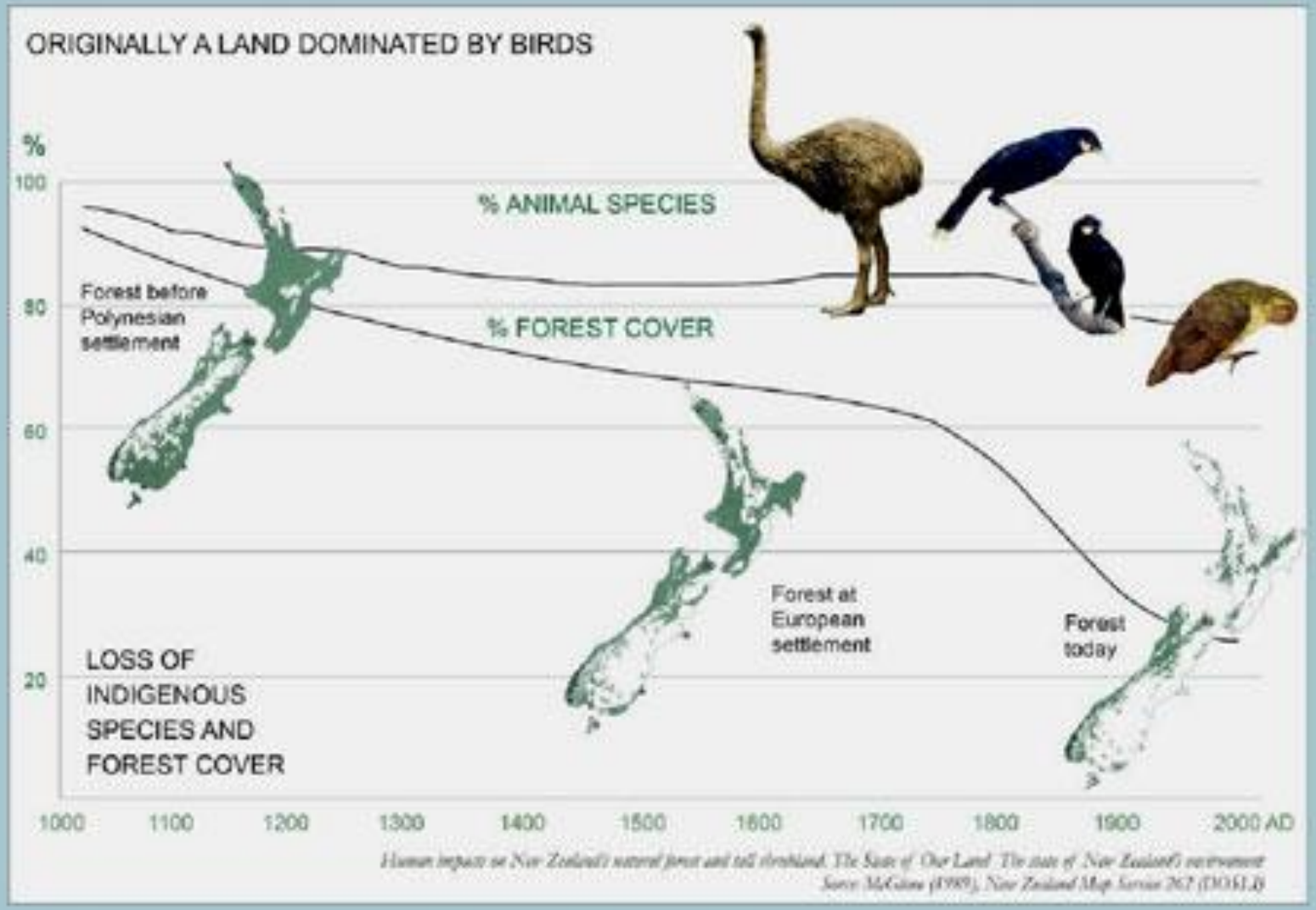
AviaNZ:

Automated Bird Song Recognition

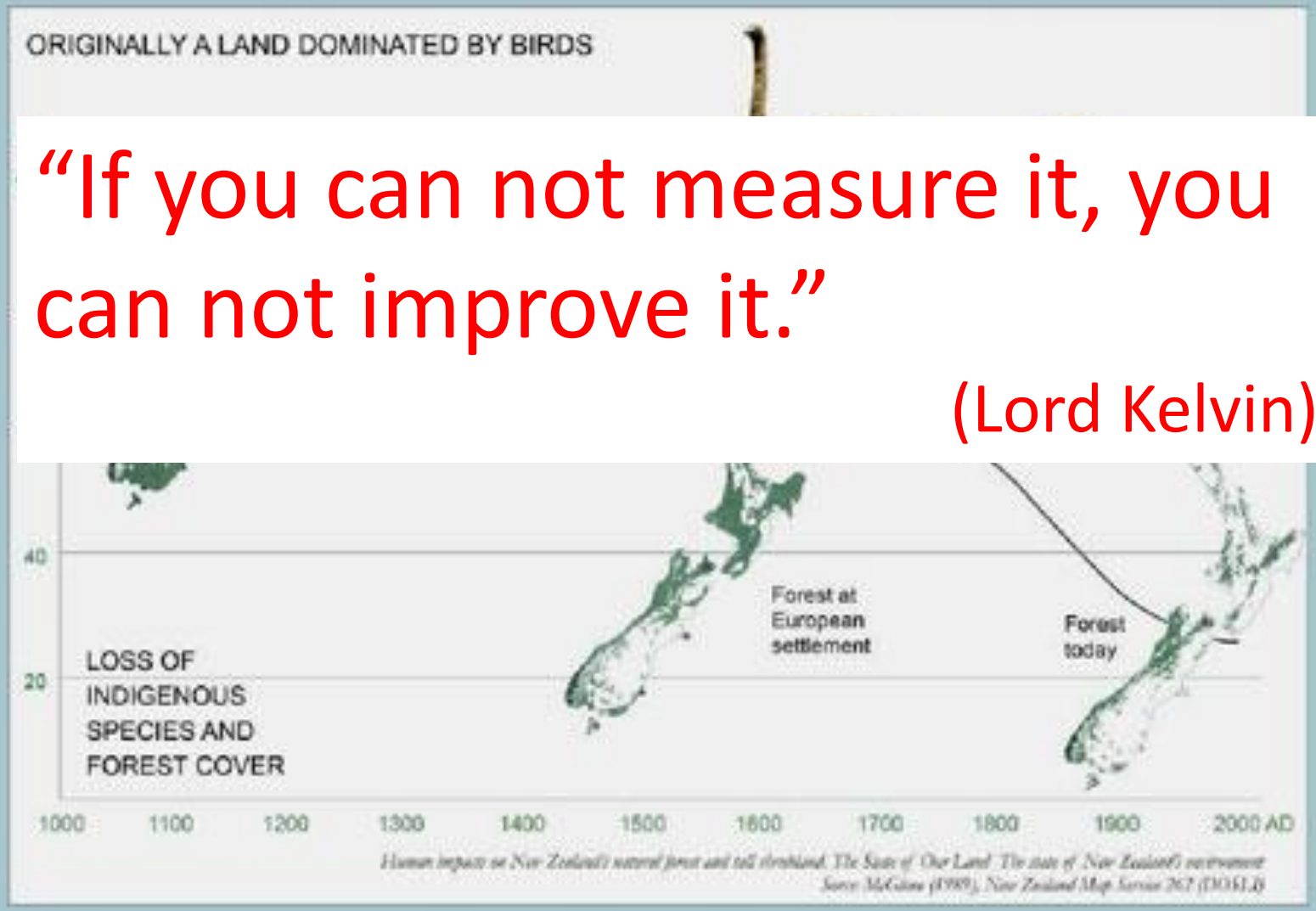
Stephen Marsland and Isabel Castro
(joint work with Nirosha Priyadarshani and Richard Witehira)

Motivation



Source: http://doc.govt.nz

Motivation



"If you can not measure it, you can not improve it."

(Lord Kelvin)

Source: http://doc.govt.nz

Measuring bird populations

Birds: cryptic beasts

- Dense vegetation
- Well-camouflaged
- Sometimes nocturnal
- Can fly, so can be in hard-to-reach areas

But they make plenty of noise!

