

# Identificando Colines (*Eleutherodactylus eileenae*) a partir de audios desordenados

## Sincronización y Procesamiento de Audios

Daniel Machado Pérez

Tutores: Dr. Roberto Mulet, Dr. Milton Borroto

Facultad de Matemática y Computación, Universidad de La Habana

daniel.machado.0206@gmail.com

### Resumen

Este estudio analiza las secuencias de cantos de la rana *Eleutherodactylus eileenae*, obtenidos mediante grabaciones de campo realizadas con nueve micrófonos distribuidos alrededor de los especímenes. Para identificar las secuencias probables de los cantos se propone un método que emplea mel-espectrogramas y análisis de las energías temporales de las grabaciones. Para garantizar la sincronización de los audios, se procedió con la obtención de los momentos de pico de energía y el cálculo de los desfases con correlación cruzada utilizando un archivo de referencia. Cada canto es conocido comúnmente como Colín (por comodidad, así llamaremos también a la especie), en imitación al sonido producido, en el que se emite una señal con frecuencia baja que denominaremos CO y una con frecuencia alta que llamaremos LIN. Los resultados revelan patrones en la estructura del coro, donde cada rana parece mantener una periodicidad consistente entre cantos y ajustar las frecuencias en respuesta a otros individuos.

### Introducción

La rana *Eleutherodactylus eileenae* (Dunn, 1926) es una especie endémica de Cuba [1]. Su canto, que se caracteriza por emisiones de frecuencias bajas (CO) y altas (LIN), representa un objeto de estudio bioacústico único. Este trabajo propone un enfoque automatizado para identificar y analizar las secuencias de cantos de Colines a partir de grabaciones realizadas con nueve micrófonos distribuidos alrededor de los especímenes. Utilizando mel-espectrogramas, análisis de energías temporales y técnicas de correlación cruzada, se logra la sincronización de los audios y la separación de los cantos individuales. Este método no solo permite un análisis eficiente de los coros, sino que también abre nuevas posibilidades para estudiar su dinámica y estructura, abordando preguntas fundamentales sobre el comportamiento acústico y social de la especie.

### Objetivos

1. Construir un dataset “limpio” donde sea más fácil obtener estadísticas e interpretar los resultados.
2. Sincronizar los archivos de audio.
3. Distinguir los especímenes más cercanos a cada micrófono.
4. Identificar las secuencias probables de los cantos de Colines.
5. Analizar el comportamiento de las frecuencias en los coros.
6. Validar la consistencia del método propuesto.

### Materiales y Métodos

Se cuenta con el dataset de las grabaciones hechas por los 9 micrófonos, cuya ubicación geográfica se representa en la Figura 1.

