

# Detección automatizada de cantos de aves continúa siendo un desafío: el caso de warbleR y *Megascops centralis* (búho del Chocó)

## Automated detection of bird songs continues to be a challenge: the case of warbleR and *Megascops centralis* (Chocó owl)

Laura Andrea Hoyos Cardona  , Juan Sebastián Ulloa  , Juan Luis Parra Vergara  

---

### Resumen

El monitoreo acústico permite evaluar cambios espacio-temporales en poblaciones animales. Sin embargo, analizar grandes volúmenes de información es desafiante. Se evaluó el desempeño de una técnica de detección (función autodetec del paquete warbleR de R) para identificar vocalizaciones de *Megascops centralis*, utilizando 6877 grabaciones de un minuto provenientes de grabadoras ubicadas en 21 sitios alrededor del embalse Jaguas, Andes de Antioquia, Colombia. Las vocalizaciones se anotaron manualmente y se seleccionaron dos sitios (597 grabaciones) con el mayor número de registros (49 y 34) para la evaluación del algoritmo. La función fue utilizada con audios a dos tasas de muestreo (44 100 Hz y 22 050 Hz) y tres umbrales de amplitud (5, 10 y 20). Se evaluó el desempeño de la función en términos de su sensibilidad y especificidad, y se estimó la probabilidad de detección de una vocalización según su calidad. La sensibilidad y especificidad presentaron gran variación (0-0.48 y 0.5-0.98 respectivamente). La probabilidad de detección de una señal aumentó con su calidad (mala: 0.12, media: 0.27 y buena: 0.64). El monitoreo acústico tiene gran potencial, y parte de su éxito depende de herramientas de reconocimiento automático, de acceso abierto y fácil implementación. Este desarrollo puede acelerarse fortaleciendo nuestras colecciones sonoras.

**Palabras clave.** Autodetec. Bioacústica. Detección automática. Monitoreo. WarbleR.

### Abstract

Acoustic monitoring allows the evaluation of spatio-temporal changes in animal populations. However, analyzing large volumes of information is challenging. We evaluate the performance of a detection technique (autodetec function of the R warbleR package) to identify vocalizations of *Megascops centralis*, using 6877 one-minute recordings from 21 sites in the vicinity of the Jaguas dam, Andes of Antioquia Colombia., All vocalizations were manually annotated and two sites (597 recordings) with the highest number of records (49 and 34) were selected to evaluate the algorithm. The function was implemented with audios at two sampling rates (44 100 Hz and 22 050 Hz) and three amplitude thresholds (5, 10, and 20). We assessed the performance of this function in terms of its sensitivity and specificity, and we estimate the probability of detection of a signal according to its quality. Sensitivity and specificity showed great variation (0-0.48 and 0.5-0.98 respectively) and the probability of detection of a signal increased with its quality (poor: 0.12, medium: 0.27 and high: 0.64). Acoustic monitoring has an enormous potential, and its success depends, in part, on the availability of automatic recognition tools, that are open access and can be easily implemented. This development can be achieved by strengthening acoustic collections.

**Keywords.** Autodetec. Bioacoustics. Automatic detection. Monitoring. WarbleR.

## Introducción

Hoy en día es clara la importancia de las herramientas bioacústicas para identificar las especies presentes en un lugar a partir de sus cantos. Sin embargo, por mucho tiempo se enfatizaron otras técnicas de muestreo, como el avistamiento directo y el uso de redes de niebla (Ralph *et al.*, 1993). Los equipos portátiles que facilitan la grabación en campo de los sonidos de las aves se desarrollaron durante gran parte del siglo XX, pero fue solo en 1956 cuando se estableció la biblioteca de sonidos de Cornell Macaulay (originalmente Library of Natural Sounds), considerada la primera colección bioacústica del mundo (Ranft, 2004). Uno de los personajes clave en este cambio de mentalidad fue el ornitólogo Ted Parker III, quien demostró que era posible registrar en siete días, por medio de herramientas acústicas, el 85 % de las especies de aves que originalmente fueron identificadas luego de 36 000 horas de muestreo con redes de niebla en una localidad en la Amazonia boliviana (Parker, 1991). La propuesta de Parker, junto con el aumento en volumen y disponibilidad de cantos, continúan generando un cambio en la manera como monitoreamos las poblaciones de aves.

Actualmente, las técnicas para el registro y análisis de señales acústicas se han popularizado y permiten optimizar la colecta de información de comunidades de aves, logrando registrar especies raras, evasivas y difíciles de detectar (Goyette *et al.*, 2011). La utilidad de estas técnicas se hace aún más evidente bajo condiciones en las que nuestra vista pierde agudeza, como sucede en la noche. Las especies nocturnas suelen ser crípticas, evasivas y activas en tiempos en los que usualmente es desafiante estar en campo (Goyette *et al.*, 2011), dificultando el registro, seguimiento y comprensión de estas. Por ello, se han desarrollado métodos que implican el reconocimiento de la actividad acústica para el monitoreo de estas especies por medio de sensores acústicos programables (König *et al.*, 2008).

Los sensores acústicos son dispositivos que permiten registrar automáticamente el ambiente sonoro de un lugar. El uso de estos dispositivos permite maximizar la cantidad de información acústica obtenida, mientras se minimizan los esfuerzos del personal en campo. Diversos estudios han demostrado que los sensores acústicos permiten monitorear de manera detallada patrones de actividad, y demuestran la eficiencia de estas técnicas en condiciones de campo, evitando además, la intervención de los investigadores al emplear herramientas como playback y puntos de conteo (Byrnes, 2013; Ferraz

*et al.*, 2010; Sberze *et al.*, 2010). De esta forma, se posibilita el monitoreo de las especies de una manera fidedigna, pues se espera registrar el comportamiento sin que se incurra en la perturbación de este (Deichmann *et al.*, 2018).

Aunque los sensores acústicos han facilitado la recolección de datos, el análisis del gran volumen de información generado continúa siendo un desafío para la implementación de esta técnica (Blumstein *et al.*, 2011). Idealmente, esta información debe ser procesada manualmente con el fin de identificar posibles falencias (por ejemplo, mal funcionamiento de la grabadora) y generar una base de datos anotada (quién o qué genera sonido en algún momento particular) que pueda ser utilizada en múltiples contextos. Es posible que, en la actualidad, este sea el mayor cuello de botella en el campo del monitoreo bioacústico. Históricamente, la depuración ha sido realizada por personas que tienen un gran interés o la responsabilidad de analizar sus datos; no obstante, el conocimiento de cualquier persona es limitado y cada vez son menos los interesados y capacitados para cumplir con esta tarea. De manera simultánea y con el fin de alivianar la carga para las personas, se han perpetrado diversos esfuerzos alrededor del mundo por automatizar los procesos de detección y clasificación de señales acústicas (Aide *et al.*, 2013; Ovaskainen *et al.*, 2018; Sethi *et al.*, 2020). Entre estos esfuerzos podemos encontrar varios programas informáticos como ARBIMON-II (Sieve Analytics, 2015), AviaNZ (Marsland *et al.*, 2019), monitoR (Katz *et al.*, 2016) y warbleR (Araya-Salas & Smith-Vidaurre, 2018).

