Using Animal Audio for Species Detection

Lin Schwarzkopf

JAMES COOK UNIVERSITY
AUSTRALIA

Acknowledgements

Paul Roe
Mike Towsey

Why detect species?

- We may want to
  - identify presence/absence, abundance or activity of individual species — study organism, rare, threatened
  - quantify numbers of species in an area in relation to habitat, anthropogenic disturbance — grazing, fire, urbanisation, etc.
  - Determine effects on ecosystem “health” — climate change, logging, agriculture, changes in land use etc.

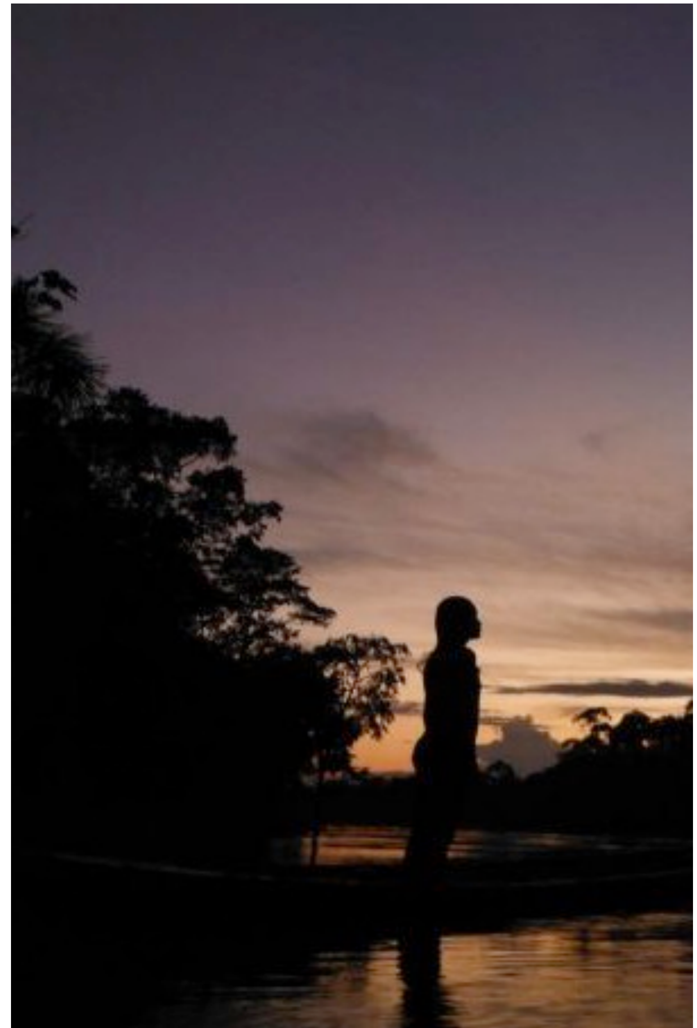




Traditional Monitoring

- Fauna & vegetation surveys

Traditional Audio Monitoring



Traditional Monitoring

Advantages:

— Provide highly accurate information on species presence/absence, activity & richness

Limitations:

— Highly spatially & very highly temporally restricted

— Expensive & time consuming to get a lot of data

— Limited to expertise that is present

— Observer bias

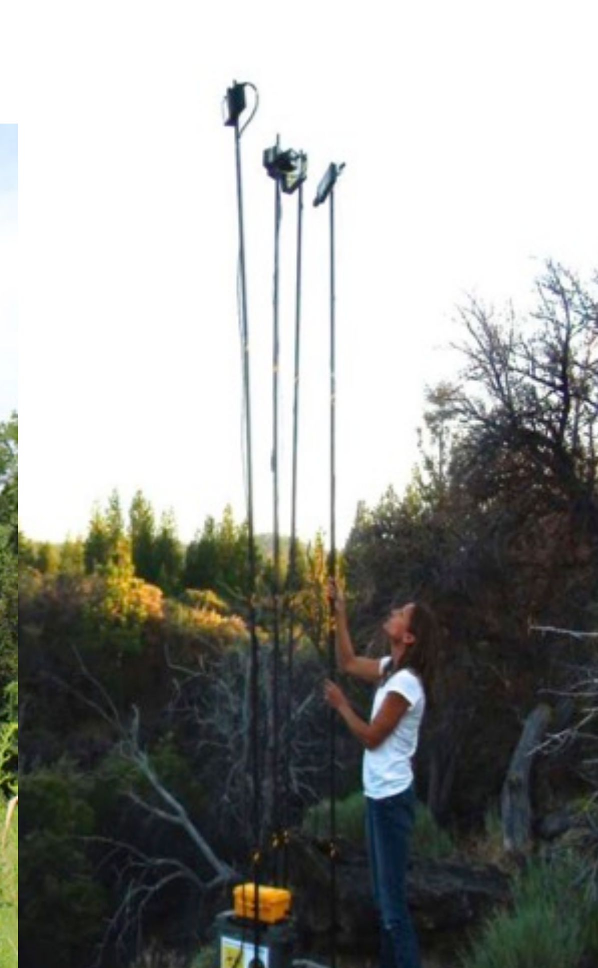




Autonomous Recording Units — Record Sound *in situ*

Advantages —

- Non-invasive
- Relatively cheap
- Collect extensive audio data
- Permanent record
- Limited only by storage capacity – which continues to increase rapidly

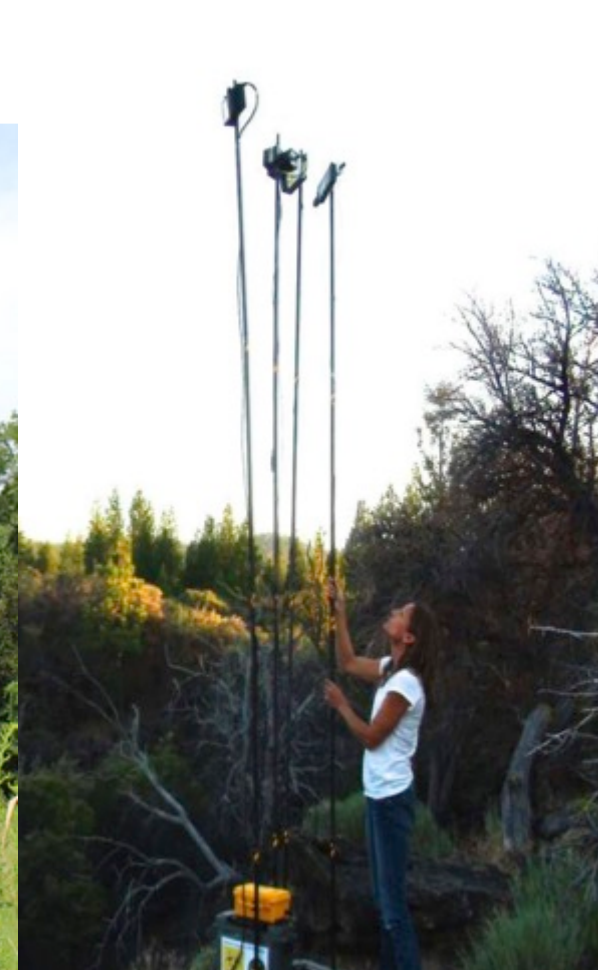


Autonomous Recording Units — Record Sound *in situ*

Disadvantages —

- Restricted to species that make some kind of noise
  - Birds, frogs, insects, some fish, some reptiles, many mammals

- There is so much data analysing it becomes a problem!