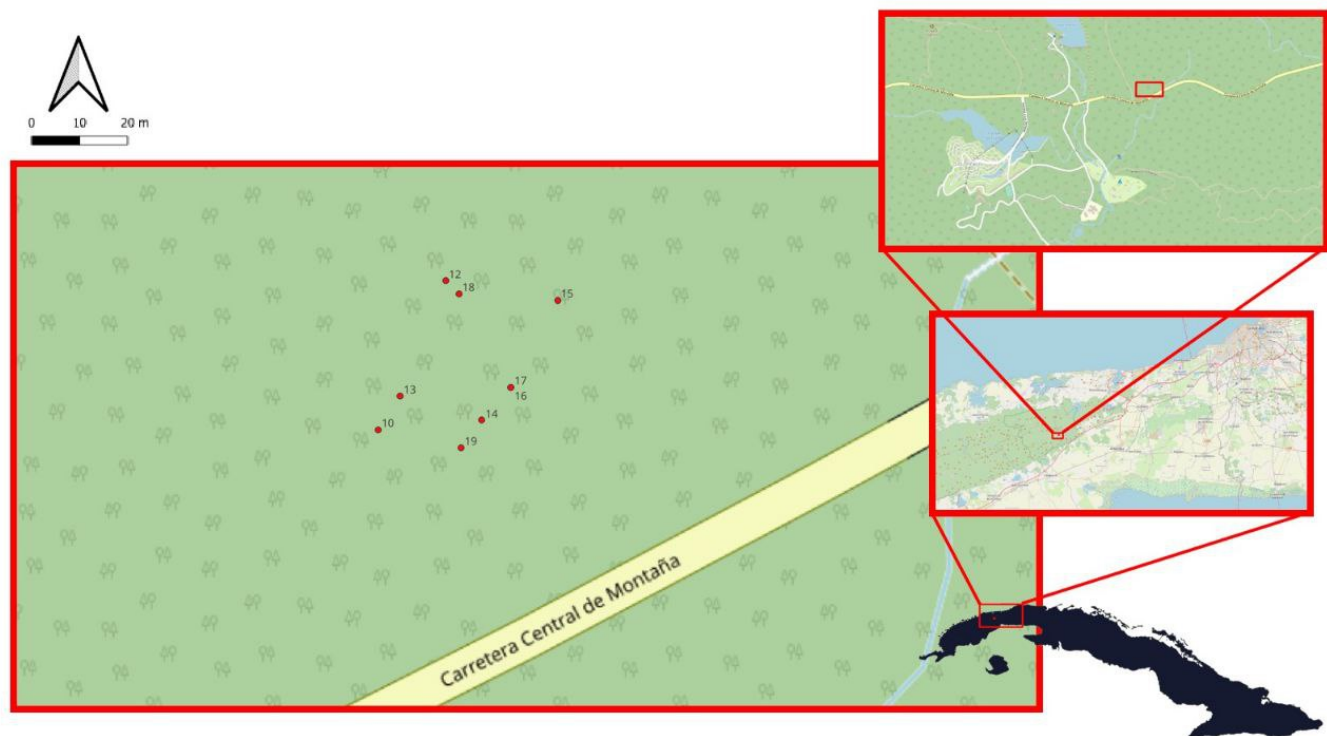


Figure 1: Distribución Geográfica de los Micrófonos.

Estas se realizaron durante 3 días, a partir de las 18:00 hasta las 06:00 horas del día siguiente, donde de cada hora se registraron 58 minutos y se descansaron 2. Los dispositivos se activaron remotamente y al mismo tiempo. Para lo que se presenta, se utilizó la información captada el día 21 de octubre de 2023 a las 19:00 horas. Cada audio se representó y procesó a través de su mel-espectrograma. [4] Un ejemplo de ello se muestra en la Figura 2.

