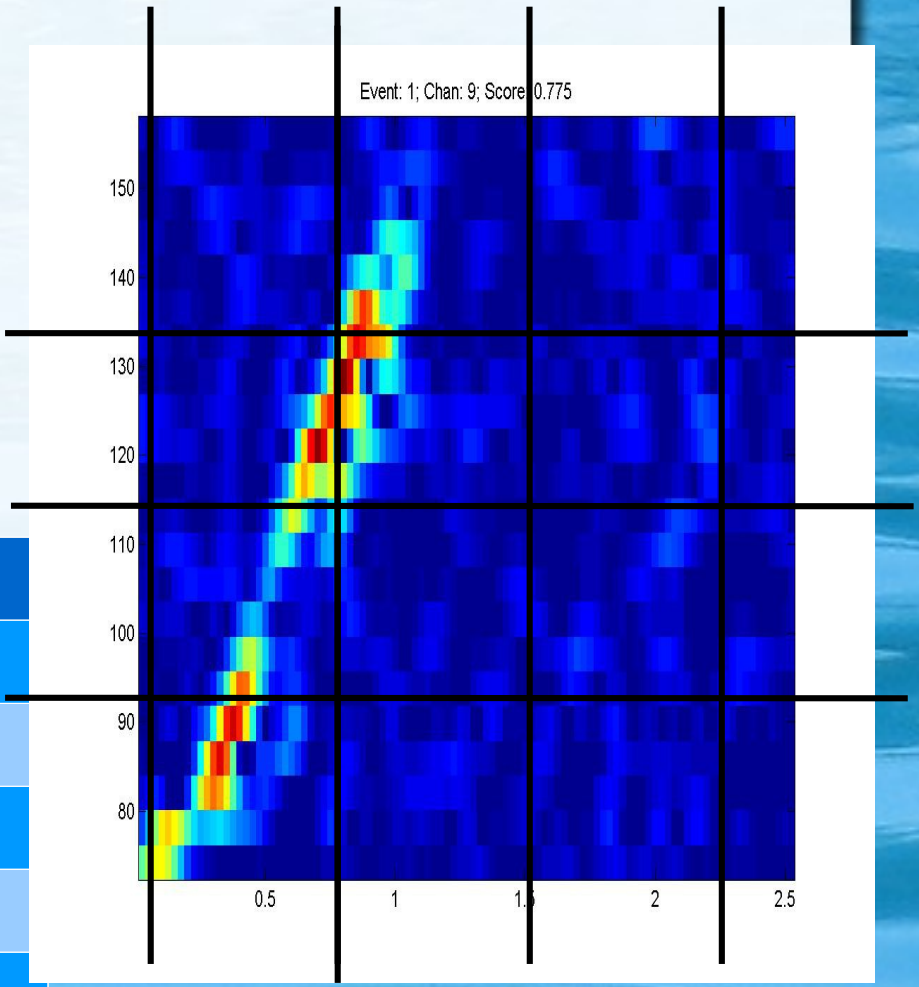


Automatic localization & density estimation



Detector Performance

- Dataset: 6-day full truth labeling of 10-channel sound from Cape Cod Bay 2009
- 3 days for training / 3 days for testing
- Feature: (i) HoD (Histogram of Direction); (ii) HoG (Histogram of gradient)



Feature	HoD	HoG	HoG'
Dim	432	216	972
ROC-AUC	0.9291	0.9009	0.8624
PR-AUC	0.9112	0.8901	0.8595
False-Pos	14.99%	14.87%	19.30%
False-Neg	12.08%	19.28%	23.38%
Pos: 43,016 vs Neg: 43,016			

Reality Check:
Building a binary classifier is not enough!