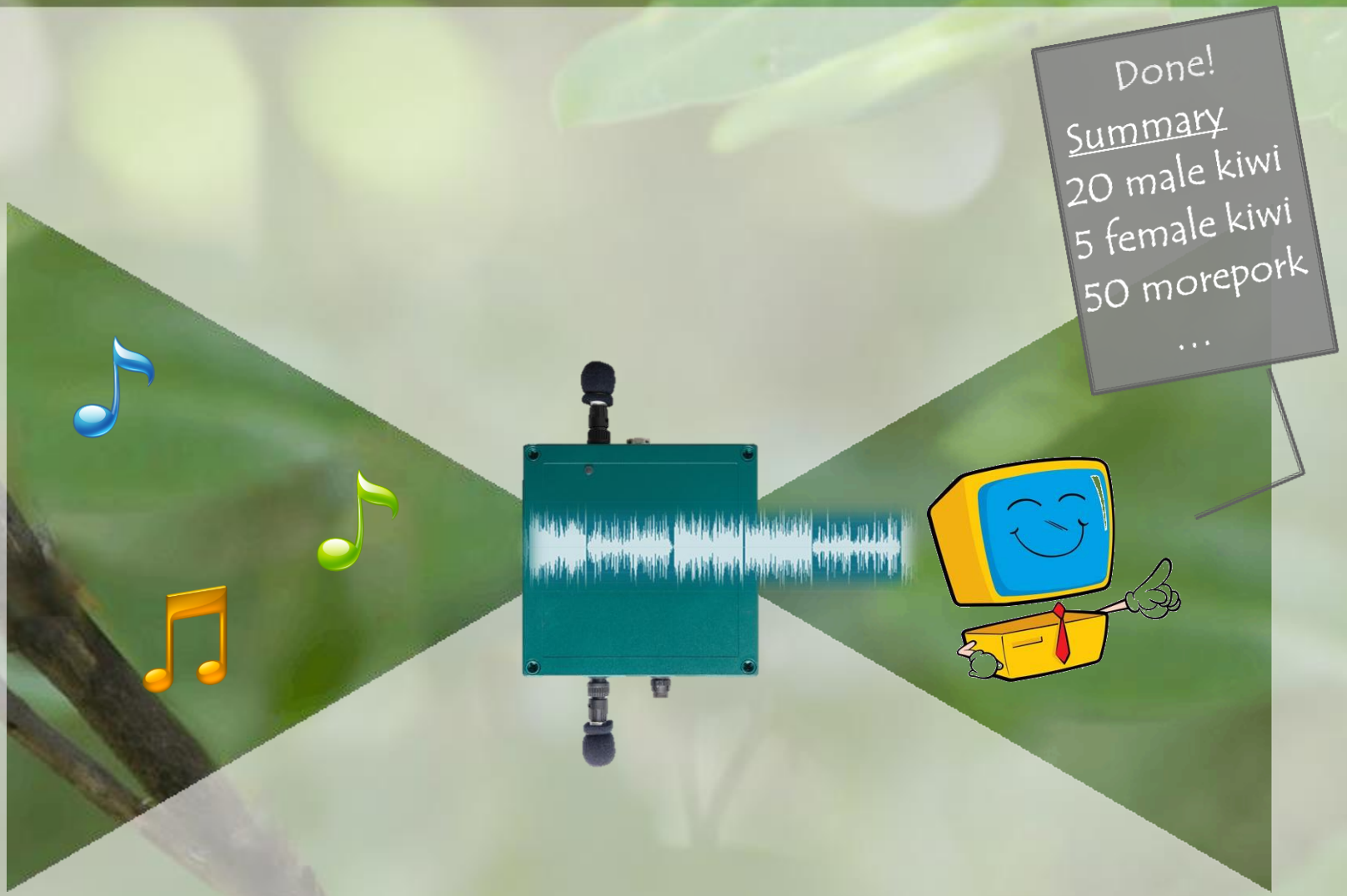
Sound courtesy of Alex Brighten

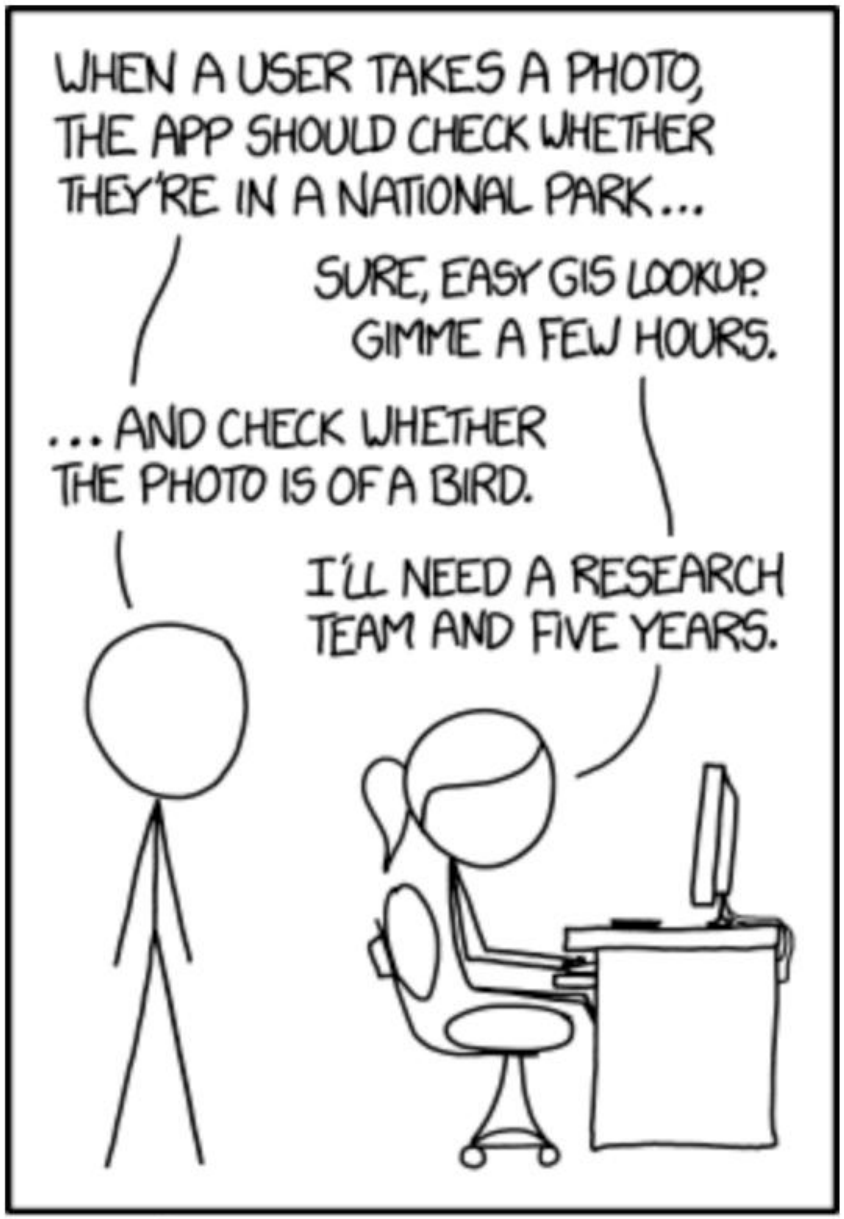
Call count surveys

8 hours a day
2 weeks
Recording is subjective and tedious

24 hours a day
3 months
Analysis is subjective and tedious

Automated Bird Recognition





IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

xkcd.com

It's kind of like speech, but…

Birds:

- Won't sing nicely straight into the microphone
- Overlap their calls
- Have large repertoires
- … and even dialects