Species Detection – Individual Species

- Humans listen & recognise calls – subsampling in time
- Songscope-type recognisers
- Human-in-the-loop combinations





What's better – ARUs or traditional methods?

- Autonomous Recording Units (ARUs) versus point counts to quantify species richness and composition of birds in temperate interior forests.
- Short-term monitoring, point counts may probably perform better than ARUs, especially to find rare or quiet species.
- Long-term (seasonal or annual monitoring) ARUs a viable alternative to standard point-count methods

Klingbeil & Willig. 2015. PeerJ 3:e973; DOI 10.7717/peerj.973

What's better – ARUs or traditional methods?

- This study used ARUs almost exactly like point counts
- Human observers at exactly the same time & place as recorders perform better – distant calls & difficult to hear calls, visual recognition
- Used Songscope^{TM} to ID calls
- Even using this method – ARUs larger samples over time produced better samples than human visits

Klingbeil & Willig. 2015. PeerJ 3:e973; DOI 10.7717/peerj.973

Species Detection – Individual Species

- Humans listen & recognise calls – subsampling in time
- Songscope-type recognisers
- Human-in-the-loop combinations

Species Detection – Individual Species

Songscope-type automated “recognisers”

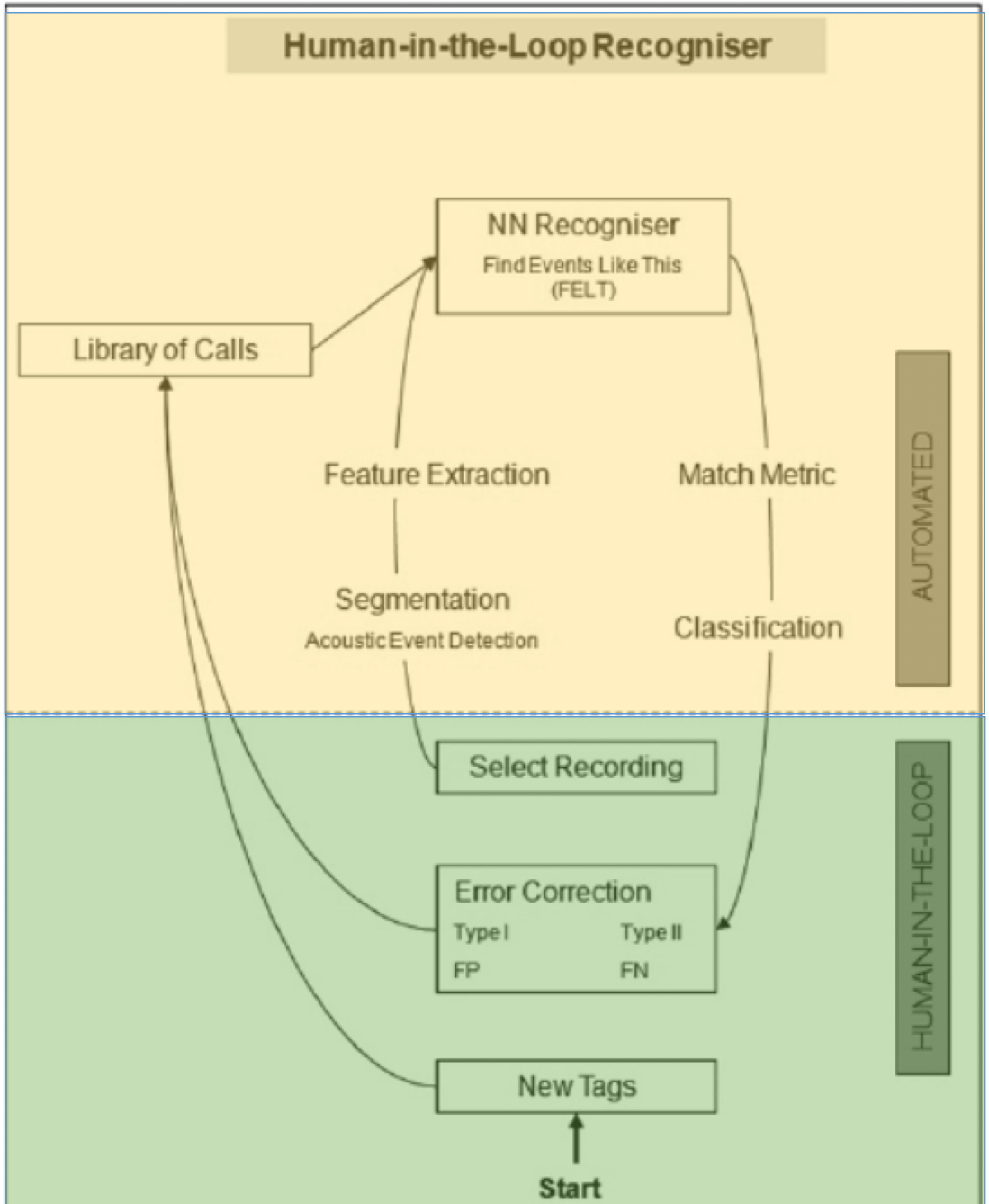
- possible based on several different kinds of algorithms: fuzzy logic, dynamic time or Hidden Markov models, oscillation detection, event or syntactic pattern recognition
- Speech recognition models are not very successful on environmental recordings because of their need for limited background noise
- Animal calls vary more than human speech
- Variable success dependent on type of background noise
- Need to be trained for call & environment





Species Detection – Individual Species

- Human-in-the-loop combinations
  - best outcomes at the moment


Human-in-the-Loop Recogniser
NN Recogniser
Find Events Like This
(FELT)
Library of Calls
Feature Extraction
Match Metric
Segmentation
Acoustic Event Detection
Classification
AUTOMATED
Select Recording
Error Correction
Type I
Type II
FP
FN
New Tags
Start
HUMAN-IN-THE-LOOP