Estos programas presentan ventajas y desventajas asociadas, por ejemplo, a sus costos económicos (p. ej. procesar un minuto de audio en ARBIMON-II tiene un precio de 0.06 USD), así como restricciones geográficas (p. ej., AviaNZ funciona para aves en Nueva Zelanda). Es por ello por lo que se hace necesario identificar algoritmos o programas que tengan un desempeño aceptable bajo las condiciones particulares de cada investigación y que los esfuerzos sean dirigidos hacia el libre acceso. La evaluación de estos algoritmos cobra importancia y es fundamental para el continuo mejoramiento de las técnicas. Sobre todo para especies crípticas y de hábitos nocturnos, para las cuales la opción más eficaz de monitoreo es a través de sus vocalizaciones.

Los búhos hacen parte de este grupo de especies y varios trabajos han demostrado la utilidad de monitoreos acústicos sobre ellos (Byrnes, 2013; Goyette *et al.*, 2011). De manera general, la mayoría de las vocalizaciones de

los búhos se caracterizan por ser sencillas, de corta duración, y de baja frecuencia (< 3000Hz; Claudino *et al.*, 2018; Dantas *et al.*, 2016; Goyette *et al.*, 2011; Krabbe, 2017; Nagy & Rockwell, 2012; Sberze *et al.*, 2010). La mayoría de búhos son nocturnos o crepusculares y, por lo tanto, sus ambientes acústicos tienen características diferentes a las encontradas durante el día (Almeira & Guecha, 2019). Debido a la naturaleza de estos cantos y a sus ambientes acústicos, esperamos que este tipo de vocalizaciones sea relativamente fácil de identificar de manera automatizada, excepto en condiciones de baja relación señal-ruido, por ejemplo, cuando existe mucho ruido a frecuencias bajas, o cuando el emisor se encuentra distante de la grabadora. Si bien existen esfuerzos en el neotrópico que buscan entender patrones ecológicos de poblaciones de aves nocturnas por medio de la bioacústica (p. ej., Baldo & Mennill, 2011; Goyette *et al.*, 2011; Sberze *et al.*, 2010), y algunos esfuerzos para evaluar la eficiencia de diferentes algoritmos en la detección de señales acústicas (Heinicke *et al.*, 2015; Kalan *et al.*, 2015; Keen *et al.*, 2017; Ulloa *et al.*, 2016), no se conocen estudios que evalúen metodologías para sistematizar, por medio de algoritmos, la revisión de grabaciones obtenidas durante la noche en el Neotrópico. Bajo el contexto anterior, planteamos como principal objetivo de este trabajo, evaluar el desempeño de la función “autodetect” del paquete warbleR en la detección de cantos de un búho nocturno (*Megascops centralis*) en grabaciones realizadas en la zona de amortiguamiento del embalse de Jaguas, en el oriente del departamento de Antioquia, en el noroeste de Colombia. Este es un algoritmo de detección que exige pocos parámetros y por lo tanto es de fácil implementación, requiriendo poco conocimiento del lenguaje de R para su ejecución. Es, además, un algoritmo de libre acceso. Nuestra hipótesis fue que, dada la sencillez de los cantos de esta especie, el algoritmo tendría un buen desempeño en general, excepto cuando la calidad de las señales fuese mala (debido a ruido en el ambiente o a distancia del emisor). Para esto, evaluamos el desempeño del algoritmo de detección variando la tasa de muestreo, el umbral de amplitud y la calidad (relación señal-ruido) de las vocalizaciones.

## Materiales y métodos

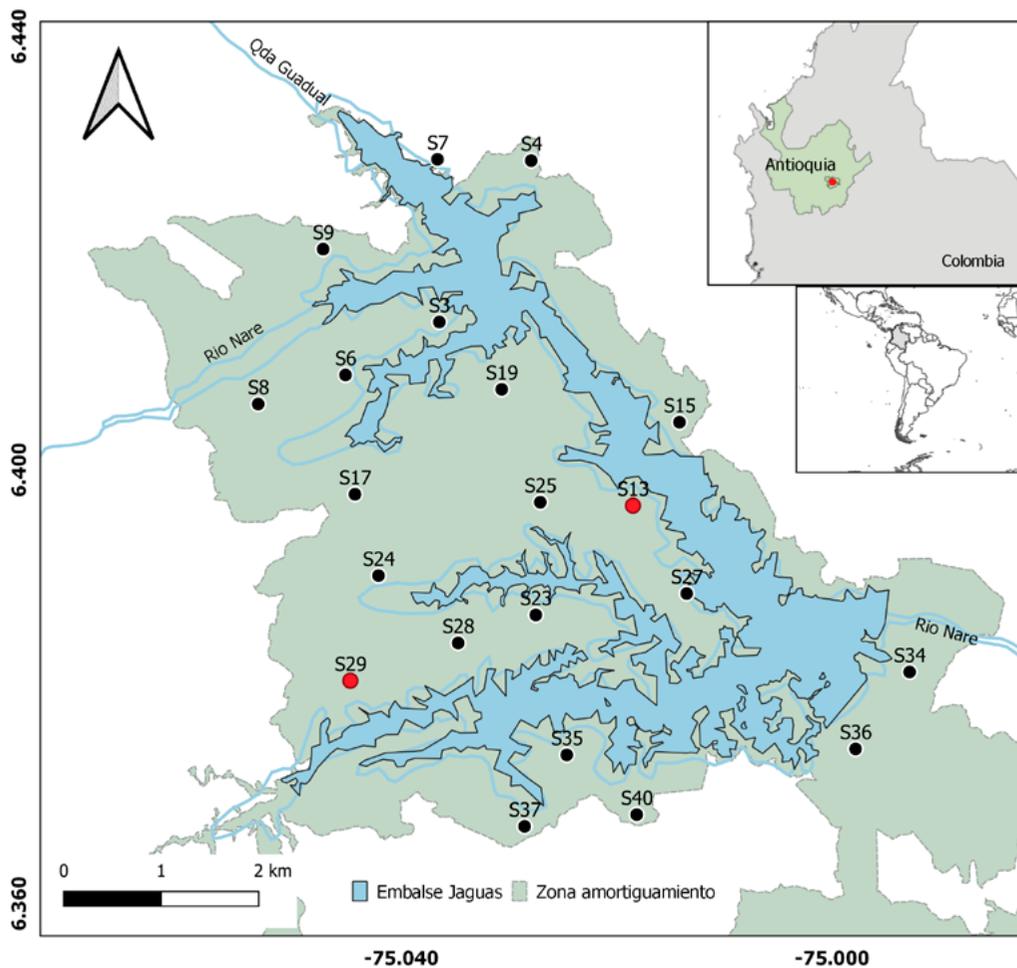
**Área de estudio.** Este estudio se llevó a cabo en el área de amortiguamiento de la hidroeléctrica Jaguas, propiedad de ISAGEN, localizada en la vertiente oriental de la cordillera Central de los Andes de Colombia, en jurisdicción de los municipios de San Rafael, Alejandría,

