Website: Elephant Listening Project

Resources:
High level description of the data: Congo Soundscapes – Public Database : Elephant Listening Project

Data:
Annotation files for Elephant Sounds
Metadata and how to use the annotation files
elephant sounds

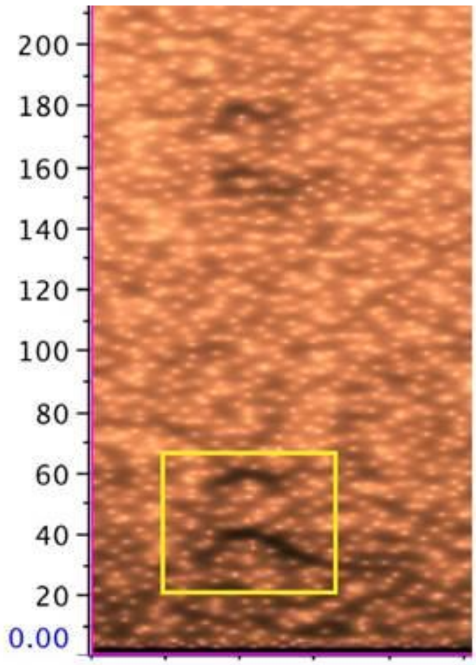
Raw data in entirety: (AWS S3 bucket )
https://registry.opendata.aws/elp-nouabale-landscape/

**Use Cases as proposed by Dr. Peter Wrege, Director Elephant Listening Project, Center for Conservation Bioacoustics, Cornell Lab of Ornithology**

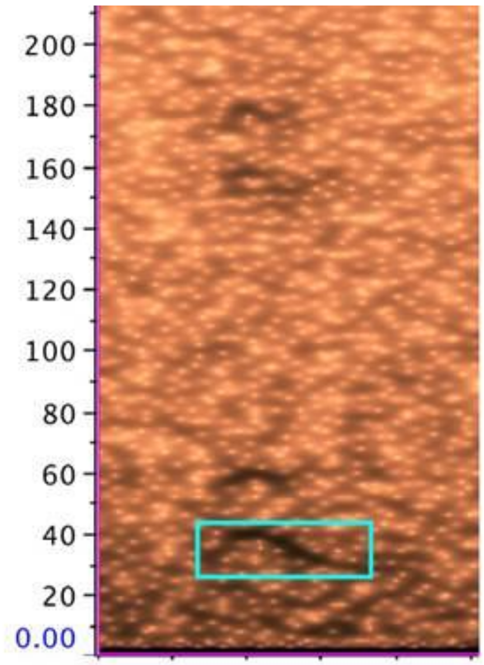
**Gunshot detection:**
The current model is very inefficient - less than .2% of tagged signals are gunshots and we typically get 10K- 15K tagged signals in a four month deployment at just one of the 50 recording sites.
The issue is we have about 200 good gunshots annotated, but because poaching is way too high, gunshots are still extremely rare in the sounds and it is extremely time consuming to create the "truth" logs where we can say that every gunshot in a 24hr file has been tagged. From our understanding this makes developing a detector more difficult.
But, can fresh minds might see a way through this?

**Determining the Demographics:**
The classification of age and sex is based on very specific measurements of the 'dominant frequency' - the frequency contour with the most energy in the elephant call. Our current detectors for elephant rumbles is not very specific about how it selects a call, and one of our detectors in development basically indicates a call in the succeeding seconds of the sound file, but gives no upper or lower bounds, and maybe not even the length.