Indices of Ecosystem Health

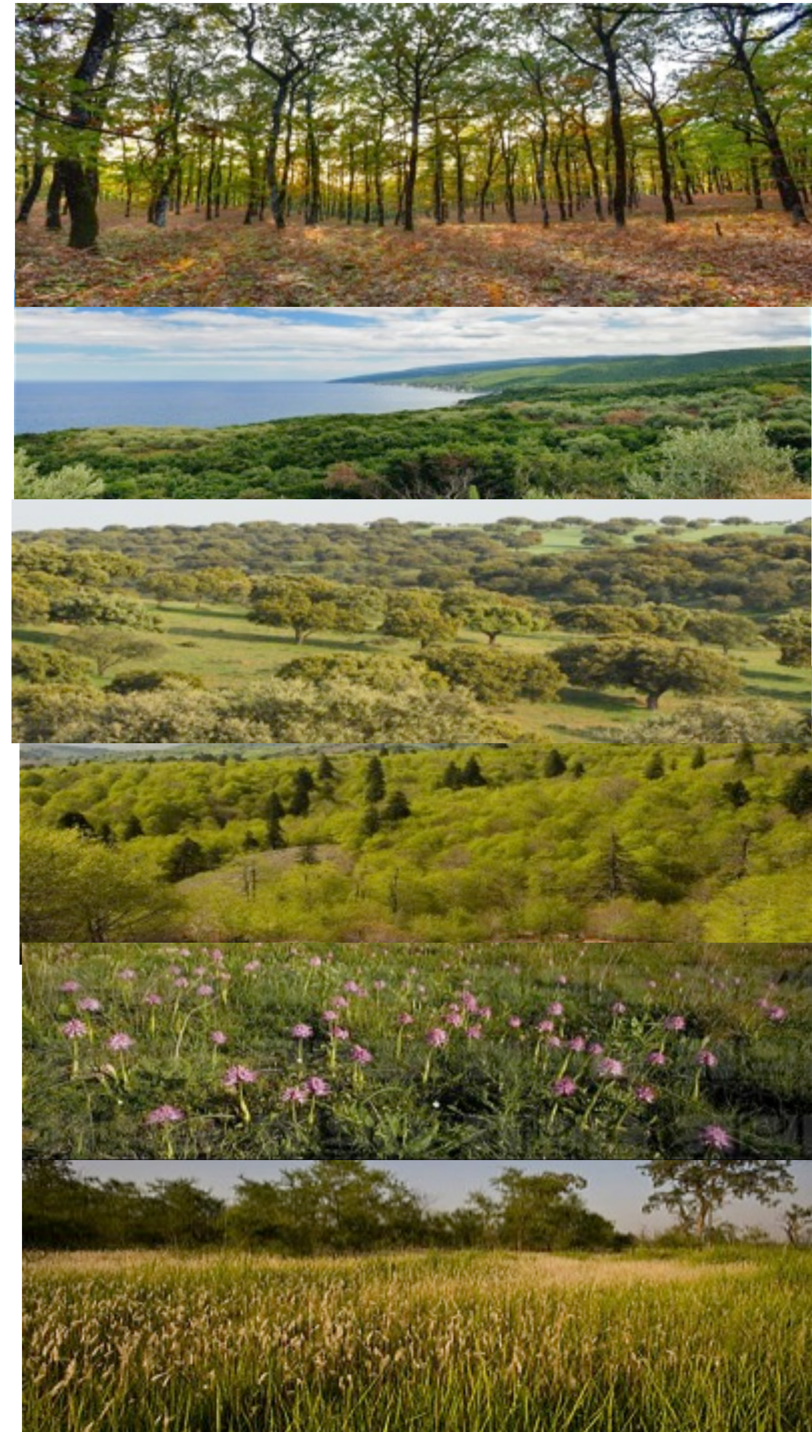
Ecoacoustics, Soundscape Ecology

— Use Acoustic Indices

— Characterise animal acoustic communities, habitats, overall ecological state

Acoustic Signatures

- Natural soundscapes should be habitat specific.
- Ambient sound in **different types of forest** was recorded
- Used digital signal techniques and machine learning algorithms
- *Even fairly similar habitat types have specific acoustic signatures distinguishable by machine*



Acoustic Complexity Index

- ACI highlights and quantifies complex biotic noise (ie. bird calls) while reducing effects of low-variability human noise (ie. airplane engines) Sueur et al. 2014. Acta Acustica 100:772-81.