Santo Domingo y San Roque, en el departamento de Antioquia (Figura 1). La zona de amortiguamiento del embalse cuenta con ~2 600 hectáreas de ecosistemas forestales entre 1250 y 1330 m s.n.m. La formación vegetal en la zona corresponde a bosque muy húmedo premontano (bmh-PM), con una temperatura entre 18 y 24 °C y un promedio anual de lluvia de 2 000 a 4 000 mm. Dentro del área de estudio, fueron ubicadas 21 grabadoras automatizadas, de las cuales seleccionamos dos para su posterior procesamiento.

**Grabadoras.** Se emplearon grabadoras Song Meter 4 (SM4-Wildlife Acoustics) localizadas en 21 sitios alrededor del área de mitigación de la represa Jaguas (una grabadora por sitio). Estas grabadoras fueron configuradas a una tasa de muestreo de 44 100 Hz y una profundidad de 16 bits, para grabar 1 minuto cada 15 minutos en estéreo entre las 19:00 y las 5:45, durante el periodo entre el 10 de marzo y el 17 de junio de 2018. Para la evaluación, se seleccionaron dos grabadoras (G13 y G29), teniendo en cuenta la cantidad de los registros (Tabla 1).

**Pre-procesamiento.** Por medio del programa RavenPro 1.5 (Bioacoustics Research Program, 2011), se revisaron 6877 archivos de audio provenientes de las 21 grabadoras (Tabla 1). La revisión fue realizada por una sola persona (LAHC) quien etiquetaba manualmente las señales que se encontraran entre los 0 Hz y los 3000 Hz, identificando las que correspondieran a la especie de interés, *M. centralis*. Las grabadoras ubicadas en los sitios G13, G29 y G25 fueron las que más grabaciones de la especie registraron, con 49, 34 y 21 archivos, respectivamente (Tabla 1). Para evaluar el clasificador/detector, se seleccionaron dos grabadoras: la que registró un mayor número de señales de *M. centralis* (G13) y la que tuvo el mayor número de señales de buena calidad (G29; ver adelante).

Con el fin de caracterizar cualitativamente la calidad de la señal en términos de la relación señal-ruido, clasificamos los audios con detección manual positiva en tres categorías, con base en cómo se percibía la señal al momento de la revisión. Las categorías utilizadas fueron: (1) buena, señales claramente visibles en el espectrograma y perfectamente diferenciables en el espectro de potencia; (2) media, señales que podían verse en el espectrograma y cuyo pico en el espectro de potencia era apreciable pero no perfectamente diferenciable; y (3) mala, señales visibles en el espectrograma, pero casi inexistentes en el espectro de potencia (Figura 2).



**Figura 1.** Disposición de las 21 grabadoras usadas para identificar las vocalizaciones de *Megascops centralis* alrededor del embalse Jaguas, Andes de Antioquia, Colombia. En rojo, las grabadoras seleccionadas para la evaluación de la función autotdetect.

**Figure 1.** Arrangement of the 21 recorders used to identify vocalizations of *Megascops centralis*, around the Jaguas dam, Andes of Antioquia Colombia. Recorders Selected to evaluate the autotdetect function are shown in red.

**Selección de parámetros para el algoritmo.** Teniendo en cuenta que el canto de los búhos está determinado genéticamente (König *et al.*, 2008) y que no se observó gran variación en las características espectrales de los cantos identificados, se eligieron cinco grabaciones y de ellas se eligieron cinco vocalizaciones con calidades entre medias y buenas. Se procedió a cuantificar la frecuencia mínima, frecuencia máxima, duración mínima y duración máxima del canto con el programa Raven Pro 1.5 (Bioacoustics Program, 2011). Para estas variables, se obtuvieron promedios, valores mínimos y máximos. Estos valores fueron utilizados como punto de partida para definir los parámetros con los que se trabajaría el algoritmo. Se realizaron ensayos de prueba

y error con el algoritmo, en los que se incluían cinco archivos de audio con señales de la especie y cinco archivos sin ellas, buscando ajustar los parámetros para una correcta identificación de la especie en el conjunto de prueba. Una vez se detectaron la mayoría de las señales, se estableció el valor final de cada uno de los parámetros para la realización de las pruebas de desempeño del algoritmo. El algoritmo evaluado cumple con las características de detector sencillo y corresponde a la función autotdetect del paquete WarbleR (Araya-Salas & Smith-Vidaurre, 2018), del software de libre acceso R, que busca señales de interés a partir de características predefinidas por los usuarios, relacionados con la duración, la banda de frecuencia de los cantos y la amplitud.

