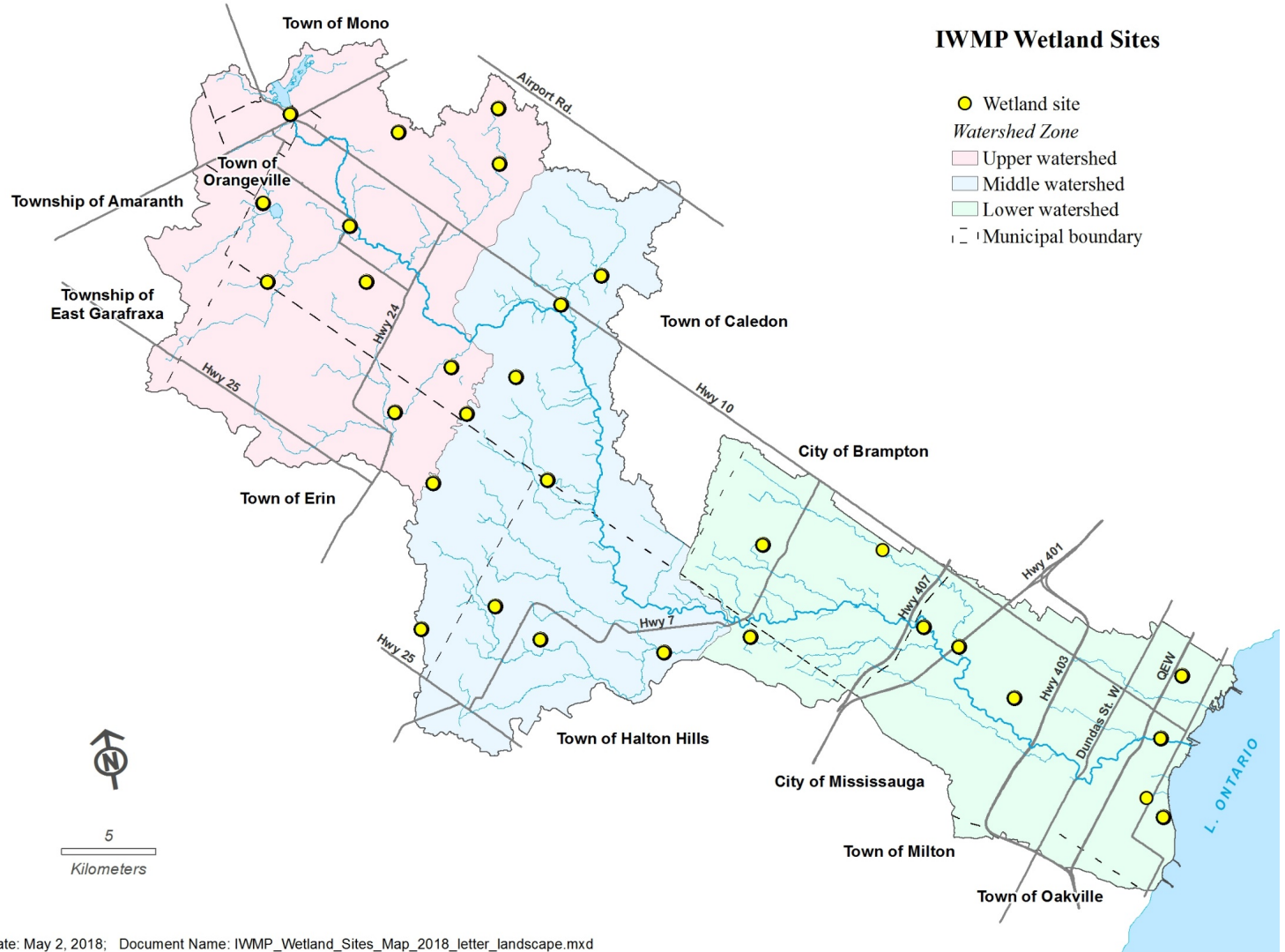
Credit Valley Conservation
inspired by nature

Applying a Novel Analytic Approach to Bioacoustic Data to Monitor and Characterize Wetland Frog Populations in Southern Ontario

