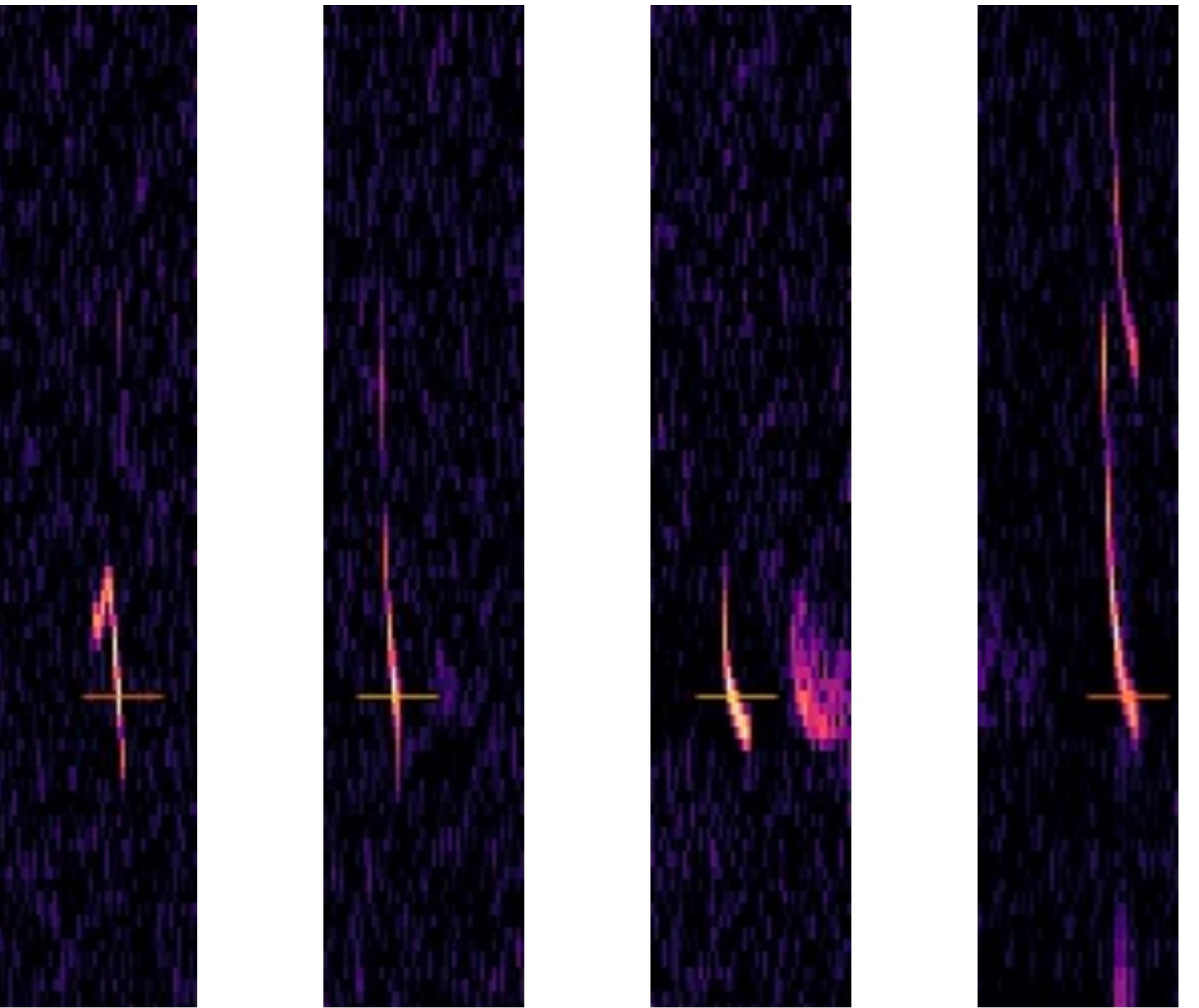


NABat Acoustic ML

Classifying North American Bats

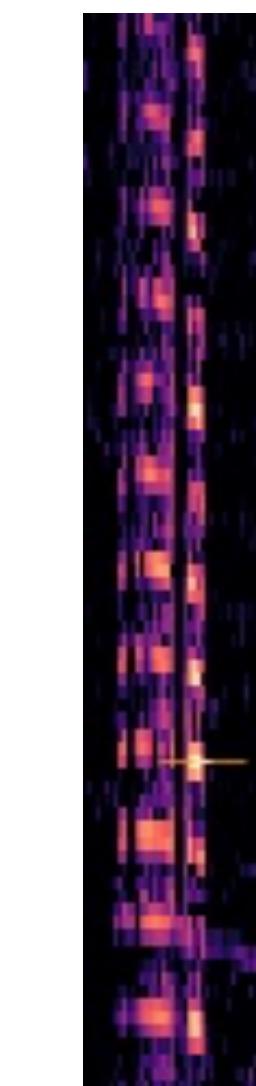
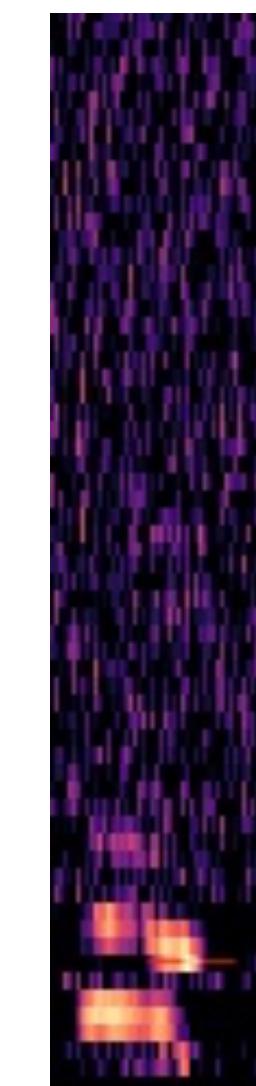
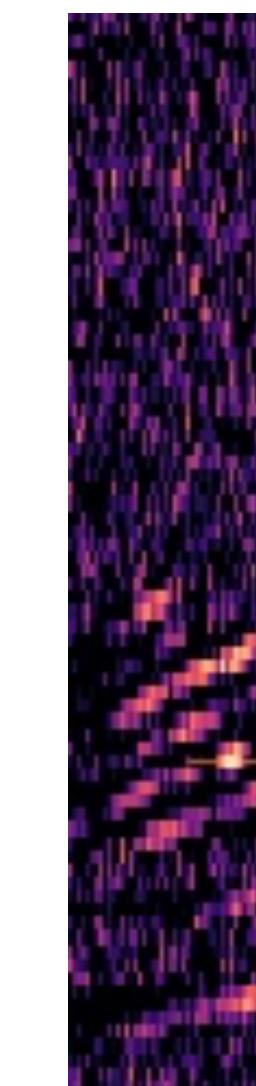
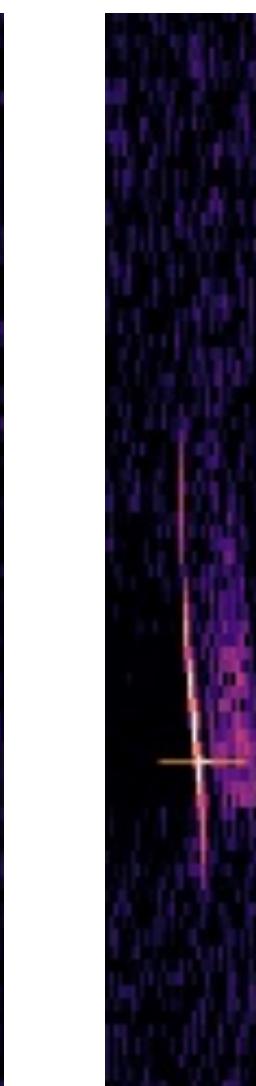
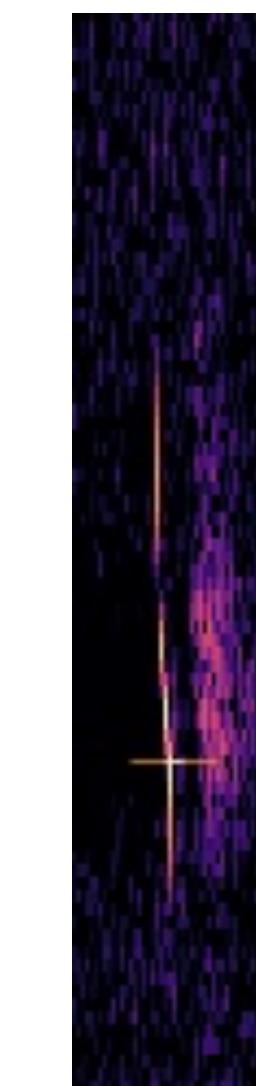
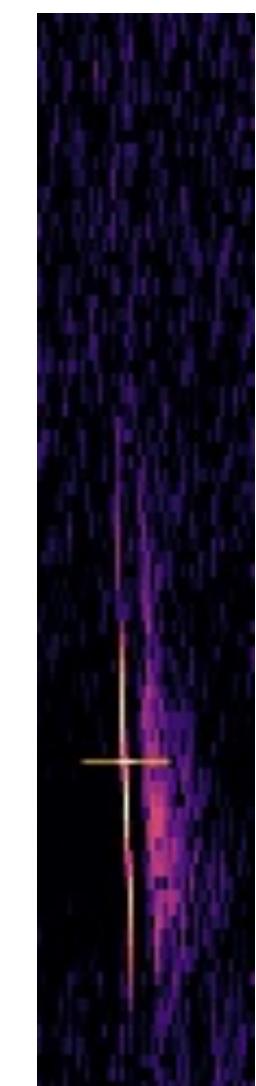
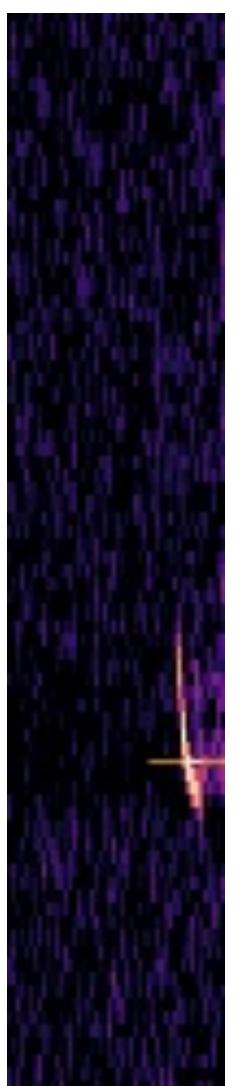


**Ben Gotthold
Ali Khalighifar
May 2021**



Objective

- Identify North American bat acoustics calls using machine learning
- Create a useful product
 - Free and open source
 - Transparent and well-documented
 - Repeatable
 - Faster than real-time
 - Works with existing NABat tools and data
 - Compatible with open source hardware like Raspberry Pi and AudioMoth



PESU

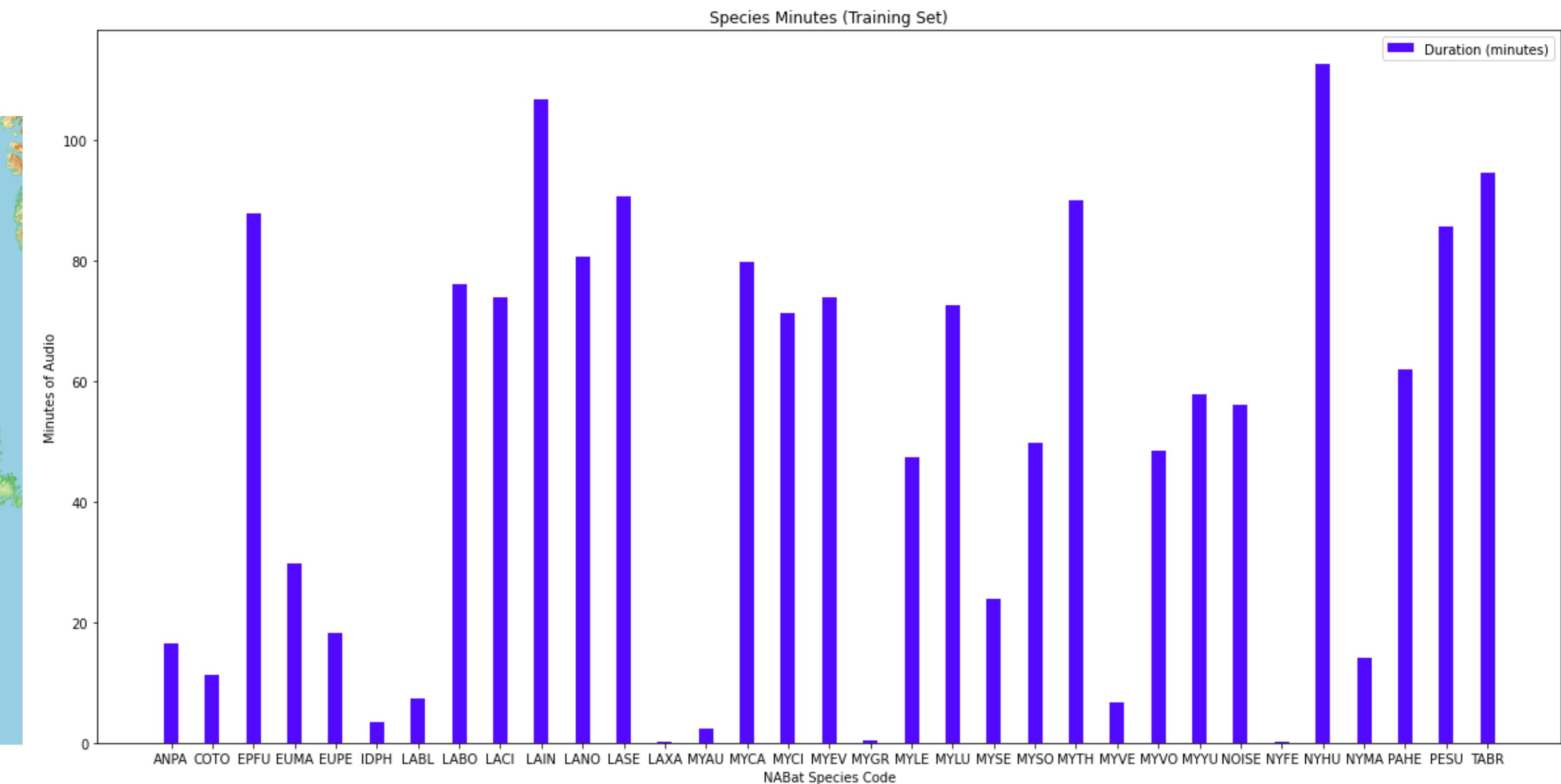
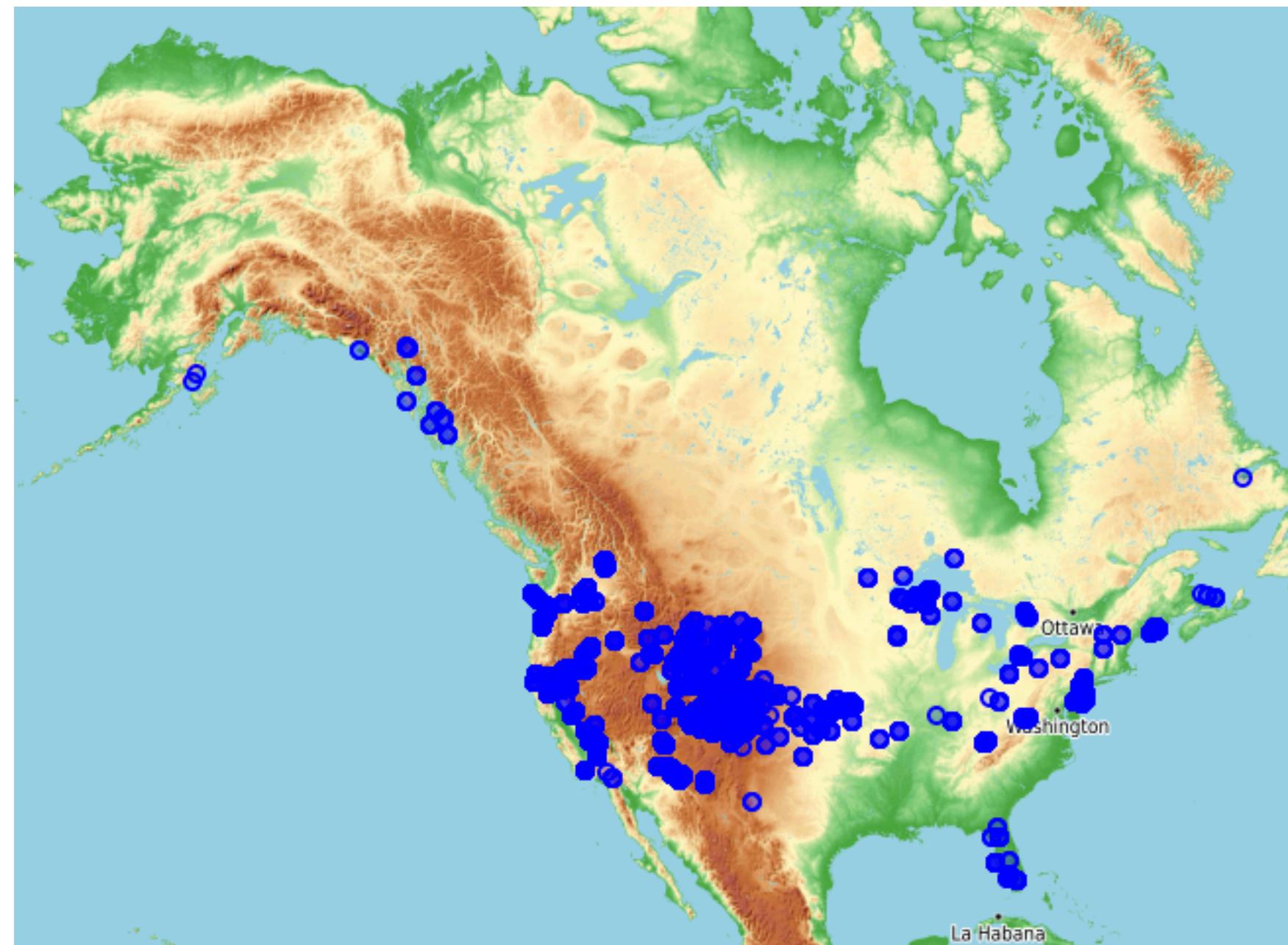
MYSE

