Acoustics 2017 Boston | ASA

Automatic Classification of Humpback Whale Social Calls

Irina Tolkova[1] and Lisa Bauer[2],
Antonella Wilby[3], Ryan Kastner[3], Kerri D. Seger[4]

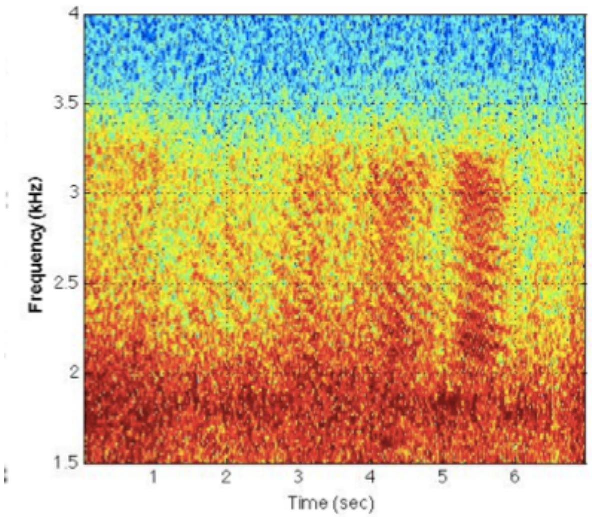
[1]Applied Math and Computer Science Departments, University of Washington, WA, USA
[2]Department of Computer Science, Johns Hopkins University, MD, USA
[3]Jacobs School of Engineering, University of California San Diego, CA, USA
[4]School of Marine Science and Ocean Engineering, University of New Hampshire, NH, USA

Cetacean Bioacoustics

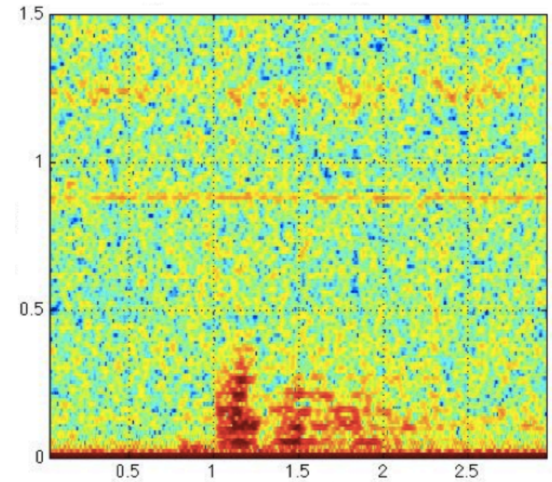
Humpback Whale Bioacoustics

While manual analysis has been applied to a variety of species, studies in automated detection of humpback whale social calls have been limited. This is largely because **their social call repertoire is very varied**, both in duration and frequency, so most classification work is manual.

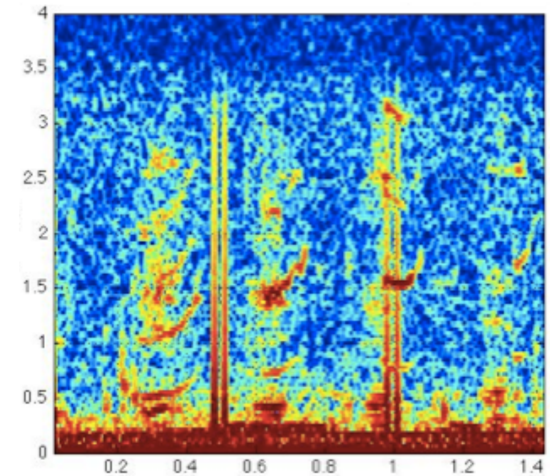
Humpback Whale Social Sound Catalog



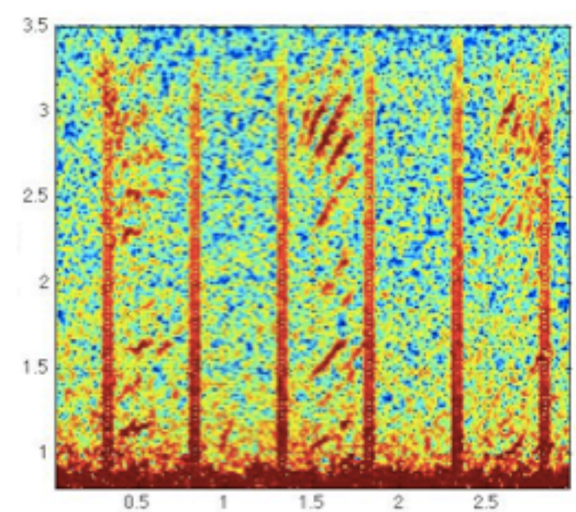
Growl



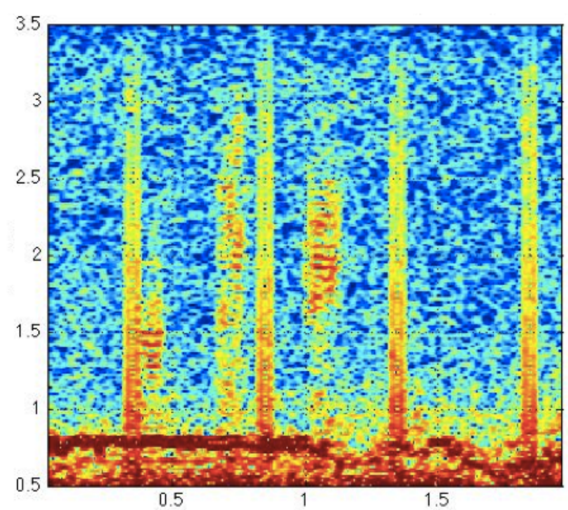
Grunt-grumble



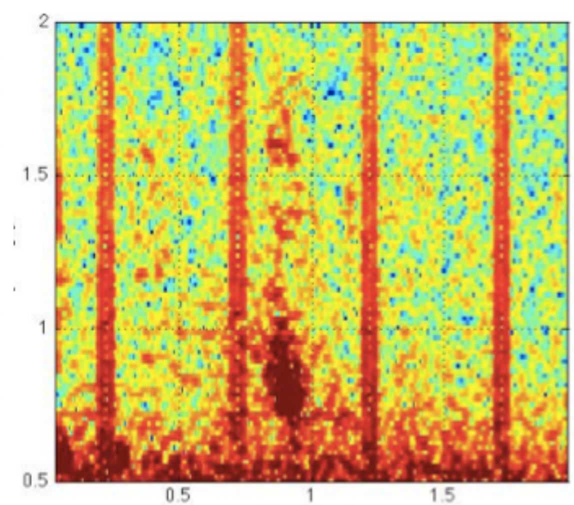
Pulses



Ascending moan



Squelch



Swop

Data Collection

The data for this project was collected by **Dr. Kerri Seger** and **Dr. Aaron Thode** of the Scripps Institute of Oceanography. **Acousonde acoustic tags** were used for recording underwater audio.