By Crystal Kelly

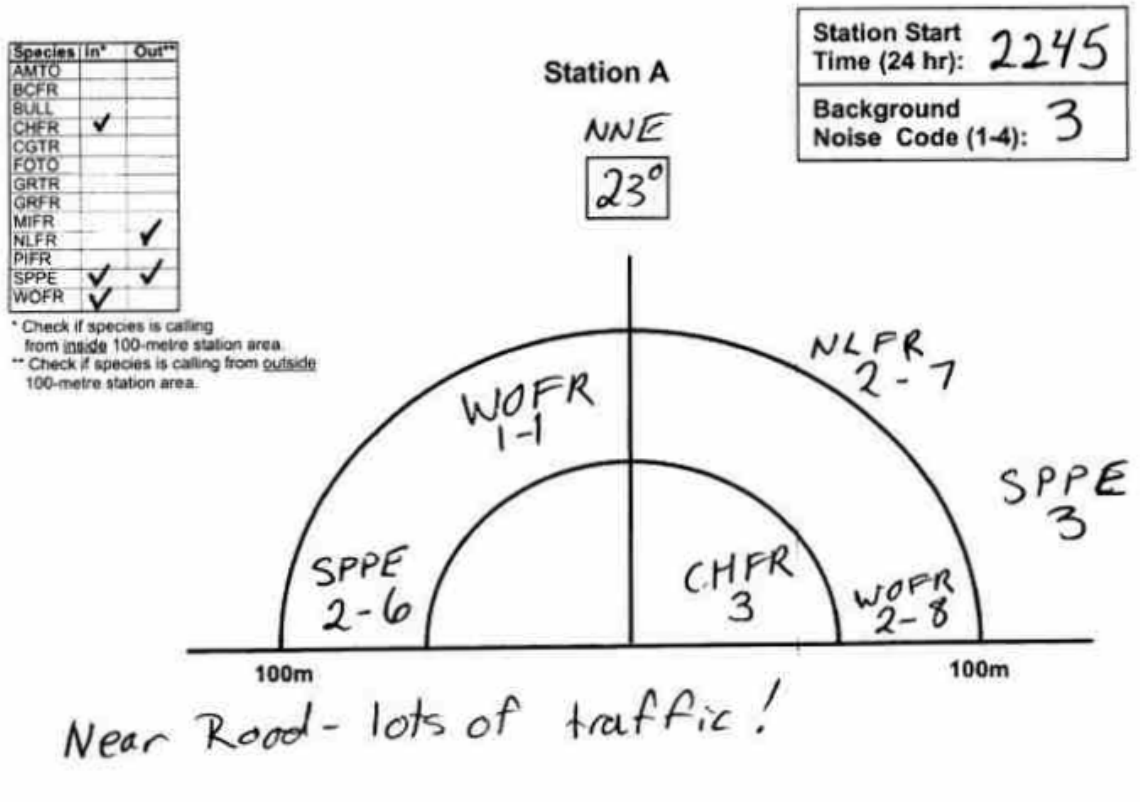
CVC's IWMP Program



Date: May 2, 2018; Document Name: IWMP_Wetland_Sites_Map_2018_letter_landscape.mxd

Acoustic Monitoring Methods

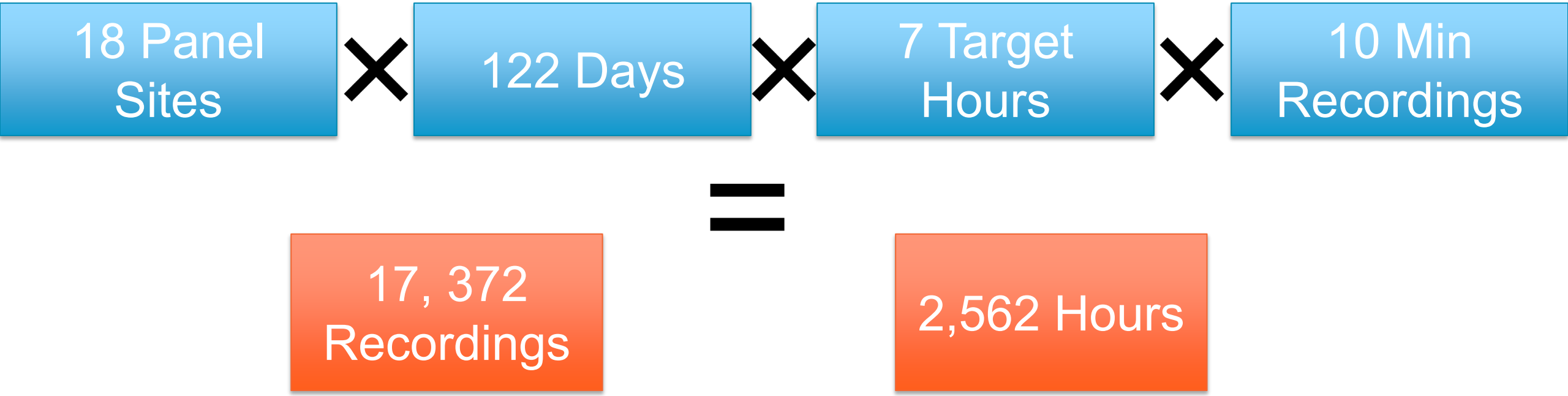
- Marsh Monitoring Program
  - Three manual call surveys/year
  - Coarse population proxy
    - L1
    - L2
    - L3