MYLU

NOISE

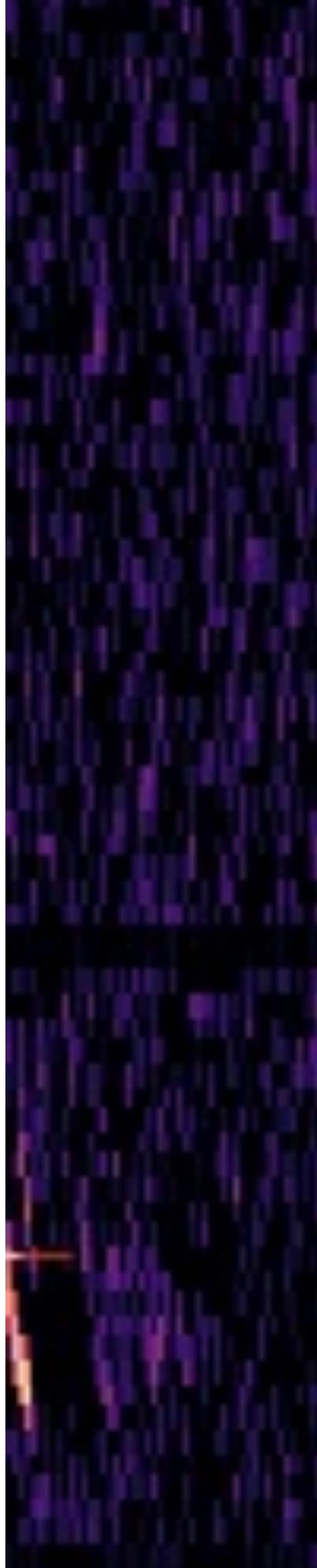
Data From NABat

- Draw up to 1000 manually reviewed acoustic files for each species class
 - Randomly order samples within each location
 - Pick samples round robin from each location until all data is exhausted or quota is met
- Set aside 80% for training
- Set aside 10% for validating
- Set aside 10% for testing

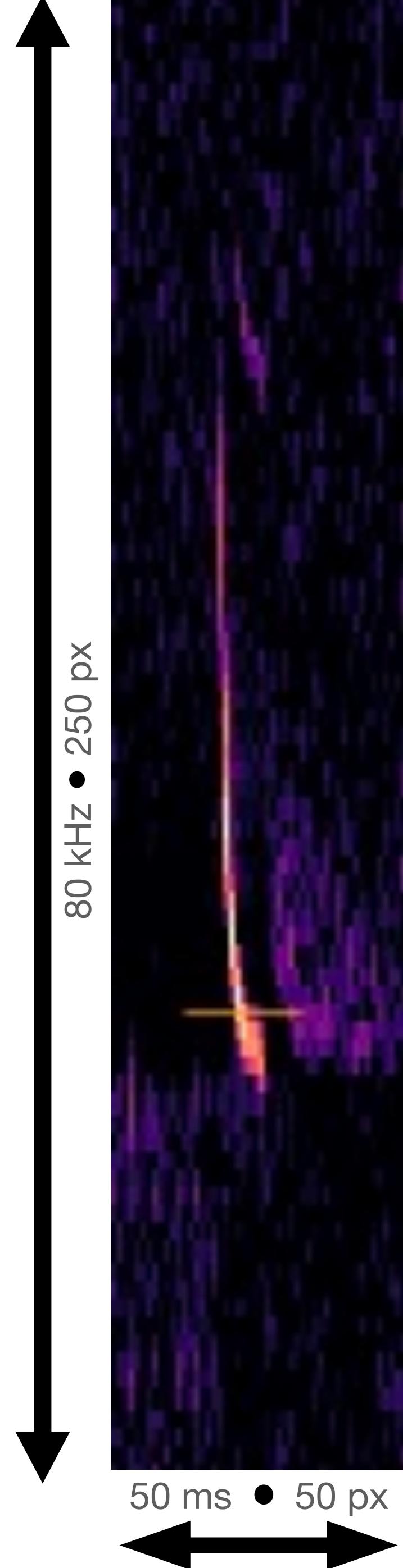
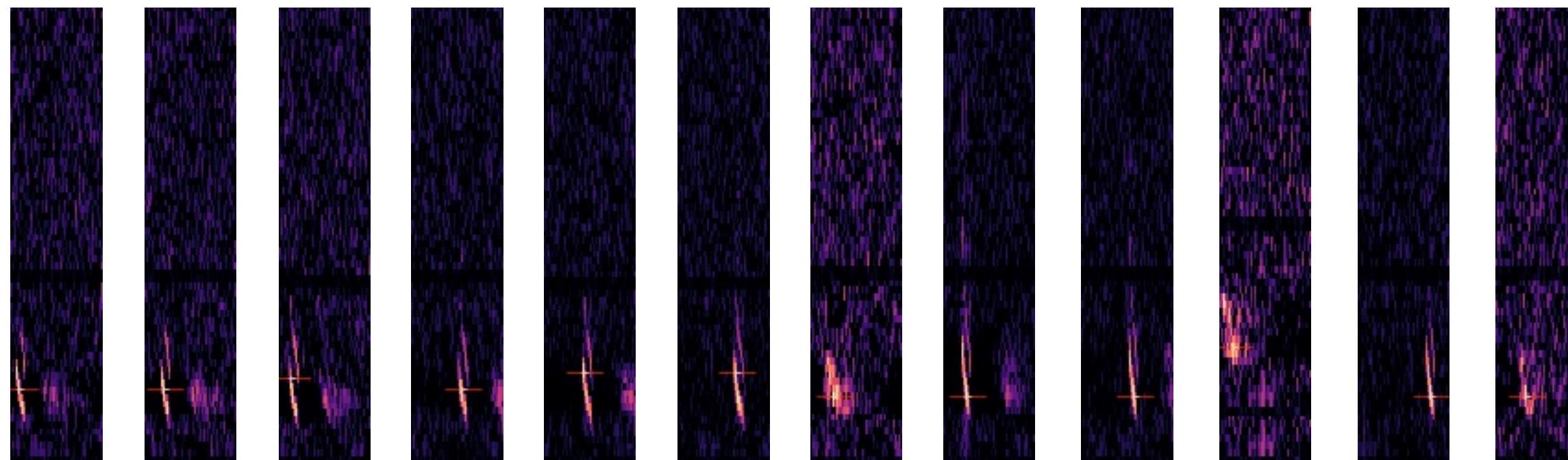
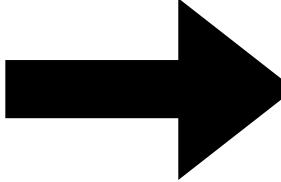


Extracting Pulses

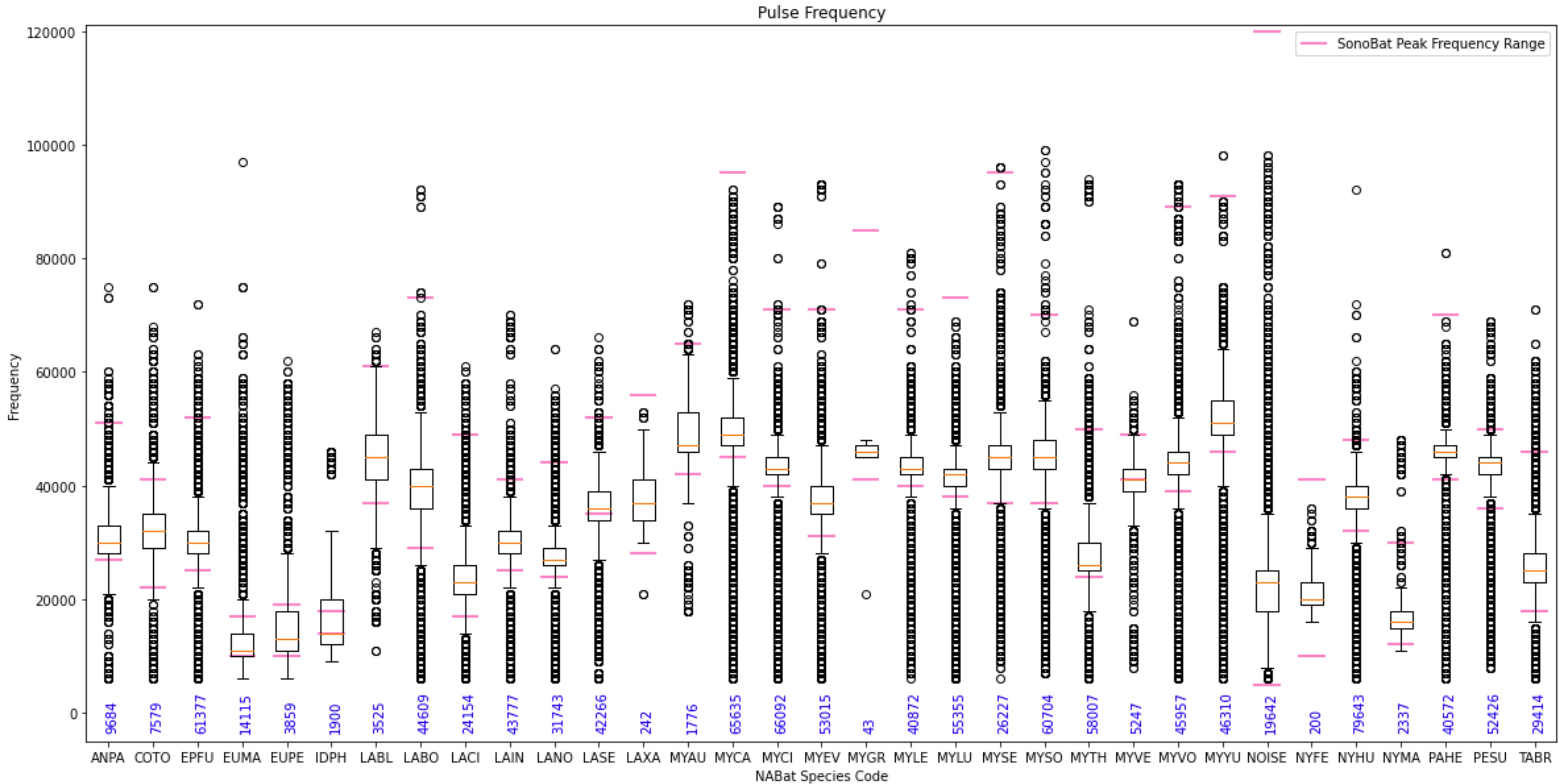
- Split wav file into 50ms slices
 - 50% overlap for training
 - 0% overlap for validation and testing
- Fast Fourier transform
 - 2D array with dimensions of frequency, time, and amplitude
- Band pass filter to remove sounds outside of 5kHz-100kHz range
- Detect peak amplitude and reject when
 - Time is 10 ms from start or end of the window
 - Amplitude $\leq 21\text{db}$
 - Ratio of peak amplitude to mean amplitude ≤ 7
- Remove noise
 - For each row, subtract the median amplitude from each value
 - For each column, subtract the median amplitude from each value
- Add horizontal targeting line centered on peak amplitude
 - 20ms long line with color scaled to frequency value
- Save as jpg image (~10kb each)



All 50ms slices of
5s recording



Extracting Pulses



Model & Training

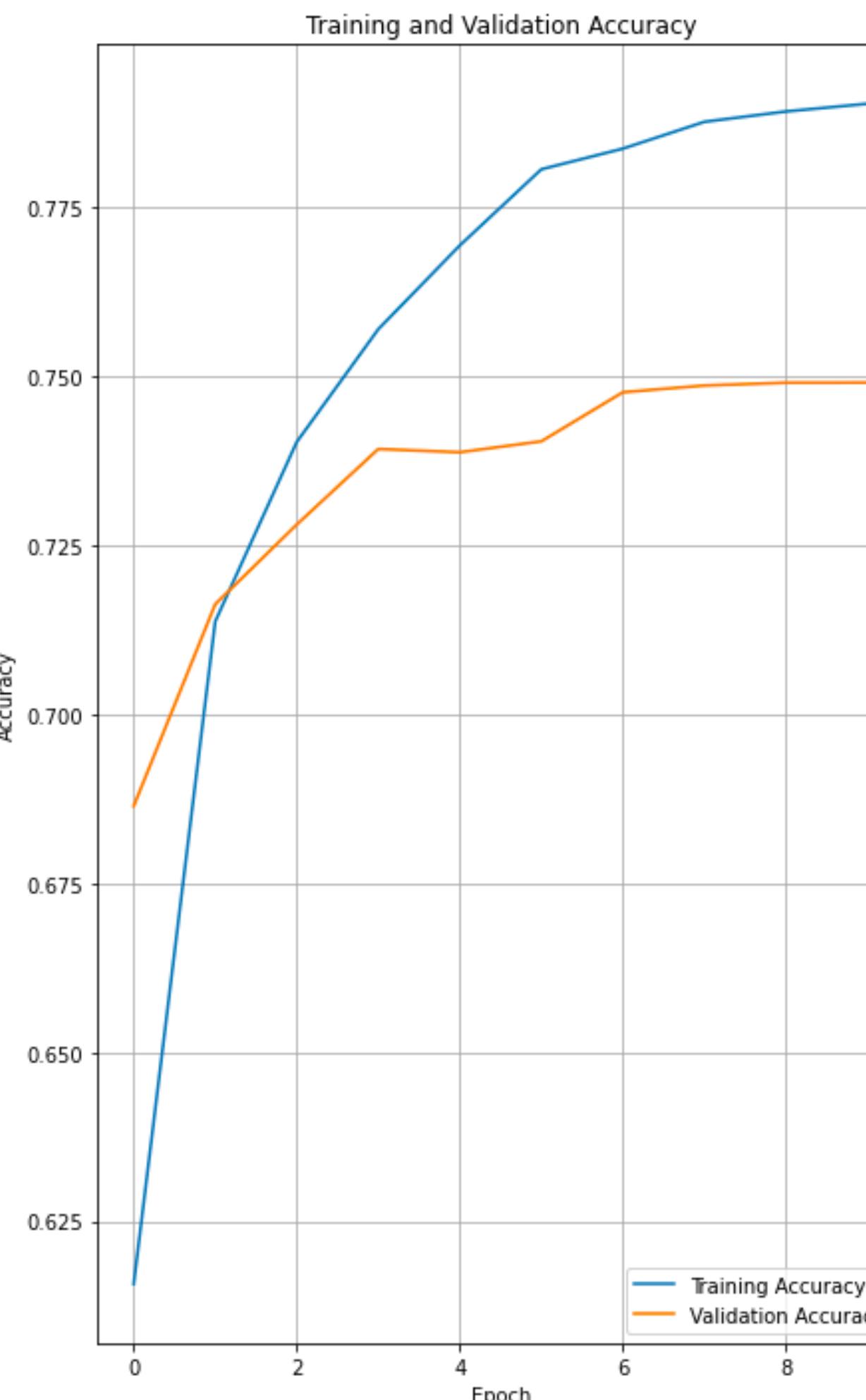
Model: "sequential"

Layer (type)	Output Shape	Param #
rescaling (Rescaling)	(None, 250, 50, 3)	0
conv2d (Conv2D)	(None, 248, 48, 32)	896
conv2d_1 (Conv2D)	(None, 246, 46, 32)	9248
max_pooling2d (MaxPooling2D)	(None, 123, 23, 32)	0
conv2d_2 (Conv2D)	(None, 121, 21, 64)	18496
conv2d_3 (Conv2D)	(None, 119, 19, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 59, 9, 64)	0
conv2d_4 (Conv2D)	(None, 59, 9, 128)	73856
conv2d_5 (Conv2D)	(None, 59, 9, 128)	147584
conv2d_6 (Conv2D)	(None, 59, 9, 128)	147584
max_pooling2d_2 (MaxPooling2D)	(None, 29, 4, 128)	0
dropout (Dropout)	(None, 29, 4, 128)	0
flatten (Flatten)	(None, 14848)	0
dense (Dense)	(None, 256)	3801344
dense_1 (Dense)	(None, 25)	6425
<hr/>		
Total params:	4,242,361	
Trainable params:	4,242,361	
Non-trainable params:	0	

```
# define a CNN model
model = Sequential([
    layers.Conv2D(32, 3, activation='relu', input_shape=(img_height, img_width, 3)),
    layers.Conv2D(32, 3, activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(64, 3, activation='relu'),
    layers.Conv2D(64, 3, activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(128, 3, activation='relu', padding='same'),
    layers.Conv2D(128, 3, activation='relu', padding='same'),
    layers.Conv2D(128, 3, activation='relu', padding='same'),
    layers.MaxPooling2D(),
    layers.Dropout(0.4),
    layers.Flatten(),
    layers.Dense(256, activation='relu'),
    layers.Dense(num_classes, activation='softmax')
])
```