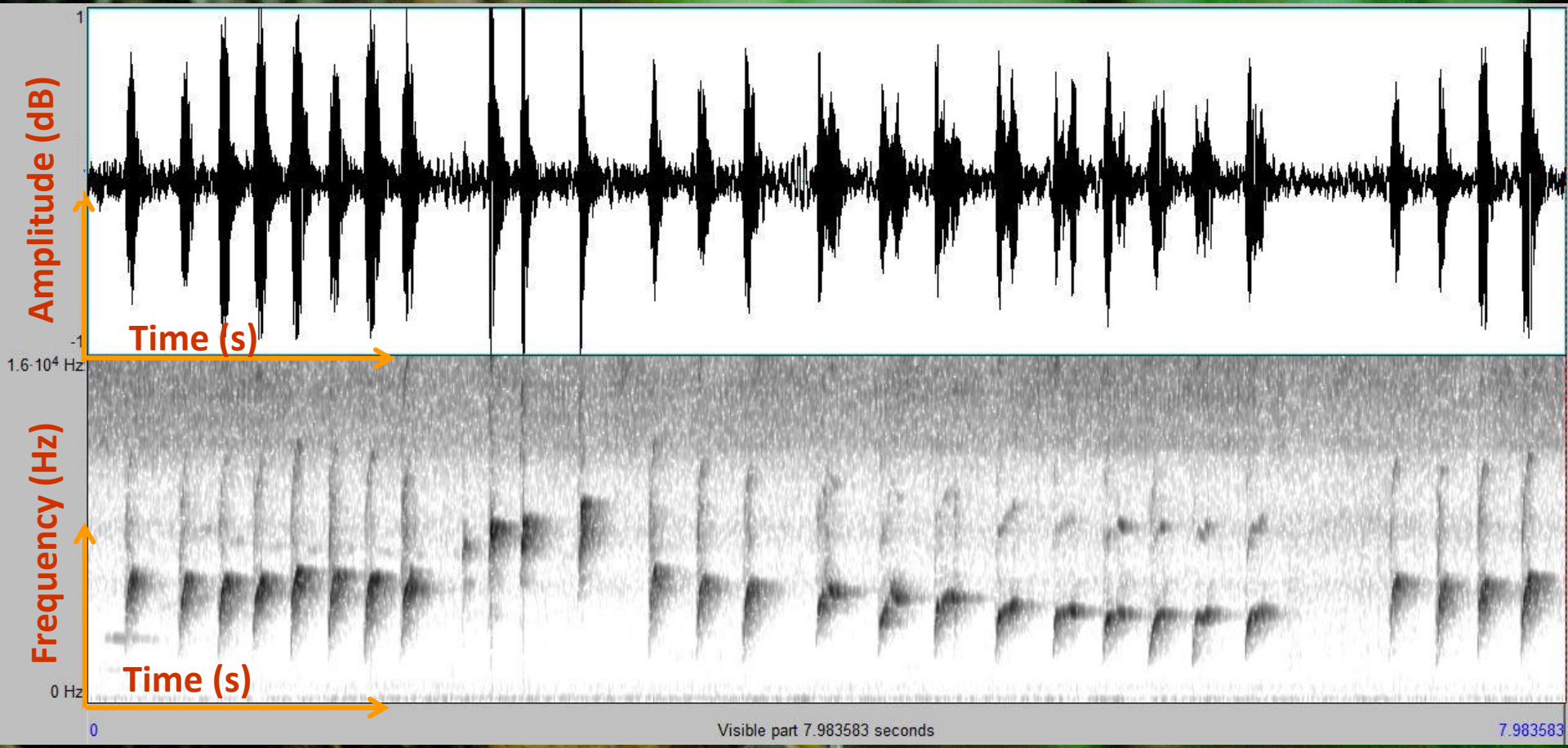
Recordings are:

- Noisy (environmental and recording noise)

“A single curve, drawn in the manner of the curve of prices of cotton, describes all that the ear can possibly hear as the result of the most complicated musical performance.... That to my mind is a wonderful proof of the potency of mathematics.”

(Kelvin again!)

It’s kind of like speech, but…



It's kind of like speech, so…

Bird	No of Ref	No of Test from same source	No of Test from different source	DTW Recognition Rate %	VQ Recognition Rate %
Kiwi	*1*	4	-	100	100
Morepork	*1*	3	7	100	100
Bellbird	*1*	22	18	97.5	77.5
Robbin	*1*	21	-	95.24	66.6
Tui	*1*	25	15	90	41.6
				96.55%	**77.1 %**

Pretty nearly perfect!

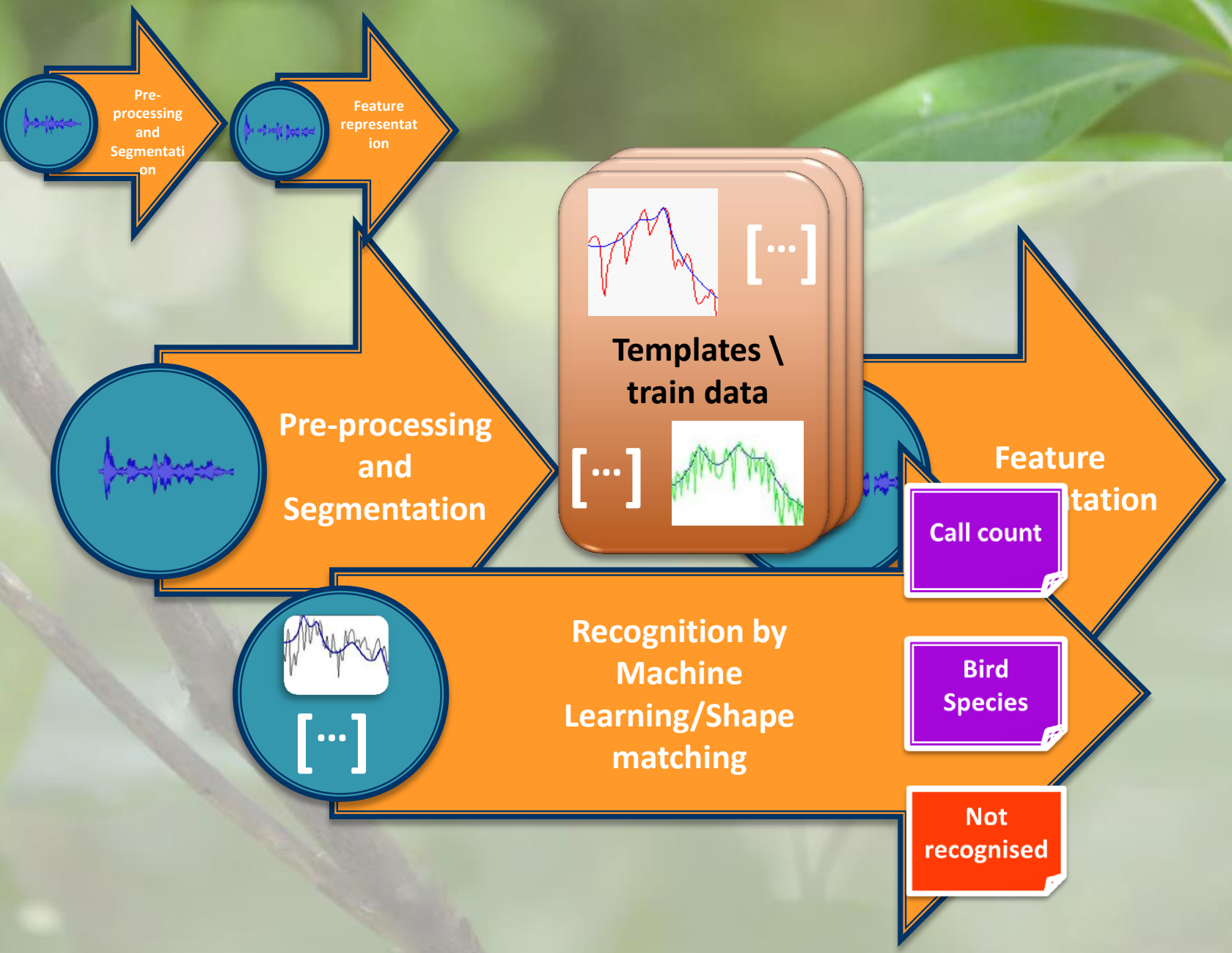
Providing numbers of birds are small, sources are clean, data chosen carefully, etc. etc.