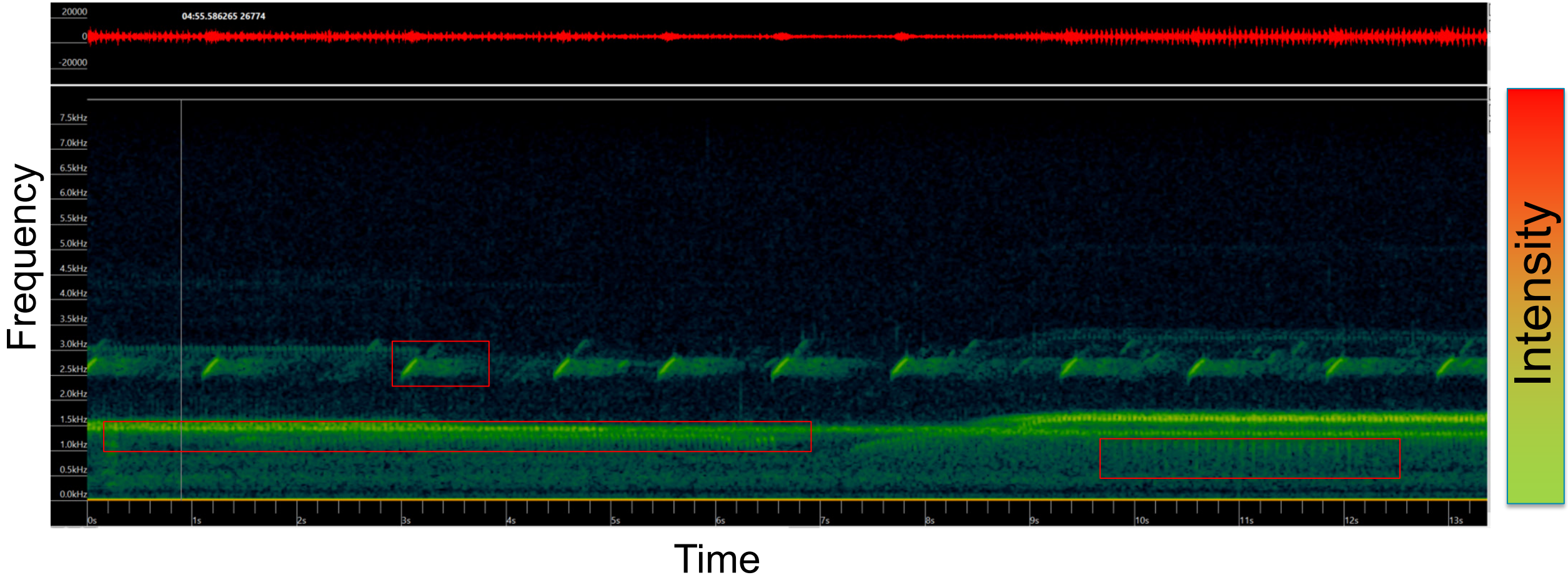
Acoustic Monitoring Methods

- Autonomous Recording Units
  - Easy and inexpensive to deploy
  - Massively increased data resolution
  - Massive amount of data