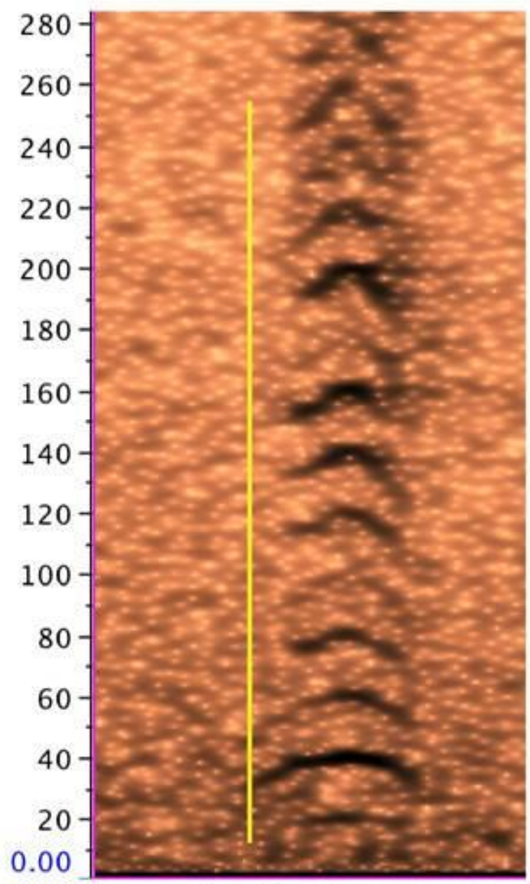
Below are a couple of spectrogram examples, on the left would be the information provided by our annotations and on the right how a measurement box would need to look like (note that we don't need a physically drawn box, of course, but begin and end time, low and high frequency. In the lower right example, there are actually at least two higher-frequency contours that have close to or even more energy than the one we target (arrows). So a solution might need to search for the highest energy from the bottom, or look for three and take the lowest with the most energy, or something like that.



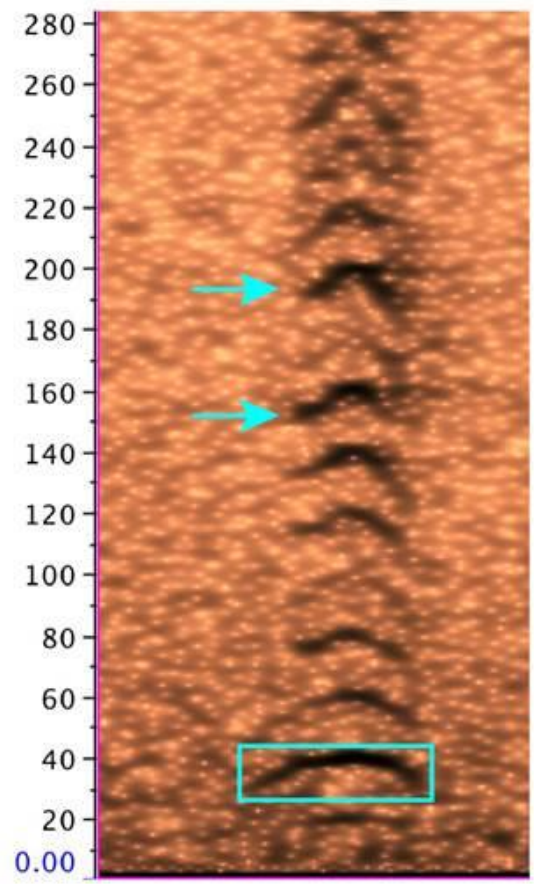

original selection bounds




tight bounds for measurement




original tagging only says
"there is an elephant call in the next 3 sec of sound"