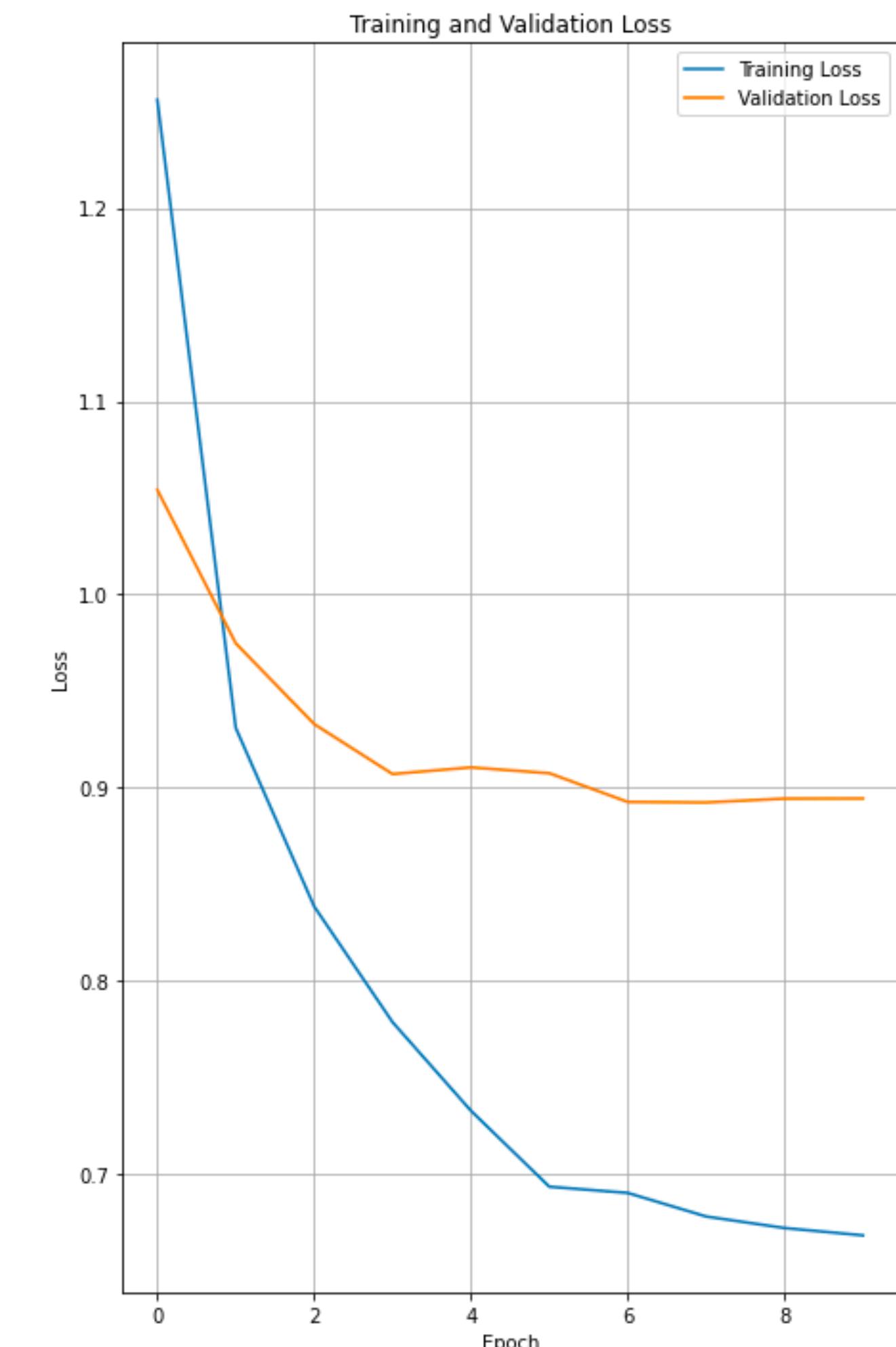
Round 1

- Epochs: 6
- Optimizer: Adam
- Learning rate: **0.0001**
- Loss: Sparse categorical cross-entropy
- 1,028,896 spectrogram images
- ~85 minutes on Amazon ml.p3.2xlarge



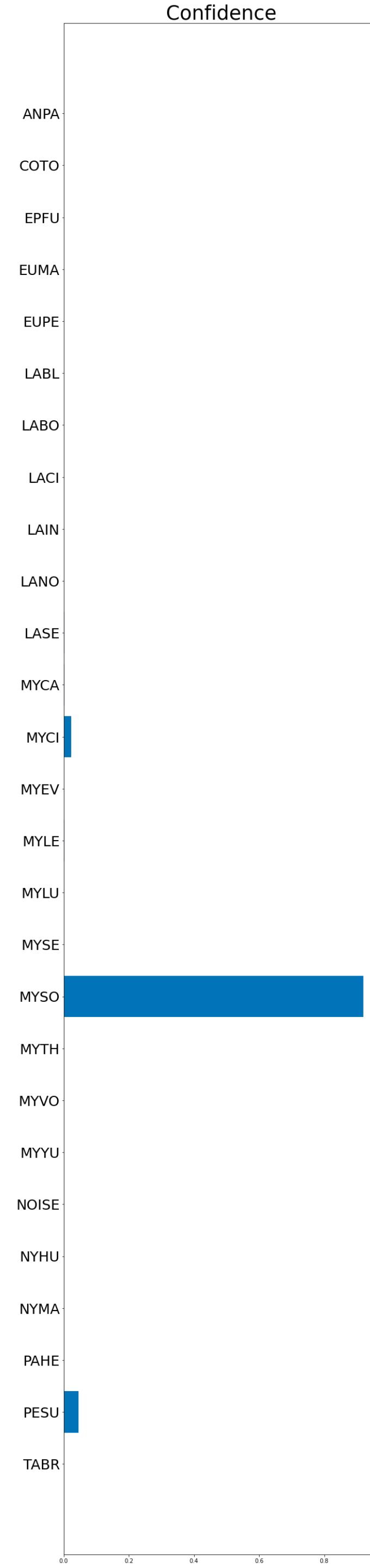
Round 2

- Epochs: 4
- Optimizer: Adam
- Learning rate: **0.000001**
- Loss: Sparse categorical cross-entropy
- 1,028,896 spectrogram images
- ~55 minutes on Amazon ml.p3.2xlarge



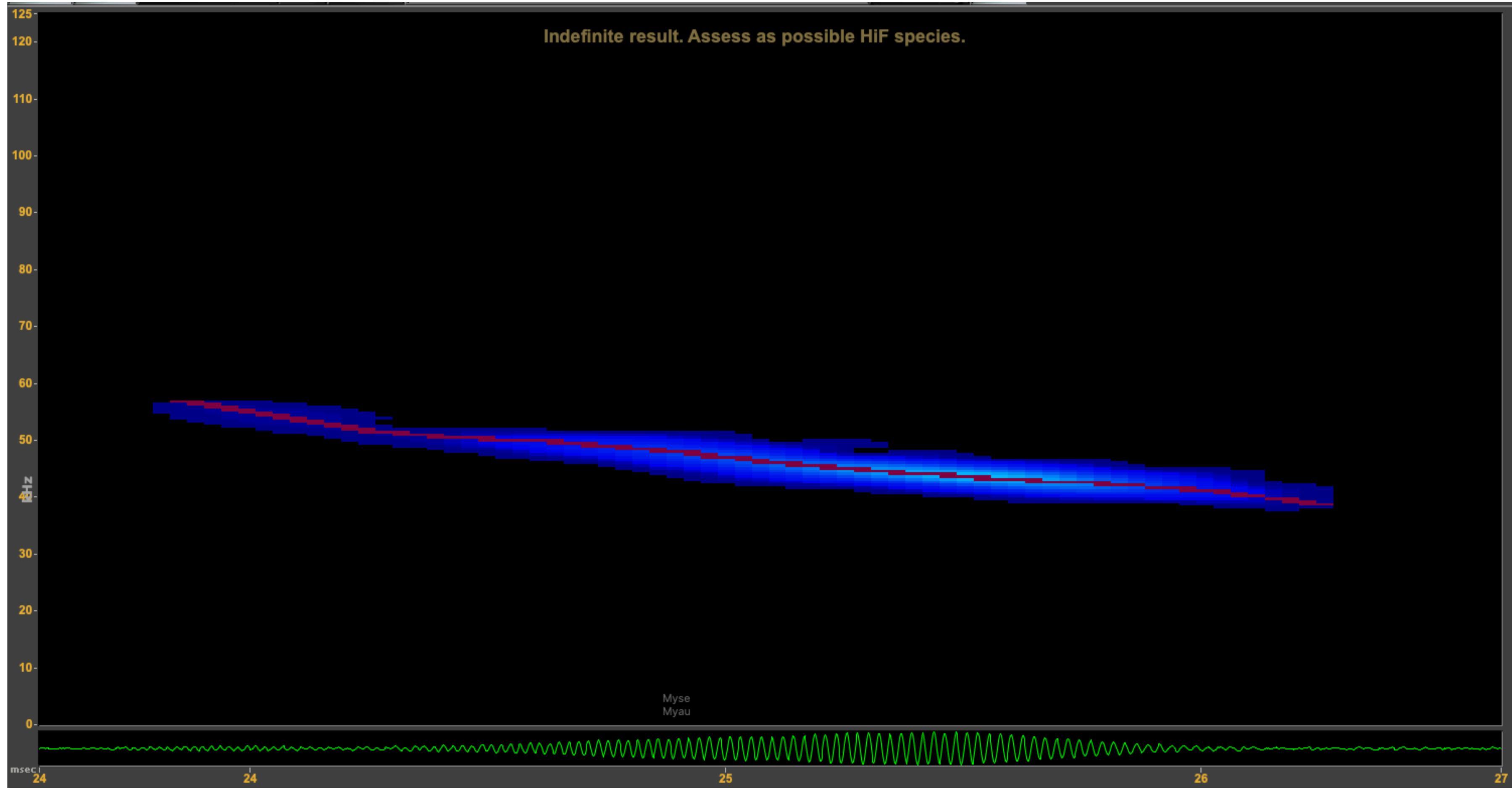
Evaluation - Pulse

260.0 | 23.2
44000.0 Hz
42.960 db
27.12 s/n
MYSO : 92.020



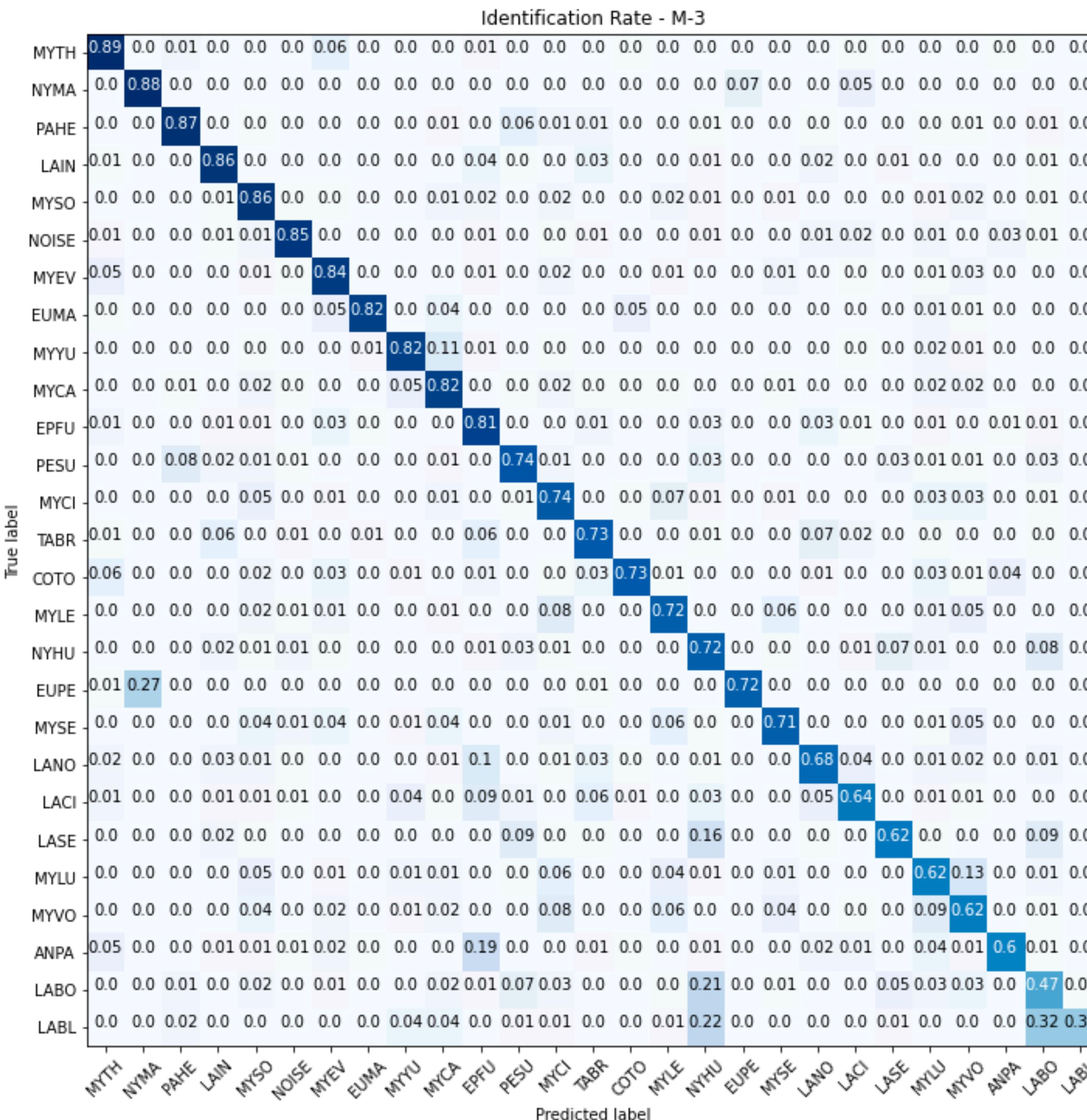
- Frequency
- Amplitude
- Time
- Signal to noise (peak amplitude / mean amplitude)
- Prediction confidence