SoundID

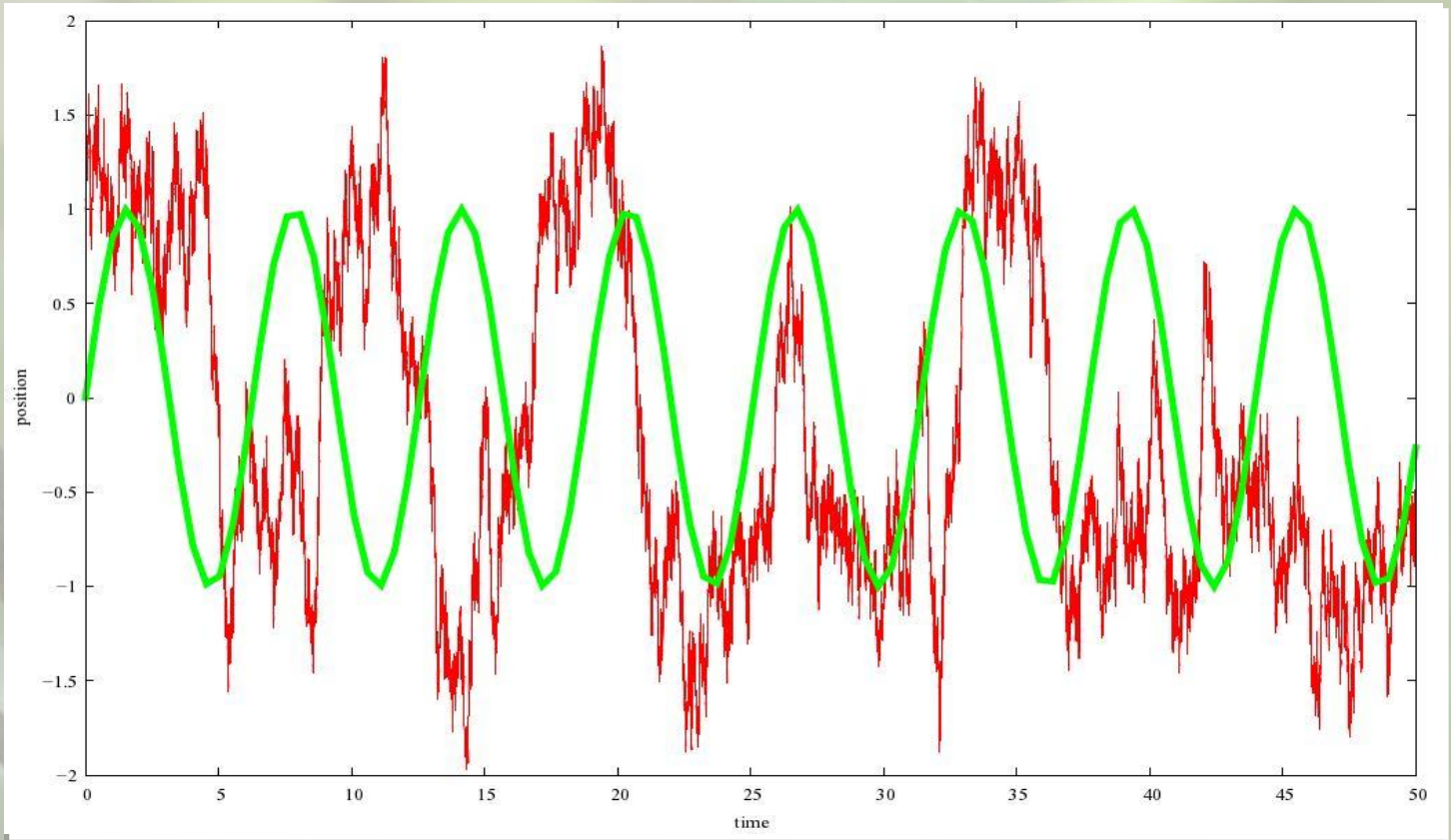
SoundID Home Page - Mozilla Firefox
File Edit View History Bookmarks Tools Help
SoundID Home Page
www.soundid.net/SoundID/index.html
Google
Most Visited Getting Started Latest Headlines
2013
SoundID for Sound Recognition and Monitoring
Home
Company Info
Software
Hardware
Field Deploys
Research Papers
Links
Training
Products
Downloads
WAV File References
FAQ's
AWSRG Information
SoundID: At last a sound recognition system that really works!!
SoundID uses all new, effective algorithms to bring you world-class sound detection and recognition . Our system requires only minor tweaking to optimise it for different sounds and it really works!!
We have the hardware and software to change the way you monitor and recognise sounds.
SoundID is all new with, no classifiers, no FFT no AI. This is not just another 80% accurate (20% inaccurate) system good for a limited number of sounds, with lots of false positives, that leaves you frustrated, but something truly professional that can recognise any sound with better than 95% accuracy . What is more it can recognise any collection of sounds of any size (that need not in any way be related).
See this Dawn Chorus Example !
See the Dawn Chorus Video
Recognise any sound - wildlife, marine, industrial, mechanical, for artificial intelligence, robotics or whatever, at better than 100 times or more faster than real time and with an accuracy comparable to a human expert!! We can in fact recognise and quantify differences in sound that a human would entirely miss.
A complete suite of software tools for bioacoustics, sound recognition, professional sound analysis for the most demanding of applications. This is a specialist non-voice sound recognition application.
SoundID 2012: A General Purpose Sound Recognition System
Program Menu

SoundID

They say:
95% accuracy
We say:
72% on six birds after 2 weeks of work
01_Advertising_01
4001 points LPC
8576 points wave
Standard Pattern
Peak Frequency : 6891Hz
01_Advertising_10
10880 points wave
Input Pattern
Peak Frequency : 6632Hz
File
Open and LPC
Sound
Geometric Distance
2.35 degrees
End
GD
Input Pattern File Name
degrees
Run Batch
Threshold of GD
Matches
out of
Average Match
No Filter
HP Filter
LP Filter
BP Filter
BE Filter
Filter Order
Cut-off Freq.1
Cut-off Freq.2
Rate of Random Noise
Normalization
Frame Width for LPC
points
LPC Order
Auto
Spread of Weighting Curve
Weighted GD by Spectrum Energy

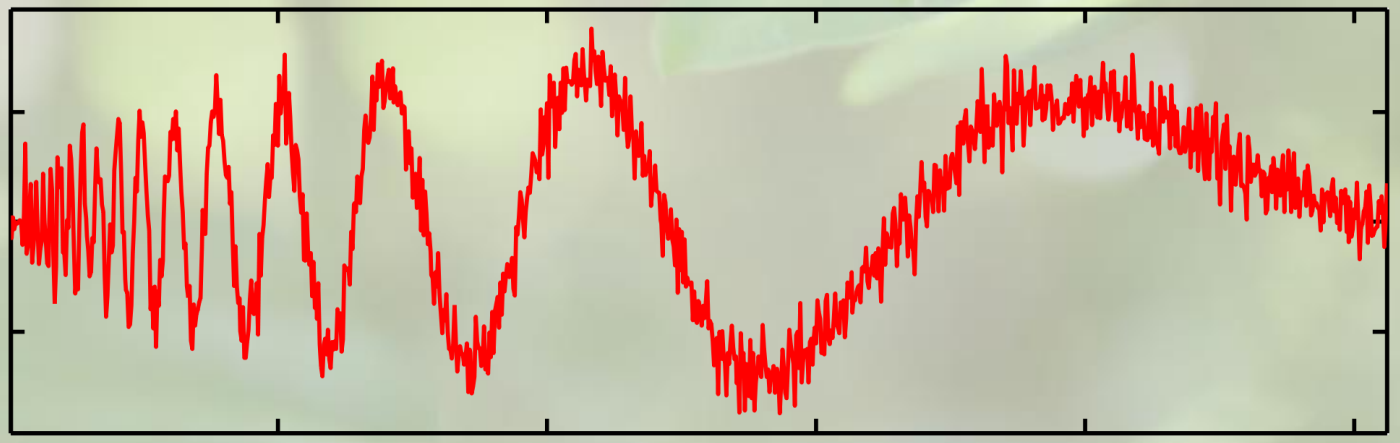

Pre-processing and Segmentation
Feature representation
Templates \ train data
[…]
[…]
Pre-processing and Segmentation
Feature tation
Call count
Recognition by Machine Learning/Shape matching
[…]
Bird Species
Not recognised