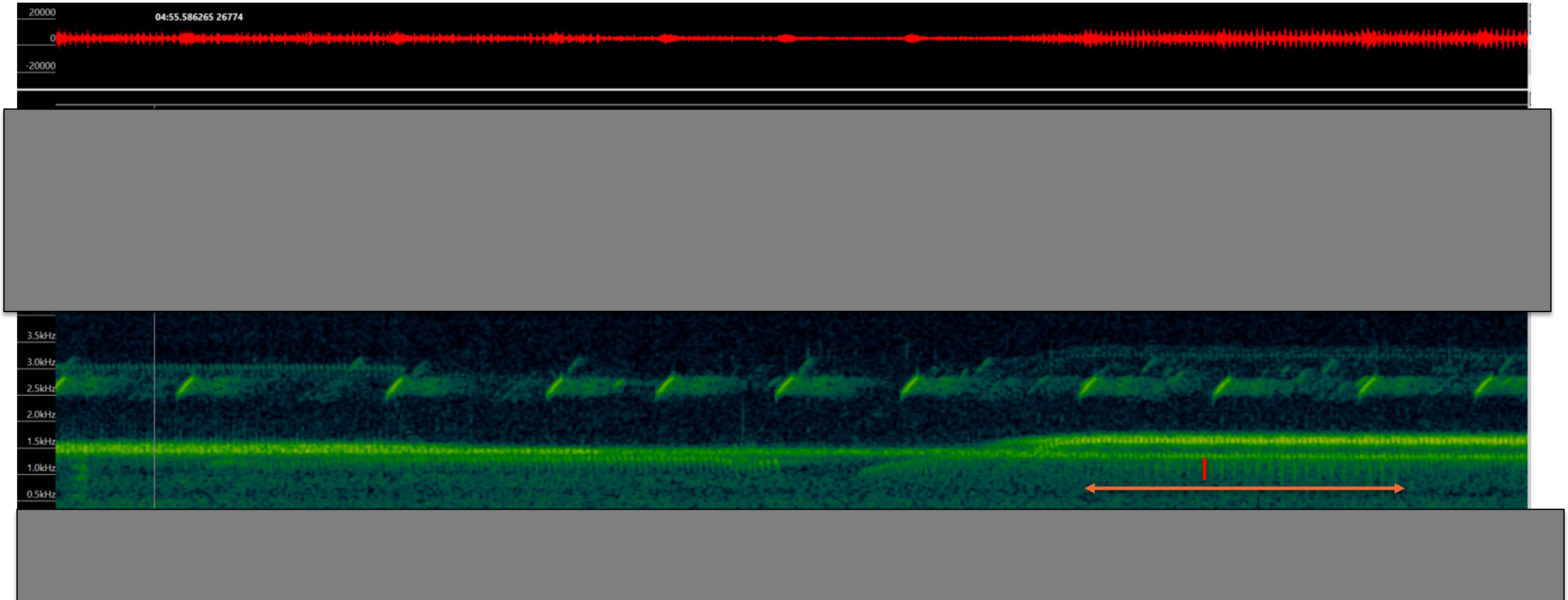
The Problem With Big Data



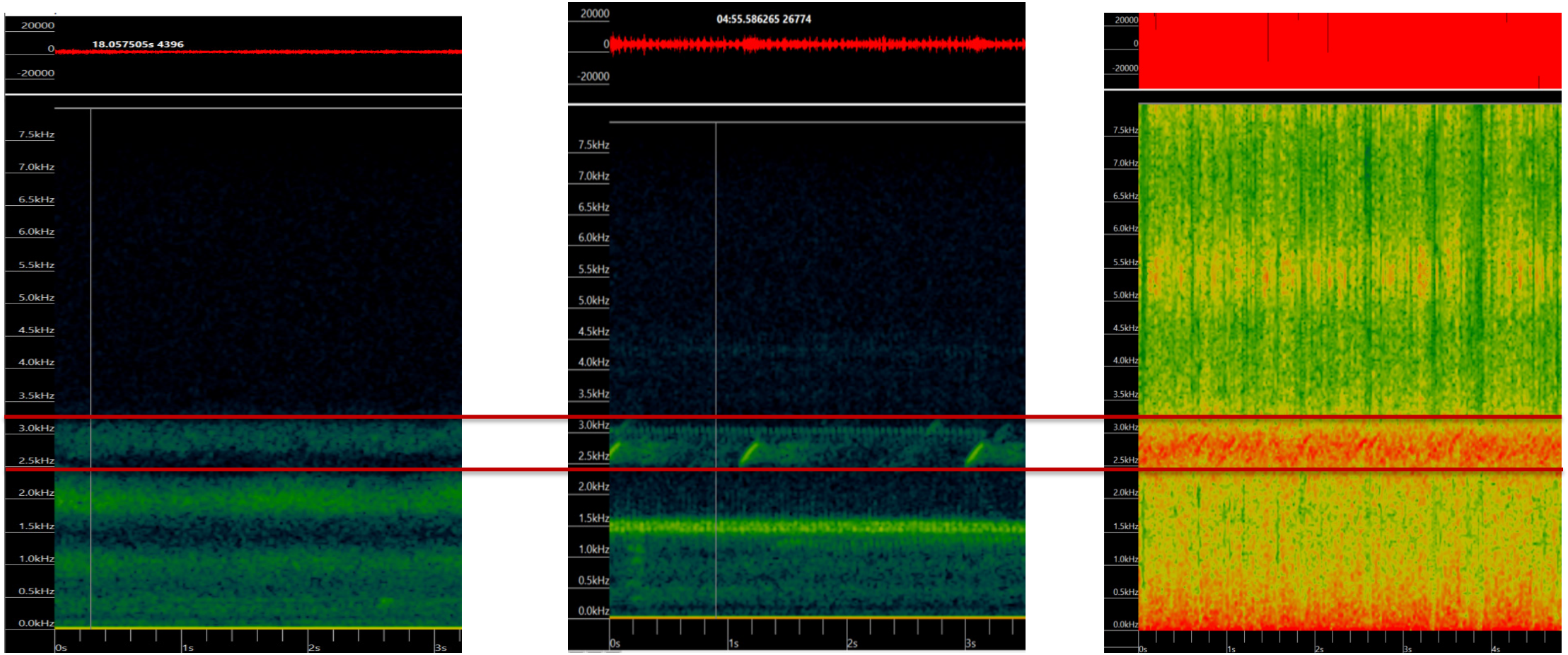
Kaleidoscope



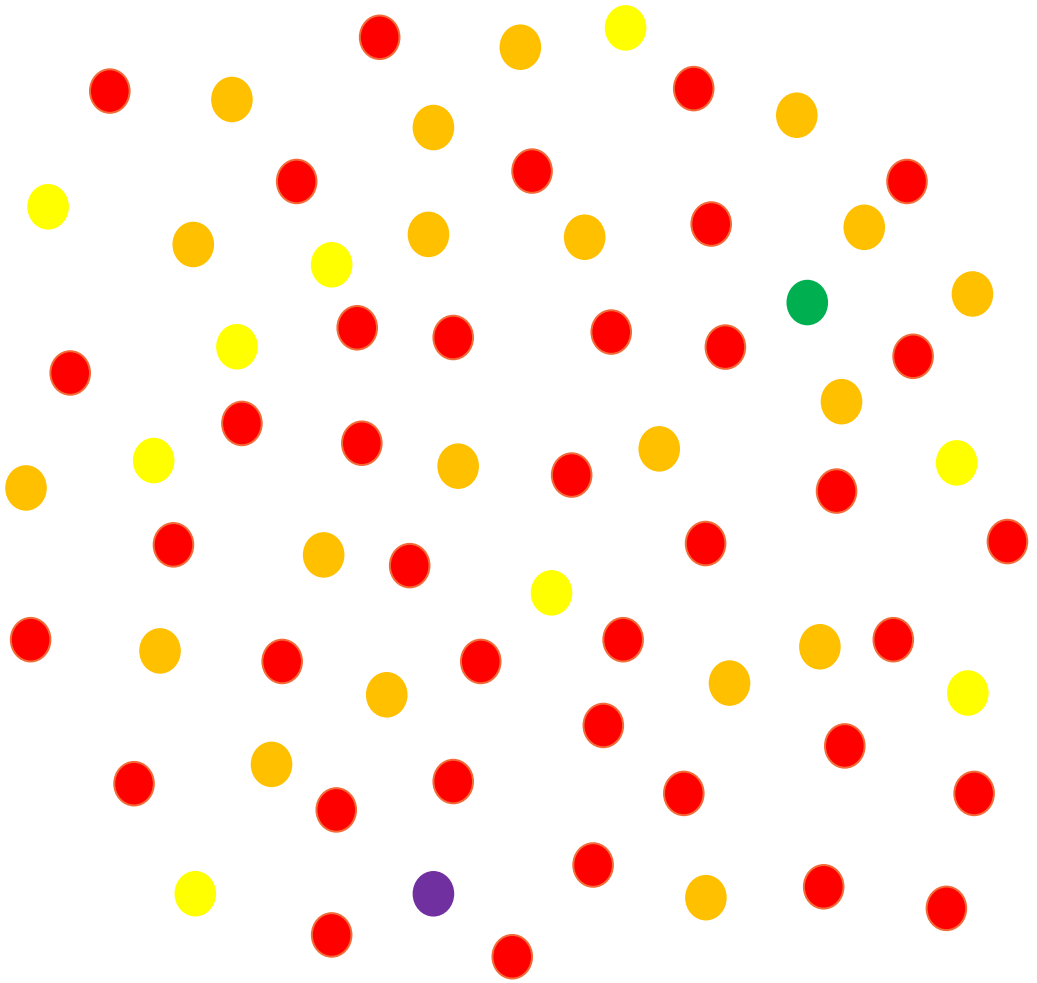
Kaleidoscope