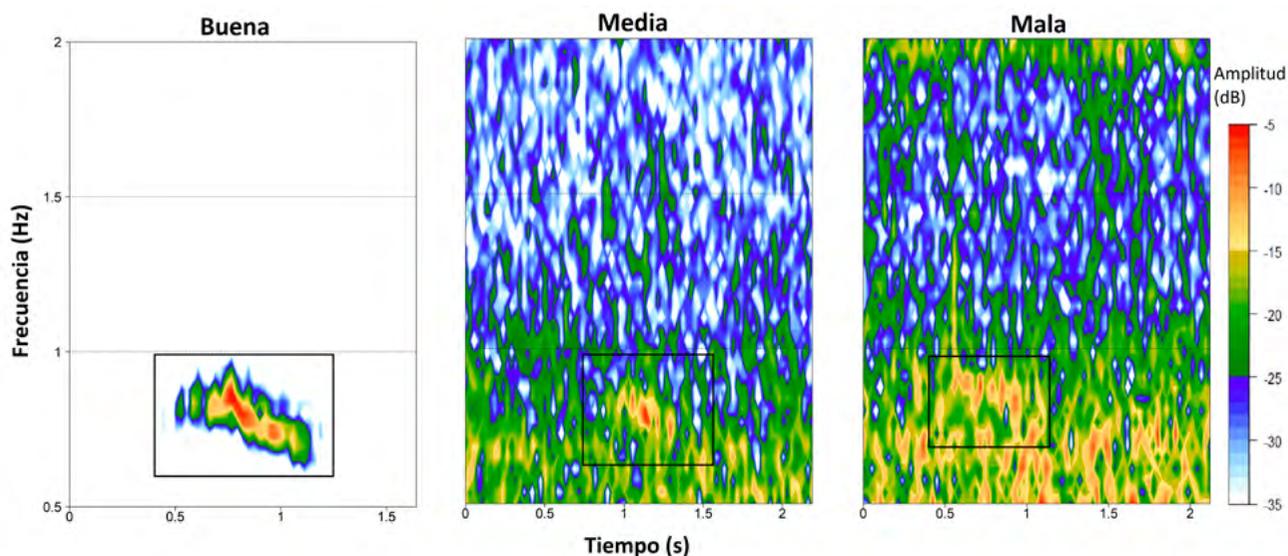
**Tabla 1.** Número total de grabaciones de 1 minuto por sitio de muestreo, fechas de actividad de cada grabadora y número de registros de *Megascops centralis* por sitio, en los alrededores del embalse Jaguas, Andes de Antioquia, Colombia.

**Table 1.** Total number of 1-minute recordings per site, their corresponding recording dates and the records of *Megascops centralis* in each site, around the Jaguas dam, Andes of Antioquia Colombia.

Grabadora	Num archivos	Fecha inicio	Fecha Final	Num registros
G03	276	11-may	18-may	0
G04	205	11-may	18-may	0
G06	459	18-may	27-may	3
G07	322	10-may	18-may	0
G08	322	10-may	18-may	0
G09	460	19-may	28-may	0
G13	321	28-may	4-jun	49
G15	456	19-may	27-may	6
G17	433	19-may	27-may	6
G19	456	19-may	28-may	0
G23	277	29-may	4-jun	7
G24	276	29-may	4-jun	0
G25	322	4-jun	11-jun	21
G27	322	28-may	4-jun	0
G28	276	5-jun	11-jun	5
G29	276	5-jun	11-jun	34
G34	324	18-jun	25-jun	0
G35	322	4-jun	11-jun	0
G36	277	12-jun	18-jun	0
G37	278	5-jun	11-jun	2
G40	217	12-jun	18-jun	1

**Análisis estadísticos.** El desempeño del modelo se determinó mediante matrices de confusión (presencia o ausencia de vocalización observada versus identificado por el modelo) para cada archivo de 1 minuto de grabación. Si el algoritmo predecía presencia del canto de la especie en un archivo de audio donde había al menos una vocalización de la especie, esto se registraba como un verdadero positivo. Es decir, que solo evaluamos el desempeño del algoritmo en términos de presencia/ausencia por archivo y no de la cantidad

absoluta de vocalizaciones en cada grabación. Los valores de sensibilidad, definida como la capacidad del algoritmo para detectar las señales de interés cuando realmente están presentes (verdaderos positivos) y la especificidad, definida como la capacidad del algoritmo para indicar la ausencia de las señales de interés cuando realmente no están (verdaderos negativos), se obtuvieron a partir de la matriz de confusión para cada parametrización del modelo, mediante el paquete *caret* (Kuhn, 2011) de R.



**Figura 2.** Ejemplos de vocalizaciones de *Megascops centralis* alrededor del embalse Jaguas, Andes de Antioquia, Colombia, correspondientes a cada una de las categorías de calidad identificadas, realizados en la banda de 0.5-2 kHz. El recuadro negro indica la ubicación de la vocalización y los colores representan la amplitud de la señal. Para la generación de los espectrogramas se empleó un tamaño de ventana (wl) de 1200 y un traslape (ovlp) del 5 %. Estas señales fueron obtenidas de los dos sitios elegidos para la evaluación: G29 (calidad buena y media) y G13 (mala calidad).

**Figure 2.** Examples of vocalizations of *Megascops centralis* around the Jaguas dam, Andes of Antioquia Colombia, corresponding to each of the signal quality categories in the 0.5-2 kHz band. The black box indicates the location of the vocalization and the colors represent the amplitude of the signal. For the generation of the spectrograms, a window size (wl) of 1200 and an overlap (ovlp) of 5% were used. These signals were obtained from the two chosen sites for evaluation: G29 (good and medium quality) and G13 (bad quality).

Además de la caracterización espectral del canto que se pretende identificar, el modelo requiere un parámetro relacionado con el umbral de amplitud que permite diferenciar el ruido de las verdaderas señales de interés (umbral de amplitud). Se evaluó la eficiencia del algoritmo, bajo todas las posibles combinaciones de los siguientes parámetros: umbrales de 5 %, 10 % y 20 %, y tasas de muestreo de 44 100 Hz (original, de ahora en adelante 44 kHz) y 22 050 Hz (de ahora en adelante, 22 kHz). Resultados preliminares evidenciaron que, en un número considerable de ocasiones, la función detectó un alto número de cantos dentro de un archivo (i.e., más de 10), lo cual nunca fue registrado durante la revisión manual en las grabaciones (máximo cuatro señales por audio). Por lo tanto, decidimos introducir como postprocesamiento de las detecciones, un nuevo parámetro: *Th\_seleccion*, el cual establece el número de selecciones por archivo de audio a partir del cual se consideraría el archivo como positivo o negativo para la evaluación del desempeño. Por ejemplo, para un umbral de selecciones de 30, si la función registró 30 o más

cantos de *M. centralis* en una grabación, se interpreta como un negativo. Si para el mismo valor del umbral de selecciones, un audio contiene 25 vocalizaciones según la función autototec, entonces se interpreta como un positivo. Se evaluaron todos los posibles valores (0-57 de este parámetro en los diferentes experimentos).

De esta manera se identificaron los modelos con mejor desempeño bajo cada criterio, buscando (1) los valores de los parámetros que igualaran –o minimizaran la diferencia entre– los valores de sensibilidad y especificidad y, (2) los valores de los parámetros que maximizaran el valor de la suma de ambos criterios.

Por último, se evaluó si la probabilidad de que el modelo acertara en la identificación de un canto variaba en función de la calidad de la señal (buena, media y mala), mediante un modelo lineal mixto generalizado donde la variable respuesta fue acierto o desacierto (binaria), la variable predictora fue la calidad de las señales, incluyendo como variables aleatorias los parámetros de

frecuencia de muestreo y el umbral de amplitud. El modelo fue implementado con la función `glmer` mediante el paquete `lme4` de R (Bates et al., 2020), utilizando la función de enlace binomial-logit.