data/test/wav/MYSO/p000_g68860_i136.wav



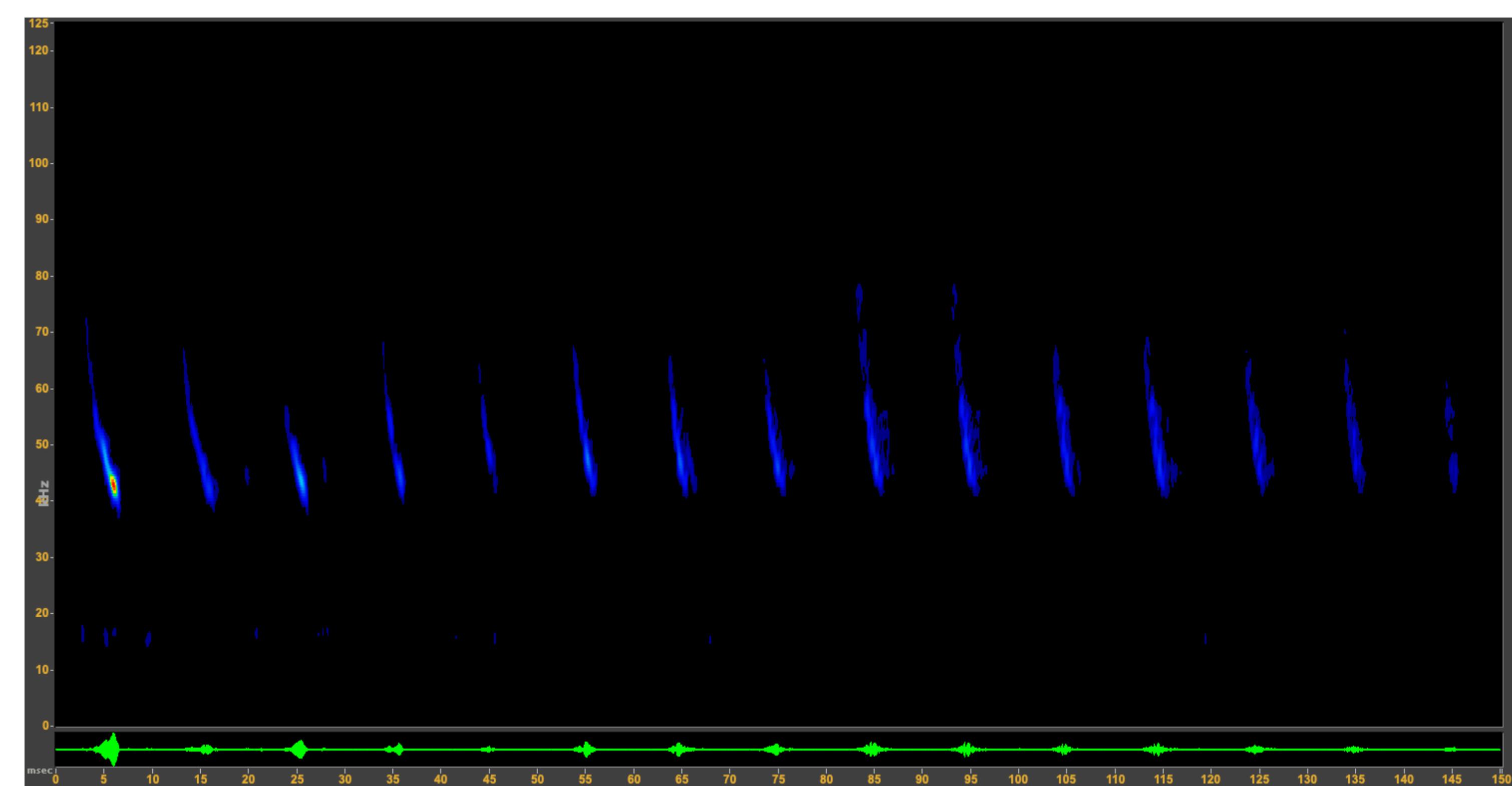
Validation Set - Pulse

- 2051 files in validation set
 - 64,388 pulses

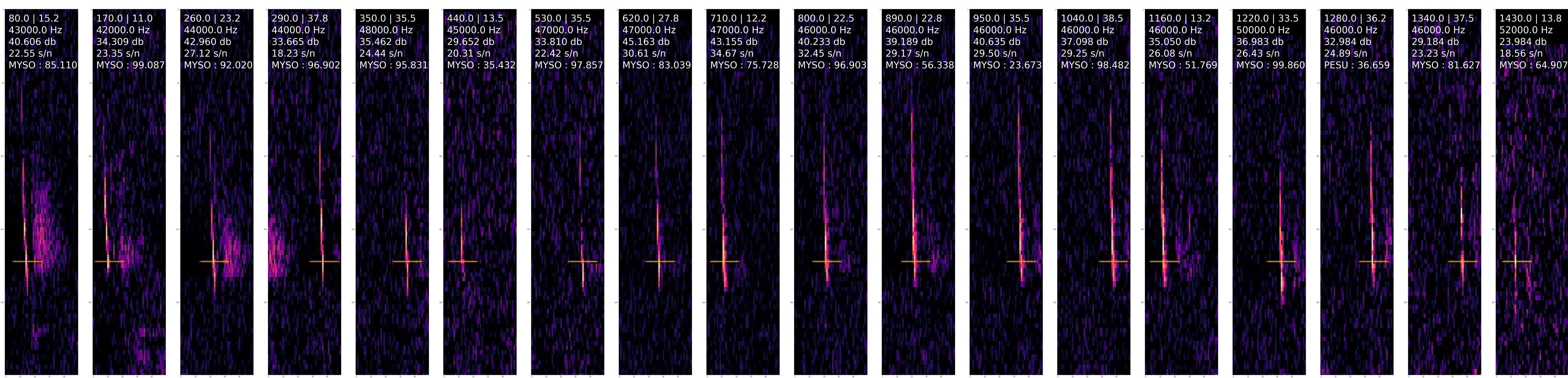


	precision	recall	f1-score	support
ANPA	0.76	0.60	0.67	776
COTO	0.71	0.73	0.72	448
EPFU	0.75	0.81	0.78	3847
EUMA	0.83	0.82	0.83	640
EUPE	0.85	0.72	0.78	157
LABL	0.53	0.33	0.40	122
LABO	0.54	0.47	0.50	2643
LACI	0.73	0.64	0.68	1517
LAIN	0.80	0.86	0.83	2318
LANO	0.74	0.68	0.71	2263
LASE	0.75	0.62	0.68	3005
MYCA	0.82	0.82	0.82	4503
MYCI	0.76	0.74	0.75	4586
MYEV	0.80	0.84	0.82	3471
MYLE	0.68	0.72	0.70	2629
MYLU	0.68	0.62	0.65	3168
MYSE	0.68	0.71	0.70	1531
MYSO	0.77	0.86	0.81	3995
MYTH	0.88	0.89	0.88	3762</td

Evaluation - File

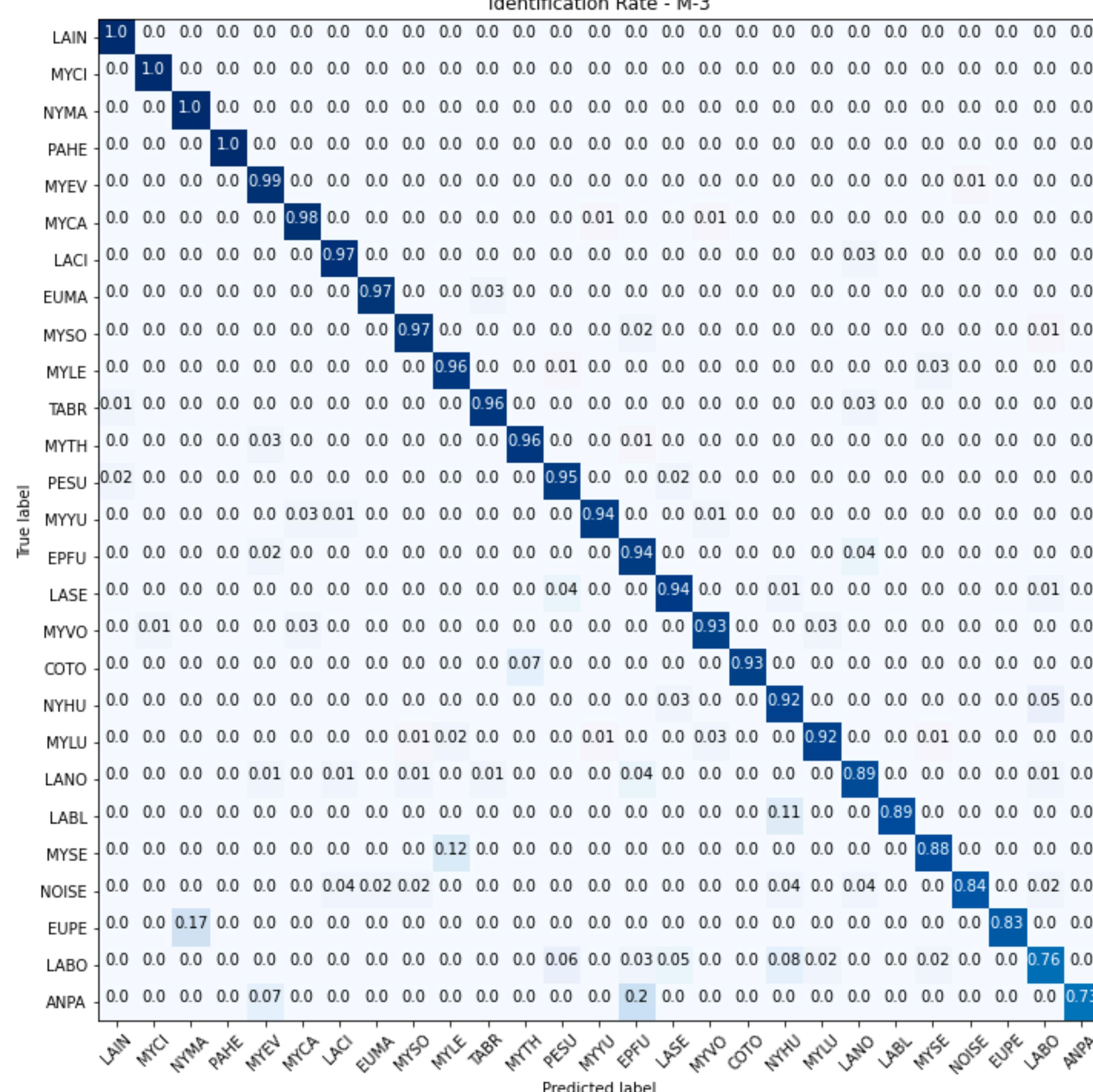


- Number of pulses
- Number of empty frames
- Number of noisy frames
- Distance between pulses
- Combined confidence



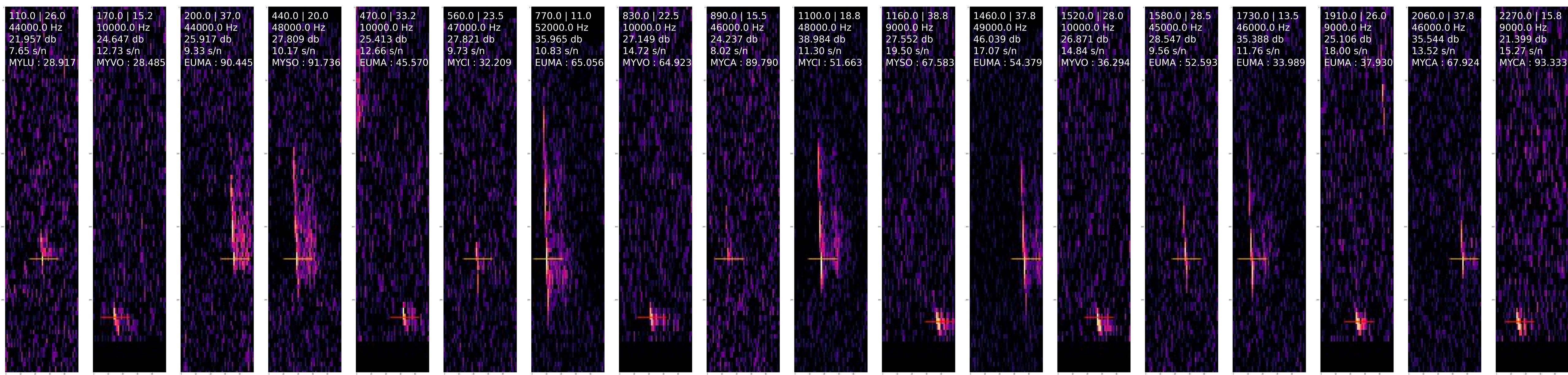
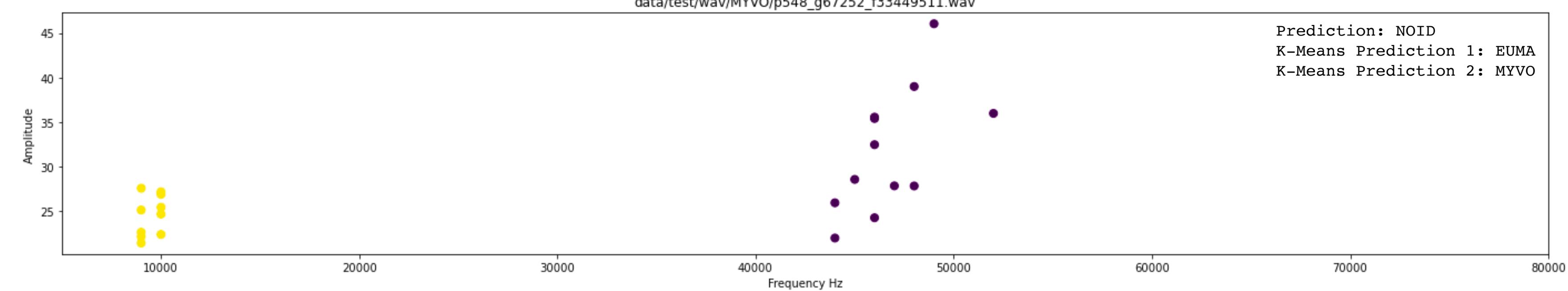
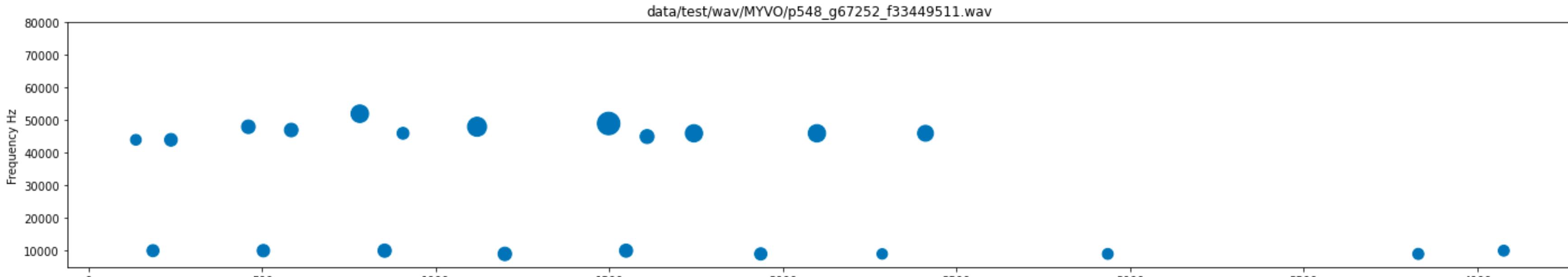
Test Set - File

- 2,052 files in test set
 - 65020 pulses
- 19 non-noise files (0.9%) were labeled NOID
 - No pulses



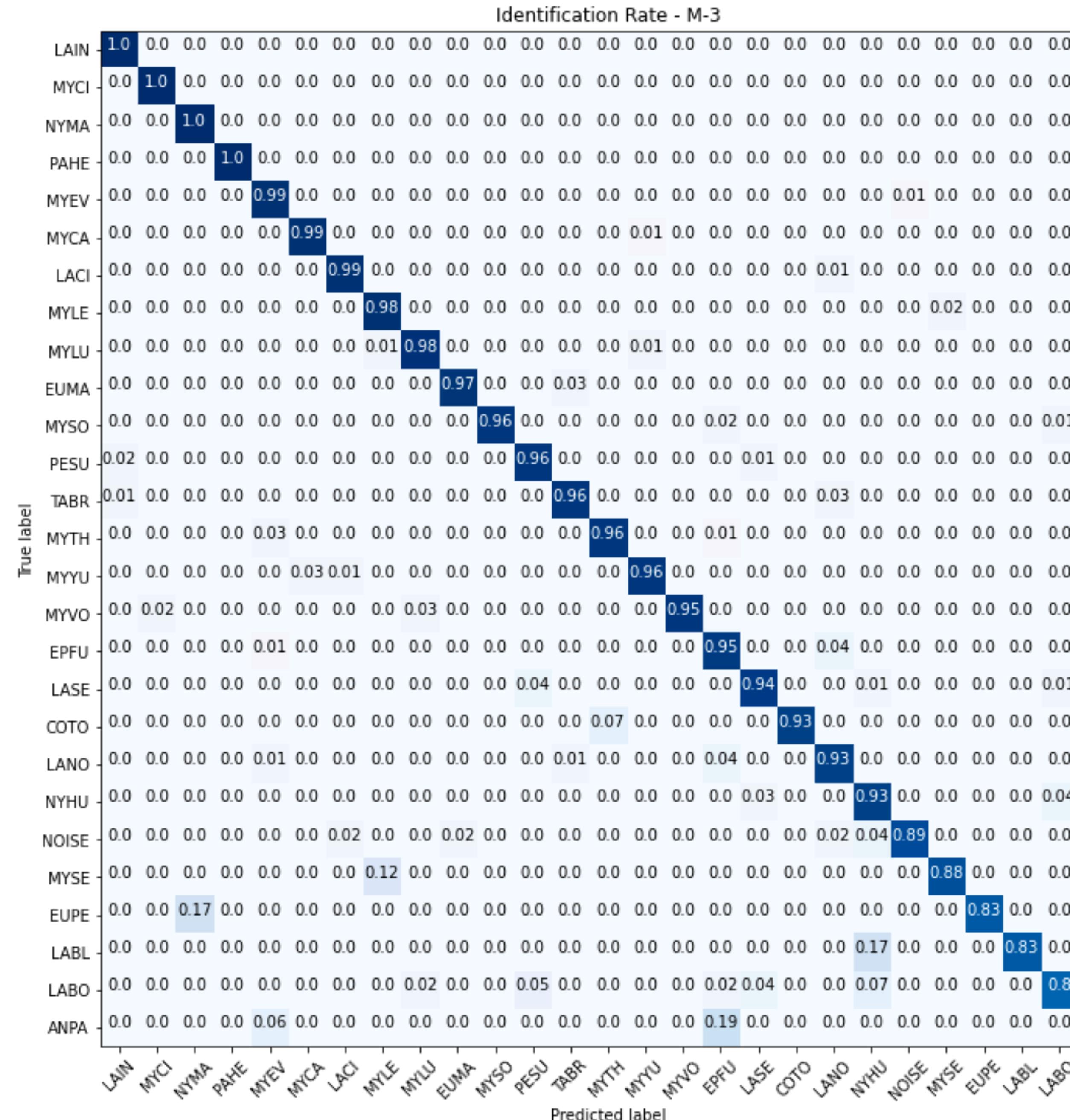
	precision	recall	f1-score	support
ANPA	0.87	0.65	0.74	20
COTO	1.00	0.93	0.96	14
EPFU	0.85	0.93	0.89	107
EUMA	0.88	0.95	0.91	37
EUPE	0.71	0.83	0.77	6
LABL	1.00	0.55	0.71	11
LABO	0.86	0.69	0.76	86
LACI	0.94	0.94	0.94	85
LAIN	0.94	1.00	0.97	94
LANO	0.86	0.79	0.82	89
LASE	0.85	0.91	0.88	92
MYCA	0.91	0.96	0.94	100
MYCI	0.91	0.97	0.94	99
MYEV	0.93	0.97	0.95	115
MLE	0.90	0.89	0.90	117
MYLU	0.93	0.86	0.89	108
MYSE	0.76	0.82	0.78	38
MYSO	0.92	0.97	0.95	89
MYTH	0.93	0.93	0.93	95
MYVO	0.83	0.87	0.85	84
MYU	0.89	0.91	0.90	79
NOISE	0.95	0.66	0.78	87
NYHU	0.85	0.90	0.88	109
NYMA	0.94	1.00	0.97	16
PAHE	0.96	1.00	0.98	69
PESU	0.88	0.91	0.90	92
TABR	0.91	0.95	0.93	76
accuracy			</td	

Evaluation - Post Processing



- Require minimum of 2 pulses
- If frequency spread > 5kHz and number of pulses >=4
 - Each k-means group average must agree
- Else
 - Average all predictions to get a single prediction
- Confirm that the predicted species buffered range overlaps the recording location
- Require a minimum confidence of 35%