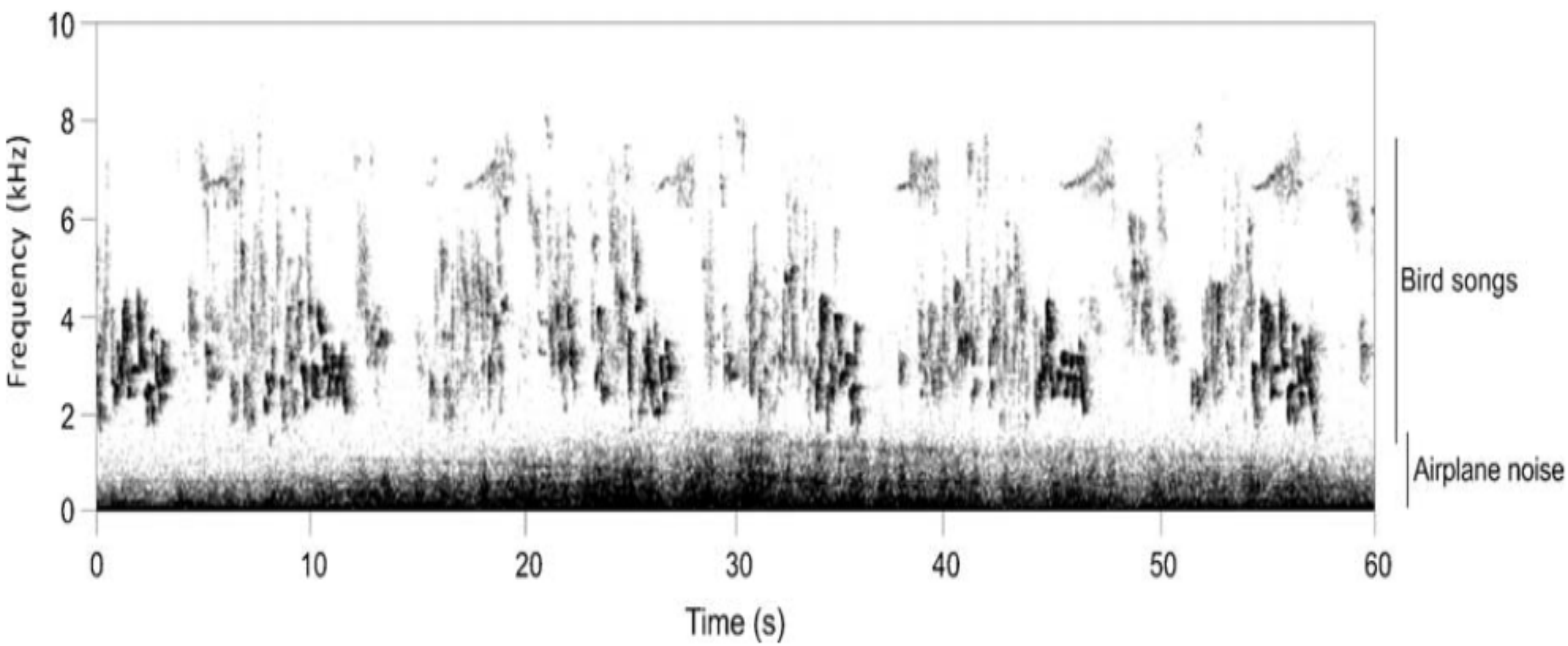


Fig. 3. Spectrogram representing a typical scene of the airplane noise overlapping the natural soundscape.

Can soundscape reflect landscape condition?

- Soundscape patterns vary with landscape configuration and condition
- 19 forest sites in Eastern Australia
- 3 indices soundscape = landscape characteristics, ecological condition, and bird species richness
- acoustic entropy (H), acoustic evenness (AEI), normalized difference soundscape index (NDSI)
- **Anthrophony** was inversely correlated with biophony and ecological condition
- **Biophony** positively correlated with ecological condition

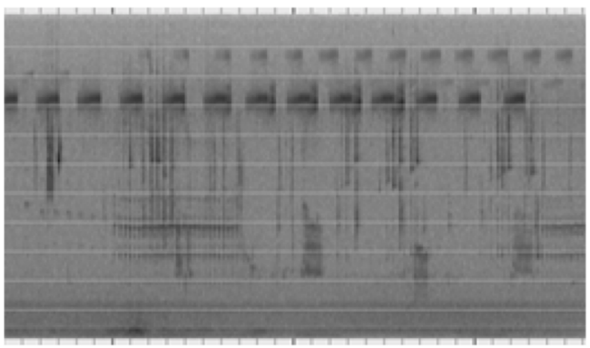
Fuller et al. 2015. Ecological Indicators 58:207-15

Overall Signatures *Not* For Species Detection

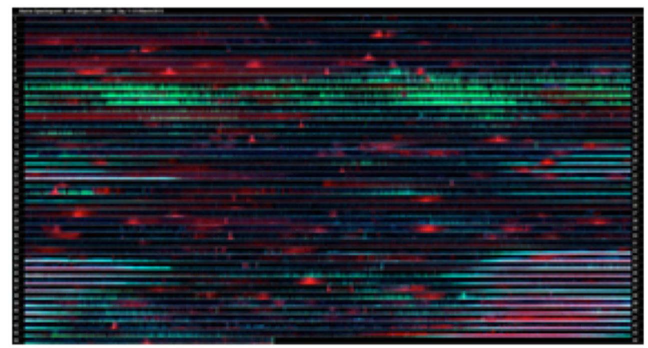
Species Richness Applications

- We want to know not only that a system is rich or diverse, or different from other systems, but **which species** are present…

How to bridge the gap?



Time scale = seconds



Time scale = days > months > years

BIO-ACOUSTICS
Single vocalisations
Species recognition

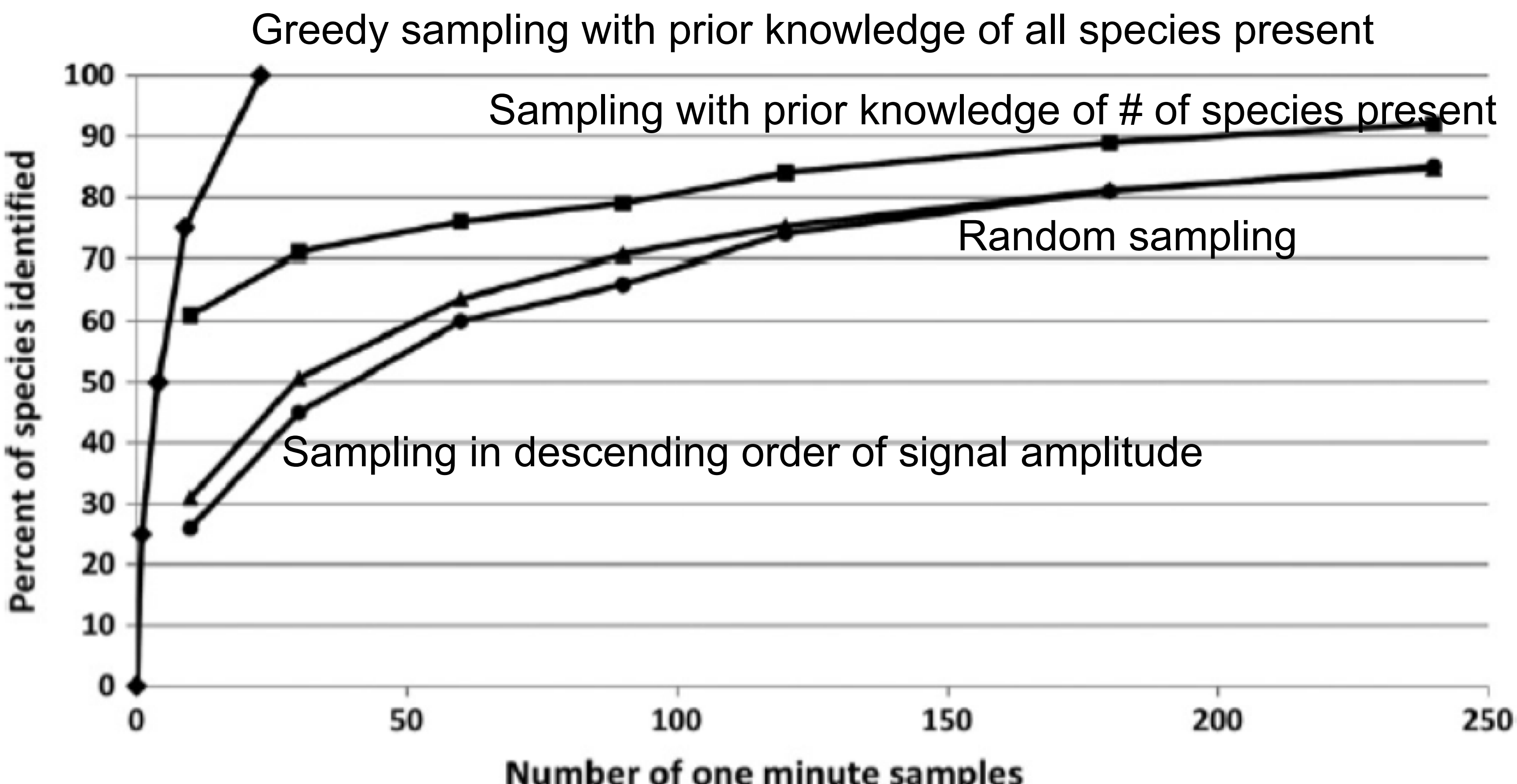
ECO-ACOUSTICS
Soundscape ecology
Ecosystem processes