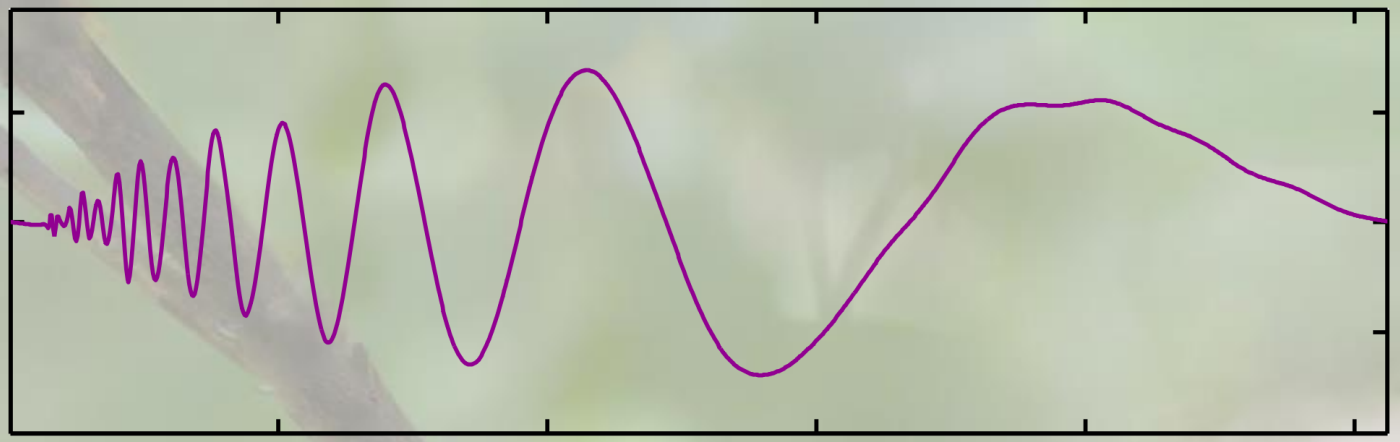
Noise



Signal= Pure Signal + Noise

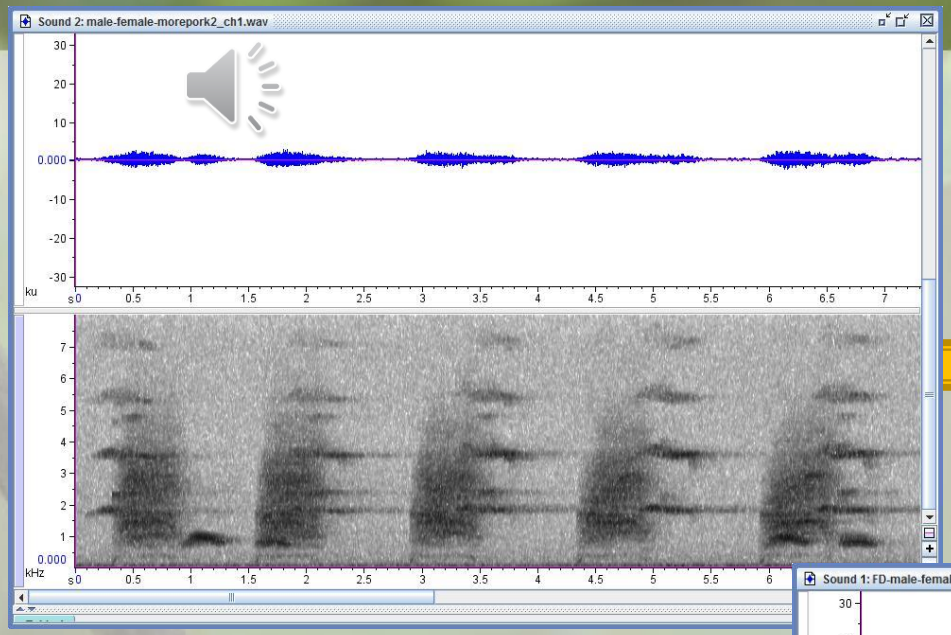
Wavelets

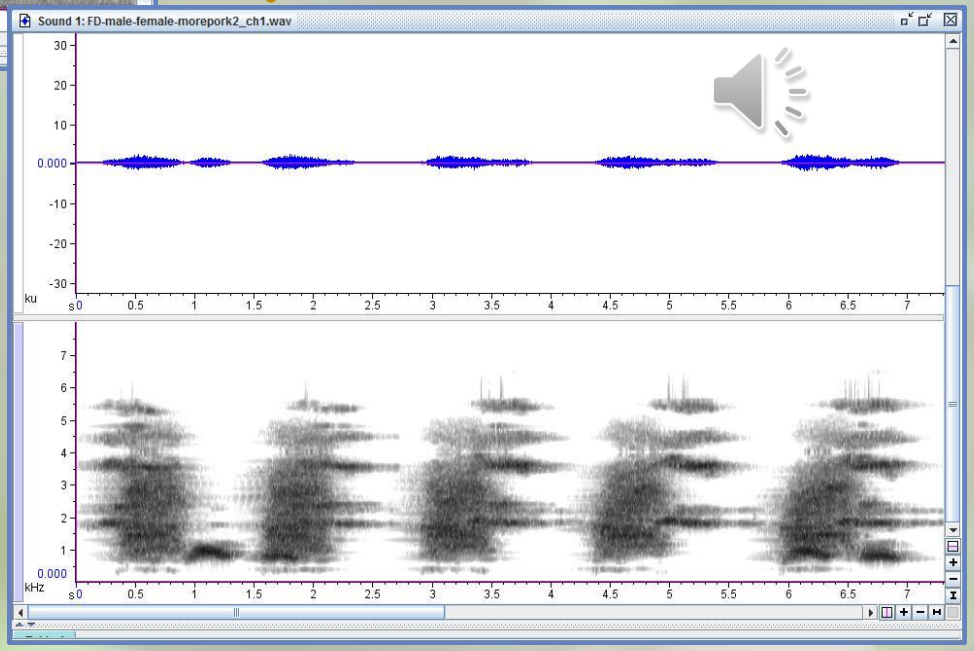




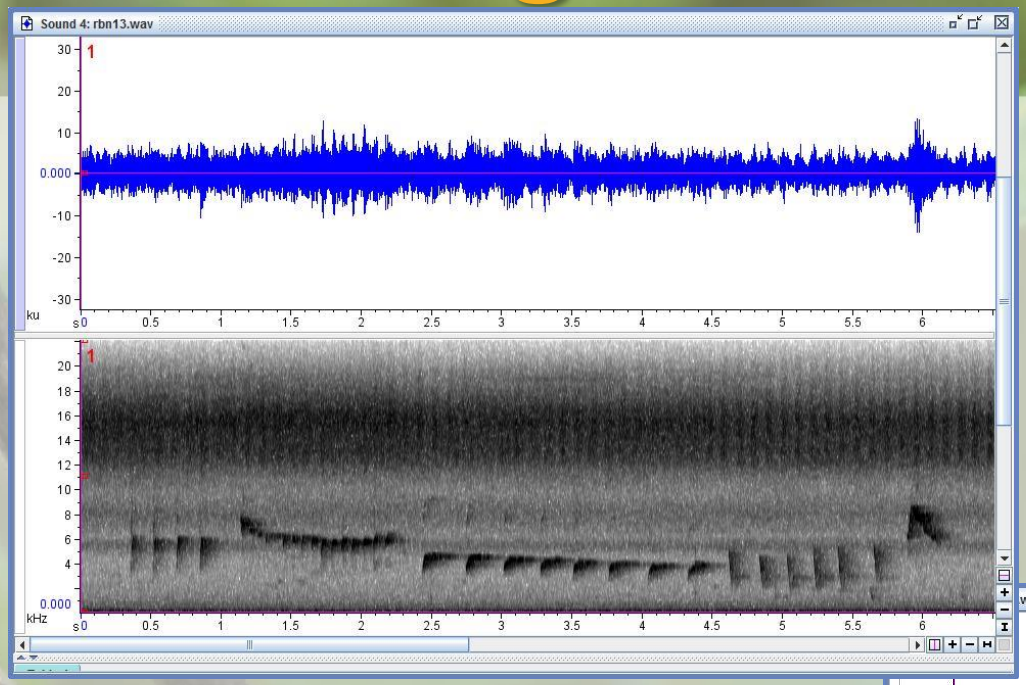
Denoising

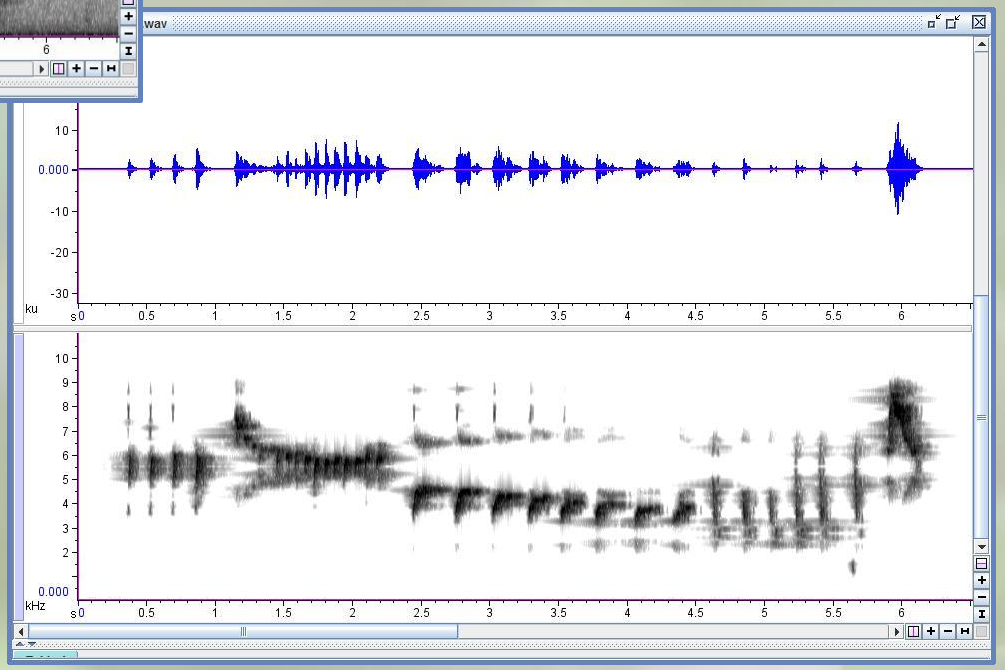
I. Castro


Sound 2: male-female-morepork2_ch1.wav
ku
kHz
s

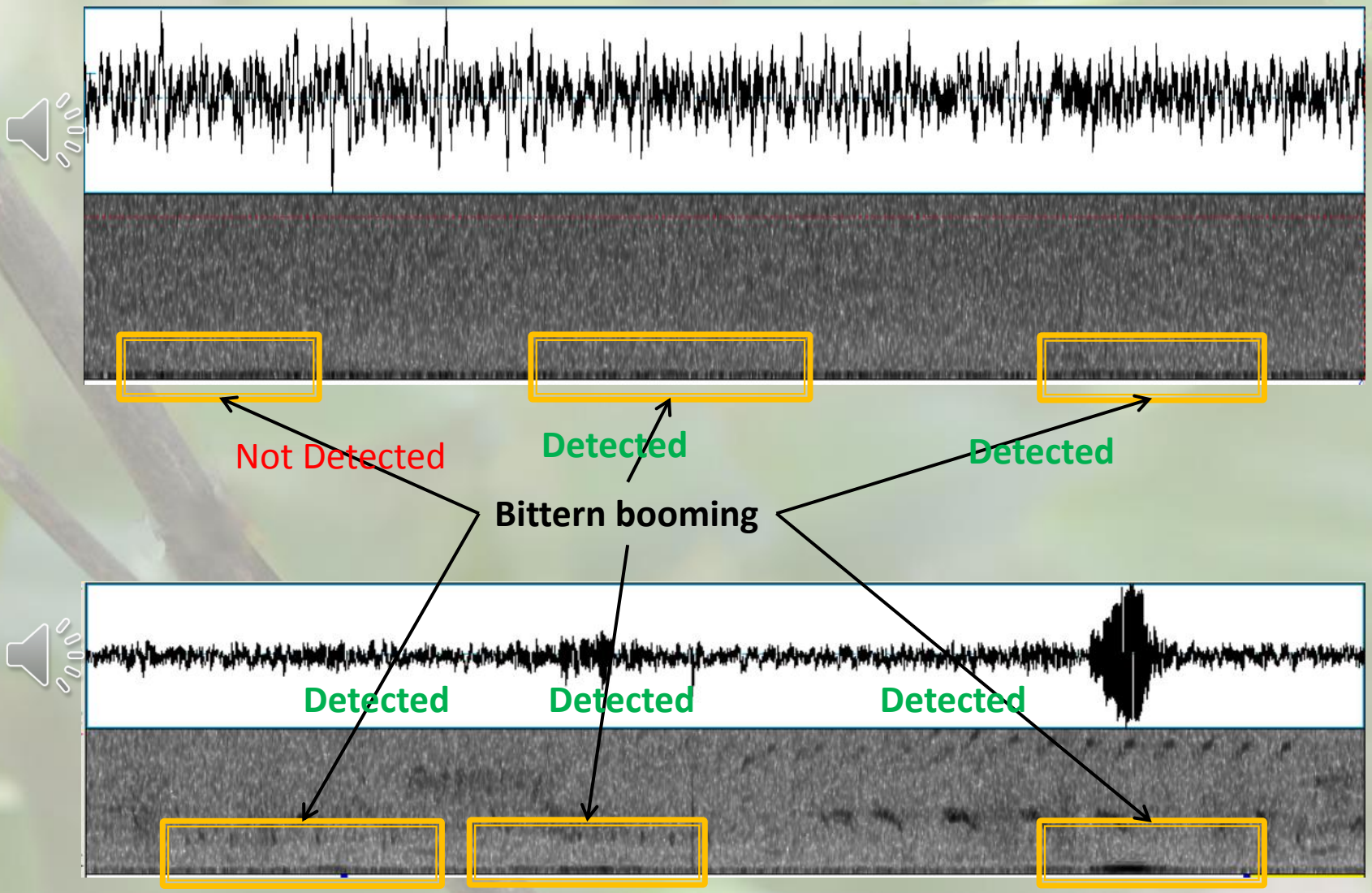