Test Set - Post Processing



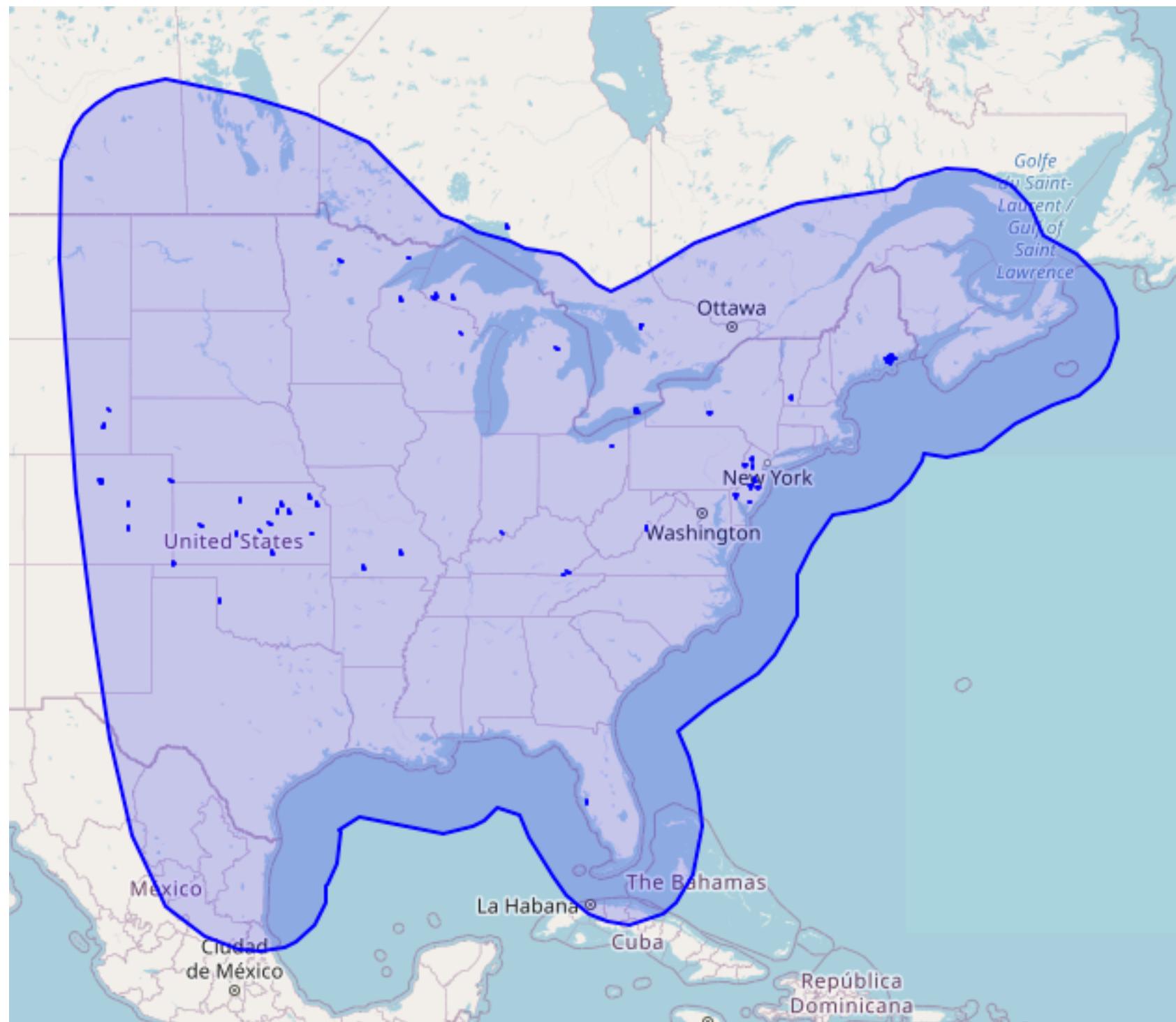
- 2,052 files in test set
- 250 non-noise files (12.2%) were labeled NOID
 - Not enough pulses (2): 35 files (1.7%)
 - Prediction outside of range: 45 files (2.2%)
 - K-means disagreement: 106 files (5.2%)
 - Not enough confidence (0.35): 64 files (3.1%)

	precision	recall	f1-score	support
ANPA	1.00	0.75	0.86	16
COTO	1.00	0.93	0.96	14
EPFU	0.90	0.95	0.93	100
EUMA	0.97	0.97	0.97	35
EUPE	1.00	0.83	0.91	6
LABL	1.00	0.83	0.91	6
LABO	0.88	0.80	0.84	55
LACI	0.97	0.99	0.98	71
LAIN	0.97	1.00	0.98	93
LANO	0.89	0.93	0.91	70
LASE	0.93	0.94	0.93	82
MYCA	0.98	0.99	0.98	90
MYCI	0.99	1.00		

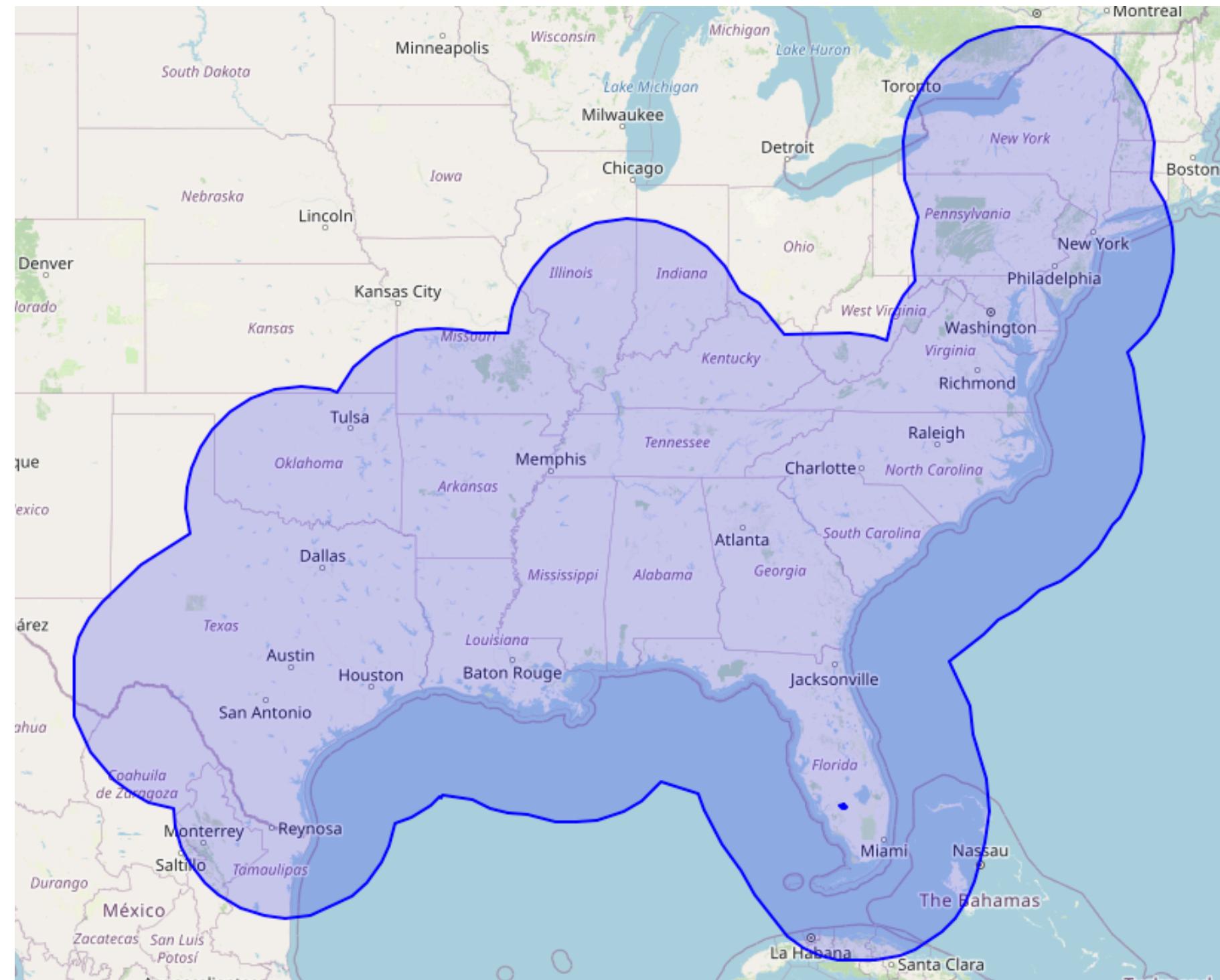
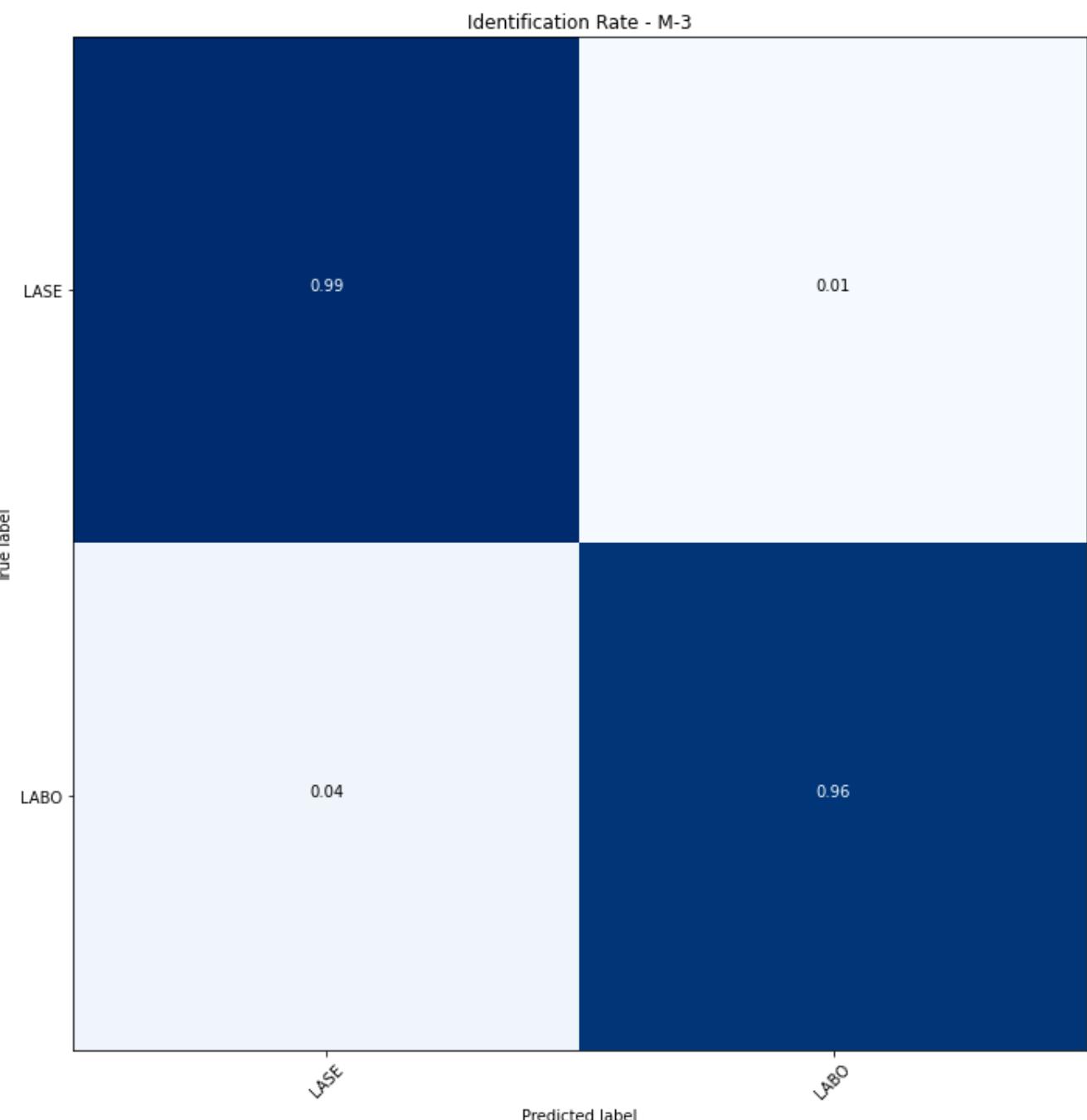
LABO - LASE

	precision	recall	f1-score	support
LABO	0.88	0.80	0.84	55
LASE	0.93	0.94	0.93	82
micro avg	0.91	0.88	0.90	137
macro avg	0.90	0.87	0.89	137
weighted avg	0.91	0.88	0.90	137

- LABO
 - 55 Manual Id Files
 - Geographically diverse sample (58 GRTS)
- LASE
 - 82 Manual Id Files
 - Geographically similar sample (2 GRTS)



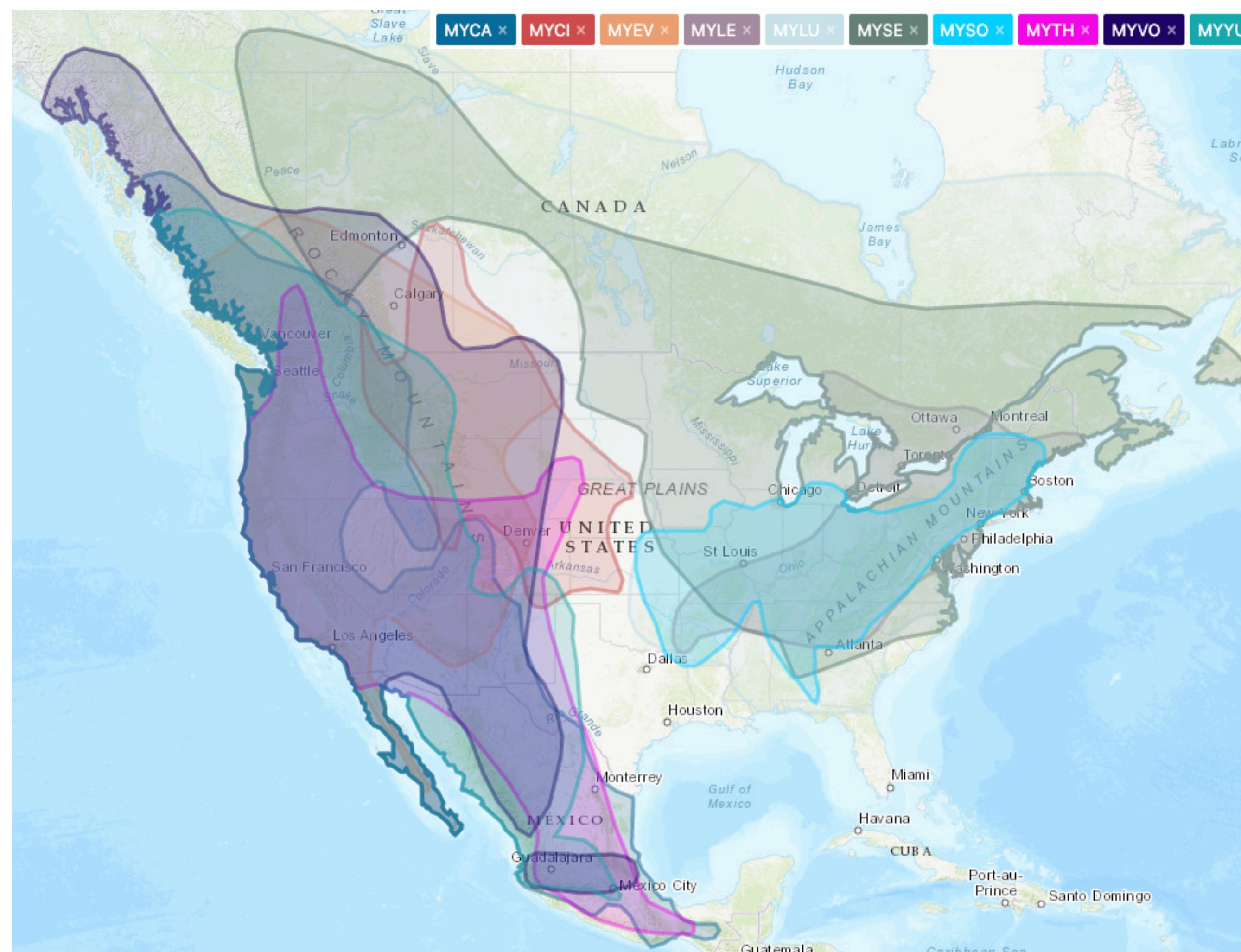
LABO - Eastern Red Bat



LASE - Seminole Bat

Myotis

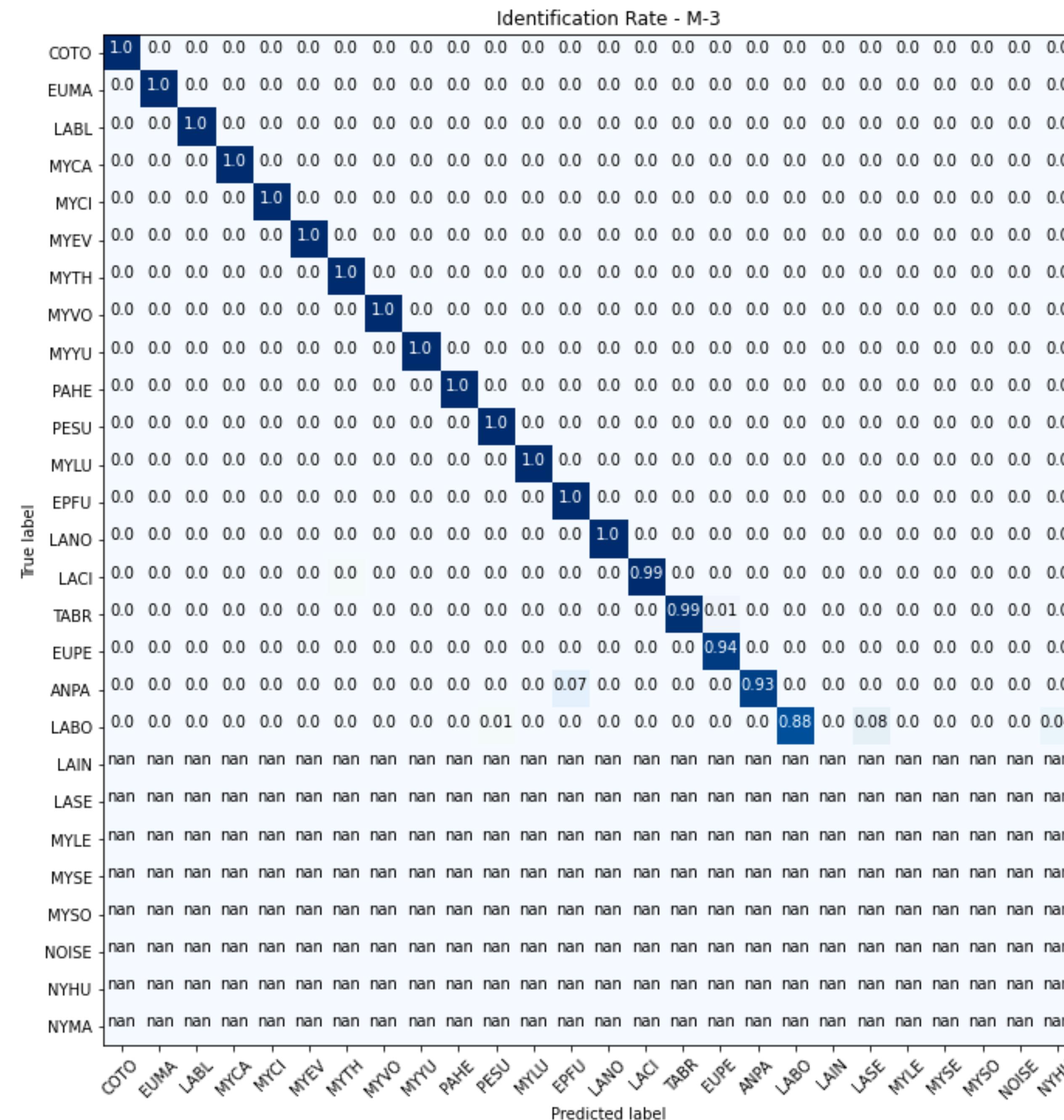
- MYCA *Californicus* - California Myotis
- MYCI *Ciliolabrum* - Western Small-footed Myotis
- MYEV *Evotis* - Western long-eared Myotis
- MYLE *Leibii* - Eastern Small-footed Myotis
- MYLU *Lucifugus* - Little Brown Bat
- MYSE *Septentrionalis* - Northern long-eared Bat
- MYSO *Sodalis* - Indiana Bat
- MYTH *Thysanodes* - Fringed Myotis
- MYVO *Volans* - Long-legged Myotis
- MYYU *Yumanensis* - Yuma Myotis