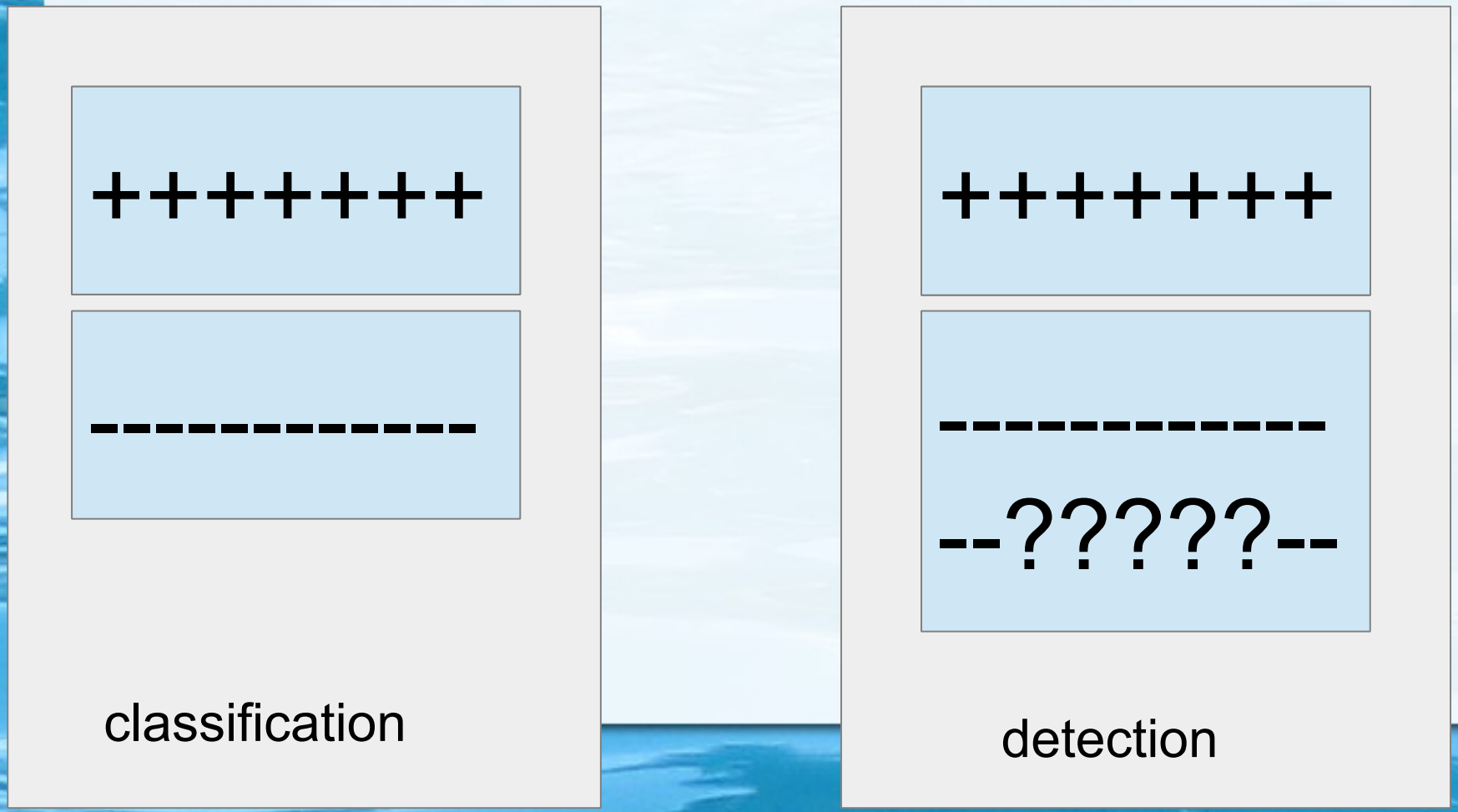
	TP	FP	FN	Precision	Recall	F1
20090221	2,434	4,932	2,487	0.3304	0.4946	0.3962
20090222	2,884	7,835	2,768	0.2691	0.5103	0.3523
20090417	4,789	15,988	4,922	0.2305	0.4932	0.3142