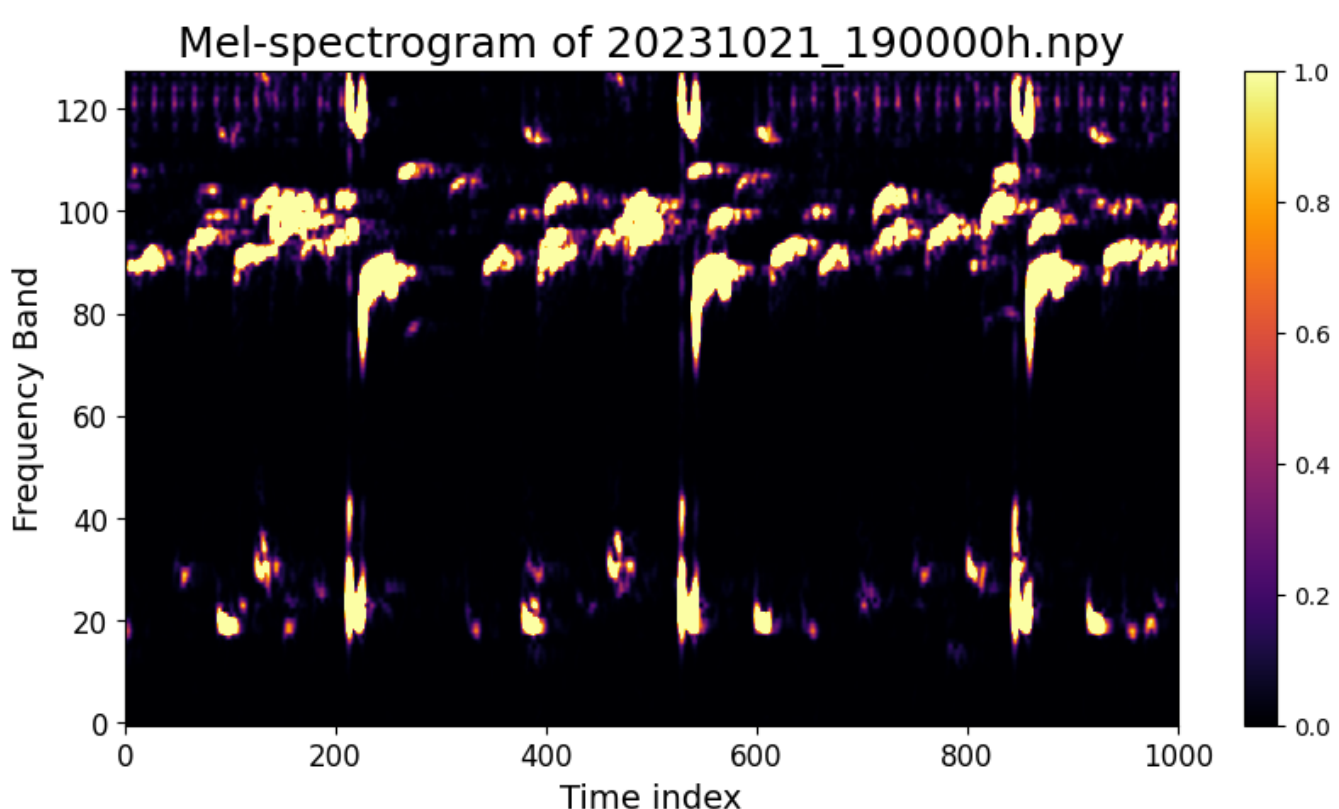


Figure 2: Mel-espectrograma del archivo *h*

El proceso de sincronización comienza con la identificación de un archivo pivote, que debe contener una señal que hayan captado todos los micrófonos y en un rango de bandas de frecuencias lo más ‘limpio’ posible. Luego se procede con el cálculo de las energías temporales con la siguiente fórmula, donde  $E_j$  es el valor de la energía en el instante de tiempo  $j$ ,  $a_{ij}$  es la intensidad en el instante de tiempo  $j$  y la banda de frecuencia  $i$  en el mel-espectrograma correspondiente.:

$$E_j = \sum_{i=0} a_{ij}^2, \forall j[3]$$

Luego se procede con la eliminación de ruido y la identificación de picos de energía. Sobre esa data se aplica **Correlación Cruzada**, entre cada archivo y el archivo pivote. El valor de *lag* o desfase es el que maximiza la Correlación:

$$CC = \sum_{i=0}^n a_i b_i, [2]$$

Donde  $a_i$  y  $b_i$  son los valores en el índice  $i$  de los arrays sobre los que se aplica la Correlación. Finalmente el resultado es como el que se muestra en la Figura 3.