Identification Rate - M-3										
	MYCI	MYEV	MYCA	MYLE	MYLU	MYSO	MYTH	MYYU	MYVO	MYSE
True label	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MYCI	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MYEV	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MYCA	0.0	0.0	0.99	0.0	0.0	0.0	0.0	0.01	0.0	0.0
MYLE	0.0	0.0	0.0	0.98	0.0	0.0	0.0	0.0	0.0	0.02
MYLU	0.0	0.0	0.0	0.01	0.98	0.0	0.0	0.01	0.0	0.0
MYSO	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
MYTH	0.0	0.03	0.0	0.0	0.0	0.0	0.97	0.0	0.0	0.0
MYYU	0.0	0.0	0.03	0.0	0.0	0.0	0.0	0.97	0.0	0.0
MYVO	0.02	0.0	0.0	0.0	0.03	0.0	0.0	0.0	0.95	0.0
MYSE	0.0	0.0	0.0	0.12	0.0	0.0	0.0	0.0	0.0	0.88

	precision	recall	f1-score	support
MYCA	0.98	0.99	0.98	90
MYCI	0.99	1.00	0.99	92
MYEV	0.95	0.99	0.97	107
MYLE	0.95	0.98	0.97	100
MYLU	0.97	0.98	0.97	85
MYSE	0.93	0.88	0.90	32
MYSO	1.00	0.96	0.98	85
MYTH	0.99	0.96	0.97	90
MYVO	1.00	0.95	0.98	66
MYYU	0.97	0.96	0.96	67
micro avg	0.97	0.97	0.97	814
macro avg	0.97	0.96	0.97	814
weighted avg	0.97	0.97	0.97	814

Other Datasets - AAA



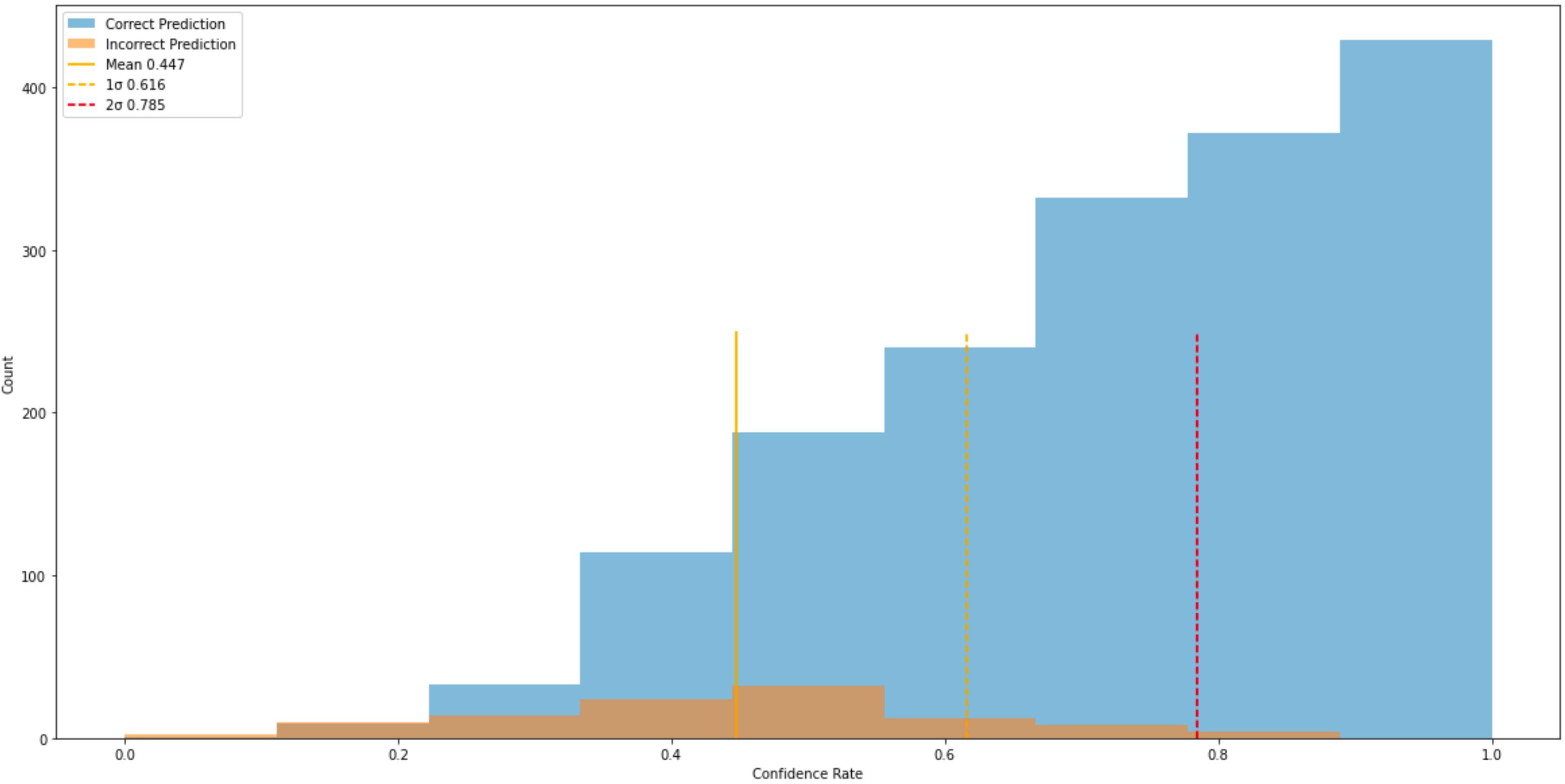
- 240 files (6.6%) were labeled NOID

- Not enough pulses
- Prediction or label not in species range
- K-means predictions disagree
- Not enough confidence

	precision	recall	f1-score	support
ANPA	1.00	0.93	0.96	14
COTO	1.00	1.00	1.00	11
EPFU	0.99	1.00	1.00	563
EUMA	1.00	1.00	1.00	4
EUPE	0.94	0.94	0.94	18
LABL	1.00	1.00	1.00	2
LABO	1.00	0.88	0.93	137
LACI	1.00	0.99	1.00	715
LAIN	0.00	0.00	0.00	0
LANO	1.00	1.00	1.00	709
LASE	0.00	0.00	0.00	0
MYCA	1.00	1.00	1.00	271
MYCI	1.00	1.00	1.00	64
MYEV	1.00	1.00	1.00	230
MLE	0.00	0.00	0.00	0
MLU	1.00	1.00	1.00	350
MYSE	0.00	0.00	0.00	0
MSO	0.00	0.00	0.00	0
MYTH	0.96	1.00	0.98	65
MYVO	1.00	1.00	1.00	23
MYYU	1.00	1.00	1.00	54
NOISE	0.00	0.00	0.00	0
NYHU	0.00	0.00	0.00	0
NYMA	0.00	0.00	0.00	0
PAHE	1.00	1.00	1.00	60
PESU	0.96	1.00	0.98	25
TABR	1.00	0.99	0.99	68
micro avg	0.99	0.99	0.99	3383
macro avg	0.70	0.69	0.70	3383
weighted avg	1.00	0.99	0.99	3383

Limitations - Dataset

- Incorrect predictions are minimally skewed
 - Some test examples are hard to predict correctly
- There exists data with incorrect labels
 - Lots of steps before reaching NABat
 - Manual identification is difficult

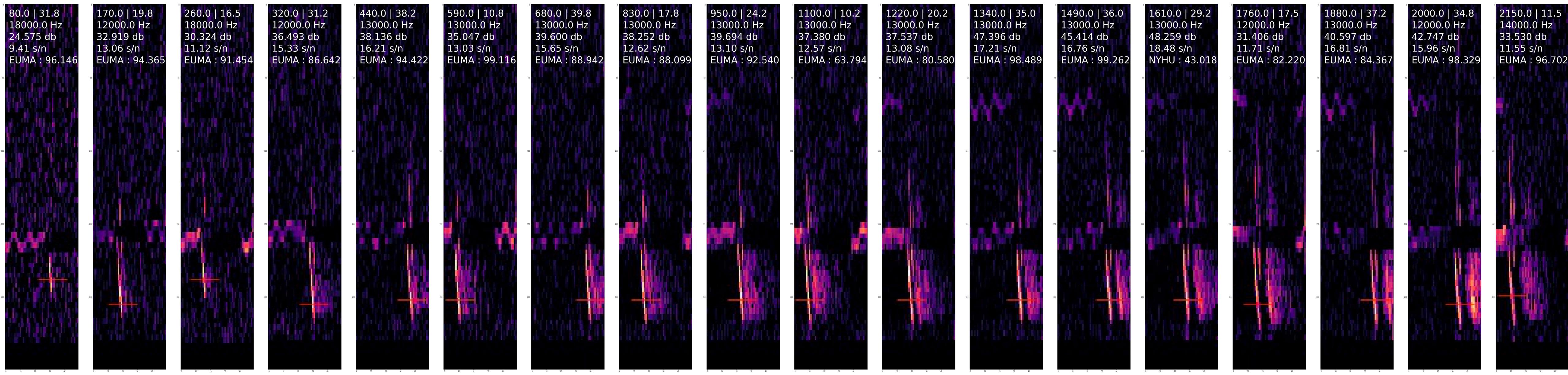


Label: NOISE

Prediction: EUMA

data/test/wav/NOISE/p548_g67252_f33447105.wav

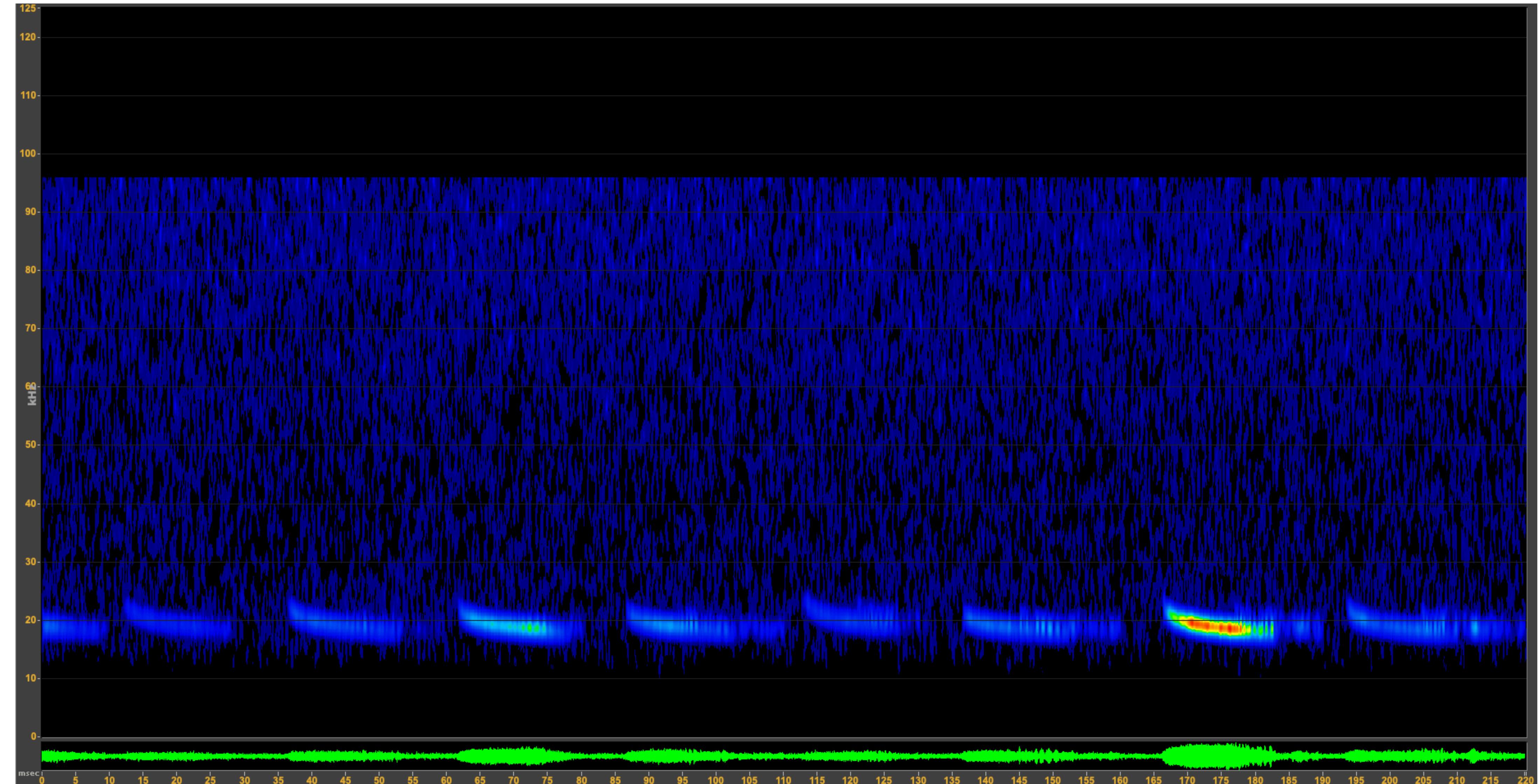
80.0 31.8 18000.0 Hz 24.575 db 9.41 s/n EUMA : 96.146	170.0 19.8 12000.0 Hz 32.919 db 13.06 s/n EUMA : 94.365	260.0 16.5 18000.0 Hz 30.324 db 11.12 s/n EUMA : 91.454	320.0 31.2 12000.0 Hz 36.493 db 15.33 s/n EUMA : 86.642	440.0 38.2 13000.0 Hz 38.136 db 16.21 s/n EUMA : 94.422	590.0 10.8 13000.0 Hz 35.047 db 13.03 s/n EUMA : 99.116	680.0 39.8 13000.0 Hz 39.600 db 15.65 s/n EUMA : 88.942	830.0 17.8 13000.0 Hz 38.252 db 12.62 s/n EUMA : 88.099	950.0 24.2 13000.0 Hz 39.694 db 13.10 s/n EUMA : 92.540	1100.0 10.2 13000.0 Hz 37.380 db 12.57 s/n EUMA : 63.794	1220.0 20.2 13000.0 Hz 37.537 db 13.08 s/n EUMA : 80.580	1340.0 35.0 13000.0 Hz 47.396 db 17.21 s/n EUMA : 98.489	1490.0 36.0 13000.0 Hz 45.414 db 16.76 s/n EUMA : 99.262	1610.0 29.2 13000.0 Hz 48.259 db 18.48 s/n NYHU : 43.018	1760.0 17.5 12000.0 Hz 31.406 db 11.71 s/n EUMA : 82.220	1880.0 37.2 13000.0 Hz 40.597 db 16.81 s/n EUMA : 84.367	2000.0 34.8 12000.0 Hz 42.747 db 15.96 s/n EUMA : 98.329	2150.0 11.5 14000.0 Hz 33.530 db 11.55 s/n EUMA : 96.702
---	---	---	---	---	---	---	---	---	--	--	--	--	--	--	--	--	--