False-Positive Re-training

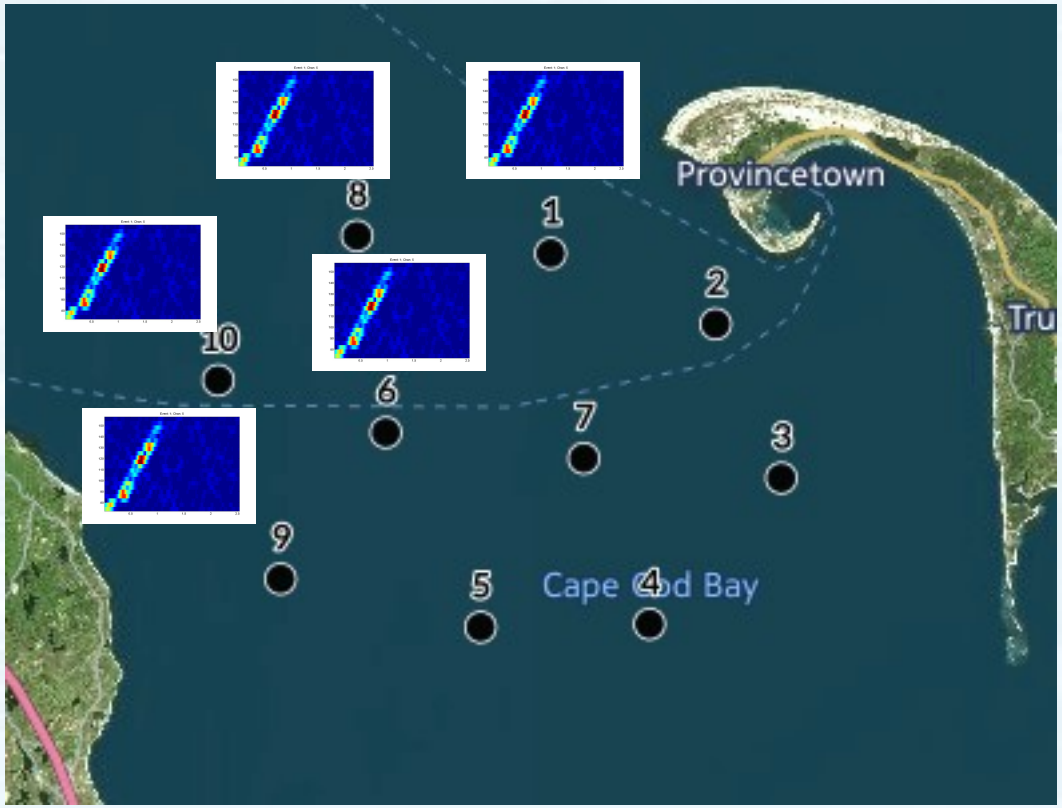
	TP	FP	FN	Precision	Recall	F1
20090221	2244	1401	2657	**0.6156**	**0.4579**	**0.5252**
20090222	2572	2007	3062	**0.5617**	**0.4565**	**0.5037**
20090417	4105	6839	5543	**0.3751**	**0.4255**	**0.3987**

I shouted: “Why?”

What caused the performance drop between classification dataset and detection dataset?



Call association is the key to L&DE



Automatic Localization needs the information of first arrival call as the input

Automatic SECR-based density estimation needs detection history of an individual call in multiple detector locations

Is it realistic to achieve call association when the detection is imperfect?

- False positive (false-alarm)
  - Hard drive mechanical sound
  - Humpback whale calls
- Solution (1): increase the detection threshold in the call detector;
- Solution (2): set up a threshold on the localization error / uncertainty

- False negative (miss)
  - Match filter (Spectrogram Correlation) can save the miss.
  - Density estimation method is another insurance

Wait! Why not “full automatic everything”?