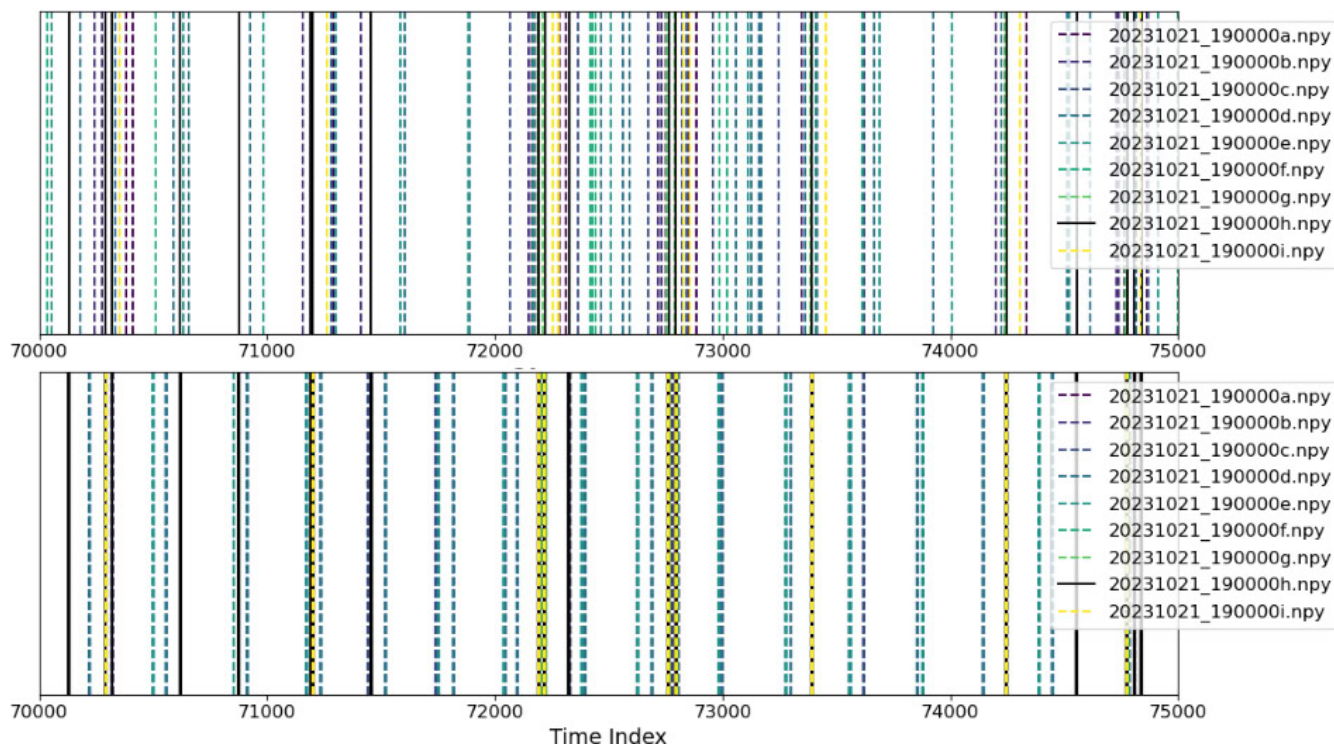


Figure 3: Resultado de la Sincronización.

Luego se procede con la distinción del Colín más cercano a cada micrófono. Para ello se propone un método que se sustenta en las siguientes hipótesis:

- Para cada micrófono, los cantos del individuo más cercano deben ser registrados con las mayores energías relativas a dicho micrófono.
- Si en cada uno se identifican las energías grandes y se elimina la información de esos cantos de las demás grabaciones, se obtendrán por separado los datos que se buscan.