Sound 1: FD-male-female-morepork2_ch1.wav
ku
kHz
s

Denoising

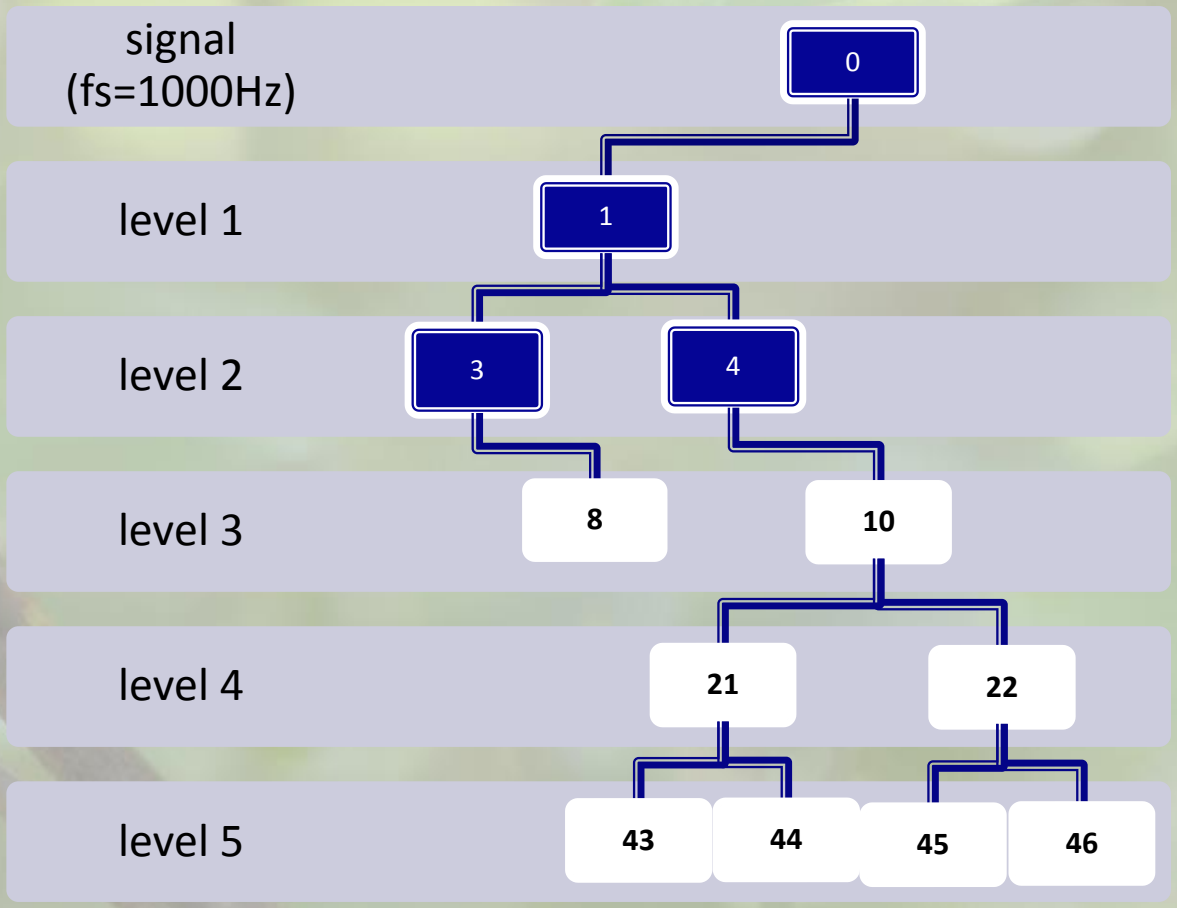




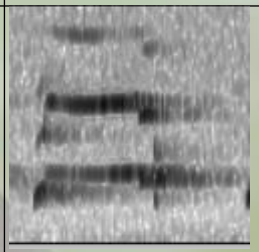
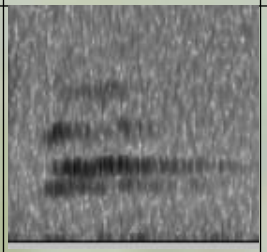
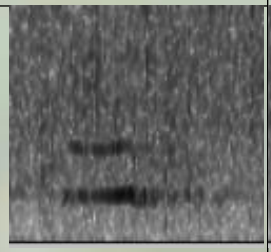
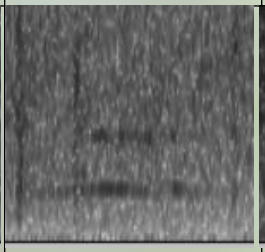
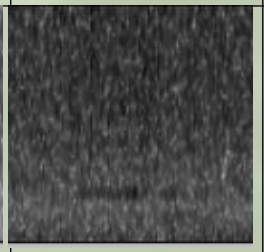
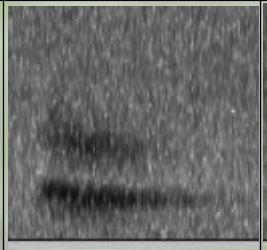
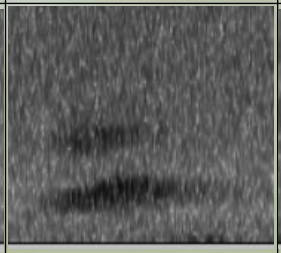