## Resultados

De los 21 sitios de muestreo, *M. centralis* fue encontrado en 10 (Tabla 1). En total, se encontraron 134 archivos con vocalizaciones de la especie de interés (1.94 % de las 6877).

De la totalidad de audios de *M. centralis* (134), 16 fueron clasificados de buena calidad, 29 de calidad media y 89 de mala calidad. La grabadora G29 incluyó un mayor número de vocalizaciones de buena calidad (10 buenas, 11 medias y 13 malas) con respecto a la grabadora G13 (0 buenas, 4 medias y 45 malas).

La banda de frecuencia del canto de *M. centralis* estuvo entre 540.9 Hz y 1030.4 Hz y la duración de los cantos evaluados osciló entre los 0.568s y 1.021s (Tabla 2). Los parámetros elegidos para la evaluación del algoritmo sobre la totalidad de los audios se encuentran en la última fila de la Tabla 2.

La función `autodetec` retorna una tabla con el número de selecciones encontradas por el detector en cada archivo. El número de selecciones en los archivos evaluados varió entre 0 y 57. En la gran mayoría de archivos, la

función `autodetec` identificó menos de 10 selecciones, dependiendo de la parametrización escogida (Figura 3).

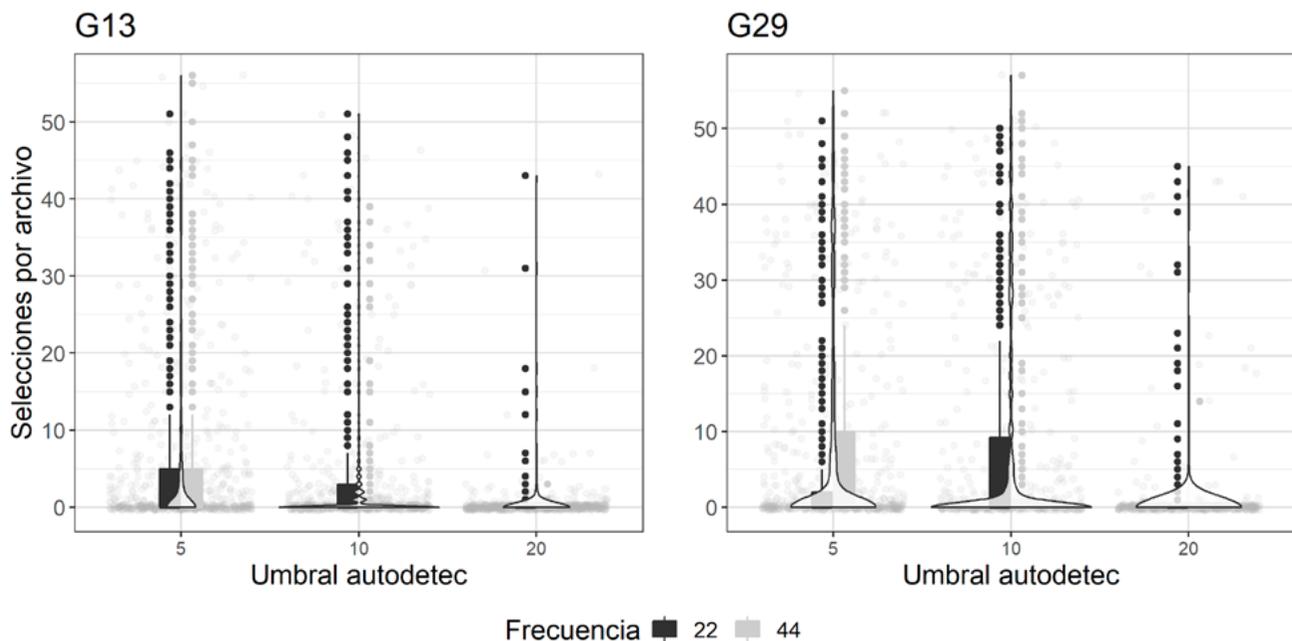
El desempeño de la función `autodetec` en la identificación de la presencia de cantos de *M. centralis* fue altamente variable en función de la parametrización del modelo (Tasa de muestreo, umbral de amplitud y `Th_seleccion`). En general, la sensibilidad (capacidad para detectar la presencia de cantos en un audio) disminuyó con un aumento del umbral de amplitud empleado en la función (Figura 4). Este patrón no se cumple a 22 kHz en G29, donde se aprecia un aumento de la sensibilidad entre 5 % (promedio  $0.45 \pm 0.05$ ) y 10 % (promedio  $0.58 \pm 0.07$ ) y una drástica caída al emplear un umbral de amplitud del 20 % (promedio  $0.1 \pm 0.01$ ). El valor máximo de sensibilidad alcanzado al usar las grabaciones del G29 está sobre 0.8, mientras que la sensibilidad máxima en el G13 es aproximadamente 0.3 (Figura 4). Por el contrario, la especificidad (capacidad de detectar ausencia de cantos en un audio cuando realmente no hay) aumentó con la tasa de muestreo y con el umbral de amplitud (Figura 4). La variación observada dentro de las cajas, en las gráficas de especificidad y sensibilidad (Figura 4) se debe al parámetro `Th_seleccion`.

Es claro que, para todos los experimentos realizados, la sensibilidad –mejor capacidad de identificar un canto de *M. centralis*– aumenta al ser más permisivos con el número de selecciones a partir del cual se considera el audio como negativo (un aumento en el valor de `Th_seleccion`), mientras que la especificidad disminuye (Figura 5).

**Tabla 2.** Valores mínimos, máximos y promedio de frecuencia mínima (Hz), frecuencia máxima (Hz) y duración de los cantos(s) de *M. centralis*, alrededor del embalse Jaguas, Andes de Antioquia, Colombia. La última fila de la tabla muestra los parámetros seleccionados para implementar la función `autodetec` que requiere especificar una duración mínima y una duración máxima. Las duraciones elegidas se muestran en la misma celda como min-max(min-max).