Data Collection

The data for this project was collected by **Dr. Kerri Seger** and **Dr. Aaron Thode** of the Scripps Institute of Oceanography. **Acousonde acoustic tags** were used for recording underwater audio.

The tags were deployed in **Los Cabos, Mexico**.
This region falls within a **migration route**, and is also a **breeding ground** of the North Pacific humpback whale subpopulation.

Data Collection

The data for this project was collected by **Dr. Kerri Seger** and **Dr. Aaron Thode** of the Scripps Institute of Oceanography. **Acousonde acoustic tags** were used for recording underwater audio.

The tags were deployed in **Los Cabos, Mexico**.
This region falls within a **migration route**, and is also a **breeding ground** of the North Pacific humpback whale subpopulation.

Acoustic data was collected during February and March of 2014 and 2015, resulting in just under 70 hours of recordings. These recordings were **manually labeled** for whale calls. Each call was **classified** into a call type category.

Project Goal

The overall goal of this project is to create an automatic system that would **detect** and **classify** humpback whale social calls.

There are two main steps to this project:

1. signal processing
2. machine learning / classification

Project Goal

The overall goal of this project is to create an automatic system that would **detect** and **classify** humpback whale social calls.

There are two main steps to this project:

1. signal processing
2. machine learning / classification

original data

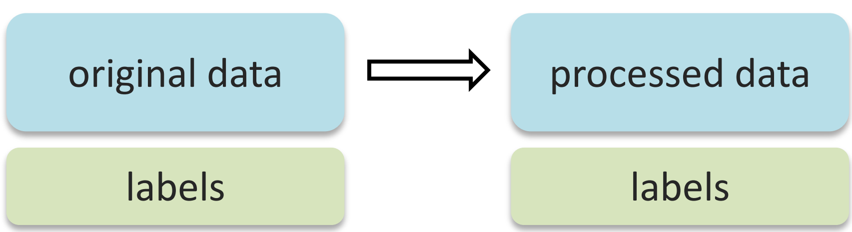
labels

Project Goal

The overall goal of this project is to create an automatic system that would **detect** and **classify** humpback whale social calls.

There are two main steps to this project:

1. signal processing
2. machine learning / classification



Project Goal

The overall goal of this project is to create an automatic system that would **detect** and **classify** humpback whale social calls.

There are two main steps to this project:

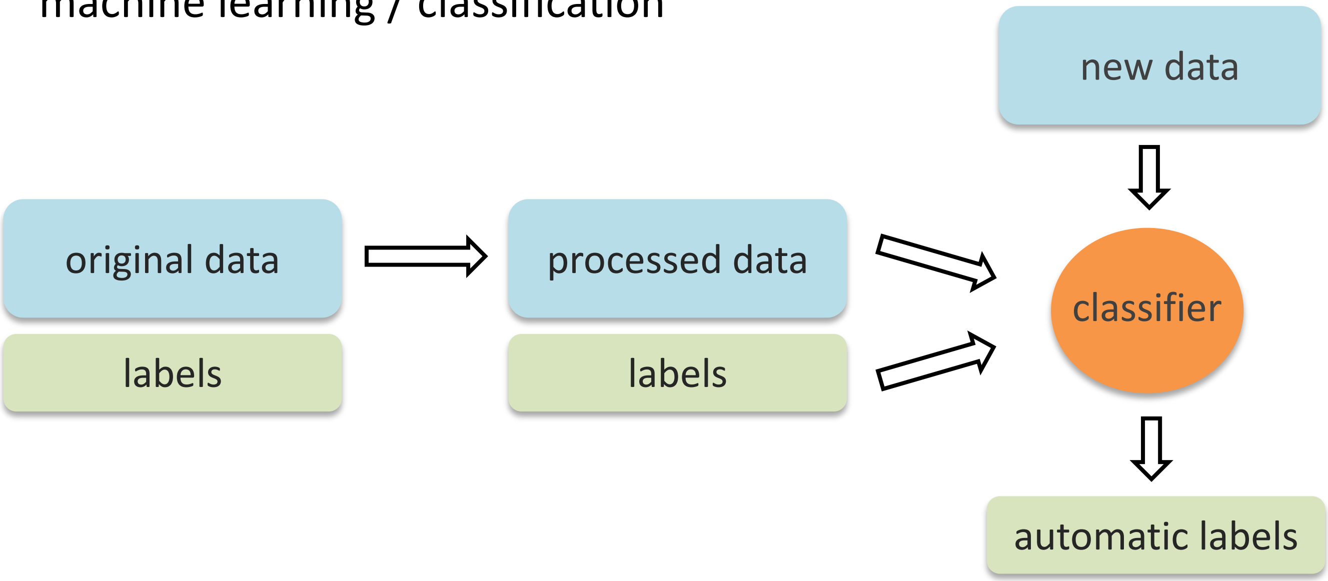
1. signal processing
2. machine learning / classification

Project Goal

The overall goal of this project is to create an automatic system that would **detect** and **classify** humpback whale social calls.

There are two main steps to this project:

1. signal processing
2. machine learning / classification

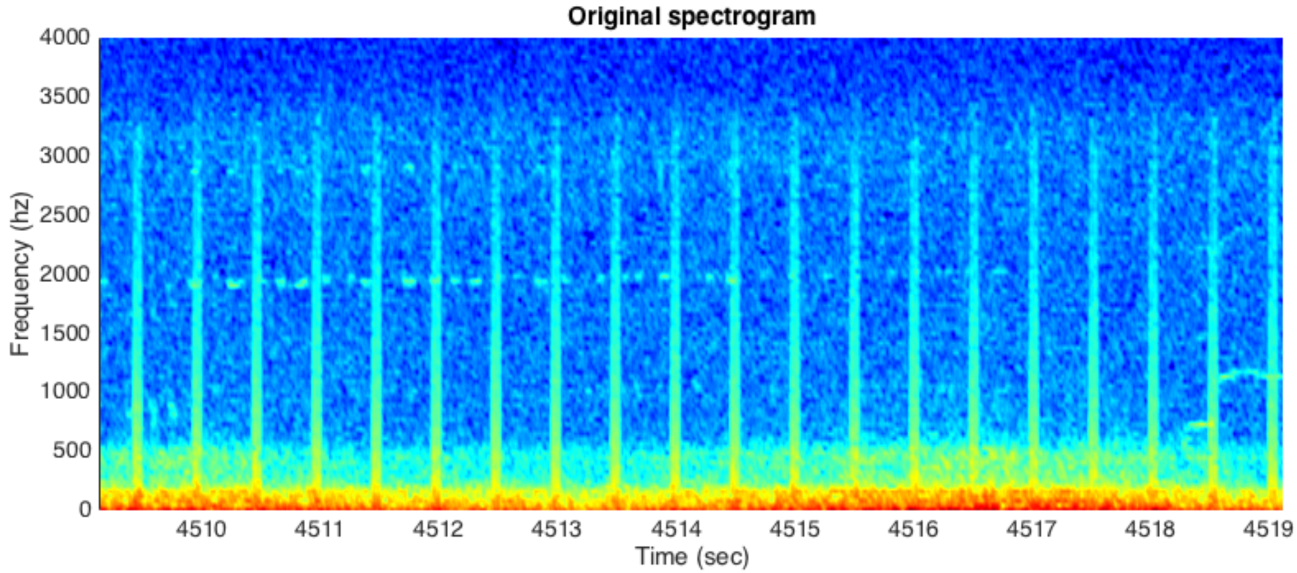


Spectrogram Processing: Pings

Here is a 10-second spectrogram of the recordings:

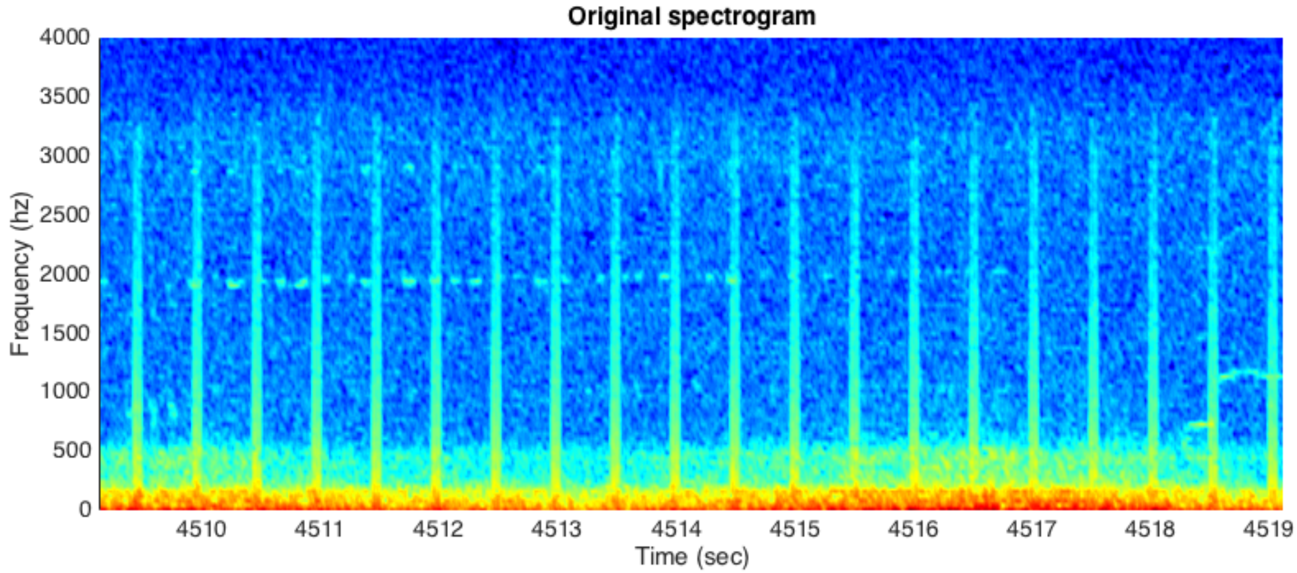
Spectrogram Processing: Pings

Here is a 10-second spectrogram of the recordings:

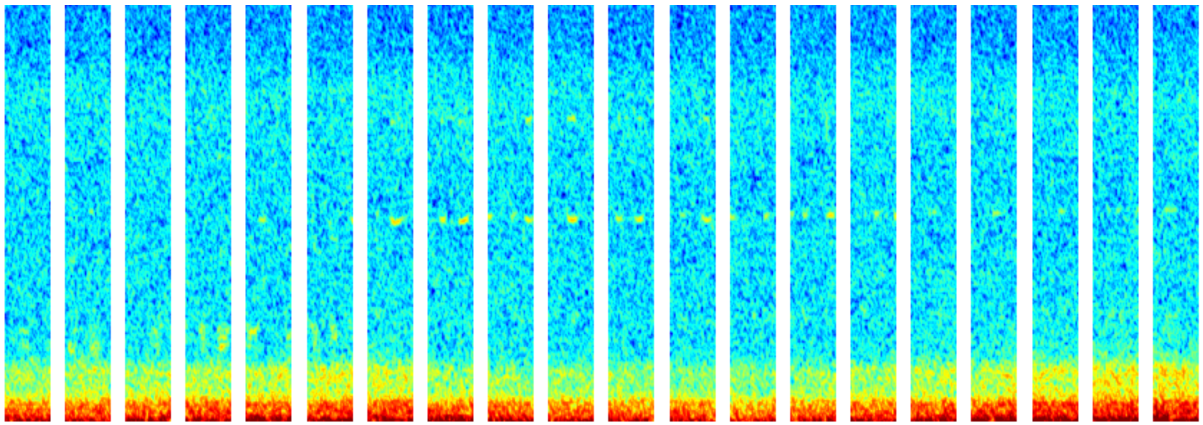


Spectrogram Processing: Pings

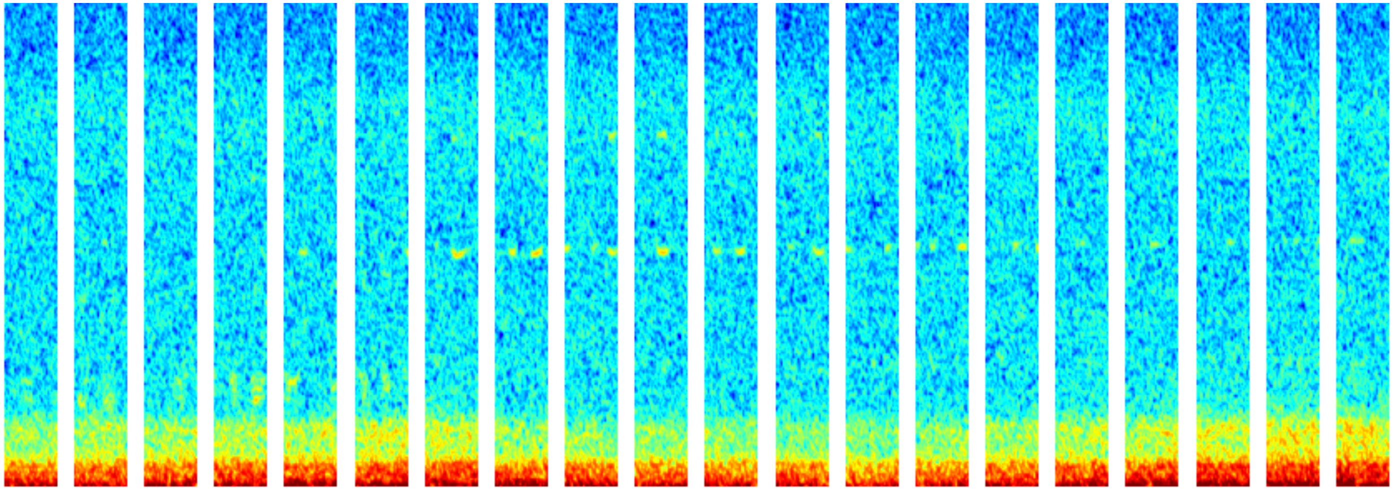
Here is a 10-second spectrogram of the recordings:



Using periodicity, remove the pings:

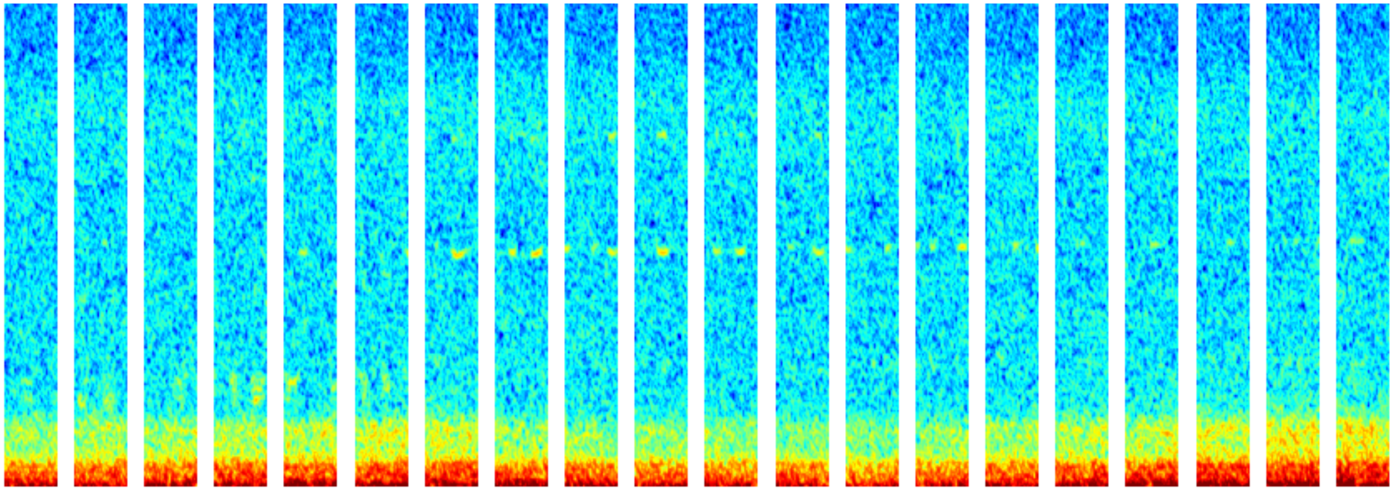


Spectrogram Processing: Background



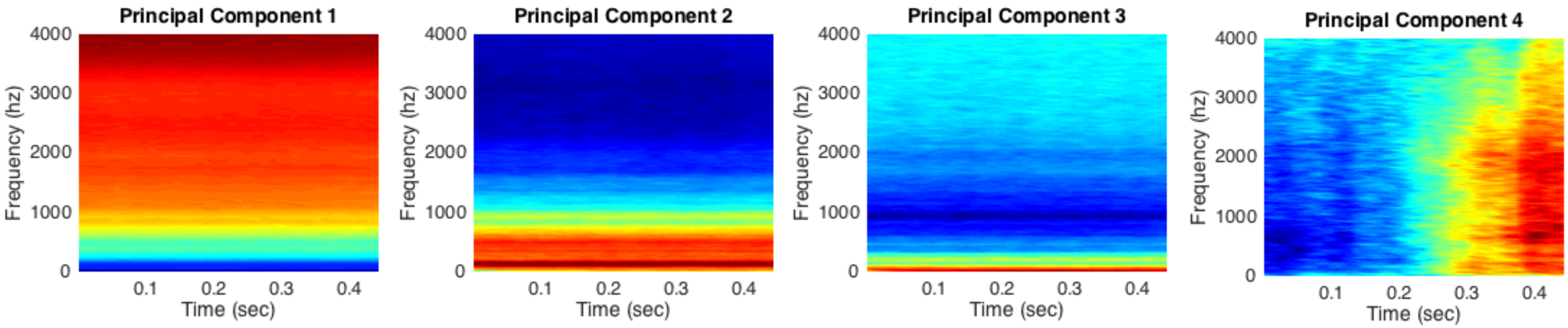
Is there a way to isolate the foreground of this spectrogram?