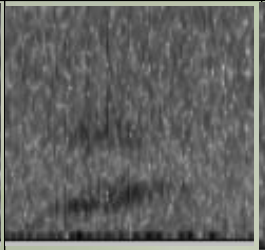
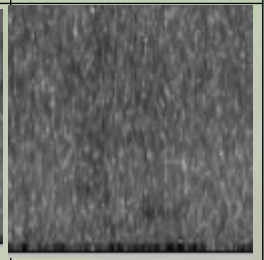
Bitterns: Booming Good Fun



Correlations in the Wavelet Tree

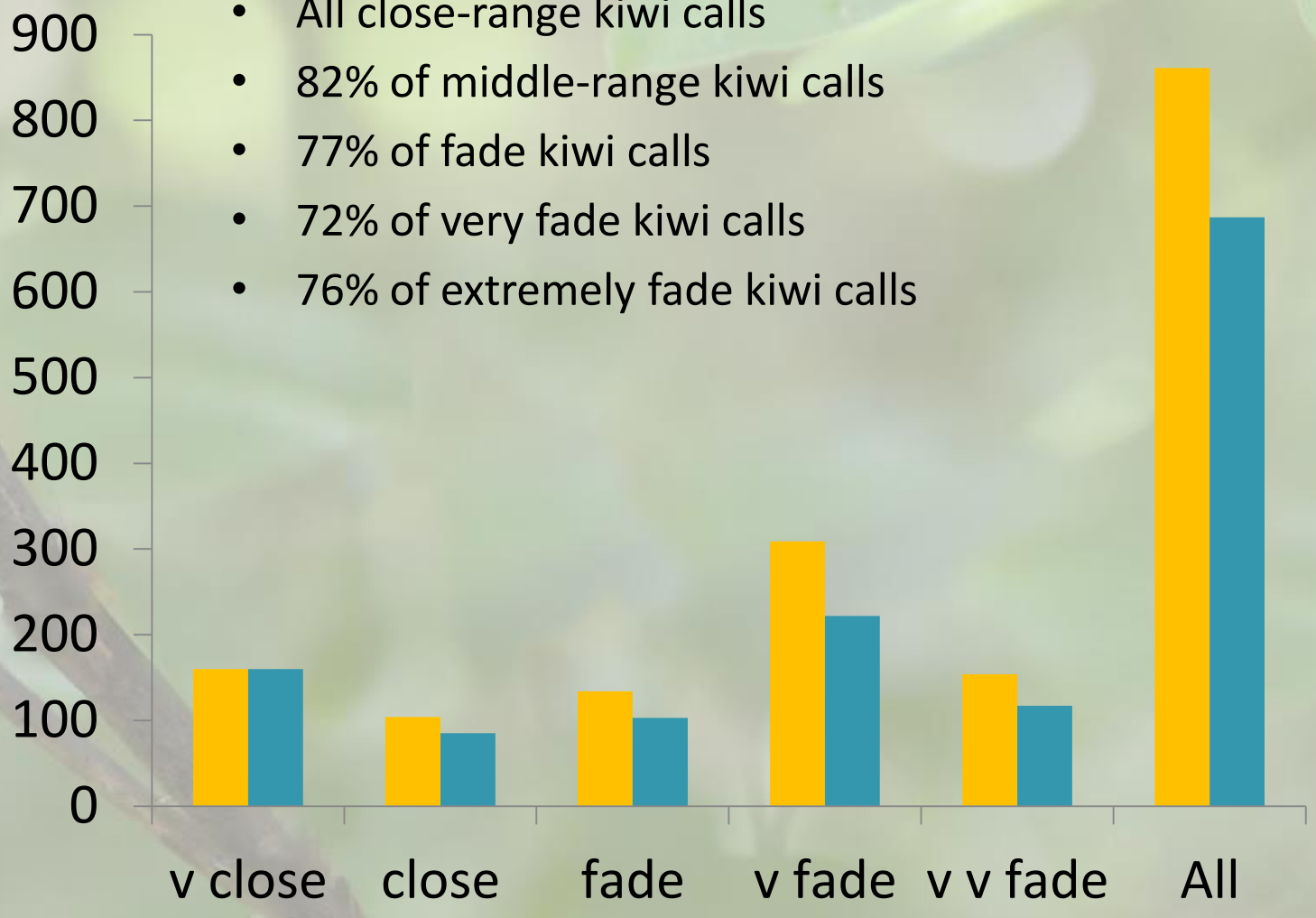


Kiwi

# Kiwi seconds (ground truth)		*v close*	*close*	*faded*	*v faded*	*v v faded*	*Total*
		160	104	134	309	154	861
e.g.	male						
	female						

Automatic (yellow) vs. human (blue)