**Table 2.** Minimum, maximum and average values of minimum frequency (Hz), maximum frequency (Hz) and signal duration (s) of *M. centralis*' vocalizations around the Jaguas dam, Andes of Antioquia Colombia. The last row of the table includes the parameter values selected to implement the `autodetec` function. The function requires specifying a minimum duration and a maximum duration, the chosen durations are shown in the same cell as min-max

	Frec min (Hz)	Frec max (Hz)	Duración (s)
Mínima (medida)	540.9	894	0.568
Máxima (medida)	765	1030.4	1.021
Promedio	627.38 (+/-17.5)	966.17 (+/- 45.7)	0.8661 (+/-0.1)
Parámetros elegidos	560	1040	0.4-1.3



**Figura 3.** Número de selecciones generadas por autodetec en cada uno de los experimentos realizados para las dos grabadoras G13 y G29, usadas para identificar las vocalizaciones de *Megascops centralis* alrededor del embalse Jaguas, Andes de Antioquia, Colombia. En el eje X se observan los umbrales de amplitud utilizados en la función, y en el eje Y el número de selecciones por archivo. Los puntos grises claros en el fondo representan el número de selecciones en cada audio, las barras corresponden al inter cuartil de los datos (percentil 25-75), las líneas incluyen el máximo para cada frecuencia evaluada (negras a 22 kHz y grises a 44 kHz) y sus respectivos valores atípicos. La trama de violín representa la densidad de los datos a lo largo del eje Y.

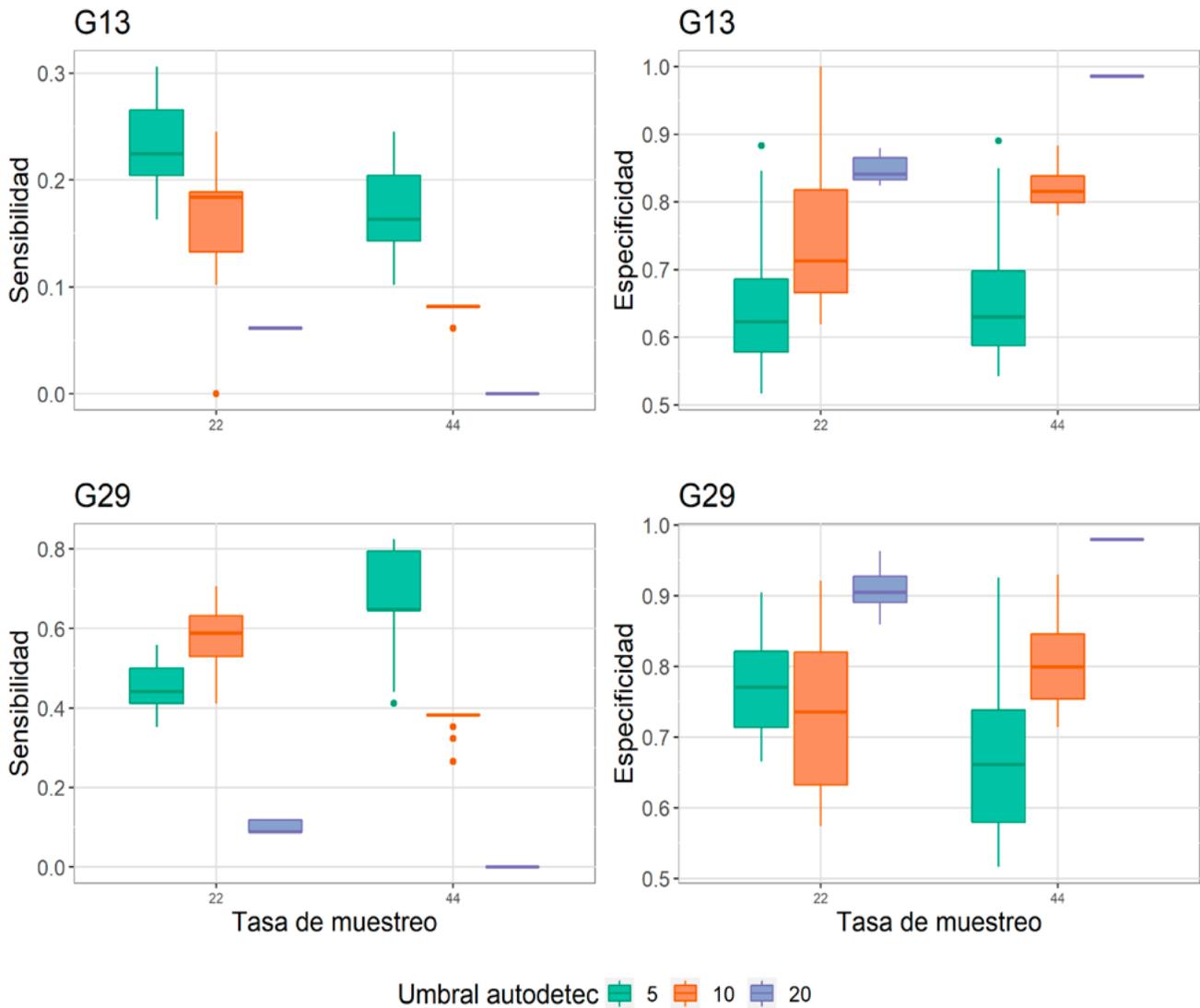
**Figure 3.** Number of selections generated by autodetec in each of the experiments carried out for the two recorders, G13 and G29 used to identify vocalizations of *Megascops centralis*, around the Jaguas dam, Andes of Antioquia Colombia. The X-axis shows the amplitude thresholds used in the function and the Y-axis shows the number of selections in each file. Light gray points in the background represent the number of selections for each audio, bars correspond to the interquartile range (25-75 percentile) for each frequency evaluated (black at 22 kHz and gray at 44kHz) and their respective outliers. Violin plots represent the density of the data along the Y axis.

Es posible observar que, para las evaluaciones de la grabadora G13, las líneas de sensibilidad y especificidad no llegan a cruzarse, mientras que en los experimentos realizados con los audios de la grabadora G29 vemos un cruce de las líneas en umbrales (de amplitud) de 5 % y 10 %.