Combination Approaches

- Estimating avian species richness from very long acoustic recordings.
- Used acoustic indices to summarise the acoustic energy information in the recording
- R**andomly sampled** 1 minute segments of 24 hour recordings - achieved a 53% increase in species recognised over traditional field surveys
- C**ombinations of acoustic indices** to direct the sampling - achieved an 87% increase in species recognized over traditional field surveys

Towsey et al. 2014. Ecological Infomatics 21: 110-119.

Sampling?



- Different sampling protocols listening to 1 minute samples of a 5-day real sound sample - Towsey et al. 2014. Ecological Infomatics 21: 110-119.

Many Indices

- Average signal amplitude
- Background noise
- Signal-to-noise ratio (SNR)
- ACI
- Acoustic activity
- Count of acoustic events
- Avg duration of acoustic events
- Entropy of signal envelope (temporal entropy = H[t])
- Mid-band activity
- Entropy of average spectrum (= H[s])
- Entropy of spectral maximum (= H[m])
- Entropy of spectral variance (= H[v])
- Spectral diversity
- Spectral persisitence

All defined in Towsey et al. 2014. Ecological Infomatics 21: 110-119.

Many Indices

- Average signal amplitude
- Background noise
- Signal-to-noise ratio (SNR)
- ACI
- Acoustic activity
- Count of acoustic events
- Avg duration of acoustic events
- Entropy of signal envelope (temporal entropy = H[t])
- Mid-band activity
- Entropy of average spectrum

(= H[s])

- Entropy of spectral maximum

(= H[m])

- Entropy of spectral variance

(= H[v])

- Spectral diversity
- Spectral persisitence

All defined in Towsey et al. 2014. Ecological Infomatics 21: 110-119.

Visualisation of Large-scale Recordings – Using Indices to Reduce "Noise"