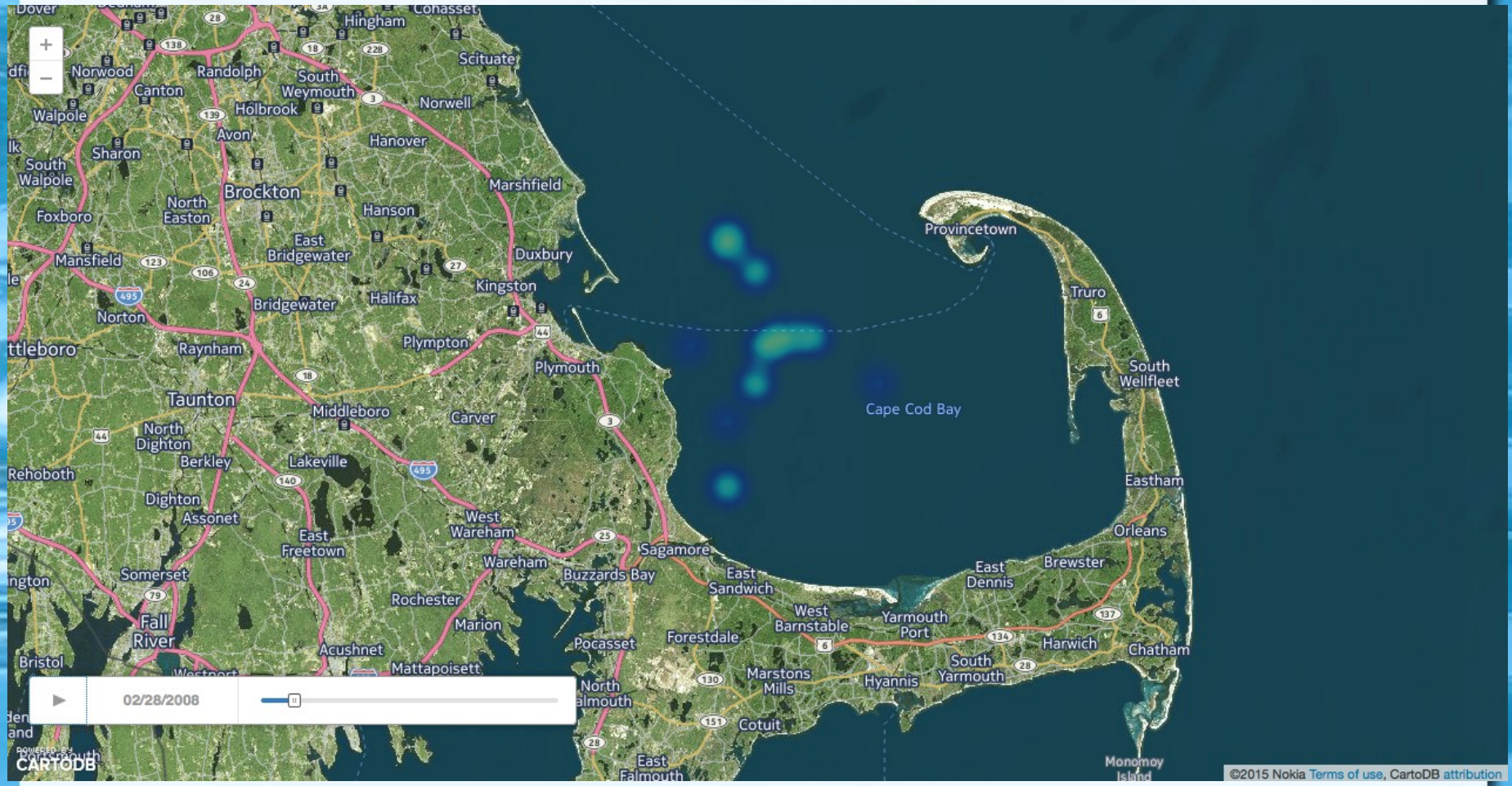
Human expert are overwhelmed in every step

Can we believe the result?