De igual manera, al juntar los datos obtenidos de ambas grabadoras, observamos que las curvas no llegan a cruzarse y que la mejor combinación de parámetros resulta de una tasa de muestreo de 44 kHz, un umbral de ruido del 5 %, y el mínimo valor del Th\_seleccion –umbral de selección– (Figura 5). Adicionalmente, al aumentar el parámetro del umbral de amplitud en la función autodetec hay una tendencia

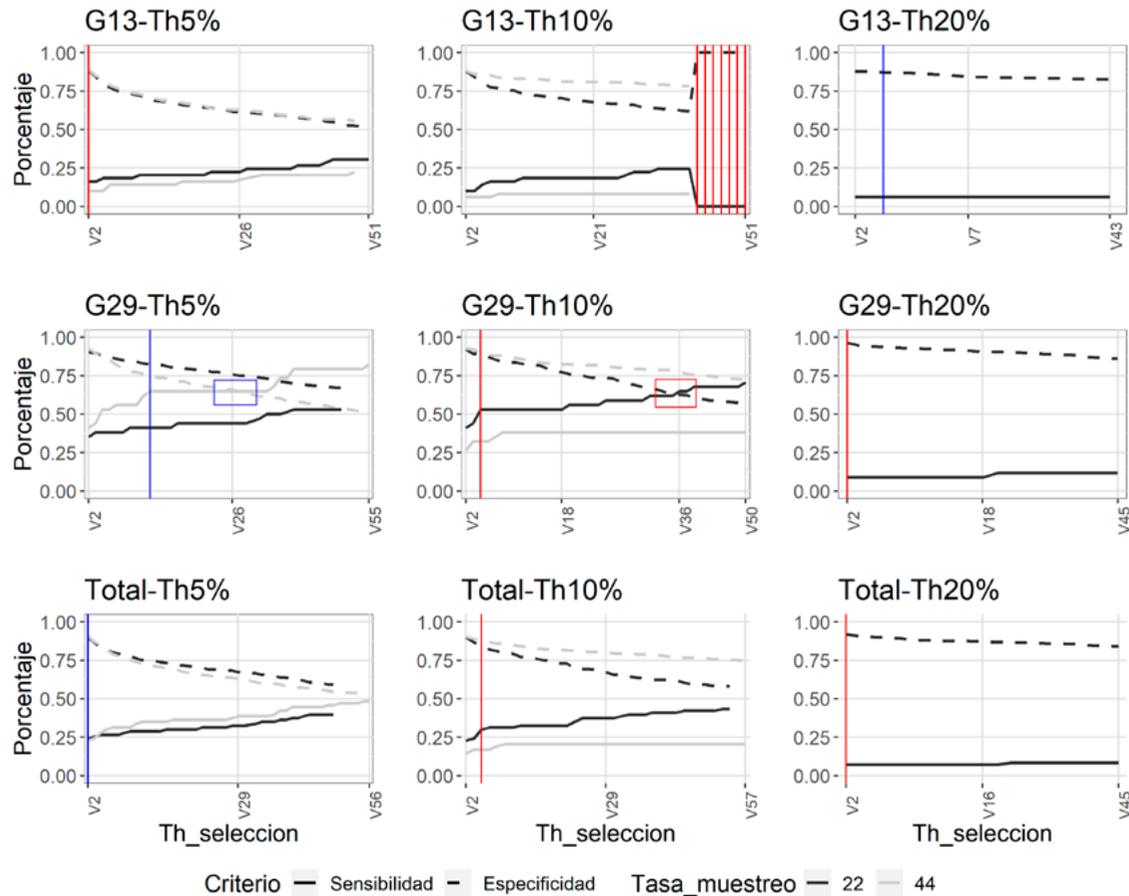
del algoritmo a identificar un menor número de cantos por audio.

Los resultados del modelo lineal generalizado mixto muestran que la probabilidad de acierto del algoritmo de autodetec se ve significativamente afectada por la calidad de la señal. El modelo que incluye la calidad como factor fijo es superior al modelo nulo donde solo se incluyen los factores aleatorios (diferencia en AIC ~ 54.4). La probabilidad de acierto estimada para audios donde las vocalizaciones son de buena calidad es 0.64, mientras que para aquellos de intermedia o mala calidad es de 0.27 y 0.12 respectivamente. El modelo completo tiene un  $R^2$  de 0.48, mientras que el  $R^2$  marginal (solo con respecto a los factores fijos) fue de 0.12.



**Figura 4.** Sensibilidad y especificidad del algoritmo autodetec para identificar presencia y ausencia de vocalizaciones de *Megascops centralis* en los audios de las grabadoras de los sitios seleccionados (G13 y G29) alrededor del embalse Jaguas, Andes de Antioquia, Colombia. La variación en la especificidad y sensibilidad para cada combinación única de parámetros se debe al parámetro Th\_seleccion.

**Figure 4.** Sensitivity and specificity of the autodetec function to identify presence and absence of *Megascops centralis* vocalizations in the recordings from the selected sites (G13 and G29) around the Jaguas dam, Andes of Antioquia Colombia. The specificity and sensitivity variation in each unique combination of parameters is generated from the Th\_selection parameter.



**Figura 5.** Desempeño de la función autotetec, en términos de sensibilidad y especificidad, bajo diferentes parámetros y con diferentes umbrales de la selección ( $Th\_seleccion$ ) para identificar vocalizaciones de *Megascops centralis* en los audios de las grabadoras de los sitios seleccionados (G13 y G29), alrededor del embalse Jaguas, Andes de Antioquia, Colombia. La línea gris corresponde a las pruebas realizadas a una frecuencia de 44 kHz, y la negra a frecuencias de 22 kHz. Las líneas punteadas muestran el comportamiento del criterio especificidad y en líneas continuas el comportamiento del criterio de sensibilidad. Las líneas verticales indican la combinación de parámetros que maximiza, por suma, el desempeño de la función. En rojo (G13-Th5 %, G13-Th10 %, G29-Th10 %, G29-Th20 %, Total-TH10 % y Total-Th20 %) se indican aquellas combinaciones de parámetros en las que el desempeño del modelo se maximiza a 22 kHz y en azul (G13-Th20 %, G29-Th5 % y Total-Th5 %) a 44 kHz. Los cuadros en las pruebas de G29 (Th5 % y Th10 %) señalan los puntos de cruce de las líneas de los parámetros (donde la sensibilidad y especificidad se igualan). La figura incluye la evaluación del algoritmo al emplear los audios de cada punto por separado (G13, G29) y al unir los audios de cada sitio (Total).

**Figure 5.** Performance of the autotetec function, in terms of sensitivity and specificity, under different parameters and different selection thresholds ( $Th\_seleccion$ ) to identify *Megascops centralis* vocalizations in the recordings from the selected sites (G13 and G29) around the Jaguas dam, Andes of Antioquia Colombia. The gray line corresponds to tests carried out at a sample rate of 44 kHz, and the black line to 22 kHz. Dotted lines show the behavior of the specificity criterion, while solid lines show the behavior of the sensitivity. The vertical lines indicate the combination of parameters that maximizes, by sum, the performance of the function. The color red (G13-Th5 %, G13-Th10 %, G29-Th10 %, G29-Th20 %, Total-TH10 %, and Total-Th20 %) indicate parameters' combination in which the performance of the model is maximized when the sample rate was 22 kHz, and in blue (G13-Th20 %, G29-Th5 %, and Total-Th5 %) when sample rate was 44 kHz. The boxes in the G29 tests (Th5 % and Th10 %) show the crossing points of the parameter lines (where the sensitivity and specificity are equal). The figure includes the evaluation of the algorithm when using the audios of each point separately (G13, G29) and when joining the audios of each site (Total).