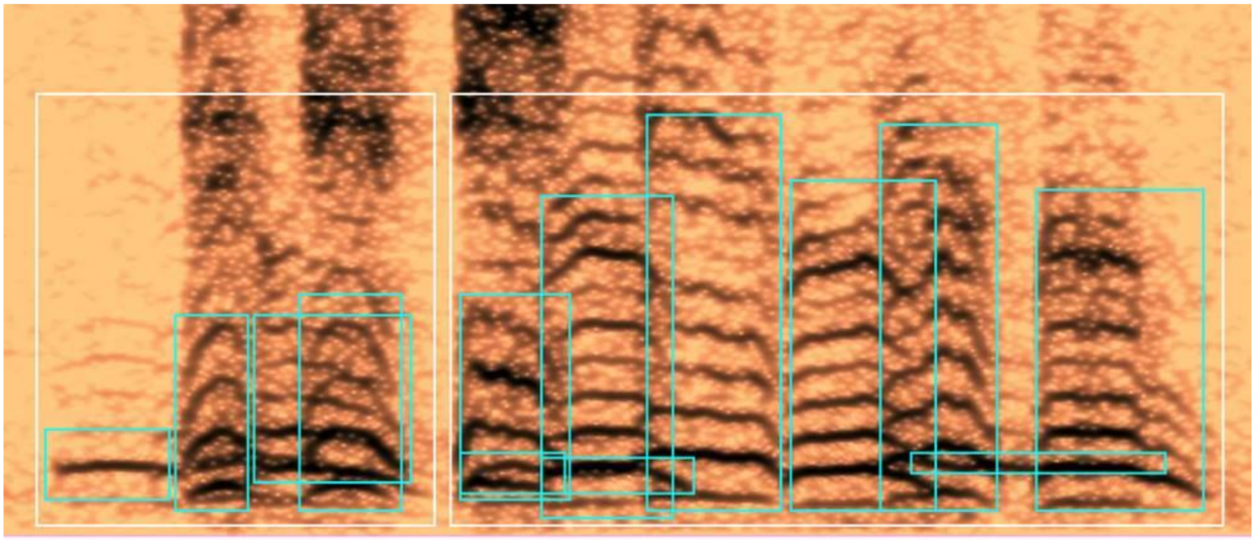



tight bounds for measurement

The basic challenge would be to 'draw' the box, given information on where a signal of interest is located. We could then put this information into a program (Raven) that would do the measuring. Of course, the measurements could be extracted at the time of making the box bounds.

**Determining more than one animal's calls:**

This is a pretty complex one. The example below is on the complex side (a very intense greeting between two sisters who had not seen one another for months), but shows the problem. Our detector would likely only make two selections here (white boxes), but the blue boxes are the actual number of different calls.



white boxes might be what our detector boxes - but blue boxes are the actual number of differenct rumbles in the spectrogram

A large proportion of these cases might only have one or two overlapping rumbles, so a simplistic description of how we make the decision:

(1) the harmonics MUST be multiples of the fundamental frequency - so they are all exactly the same distance apart (on a linear scale)

(2) an elephant can't produce two fundamental frequencies at once (as birds can), so if a 'contour' crosses another contour they are different animals.