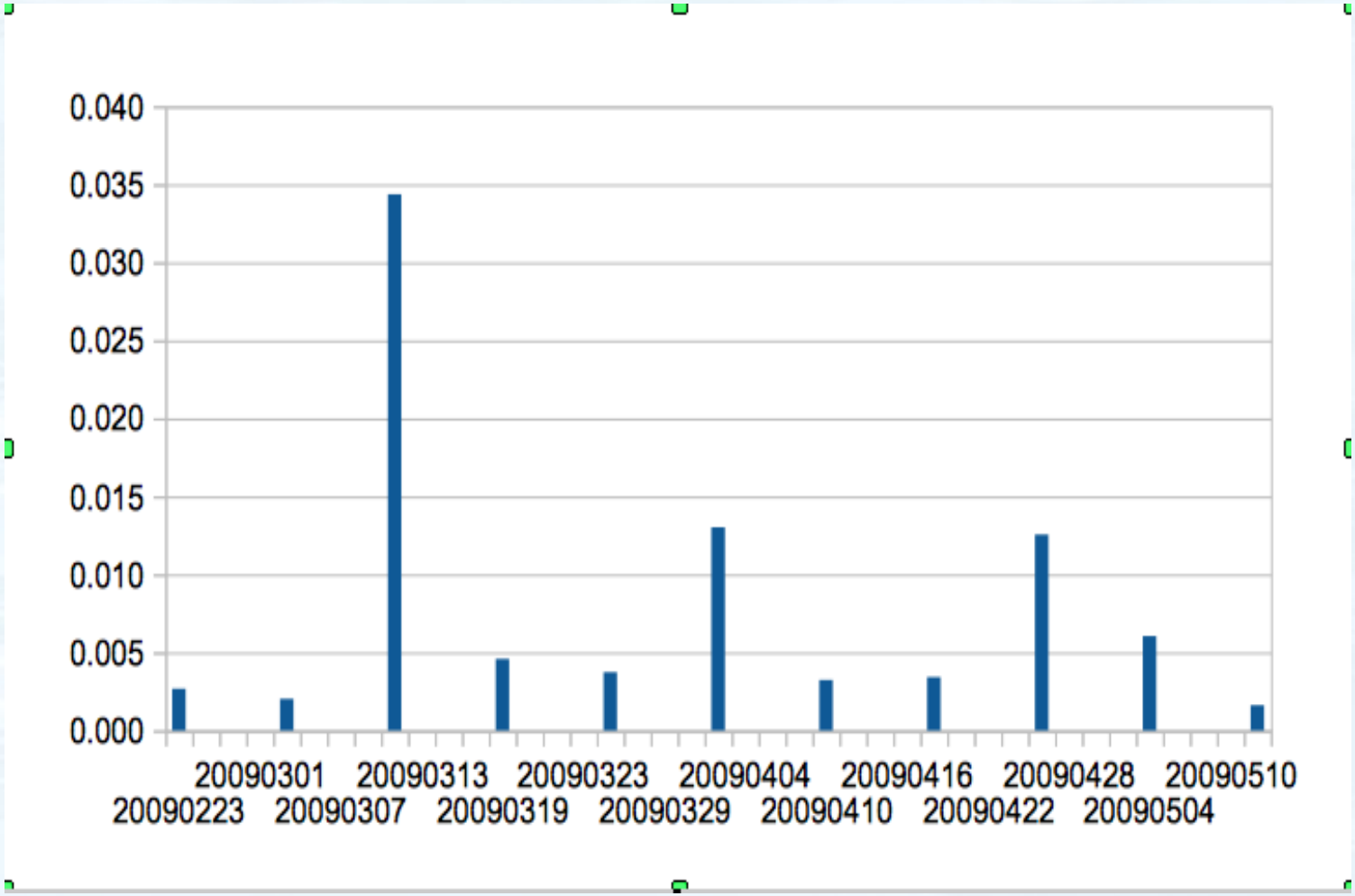
As more automatic as possible whereas human expert verified the results in the end

Case study: 45 Days in Cape Cod Bay

http://cdb.io/1HoWmFk



SECR-based Call Density



Summary

- A near-fully-automatic fixed-hydrophone-array-based information framework is presented.
- To enlarge the scale of information extraction is to engage automation as far as we can and then human expert steps in.
- Challenge to every step of automation is highlighted and solution is offered.
- Primitive results of localization and SECR-based call density estimation are presented.

Future Work

- Use the offered information to answer long-term scientific questions on animal behavior and population
- Offer more information in large scale with known certainty:
  - Call signal characteristics
  - Movement tracking via bearing change or locations

Feedback & Questions are welcome!

Thanks to Christopher W. Clark, J. Kaitlin Palmer, Dimitri W. Ponirakis, Aaron Rice & other BRP colleagues