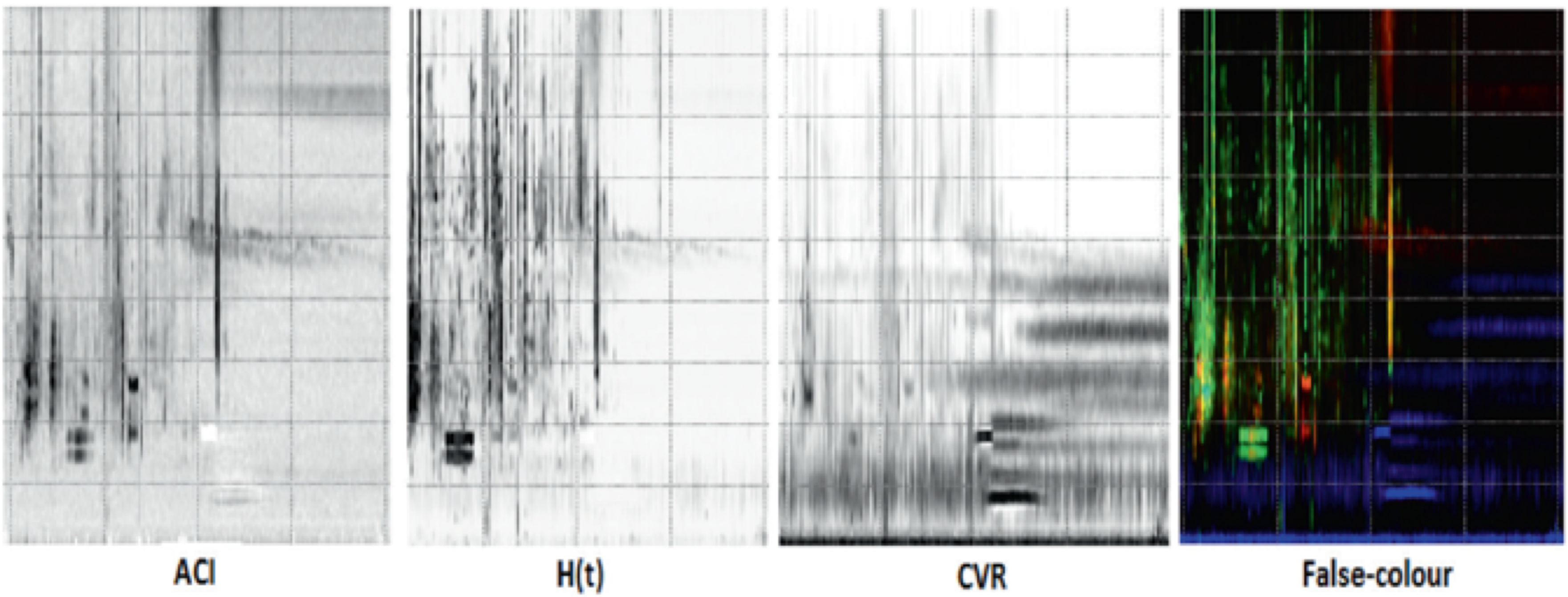


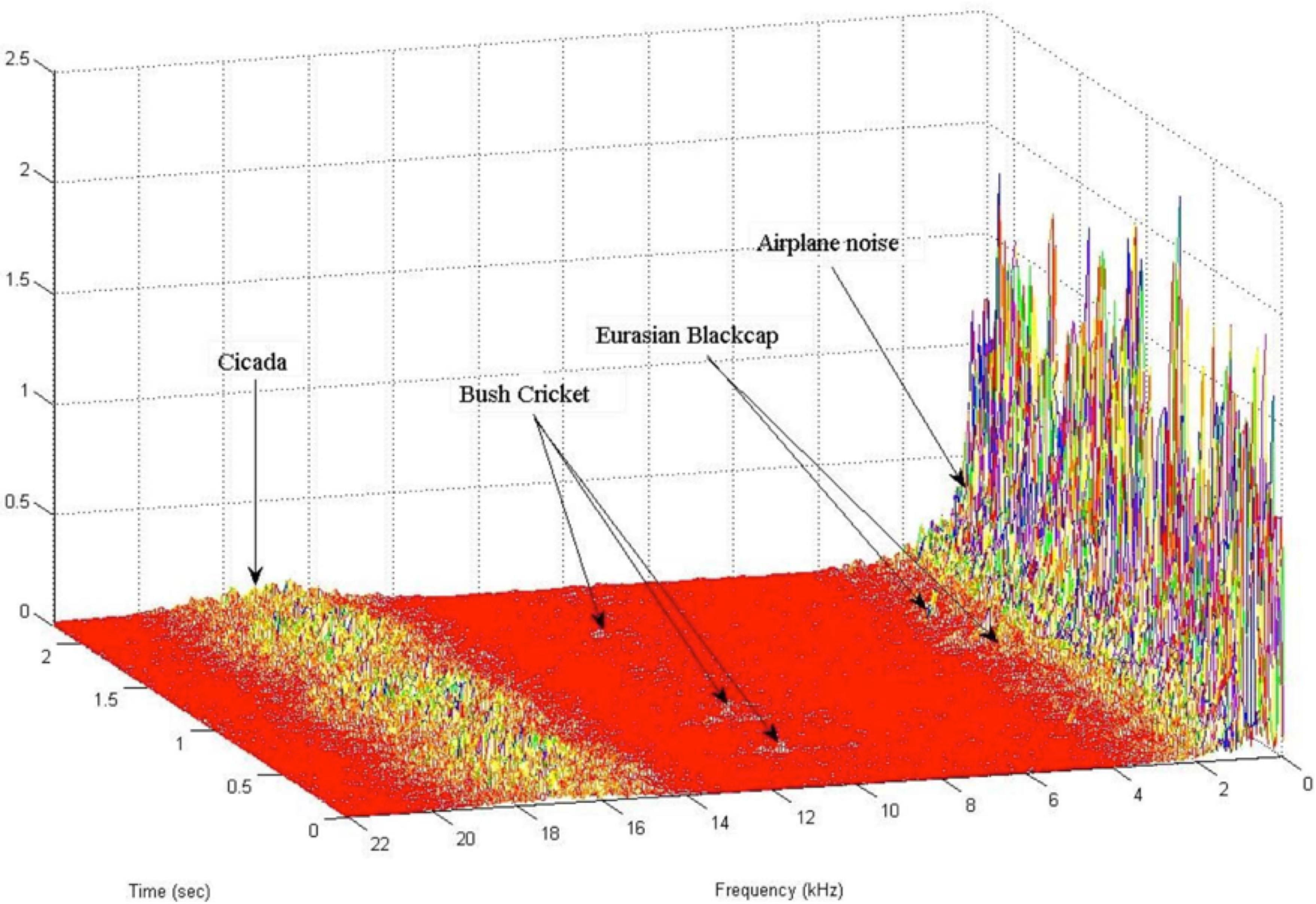
Figure 3: The false-color spectrogram on the right was obtained by combining the ACI, 1-H[t] and CVR spectrograms in red, green and blue colors respectively.

A visual approach to automatic classification from recordings in the wild

- A multi-instance, multi-label framework on bird vocalizations to detect simultaneously vocalizing birds of different species.
- Integrates novel, image-based heterogeneous features designed to capture different aspects of the spectrum.
- monitor 78 bird species, 8 insects and 1 amphibian (total = 87 species under challenging environmental conditions)
- The *classification accuracy* assessed by independent observers = 91.3% (note not compared to traditional surveys)

Potamitis, I. 2014. PLoS1 9(5):e96936

Illustration of Sound Interference



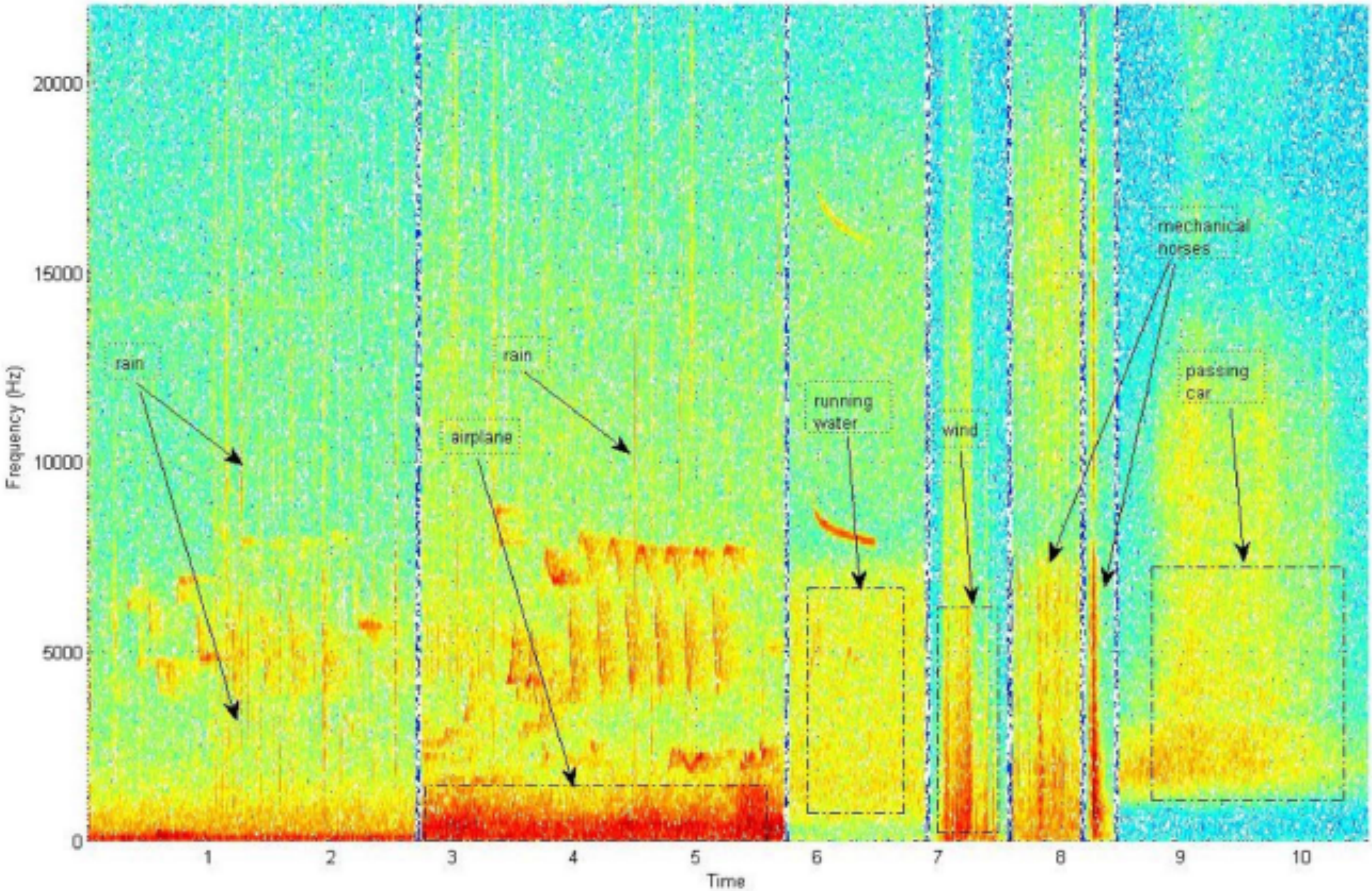


Figure 2. Types of anthropogenic and abiotic interfering sounds.
doi:10.1371/journal.pone.0096936.g002