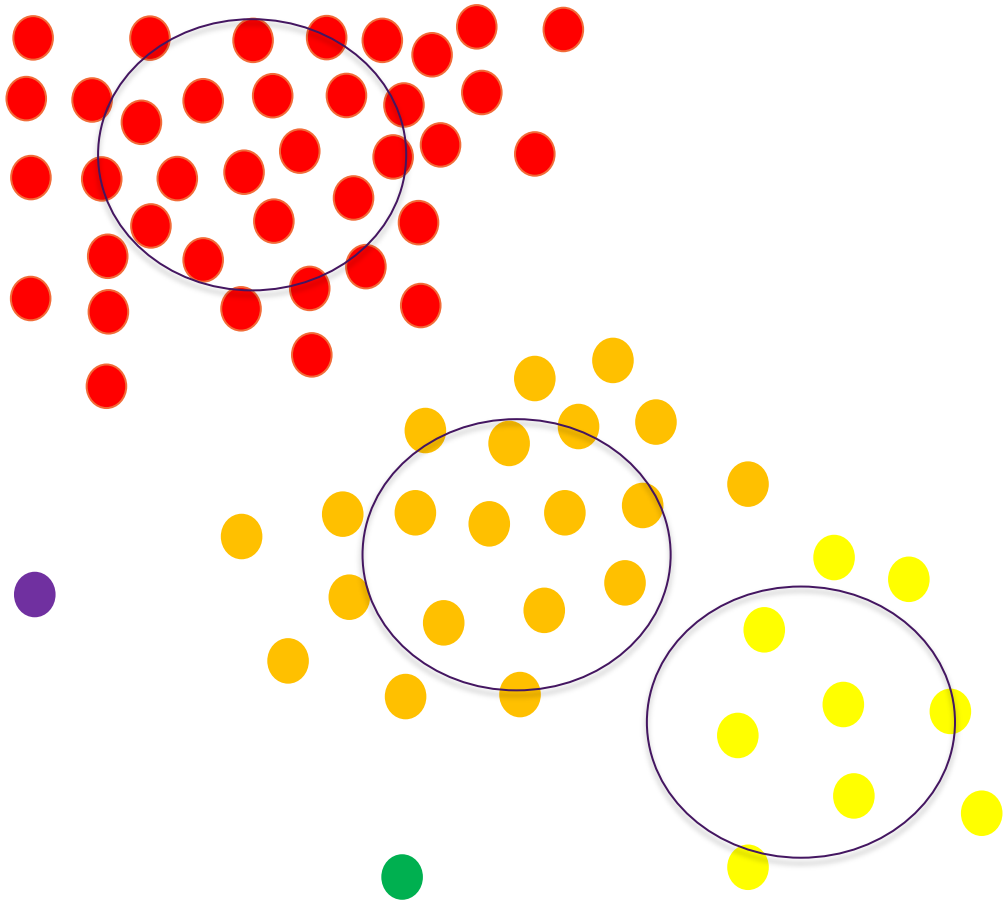


Kaleidoscope

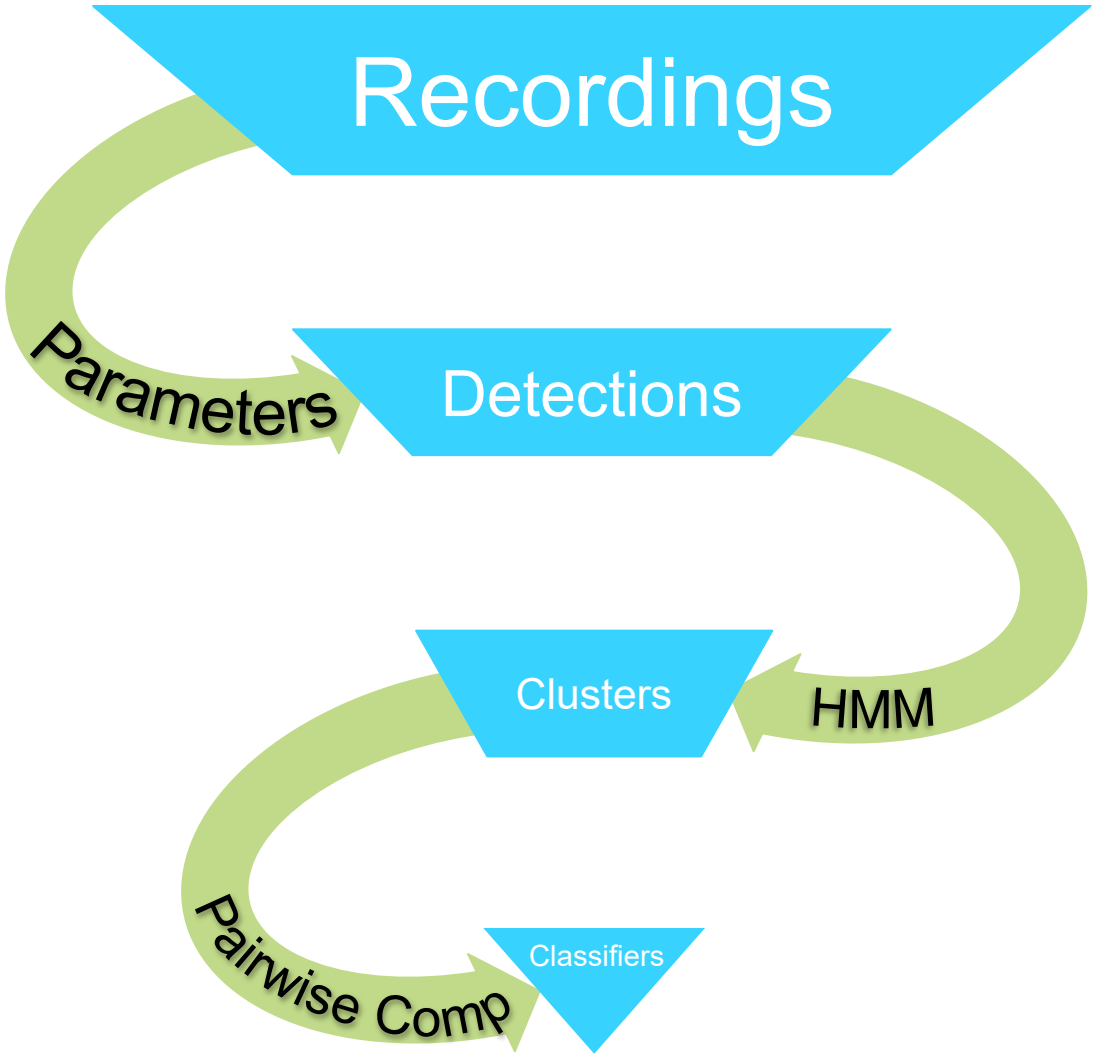


Kaleidoscope

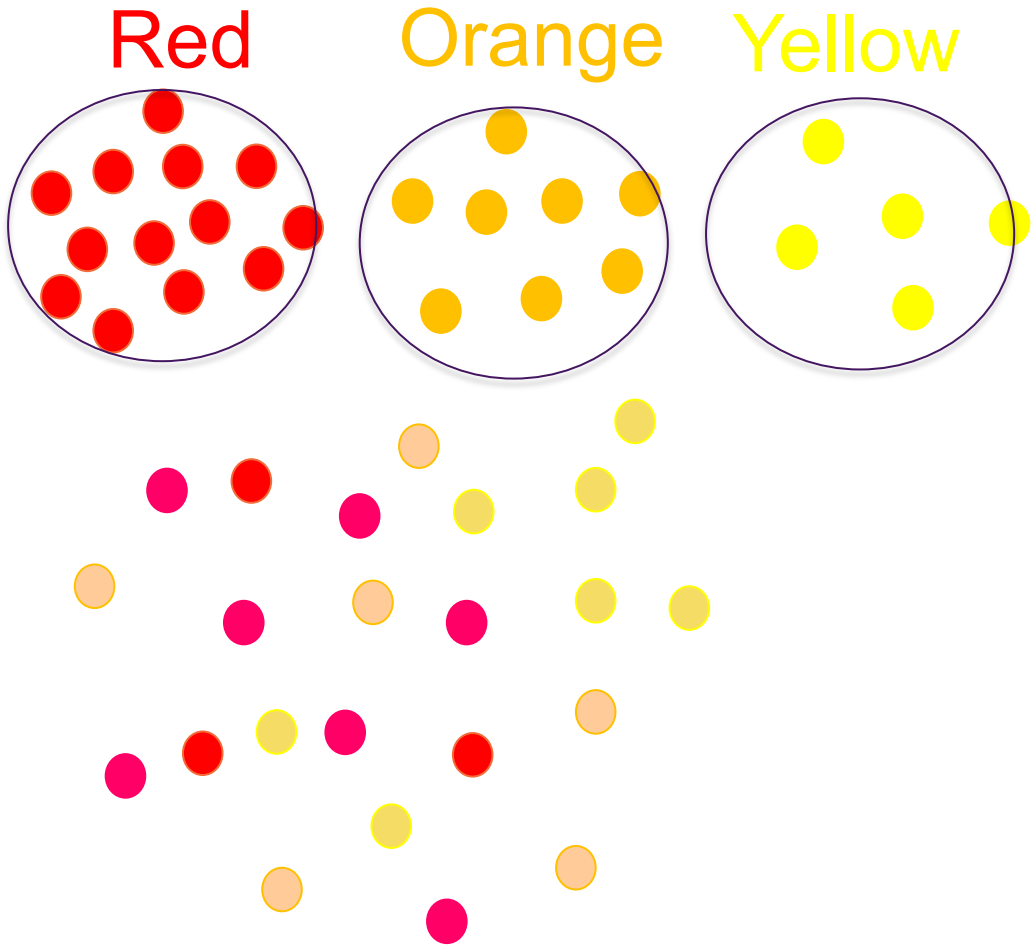




Kaleidoscope



Kaleidoscope



Kaleidoscope VS. Manual Classification

	Ken Whillans (%)	Robert Baker (%)	Total (%)
AMTO	65.5	76.5	69.6
GRFR	88.5	76.2	83.0
GRTR	88.9	17.1	31.8
SPPE	13.2	6.1	9.2
WOFR	100.0	43.8	50.0