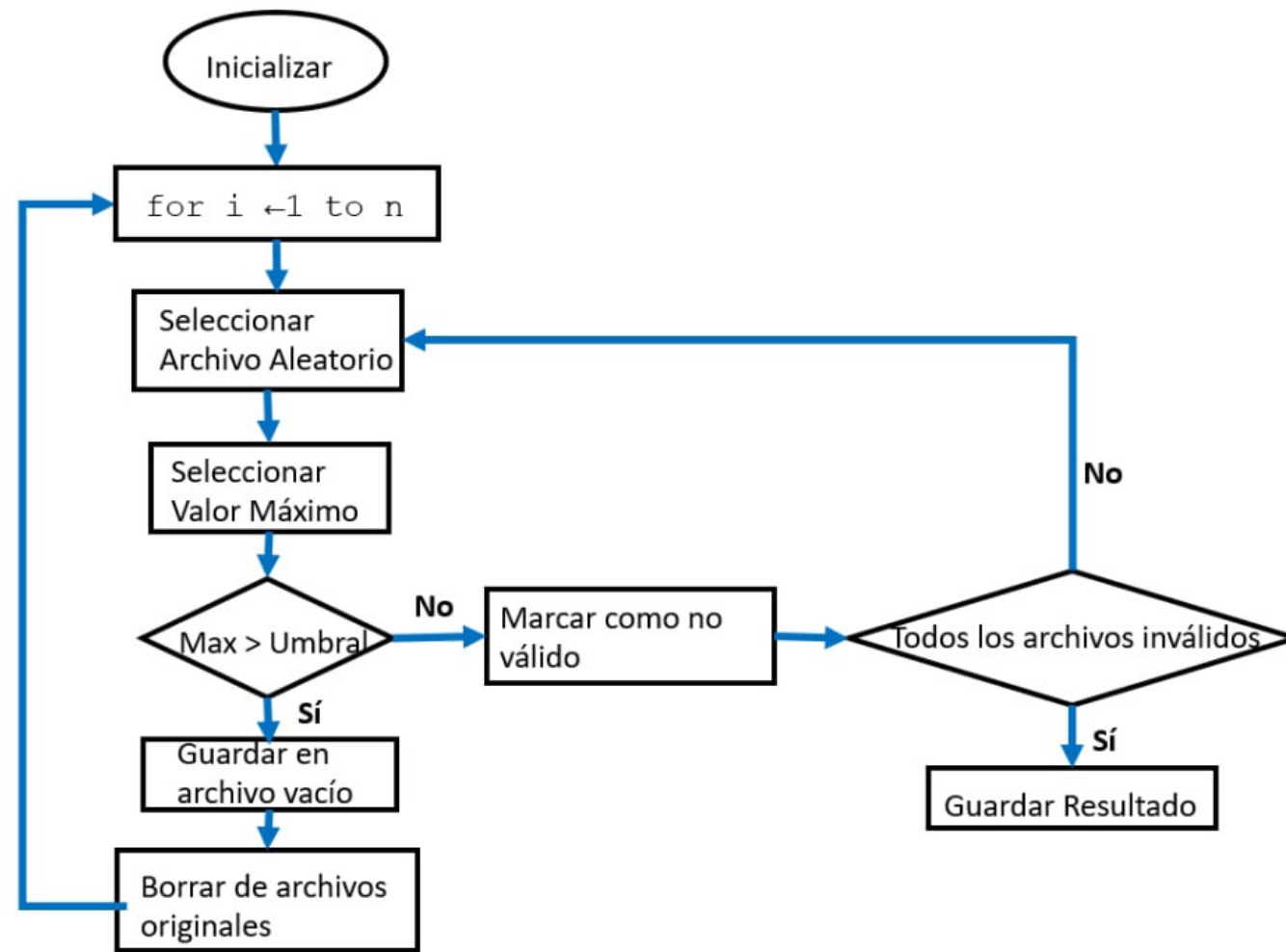


Figure 4: Esquema del método para distinción de cantos cercanos a micrófonos.

### Resultados

Luego de tener la información de los cantos del Colín cercano a cada micrófono, podemos construir los coros, como se muestra en la Figura 5.