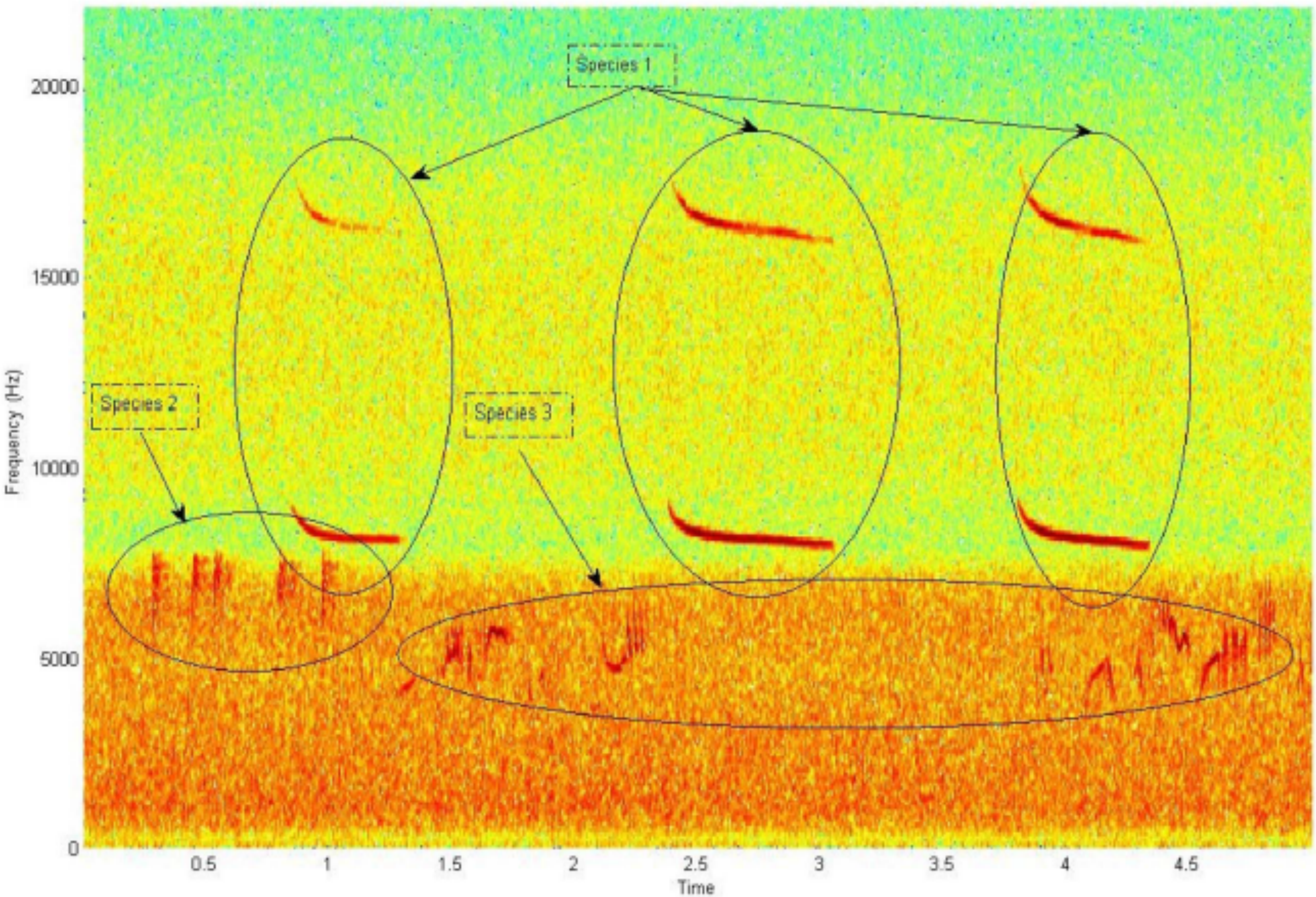


Figure 3. Spectrogram corresponding to a recording with 3 partially overlapping bird species (trainfile005 in NIPS20134B database). The lower part of the spectrum is coloured by the sound of running water and strong wind.
doi:10.1371/journal.pone.0096936.g003