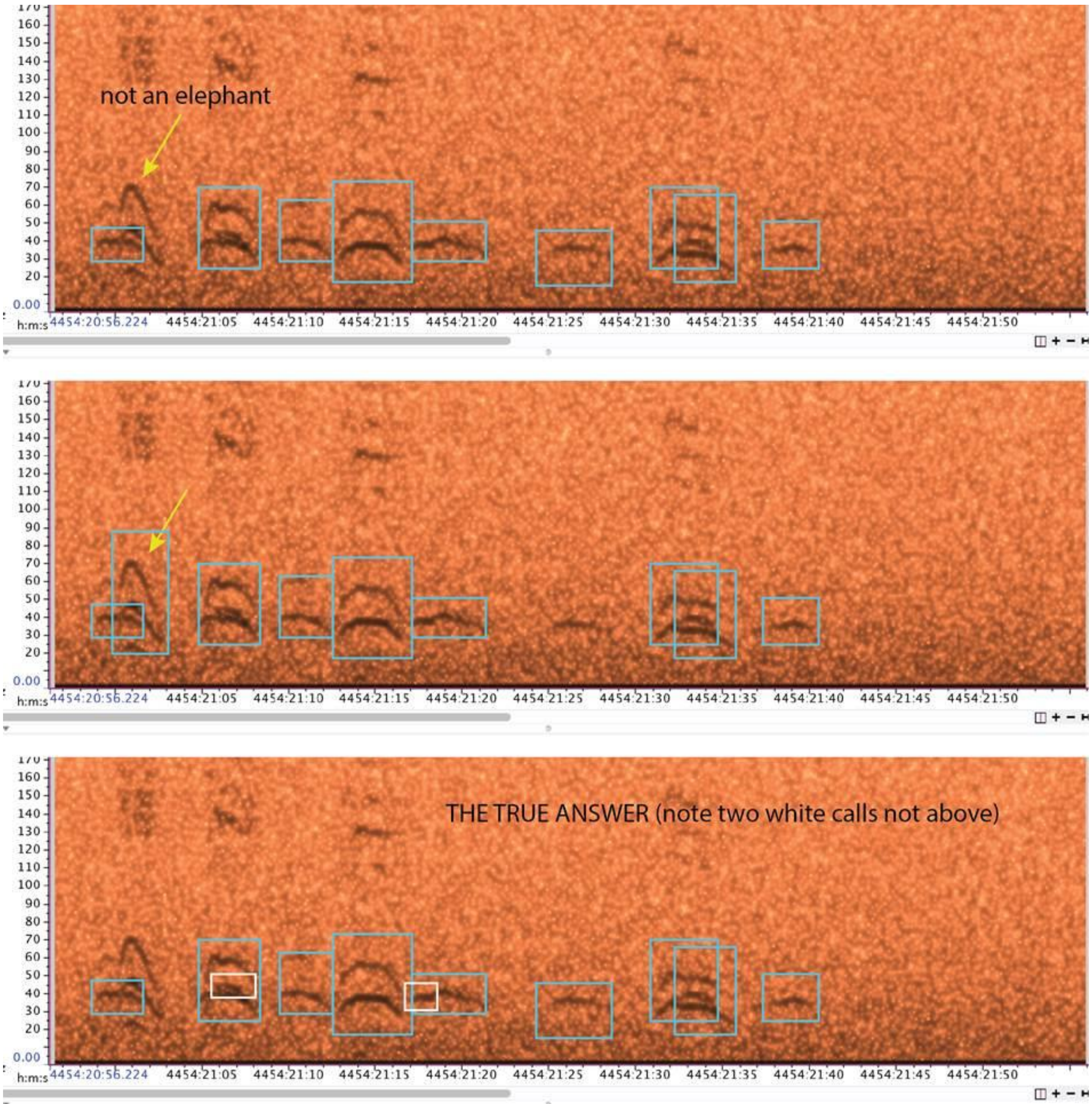
**Out of Grid Detection:**

"I don't think we have data prepared that would allow discerning whether elephants are leaving the recording grid. We are sort of getting at this by refining our analyses to estimate total numbers of elephants in the study area (involves knowing detection distance for each acoustic unit, adjustments for background noise levels that affect the detection distance, calling rate of elephants, and performance metrics of the detector used). Once we are happy with this we can, in theory, estimate numbers over shortish blocks of time and see whether we consistently get lower numbers within the grid at some times of year, but coming back up to 'steady state' sometime later."

*Sarah Maston Note: This is a good use case to imagine an expanded IoT solution to work concurrently with the sound analysis.*

**Verify calls in spectrograms:**
The other challenge  has to do with how we could gather in lots of elephant-interested citizen scientists from around the world to help verify calls in spectrograms. Even if we have a really great detector, we will still always need to do verification by hand for subsets of data. We had the beginnings of such a web-based system that could serve up clips of sound within which users could box calls. ( if you are interested, this link gets you to the starting page of this application - you do have to create an account (nothing is actually done with it), and then click on 'participate' and then, in the second paragraph, 'tutorial' to see where we were going. )

The Idea would be to give maybe 5 different participants the same clip - but then we need a way to figure out what the real number of calls in that clip might be, by getting the 'consensus' of multiple users. If some defined level of agreement is not reached, the program gives that clip to another set of users and tries again.
So presumably the problem is to compare the overlap of selection boxes (which might be sloppy?), when sometimes different signals, including incorrect ones, might be selected.
Here is a made up example:



not an elephant
0.00
h:m:s 4454:20:56.224
4454:21:05
4454:21:10
4454:21:15
4454:21:20
4454:21:25
4454:21:30
4454:21:35
4454:21:40
4454:21:45
4454:21:50
THE TRUE ANSWER (note two white calls not above)