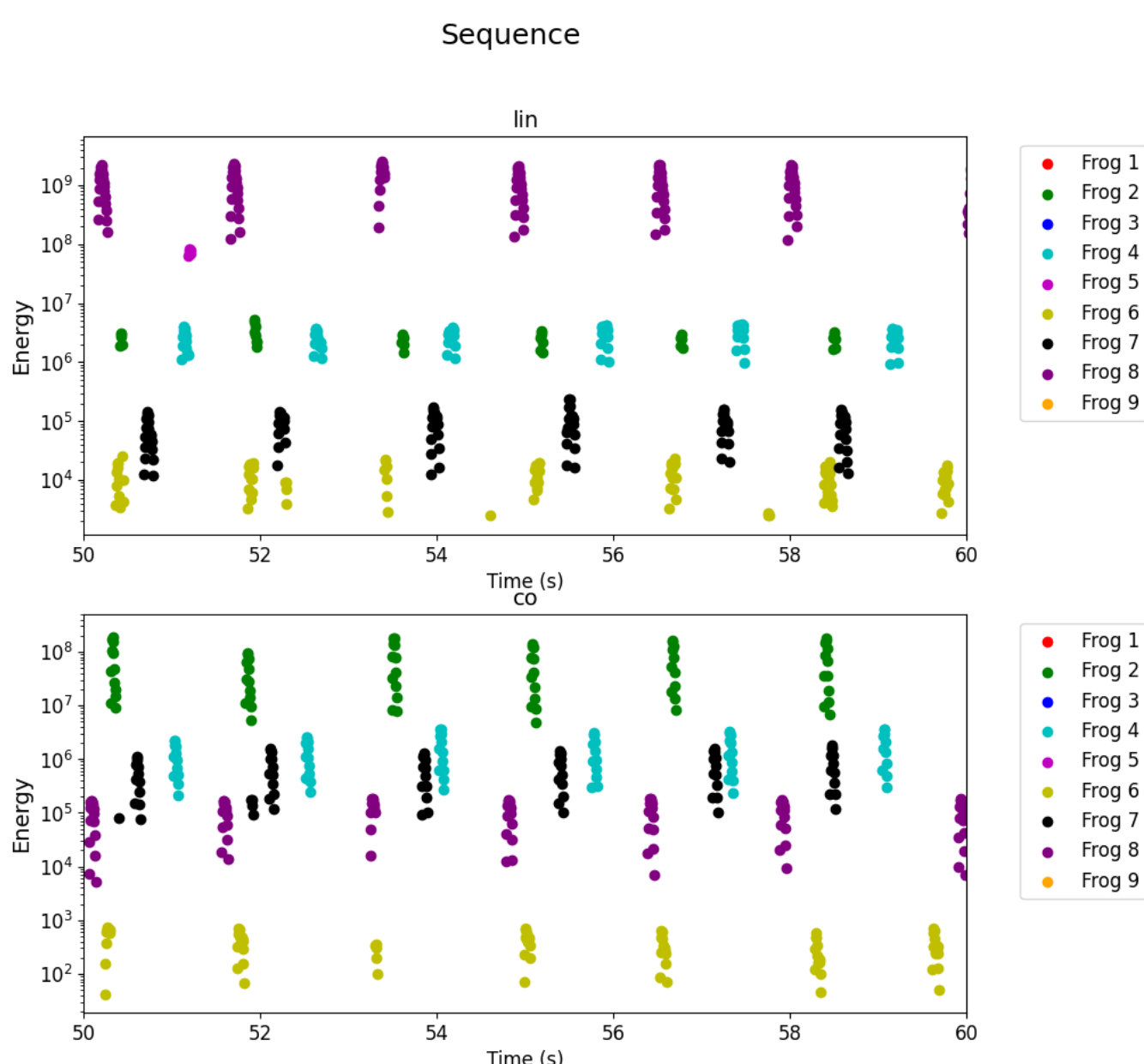


Figure 5: Secuencia de cantos.

Además podemos conocer el comportamiento en cuanto a la frecuencia que manifiesta cada rana.

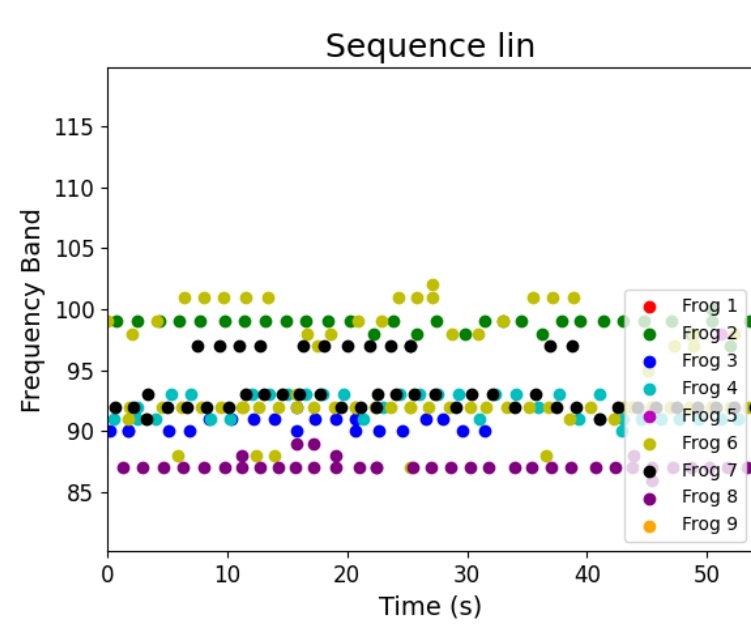


Figure 6: Frecuencia de los LIN.