Spectrogram Processing: Background

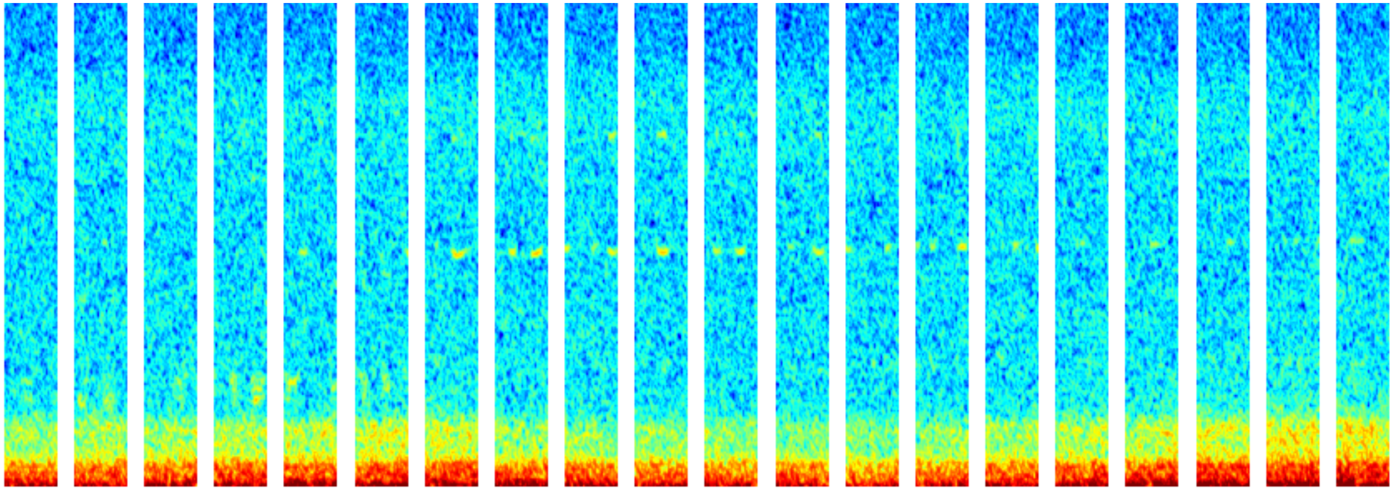


Is there a way to isolate the foreground of this spectrogram?

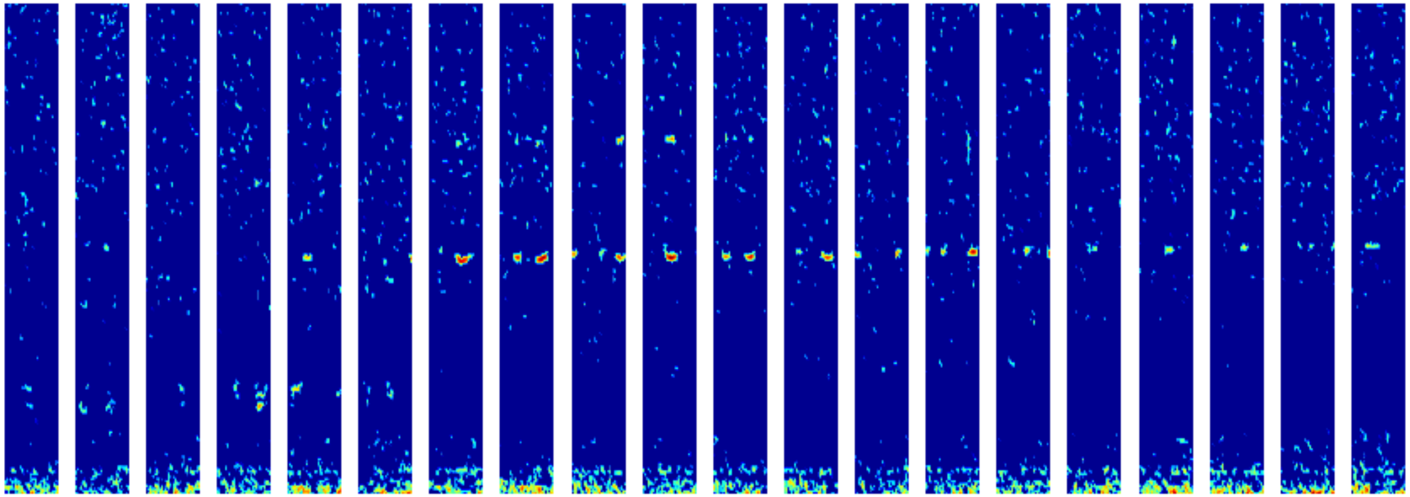
Principal component analysis (**PCA**) is a method for calculating the optimal basis of a matrix. In this case, the principal components are:



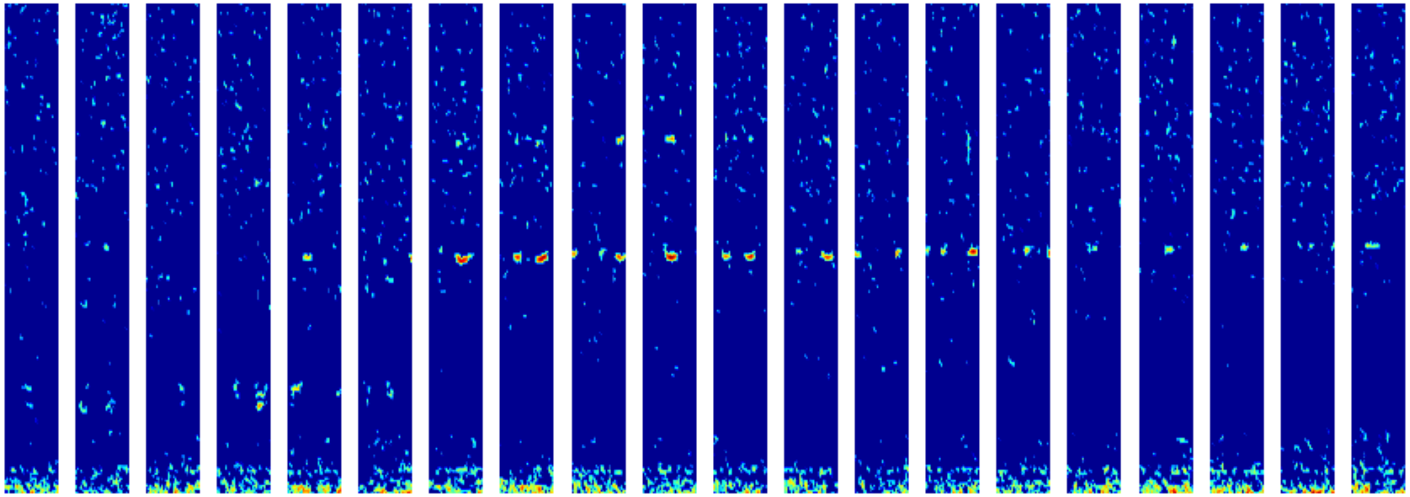
Spectrogram Processing: Background



For each window of the spectrogram, remove the contribution (projection) of the first three principal components. The spectrograms are now:

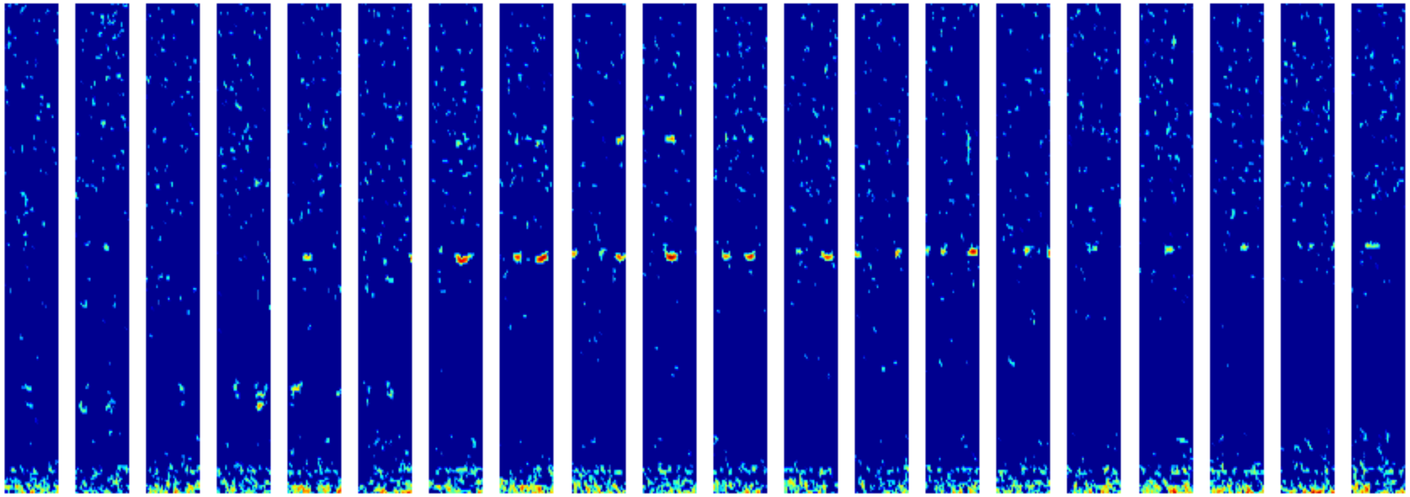


Spectrogram Processing: Background

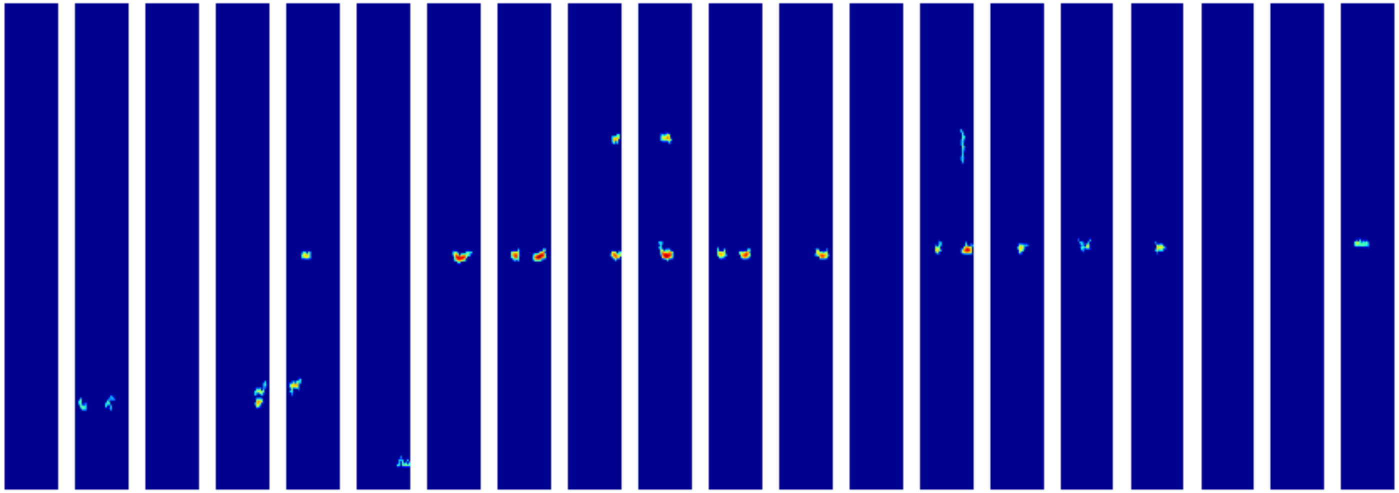


Use a connected-components approach to further process the images:

Spectrogram Processing: Background

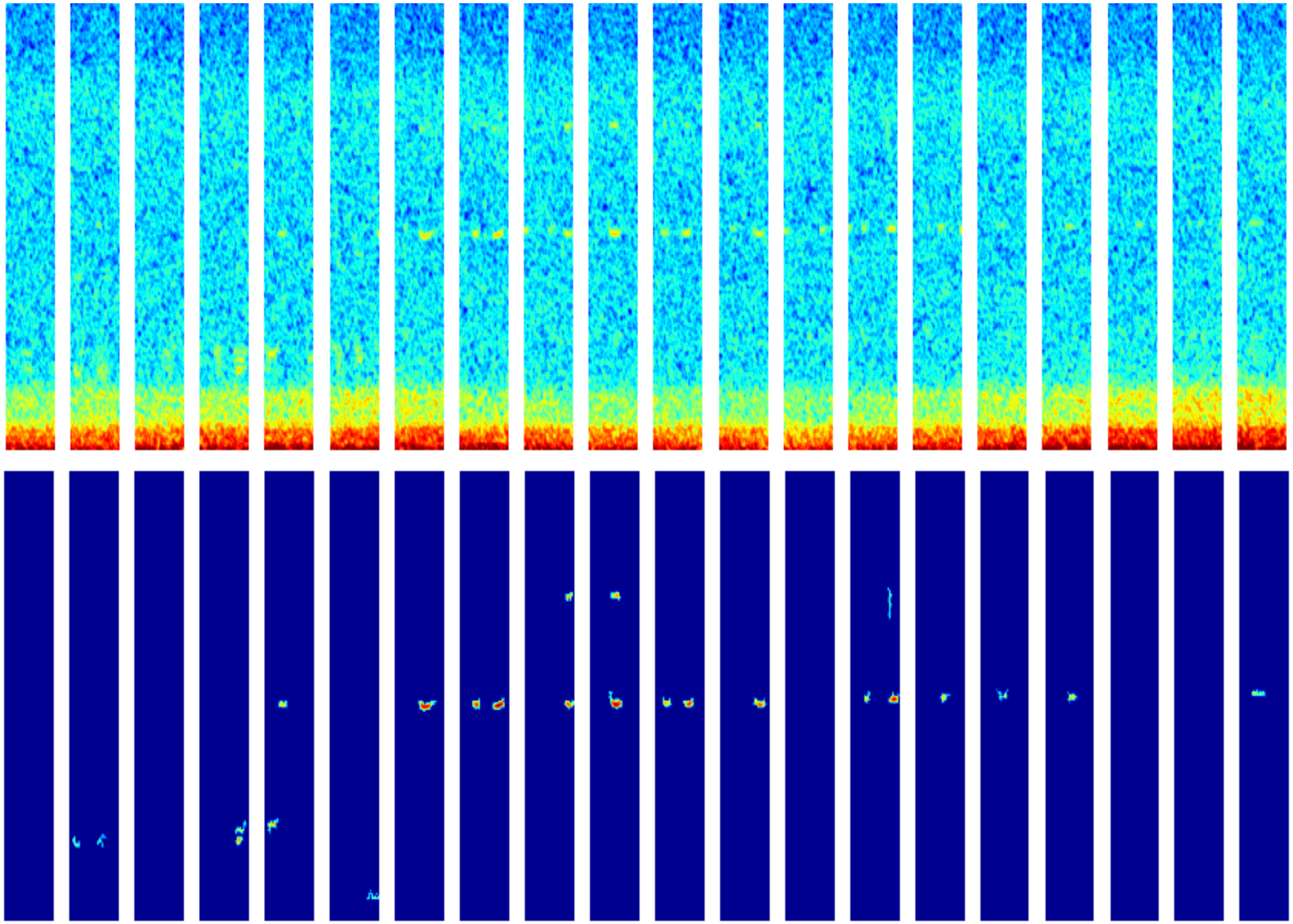


Use a connected-components approach to further process the images:



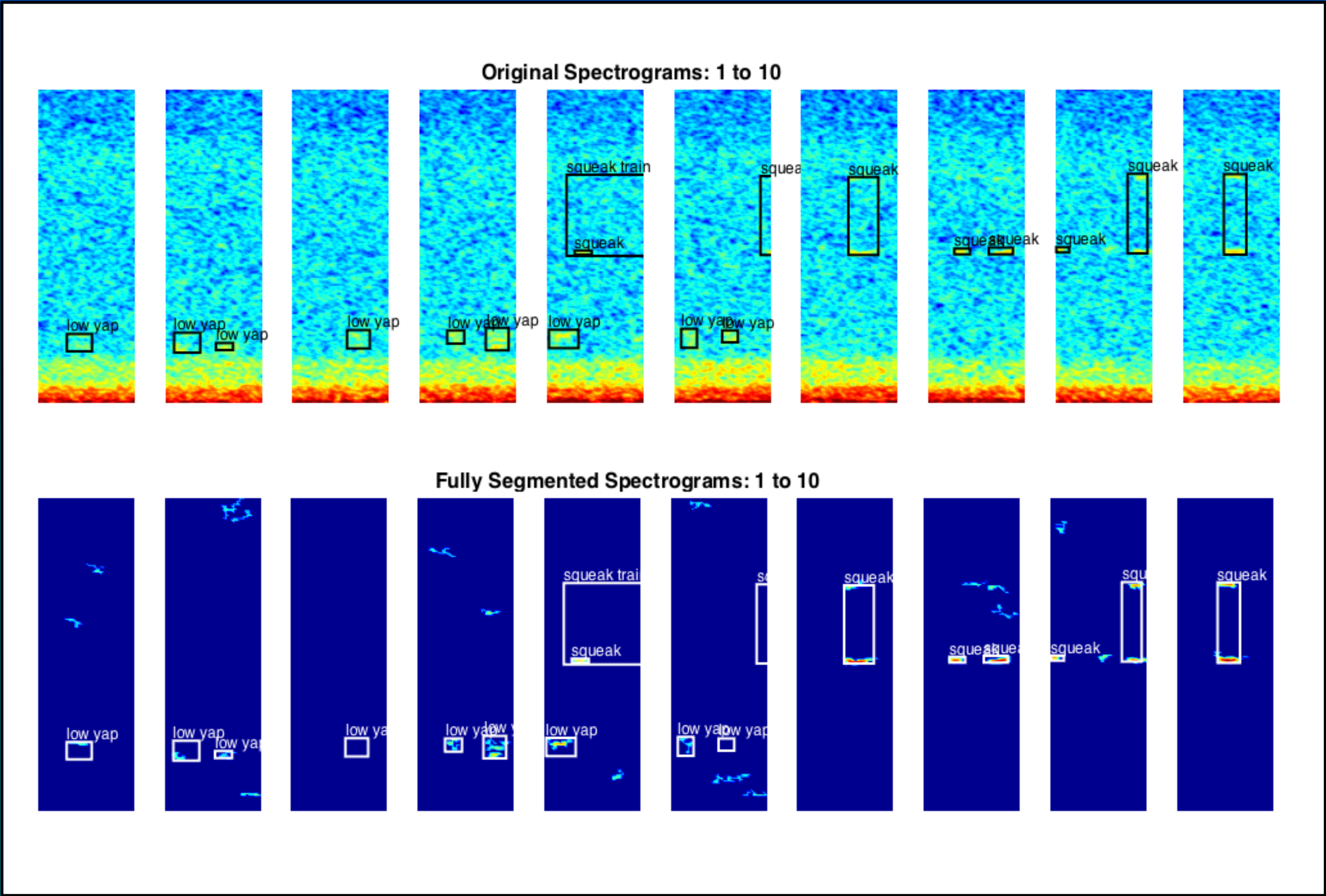
Spectrogram Processing: Cleaning

Before and after:

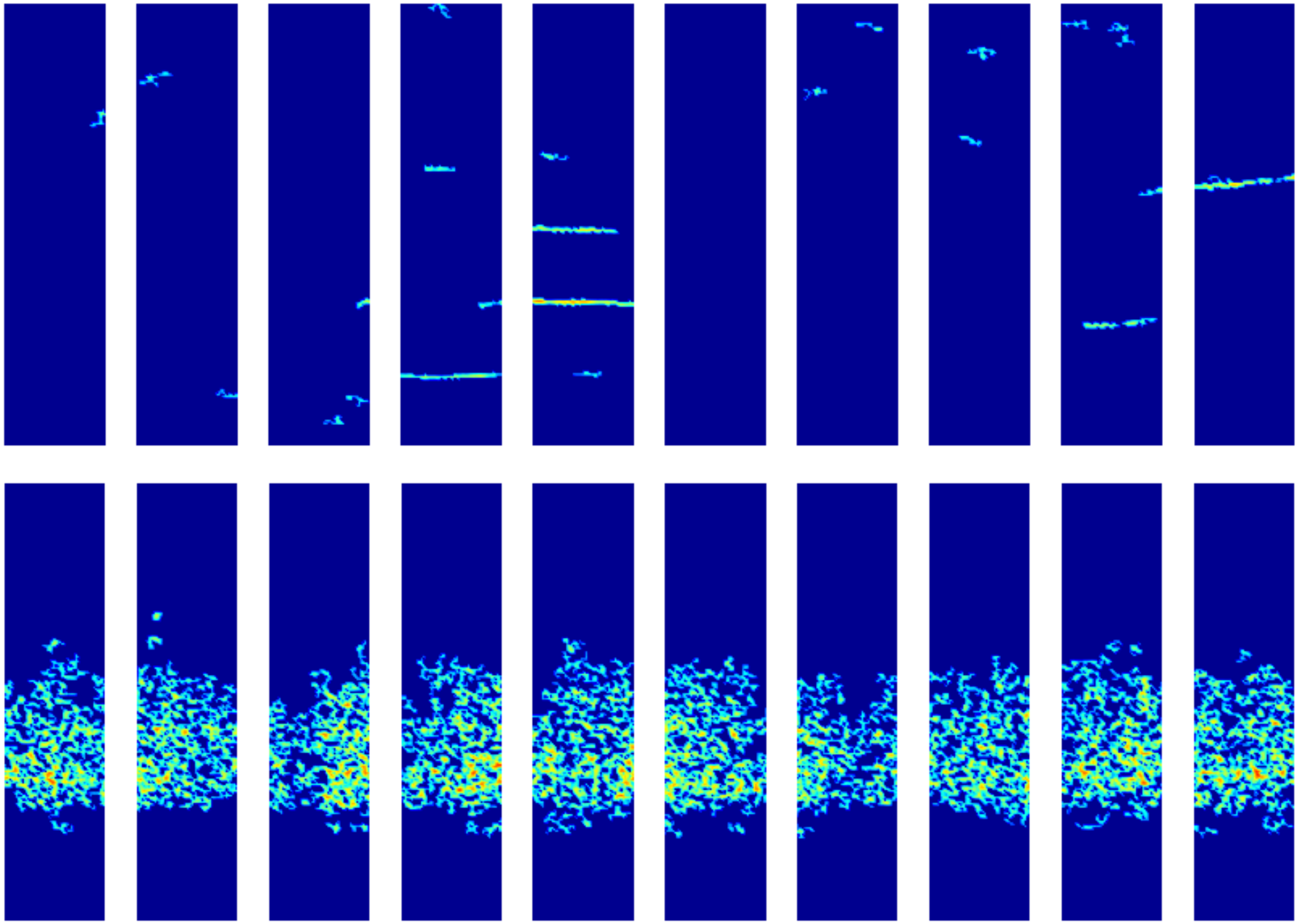


These processed spectrograms are input to the machine learning step.

A closer look at the data...



Difficulty in Classification

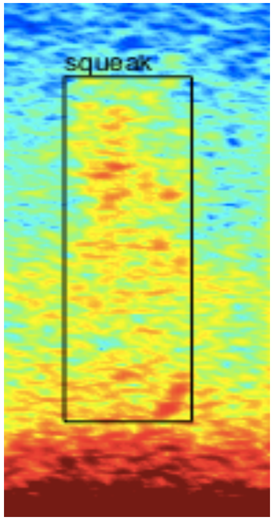


There are many difficulties in detection and classification due to differences between calls, flow noise, noise from the tag, boat noise, and so on. This limits achievable classification accuracy.

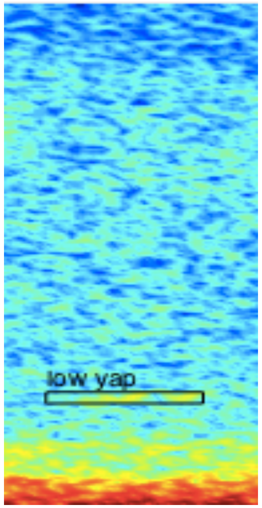
Classification Approach: SVM

We used the **Support Vector Machine** (**SVM**) classifier: a machine-learning approach for categorical classification of continuous data.

We focused on **two most common call types**: “squeaks” and “low yaps”



squeak



low yap

Training set: 2/3 of combined segments of five datasets
Testing set: 1/3 of combined segments of five datasets

To prioritize false positives over false negatives, we applied **weights** to “bias” model towards finding calls.

Classification Results

Here are the confusion matrices resulting from the classification:

	No call	**Squeak**	**Accuracy**
No call	172	14	0.925
Squeak	7	6	0.462

	No call	**Low yap**	**Accuracy**
No call	192	1	0.995
Low yap	3	3	0.5

These results are highly sensitive to the choice of class weights.

Classification error results from:

- **high-amplitude noise** from other sources that can be mistaken for calls
- **background noise** (such as flow noise) that obscures calls
- **very low ratio** of call samples to non-call samples

Conclusion

All in all, this study suggested and implemented a method for **automatically detecting and classifying humpback whale social calls** within acoustic data.

The results have much room for improvement, but are promising for future work.

Some of the error is unavoidable, due to noise in the data as mentioned previously.

We hope that classification accuracy would increase with:

- **more data** incorporated into the analysis
- **different call types** incorporated into the analysis
- **experimentation with signal processing** techniques
- **experimentation with classifiers**, such as neural networks
- **“phylogenetic” categorization** of calls?

Acknowledgements

We would like to thank:

- the **ASA**
- the **NSF**, and specifically the REU program
- our **respective institutions**
- **Greeneridge Sciences**, for producing the acoustic tags

As well as:

- **Aaron Thode**, for facilitating data collection
- **Alric Althoff**, for fantastic mentorship during this project
- **Bill Burgess**, for help with interpreting the data
- **Brooke Hawkins**, for help with interpreting the data and ecological consulting

Questions?