Limitations - Pulse Detection

- Single pulse detected
 - Low amplitude
 - Noisy

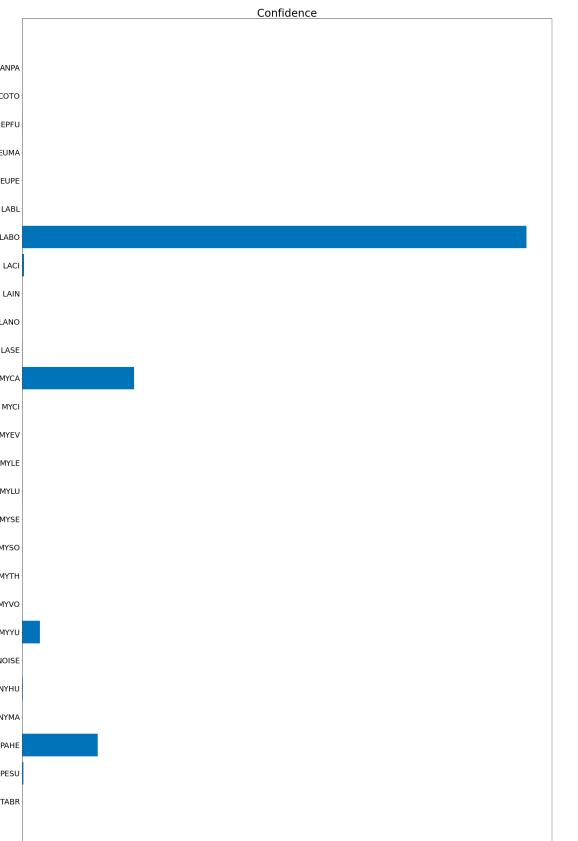
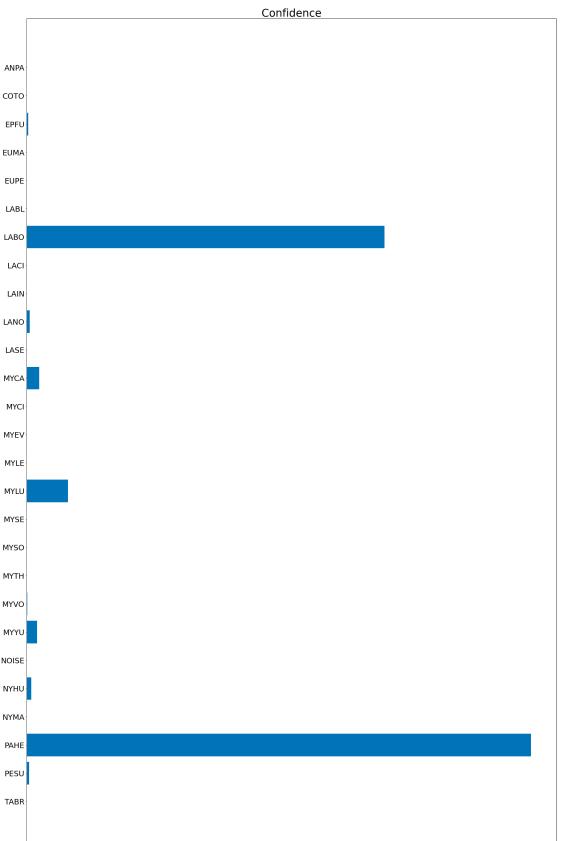
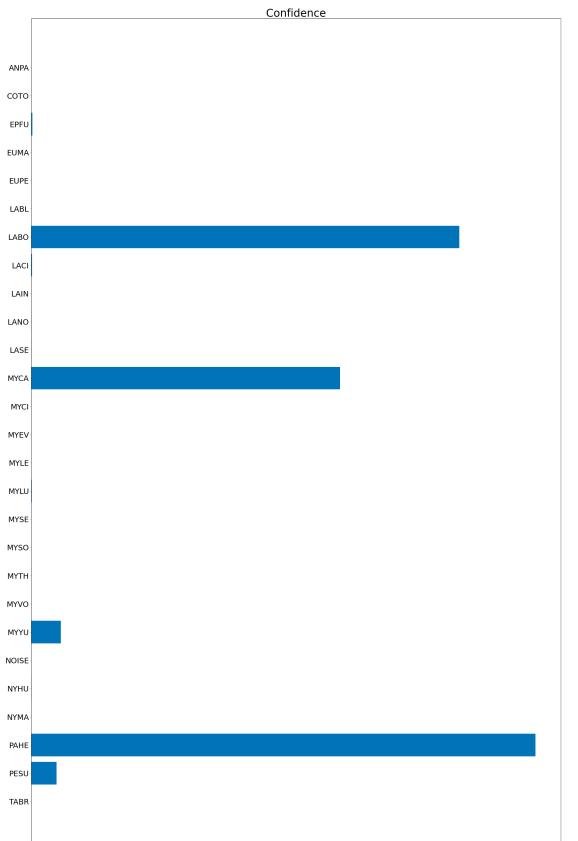
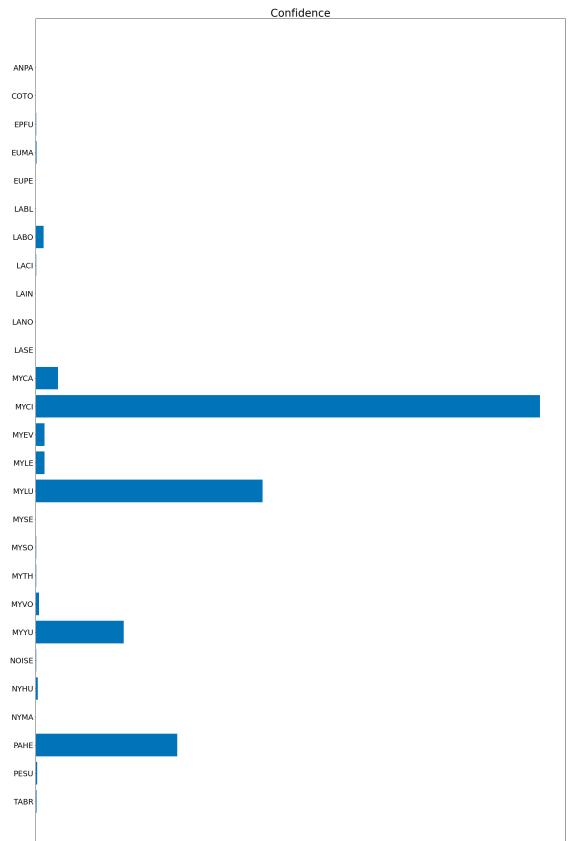
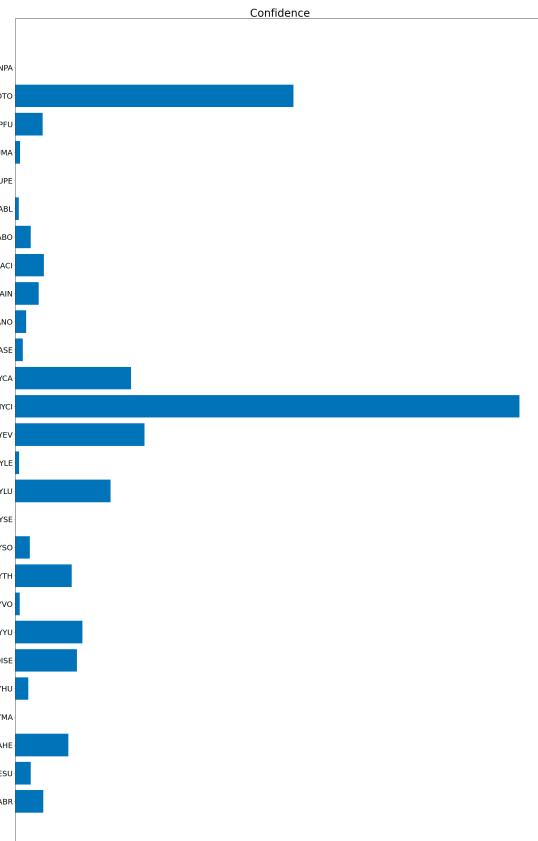
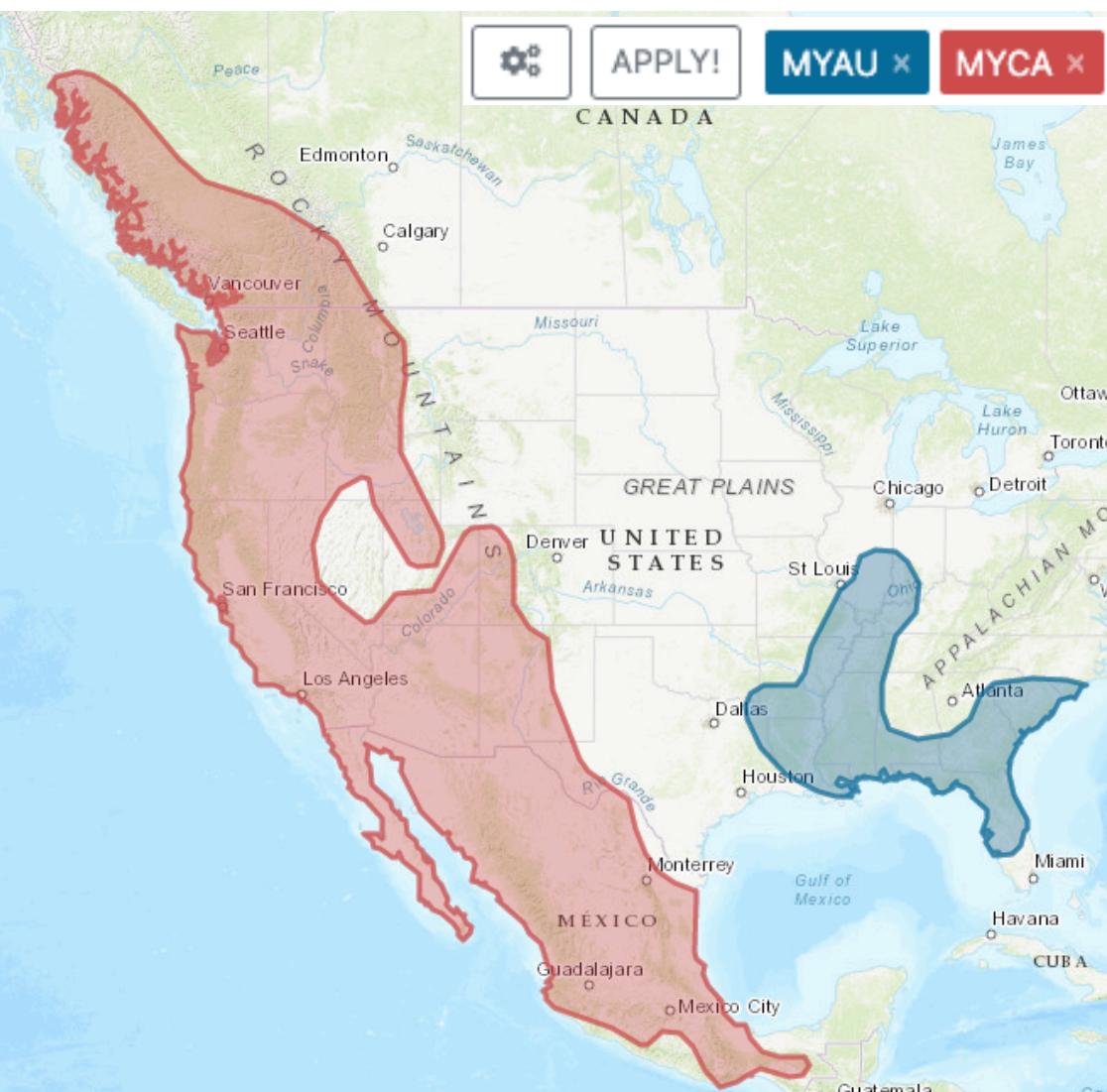
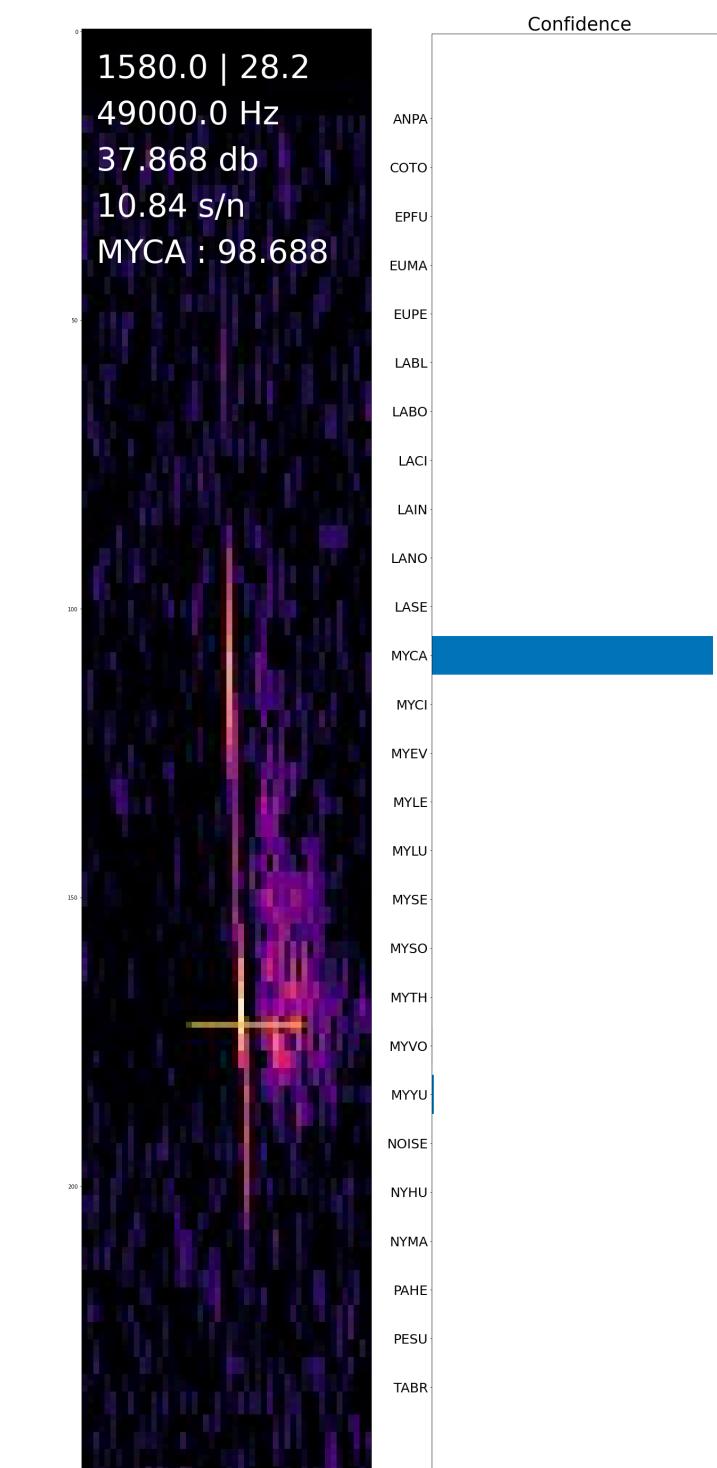
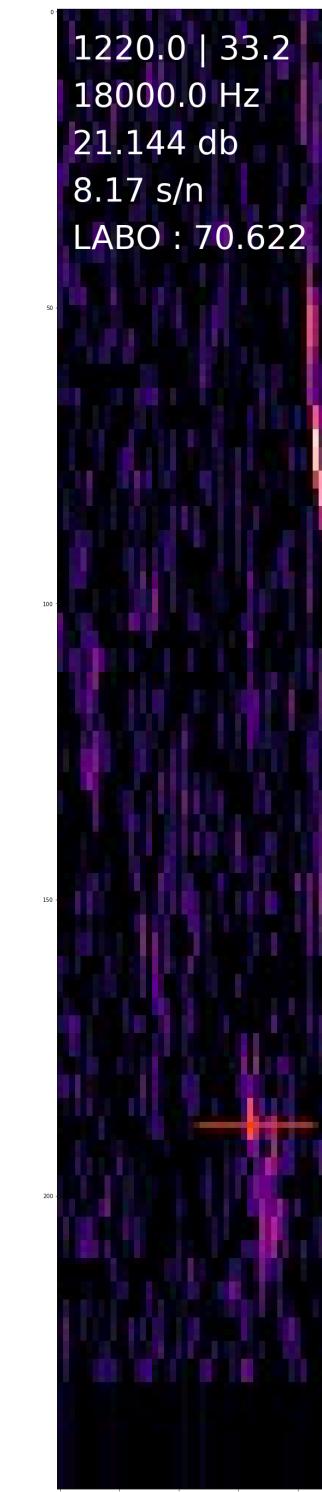
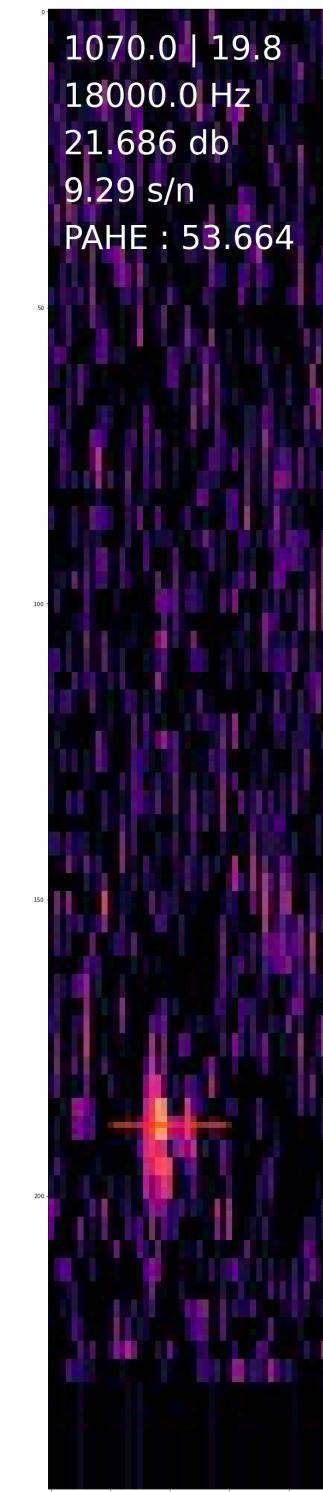
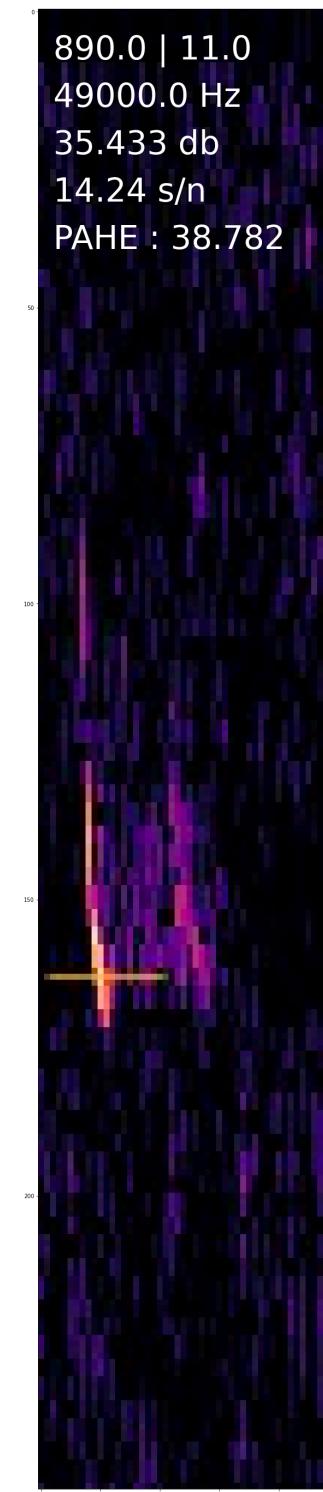
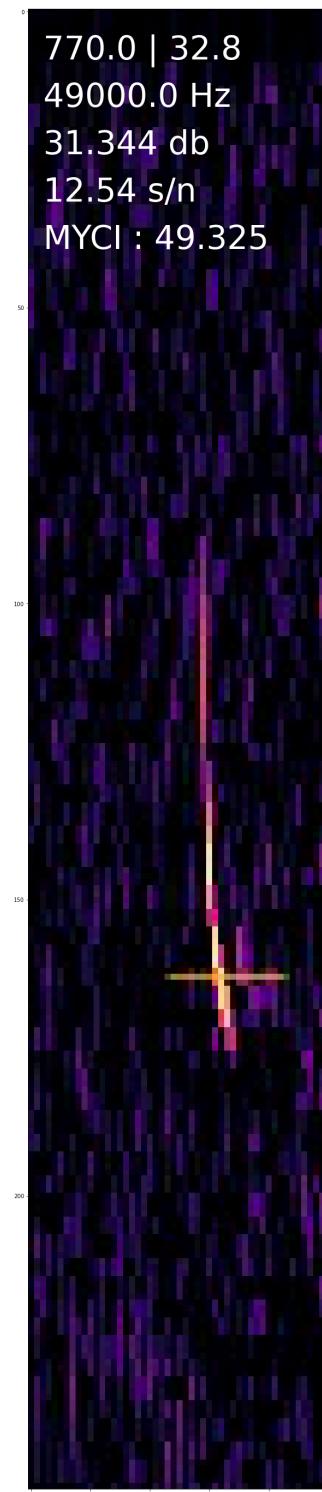
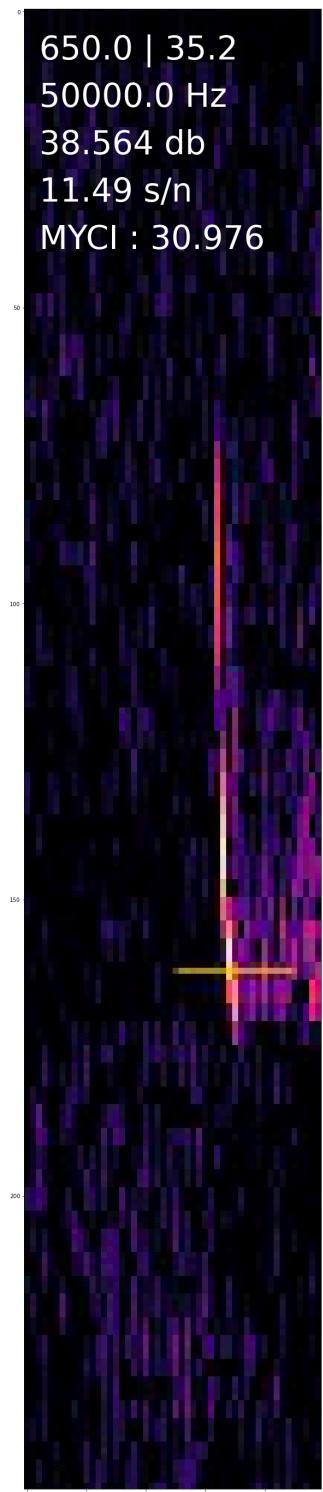
Label: LACI
Prediction: NOID
data/test/wav/LACI/p33_g81634_f29868405.wav