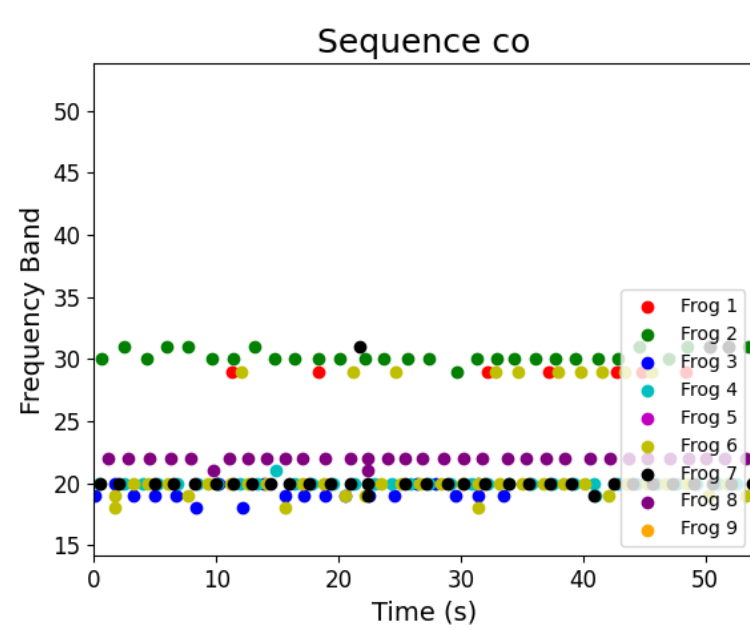


Figure 7: Frecuencia de los CO.

De lo observado en las Figuras 6 y 7 se puede formular la hipótesis de que las ranas regulan la frecuencia de sus cantos en dependencia de las frecuencias de los demás miembros del coro, buscando baja superposición y tal vez mayor distinguibilidad.

Para validar la consistencia del algoritmo, se ejecutaron 10 corridas independientes y se hizo un análisis de correlación entre cada par de arrays resultantes, también utilizando Correlación Cruzada. El resultado puede observarse en la Figura 8, donde en cada fila de la matriz se resaltan en rojo los 10 mayores valores de correlación, y los arrays están organizados primero por audio y luego por corrida. Es decir, las 10 primeras filas corresponden al array resultante del audio 1 en cada una de las 10 corridas, las siguientes 10 filas al array del audio 2 en las 10 corridas y así consecutivamente. De la misma forma están organizados en las columnas. Como se observa, la mayor correlación se concentra entre los arrays del mismo audio en las 10 corridas, probando que el algoritmo es consistente.