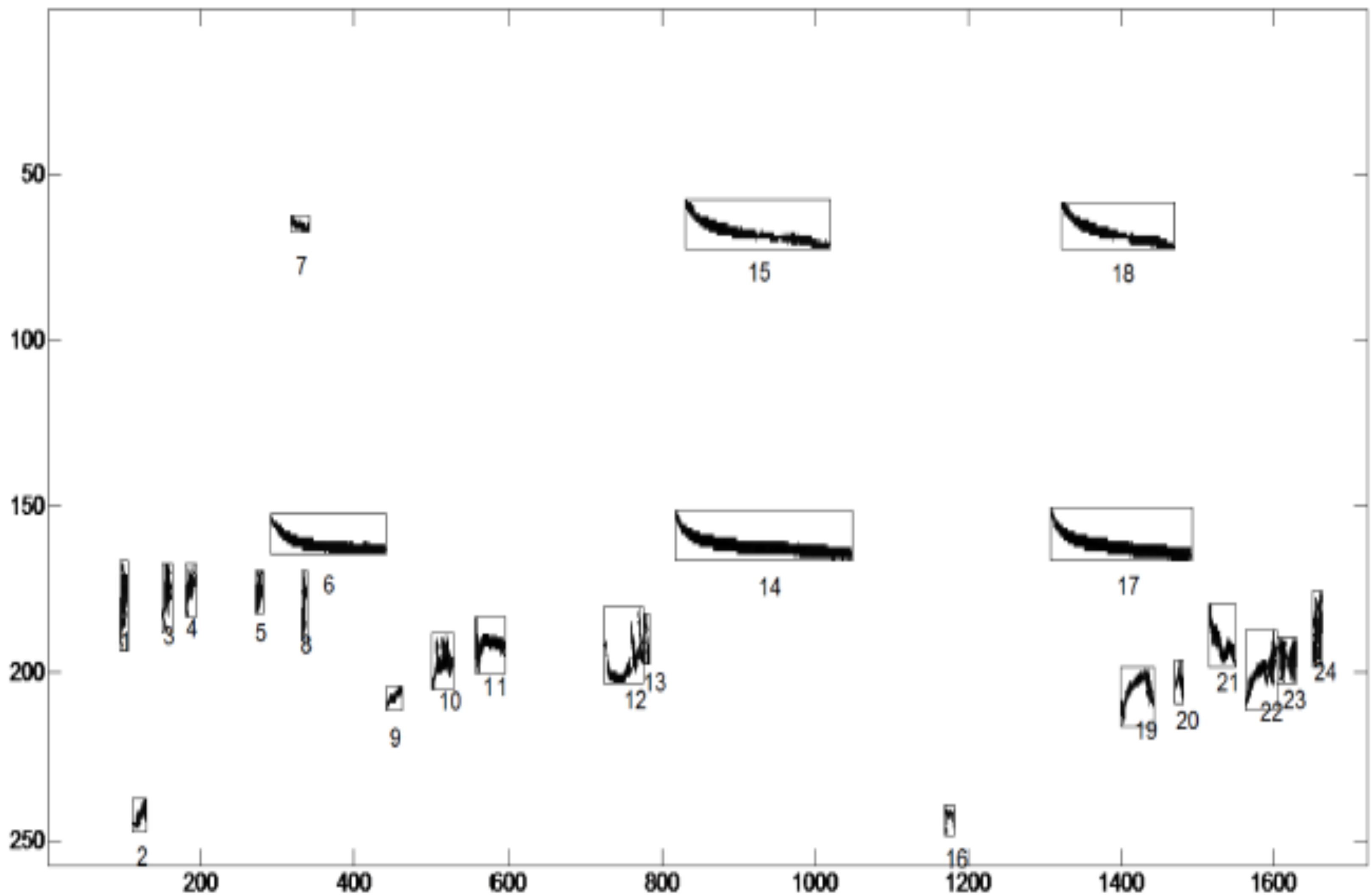


Figure 4. Detected spectrogram blobs of Fig. 3. Derivations and enumeration of the masks. Axis are enumerated according to their pixel index.
doi:10.1371/journal.pone.0096936.g004

Conclusions

- ARUs could be *extremely valuable* to collect a massive amount of data on species presence/absence, richness
- Massive amount of data is a double edged sword
- ARUs are especially good for rare or (acoustically) hard-to-detect species
- There is a *great deal* of research to be done in how best to analyse this data





One more thing

- Caller-listeners, rather than just listeners may increase the probability that a rare thing will call

- Such an invention increases the probability of calling by rare species

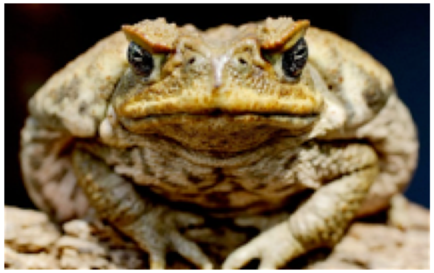
- Increases detectability of rare species, because then we know WHEN to look for their calls in long recordings





Current work: Detecting Invasive Species

- Detecting the arrival of invasive cane toads on Groote



- Listening & Calling for toads
- Working with the Anindilyakwa Land Council

- Hoping not to get an answer!

Monthly Average Spectrogram

- Averaging values of acoustic indices over consecutive days
- More ‘washed out’ appearance due to averaging
- But seasonal changes in acoustic landscape are clearly visible
- Morning chorus strongest during late winter and early spring
- Night-time Orthopteran sounds are minimal during winter months

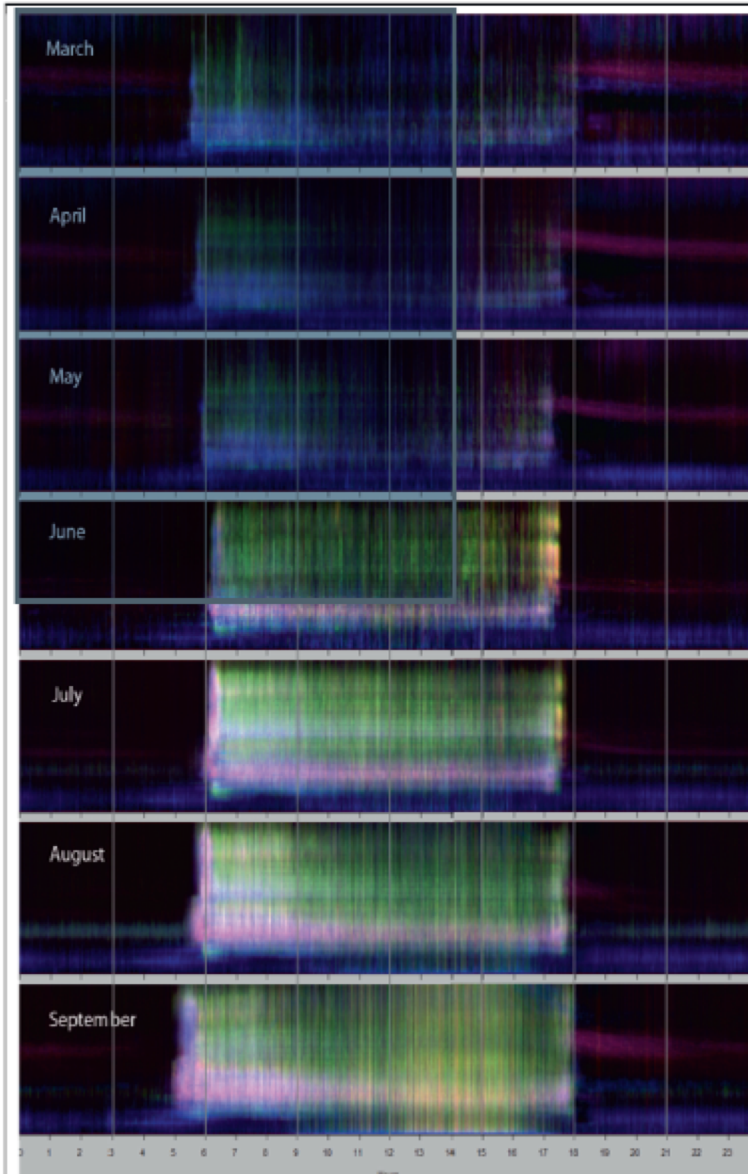


Figure 4: Monthly averaged spectrograms derived from the months March to September, 2013.