1790.0 | 31.2
19000.0 Hz
23.732 db
13.70 s/n
LACI : 95.010

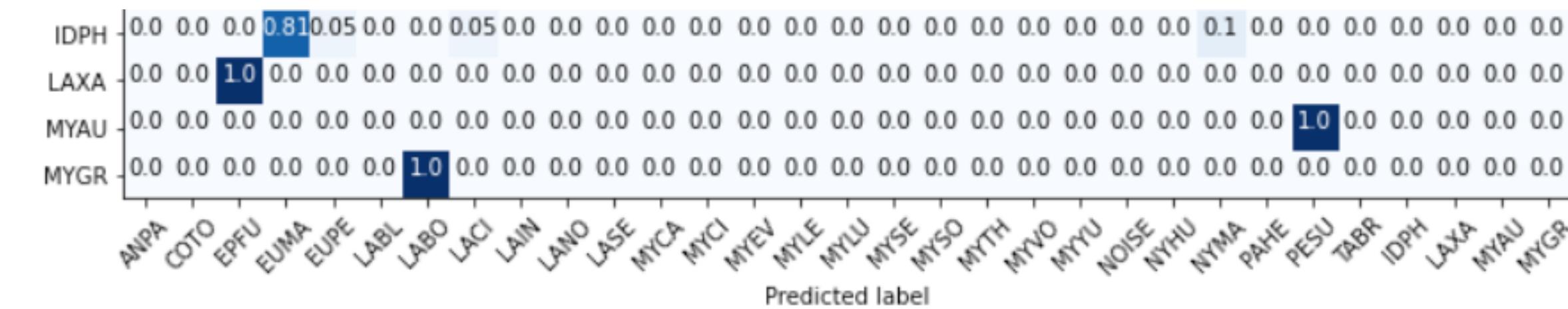
Limitations - Species Not Considered

- Attempt to classify species we have not trained on
- Generally, predictions are low confidence and spread among many classes. (NOID)
- Sometimes they are highly confident and we have to rely on our other filter methods



data/train/wav/MYAU/p303_g9178_f26623976.wav

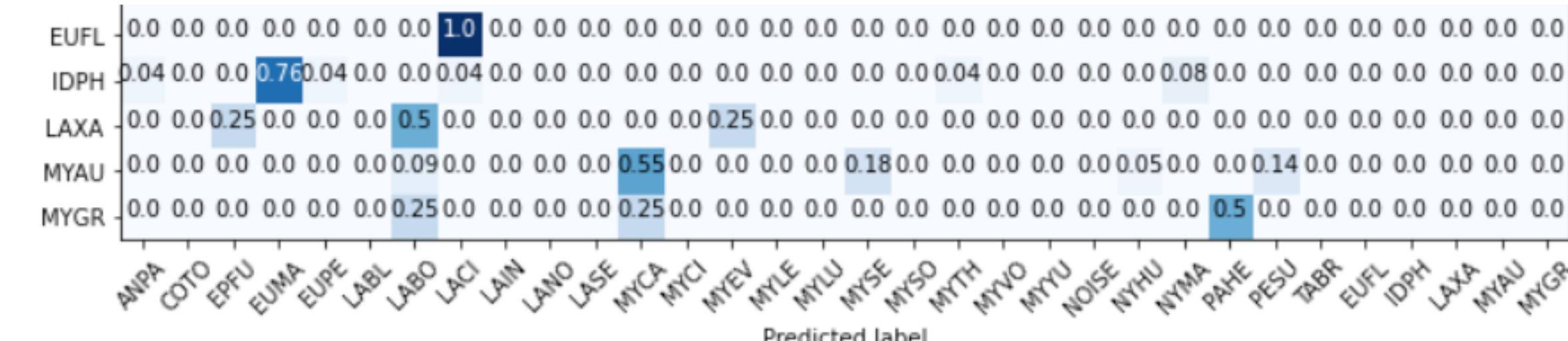
Species Not Considered



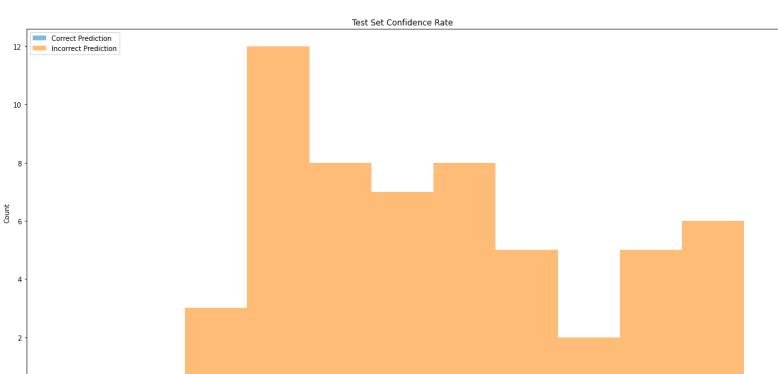
Standard post processing.

- Pulse count
- K-means
- Species range
- Minimum confidence threshold

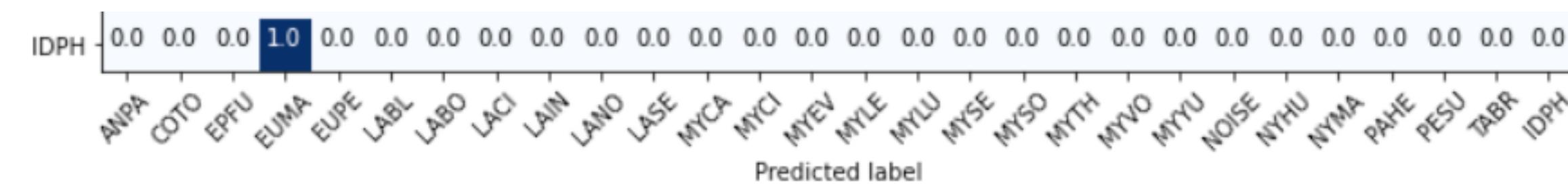
	precision	recall	f1-score	support
IDPH	0.00	0.00	0.00	21.0
LAXA	0.00	0.00	0.00	1.0
MYAU	0.00	0.00	0.00	2.0
MYGR	0.00	0.00	0.00	1.0
micro avg	0.00	0.00	0.00	25.0
macro avg	0.00	0.00	0.00	25.0
weighted avg	0.00	0.00	0.00	25.0



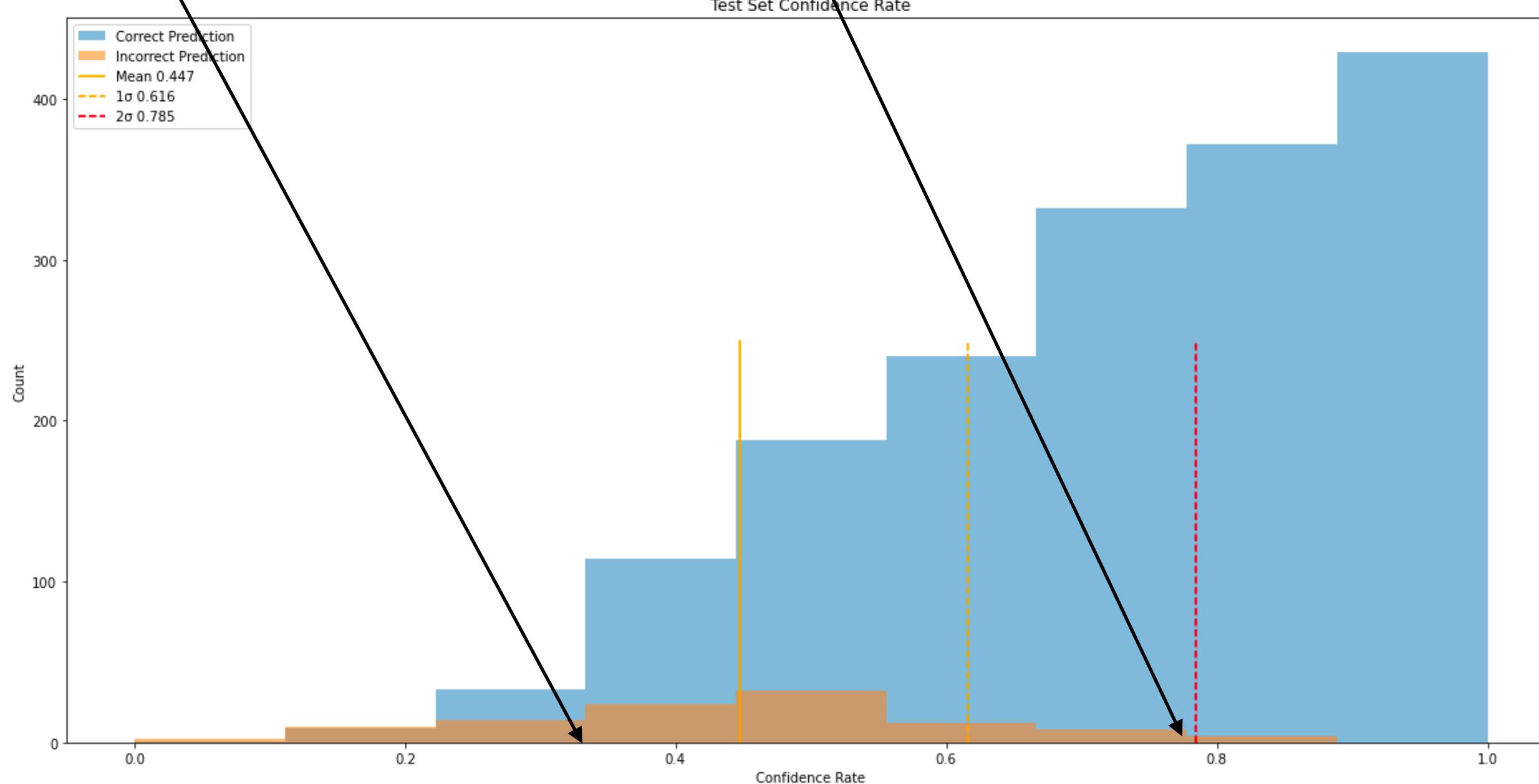
Raw output averaged to file



	precision	recall	f1-score	support
EUFL	0.00	0.00	0.00	1.0
IDPH	0.00	0.00	0.00	25.0
LAXA	0.00	0.00	0.00	4.0
MYAU	0.00	0.00	0.00	22.0
MYGR	0.00	0.00	0.00	4.0
micro avg	0.00	0.00	0.00	56.0
macro avg	0.00	0.00	0.00	56.0
weighted avg	0.00	0.00	0.00	56.0



	precision	recall	f1-score	support
IDPH	0.00	0.00	0.00	11.0
micro avg	0.00	0.00	0.00	11.0
macro avg	0.00	0.00	0.00	11.0
weighted avg	0.00	0.00	0.00	11.0



False positive reduction

- Standard method
- Increase minimum confidence threshold to 78.5%

Future



File	Autold	Pulse Count	ANPA	COTO	EPFU
/Users/bgotthold/Downloads/Example_calls1/COTO/NABAT3053-SE_20160730_023246.wav	COTO	5	0.005349322035908700	0.8714347720146180	0.005415789969265460
/Users/bgotthold/Downloads/Example_calls1/COTO/NABAT3053-SE_20160801_041814.wav	NOID	0			
/Users/bgotthold/Downloads/Example_calls1/COTO/NABAT-3069NW_20160603_221835.wav	COTO	4	0.007126843323931100	0.6546708273235710	0.006465562502853570
/Users/bgotthold/Downloads/Example_calls1/EUAMA/NABAT2717NW_20160727_231711.wav	NOISE	13	0.005762952702263230	0.005747478956786490	0.005895103757771160
/Users/bgotthold/Downloads/Example_calls1/EUAMA/NABAT2717NW_20160729_021029.wav	EUAMA	5	0.0053027537651360000	0.005302301701158290	0.0053041745908558400
/Users/bgotthold/Downloads/Example_calls1/EUAMA/NABAT221NE2_20160723_233405.wav	EUAMA	22	0.005703592821109020	0.0056739024479280800	0.005979725985195150
/Users/bgotthold/Downloads/Example_calls1/EPFU/52925_NE_20200720_212929.wav	EPFU	35	0.006104912994695560	0.005642674025148150	0.8077676181681460
/Users/bgotthold/Downloads/Example_calls1/EPFU/40637_NW_20200709_212033.wav	EPFU	44	0.024648784860883900	0.005721084149131040	0.7360636193135920
/Users/bgotthold/Downloads/Example_calls1/EPFU/NABAT-989NW_20180819_201800.wav	EPFU	23	0.005929471803424150	0.005580260965001320	0.8262667798477670
/Users/bgotthold/Downloads/Example_calls1/MYVO/NABAT-1853SW_20180802_024802.wav	NOID	2			