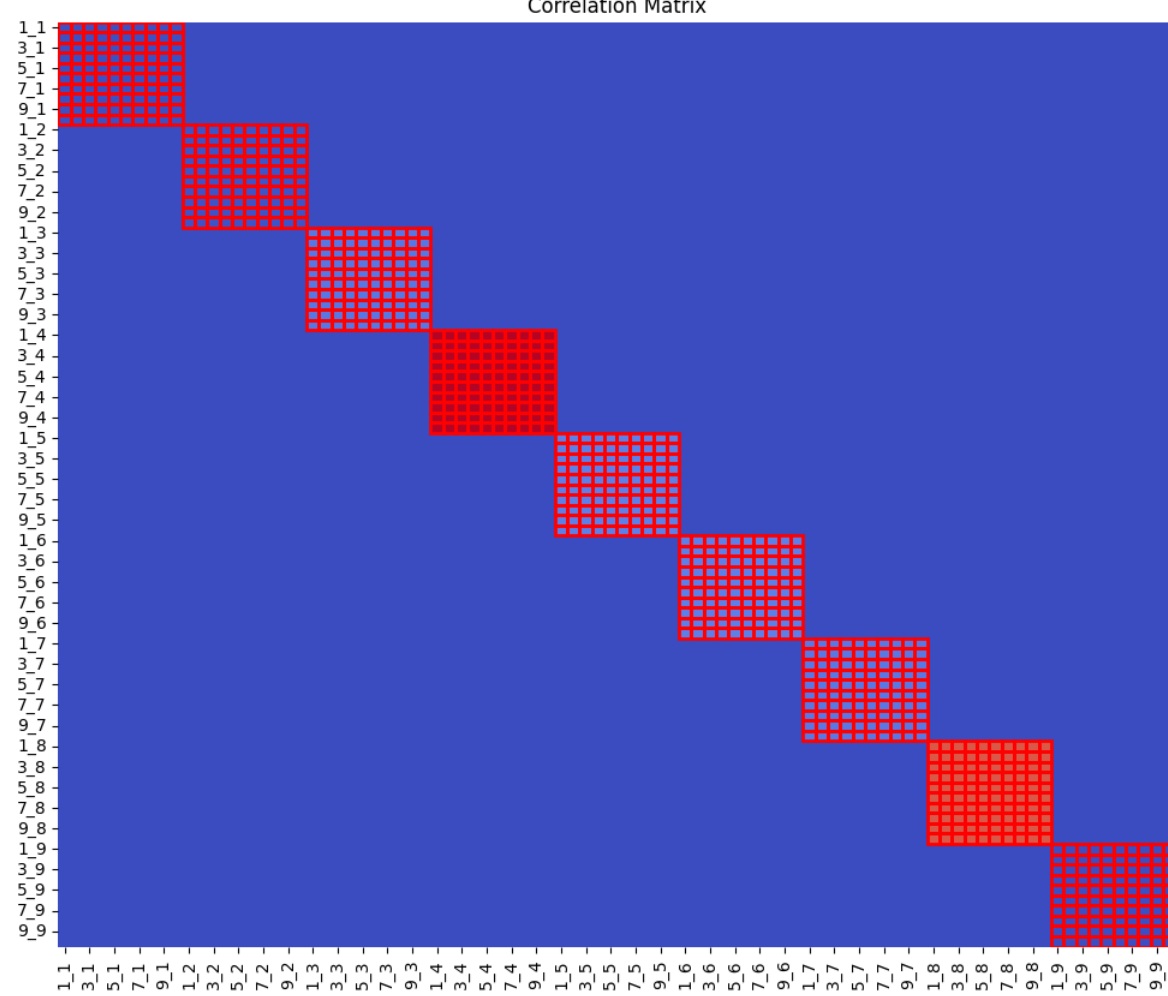


Figure 8: Matriz de Correlación entre corridas distintas del Algoritmo

### Conclusiones

- Se diseñó un método semi-automático de sincronización usando Correlación Cruzada, con resultados positivos.
- Se diseñó un método automático para distinguir especímenes cercanos a micrófonos, del cual se logró validar su consistencia y efectividad.
- Se construyó un dataset “limpio” que permitió obtener la secuencia de cantos y el comportamiento en frecuencia.

### Líneas Futuras de Investigación

- Análisis de la hipótesis de que los Colines modifican su registro de frecuencia para diferenciarse de los demás cantos del coro.
- Análisis de causalidad, para estudiar la posibilidad de que en un coro existan individuos que manifiesten un papel protagonista en su estructura.

### Referencias

- [1] Rodríguez-Gómez A. Alonso, R. and A. R. Estrada. Patrones de actividad acústica y trófica de machos cantores de *eleutherodactylus eileenae* (anura: Leptodactylidae). *Revista española de herpetología*, 15:45–52, 2001.
- [2] P. Bourke. *Cross Correlation. Cross Correlation*, *Auto Correlation—2D Pattern Identification*. Self-published, 1996.
- [3] D. Jurafsky and J. H. Martin. *Speech and Language Processing*. Pearson, 2nd edition, 2008.
- [4] Centic Murcia. Mel spectrogram. *Curso de Ciencia de Datos*, n.d.

### Agradecimientos

Al Dr. Roberto Alonso, por sus consejos sobre el comportamiento de la especie y por proveernos de los datos.