49% Missing calling dates

Kaleidoscope

Pros

- Can sort our data into something meaningful

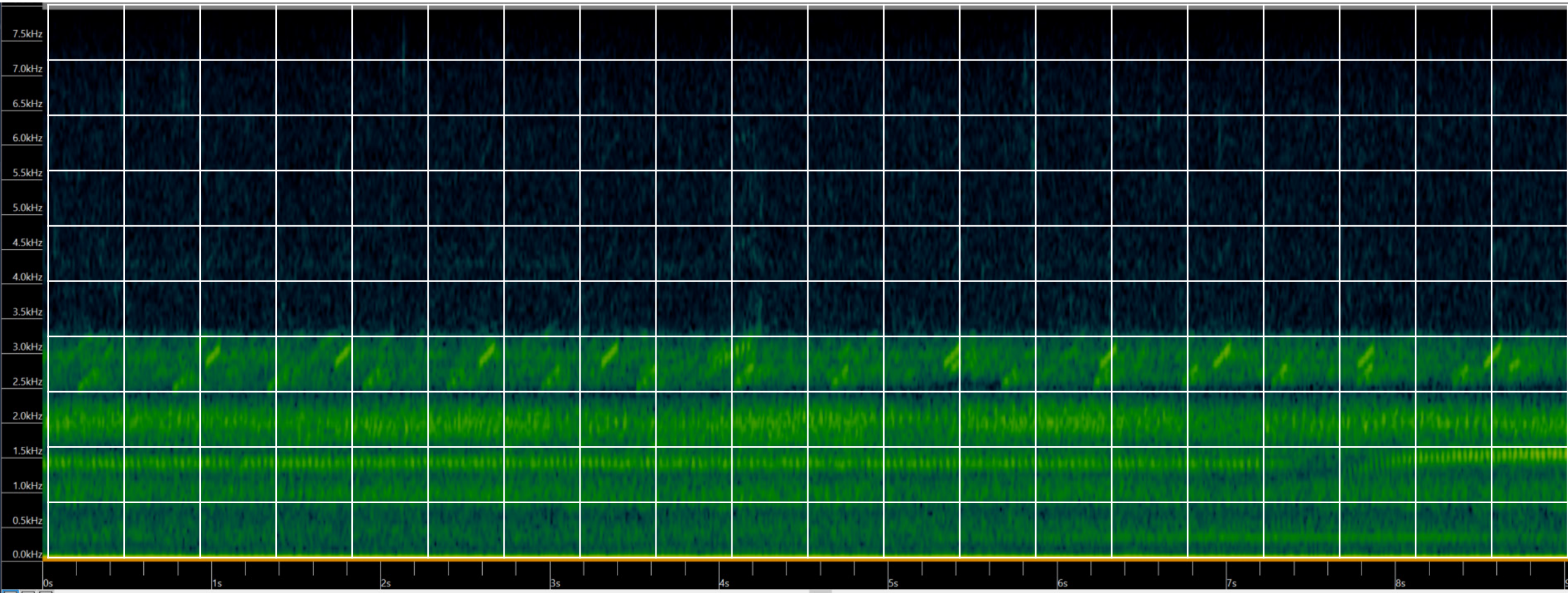
Cons

- Still requires massive amount of time
- Common species overwhelm rare species
- A problem with variation

Acoustic Indices

- Acoustic Indices
  - First applied by Towsey et al. 2014 as a way to deal with terrabytes of data
  - A way to mathematically categorize sound properties by averaging over a specific time period within a specific frequency bin.
    - Waveform: Energy and power of a sound
    - Frequency: Pitch of a sound