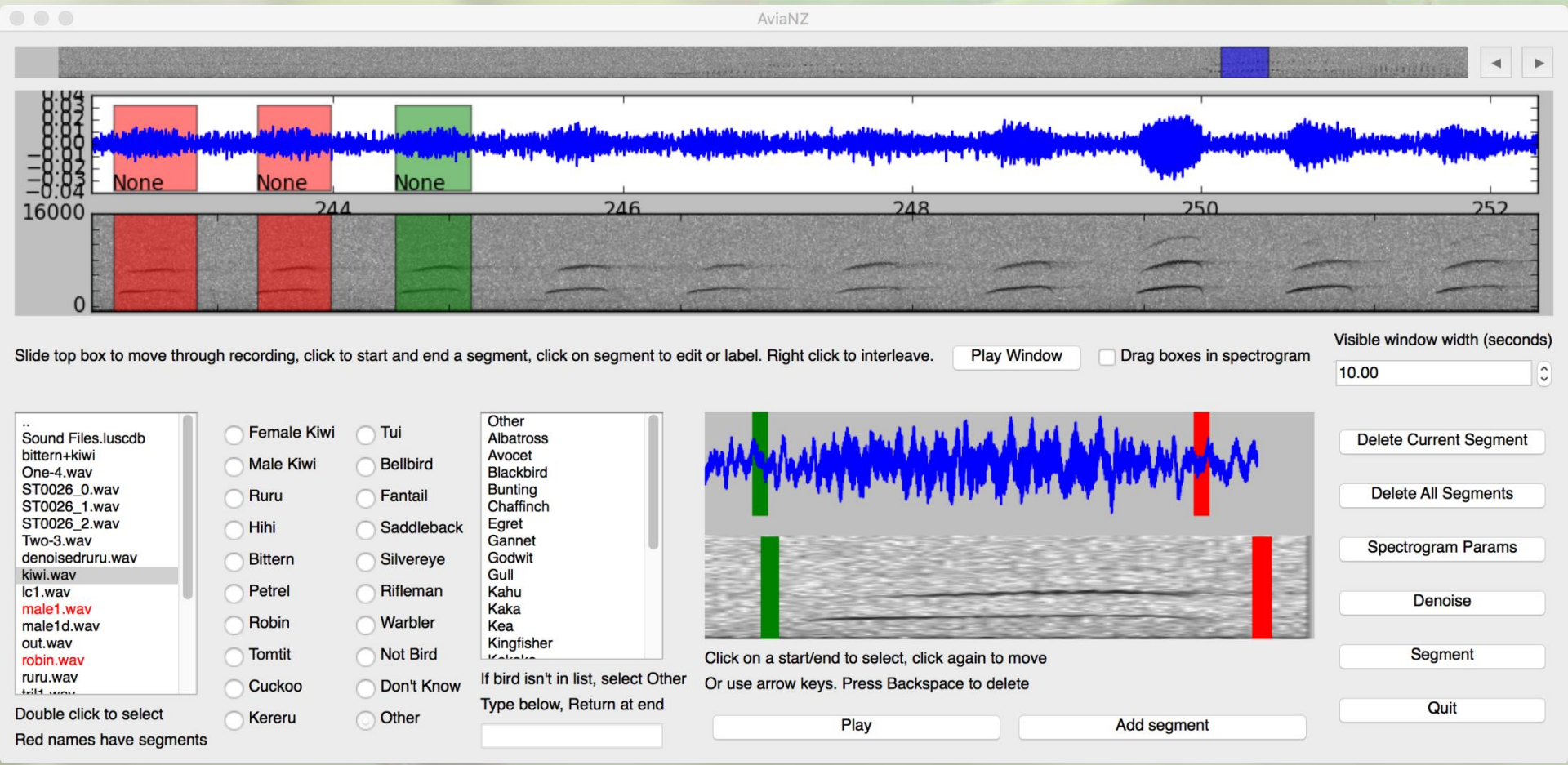
- All close-range kiwi calls
- 82% of middle-range kiwi calls
- 77% of fade kiwi calls
- 72% of very fade kiwi calls
- 76% of extremely fade kiwi calls

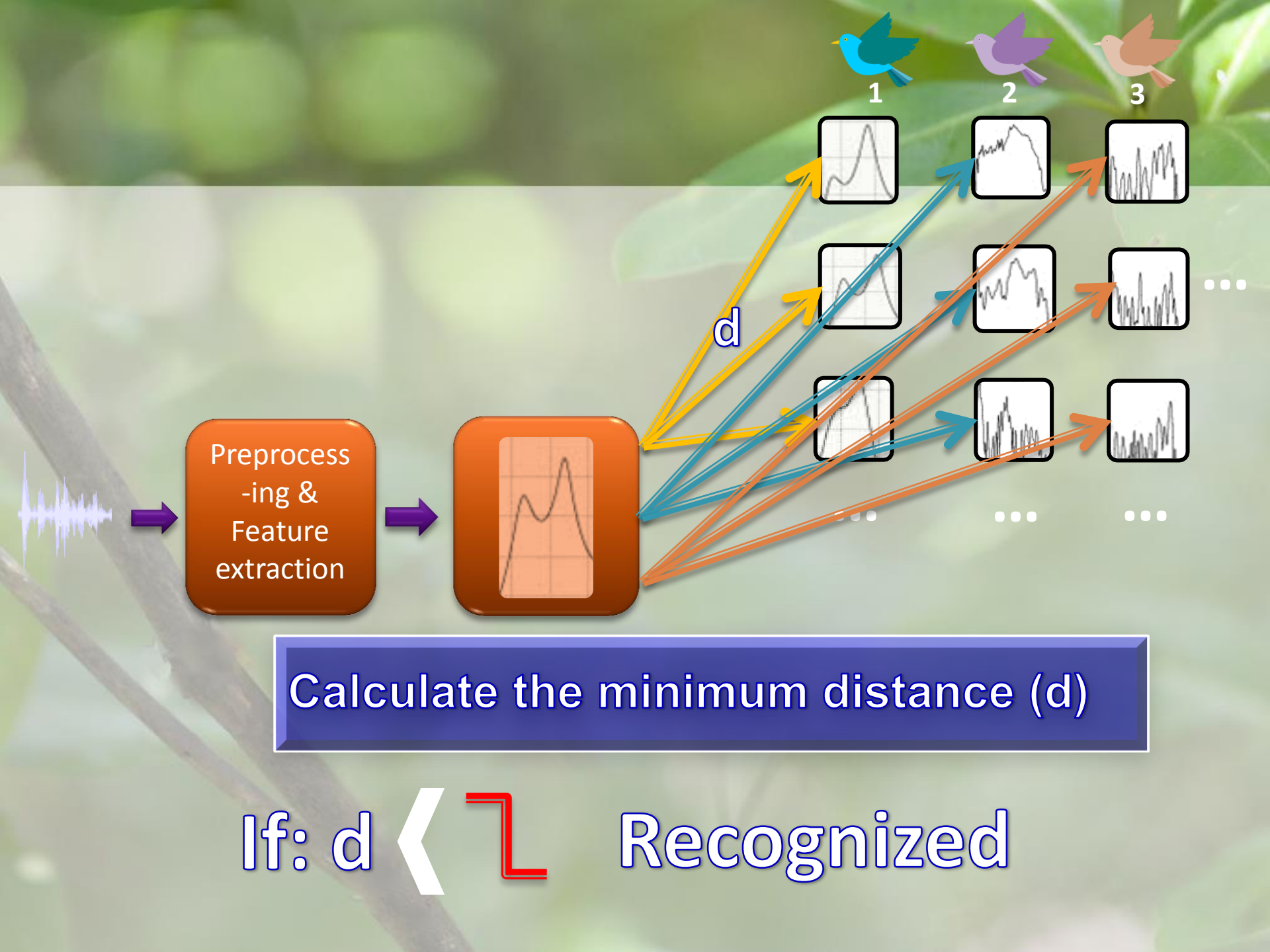


Results for a number of species

	Bittern	Kakapo	Brown Kiwi	Morepork	Overall (%)
Very Close	3/3	126/126	139/145	72/78	97%
Close	20/20	162/164	129/135	266/290	95%
Moderate	23/26	80/84	129/190	118/134	81%
Far	42/51	40/62	65/113	79/93	71%
Very Far	20/43	55/147	5/42	25/55	37%

Our Program


AviaNZ
None
None
None
16000
0
244
246
248
250
252
Slide top box to move through recording, click to start and end a segment, click on segment to edit or label. Right click to interleave.
Play Window
Drag boxes in spectrogram
Visible window width (seconds)
10.00
..
Sound Files.luscdb
bittern+kiwi
One-4.wav
ST0026_0.wav
ST0026_1.wav
ST0026_2.wav
Two-3.wav
denoisedruru.wav
kiwi.wav
lc1.wav
male1.wav
male1d.wav
out.wav
robin.wav
ruru.wav
Double click to select
Red names have segments
Female Kiwi
Male Kiwi
Ruru
Hihi
Bittern
Petrel
Robin
Tomtit
Cuckoo
Kereru
Tui
Bellbird
Fantail
Saddleback
Silvereye
Rifleman
Warbler
Not Bird
Don't Know
Other
Other
Albatross
Avocet
Blackbird
Bunting
Chaffinch
Egret
Gannet
Godwit
Gull
Kahu
Kaka
Kea
Kingfisher
If bird isn't in list, select Other
Type below, Return at end
Click on a start/end to select, click again to move
Or use arrow keys. Press Backspace to delete
Play
Add segment
Delete Current Segment
Delete All Segments
Spectrogram Params
Denoise
Segment
Quit


1
2
3
d
Preprocess
-ing &
Feature
extraction
Calculate the minimum distance (d)
If: d
Recognized

Future Work

- Protocols
  - Where and when to record
  - Linking to community groups
- More machine learning
- Comparison with humans
- Linking calls with abundance
- Interpreting 5 minute call counts
- And more, and more