Acoustic Indices

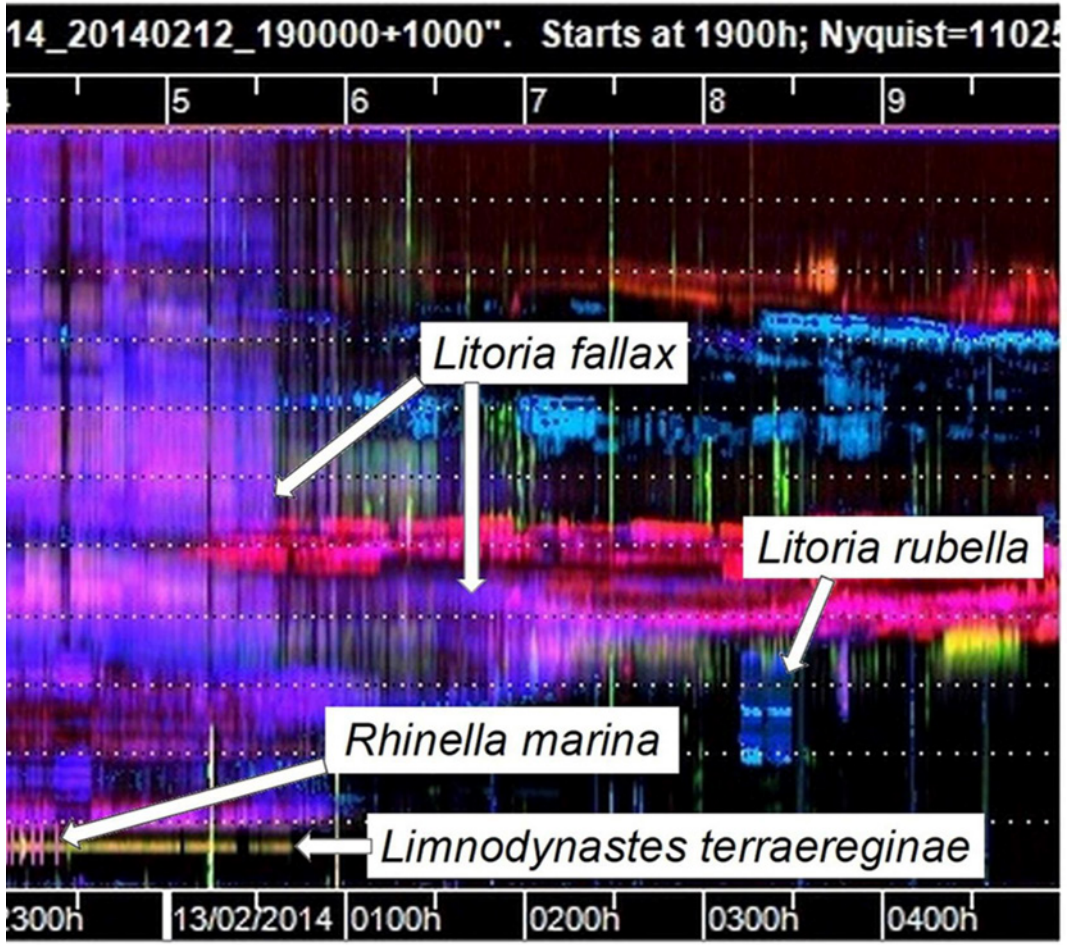


Acoustic Indices

Full name	Abbreviation	Principle	Reference
α indices			
Relative avian abundance	–	Area under spectrum in relation with an amplitude threshold	[43] [43]
Temporal Entropy	H_t	Envelope complexity	[44]
Spectral Entropy	H_f	Spectrum complexity	[44]
Acoustic Entropy Index	H	Envelope and spectrum complexity	[44]
Ratio of biophony to anthrophony	ρ	Ratio of biophony to anthrophony	[54]
Acoustic Complexity Index	ACI	Spectrogram complexity	[50]
Biophony	–	Biophony level	[45]
Biophony peak	bioPeak	Biophony level	[56]
Acoustic Entropy Index	AEI $(= H)$	Envelope and spectrum complexity	[45]
Shannon's Index	H'	Spectrum complexity	[46]
Acoustic Richness	AR	Envelope complexity and intensity	[48]
Median of amplitude envelope	M	Median of amplitude envelope	[48]
Normalised Difference Soundscape Index	NDSI	Ratio of anthrophony to biophony	[55]
Acoustic Diversity Index	ADI $(= H')$	Spectrum complexity	[47]
Sound pressure level parameters	L	Ratio of sound pressure relative to a reference value	[35] [35]
Number of peaks	NP	Spectrum complexity	[19]
Mid-band activity	–	Fration of spectrum above an amplitude threshold	[42] [42]
Entropy of spectral maxima	H_m	Spectrum composition	[42]
Entropy of spectral variance	H_v	Spectrum complexity	[42]
Spectral diversity	–	Number of clusters	[42]
Spectral persistence	–	Duration of repeated clusters	[42]
β indices			
Spectral Dissimilarity	D_f	Spectrum dissimilarity	[44]
Temporal Dissimilarity	D_t	Envelope dissimilarity	[44]
Acoustic Dissimilarity Index	D	Envelope and spectrum dissimilarity	[44]
Kolmogorov-Smirnov distance	KS	Spectrum dissimilarity	[60]
Kullback-Leibler distance	KL	Spectrum dissimilarity	[60]
Vectorial correlation coefficient	RV	Spectrum similarity	[60]
Cumulative Dissimilarity	D_{cf}	Spectrum similarity	[61]

Supervised Machine Learning

- Labelled dataset by the min
- Uses calculated acoustic indices to create predictive models about species presence/absence
- Recall rates ~80% +
  - Perform better with more data
  - Common species work better



Building The Database

- Label 5 sets of recordings in 2021 and 2022
  - Mix of urban and rural sites
  - AMTO, GRFR, GRTR, SPPE, WOFR
  - 38,000+ minutes of recording

Random Forest Modelling

- Kaleidoscope
  - 1 min recordings
  - NDSI, ACI, ADI, AEI, BGN, SNR, ACT, EVN, LFC, MFC, HFC, CENT, MEAN, SD, SEM, MEDIAN, MODE
- Use AI to predict frog P/A

Random Forest Modelling

Species	Specificity (%)	False Positive (%)	Sensitivity (%)	False Negative (%)	A:P
AMTO	99.4	0.6	46.3	53.7	10.7
GRFR	98.4	1.6	74.3	25.7	4.5
GRTR	99.3	0.7	64.0	36.0	6.2
SPPE	96.3	3.7	75.9	24.1	2.0
WOFR	99.5	0.5	62.2	37.8	25.9

Kaleidoscope vs. Random Forest

	Kaleidoscope Missed calling dates(%)	Random Forest Missed calling dates (%)
AMTO	69.6	6.5
GRFR	83.0	10.6
GRTR	31.8	4.5
SPPE	9.2	0.8
WOFR	50.0	11.1

49% vs **7%**

missed calling dates

Conclusions

- Random forest models are useful
  - 600x data resolution
  - Potential to move beyond chorus level
  - Parse data
- Next Steps
  - Sub setting by chorus level
  - Building “rare” database

inspired by nature