## Discusión

Los resultados de nuestro trabajo indican que (1) *M. centralis* es una especie difícil de registrar: solo 1.94 % del total de audios incluyeron cantos de esta especie; (2) el algoritmo autodetect del paquete warbleR es de fácil implementación y su desempeño depende de la tasa de muestreo, el umbral de ruido, y la calidad de los cantos en las grabaciones. Este algoritmo tiene un muy buen potencial para la detección automatizada de cantos sencillos, pero requiere de buena calidad de los cantos en sus grabaciones.

De las 6877 grabaciones revisadas, la especie solo fue registrada en 134 grabaciones y en solo 10 de los 21 sitios muestreados, lo que puede dar indicios de que es una especie poco común y difícil de detectar; sin embargo, poco se conoce sobre las variables que pueden estar afectando su ocupación y detectabilidad. Dado que el tiempo de muestreo de las grabadoras fue corto (mayo-junio), dudamos que la heterogeneidad entre grabadoras pueda confundirse con estacionalidad en la actividad vocal de *M. centralis*. Se cree que los patrones de actividad de este género pueden estar ligados a los ciclos lunares (Rosado-Hidalgo, 2018), y es posible que el periodo de muestreo no haya coincidido con el pico de actividad vocal de la especie, que se espera sea durante periodos de luna llena; sin embargo esta hipótesis se encuentra en evaluación. Los datos generados durante este estudio pueden ser utilizados para entender si la ocupación y la detección de esta especie están relacionadas con covariables ambientales (Campos-Cerqueira & Aide, 2016), lo cual sería de gran interés, ya que se trata de una especie recientemente separada de *M. guatemalae* (SACC, 2018), con una distribución aparentemente fragmentada. Las herramientas bioacústicas como la presentada en este trabajo cobran particular importancia para el estudio de estas especies crípticas.

Uno de los insumos principales para establecer límites taxonómicos en búhos son sus vocalizaciones, y una buena caracterización de sus vocalizaciones es, por lo tanto, esencial (Dantas *et al.*, 2016; Krabbe, 2017). Para *M. centralis*, Krabbe (2017) encontró que los cantos oscilaban entre 687 y 920 Hz y que la duración era de entre 0.7s y 1.6s. Los resultados obtenidos en el presente estudio indican un rango mayor de frecuencia (540.9Hz-1030.4 Hz) y un rango de duración menor (0.6-1.02 s). La diferencia en la frecuencia de los cantos puede deberse a que los audios analizados por Krabbe (2017) corresponden a cantos registrados en diferentes localidades de Suramérica, o puede también sugerir que existe cierta estructura

poblacional al interior de esta especie. Las diferencias existentes en las duraciones de los cantos pueden deberse a la calidad de las señales, puesto que, por efectos de atenuación y dispersión, las componentes de la señal con menos energía tienden a enmascarse con el ruido de fondo, perdiendo definición en la delimitación de sus características temporales (Marten *et al.*, 1977).

El algoritmo presentó un desempeño variable y relacionado, en parte, con la calidad de los cantos en las grabaciones. Este algoritmo fue originalmente diseñado e implementado sobre grabaciones direccionales de *Phaetornis longirostris* obtenidas de xeno-canto (Araza-Salas & Smith-Vidaurre, 2017), lo que implica niveles de ruido mucho menores y señales con una alta calidad. En nuestras condiciones –micrófonos omnidireccionales, ruido ambiental y distancias variables del emisor a la grabadora–, el algoritmo pudo alcanzar especificidades y sensibilidades altas, pero solo con un alto contenido de grabaciones con vocalizaciones de alta calidad. Como es frecuente, existe un compromiso entre la capacidad de identificar presencias y la capacidad de identificar ausencias: cuando el algoritmo logra detectar un alto número de audios con cantos (alta sensibilidad), presenta un alto número de falsos positivos (baja especificidad), y viceversa. Aunque resultados de altas especificidades (capacidad de detectar ausencia de cantos en un audio cuando realmente no hay) pueden ser útiles, por ejemplo, para descartar archivos que no contengan audios y disminuir el número de grabaciones que escuchar, requieren sensibilidades (capacidad de detectar presencia de cantos en un audio cuando realmente los hay) medias o buenas que permitan realizar un filtro efectivo. Otros trabajos en los que se han utilizado algoritmos para la detección automatizada de cantos de aves también han reportado resultados con un alto número de falsos positivos (Bardeli *et al.*, 2010), mientras que otros han encontrado resultados de puntajes de AUC entre 70% y 90% (Stowell *et al.*, 2019). Estos resultados son esperanzadores, pero usualmente se refieren a metodologías más avanzadas, como redes neuronales convencionales (CNN), que no cuentan con una interfaz de usuario, lo que requeriría un entendimiento más avanzado de los lenguajes de programación empleados. Los resultados de autodetect pueden ser poco esperanzadores para el tipo de grabaciones de campo omnidireccionales y sugieren el uso de herramientas más complejas con el fin de obtener mejores resultados.

Obtener un balance entre la sensibilidad y especificidad de un detector es ideal, aunque no resulta sencillo. En nuestros resultados, solo pudimos evidenciar

un balance en la grabadora del sitio G29, donde se encontraban el mayor número de audios con alta calidad (cruces de las líneas de sensibilidad y especificidad, [Figura 5](#)). Para esta grabadora se pudieron obtener dos equilibrios, el primero a 44 kHz con un umbral de amplitud del 5 %, con una sensibilidad de 0.74 y una especificidad de 0.6, y el segundo a 22 kHz, umbral de amplitud de 10 %, con una sensibilidad de 0.58 y una especificidad de 0.74. Ambos resultados se obtienen al ubicar el umbral de selección ( $Th_{seleccion}$ ) en 36 y 26 respectivamente (aceptando hasta 36 o 26 señales en un audio, como positivo). El resultado obtenido para los datos de G29 se vuelve interesante cuando se considera el Teorema de Nyquist-Shannon ([Nyquist, 1928](#)) que indica que para recuperar una señal por medio de un espectrograma, debe grabarse al doble de la frecuencia en la que ésta se produce. Dado que *M. centralis* vocaliza alrededor de los 1000 Hz, con una tasa de muestreo de alrededor de 2000 Hz sería suficiente para obtener la información requerida. Sería de pensar, entonces, que tasas de muestreo de 22 kHz y 44 kHz no deberían mostrar resultados diferentes; sin embargo, nuestros resultados sugieren que la manera como funciona el detector de esta función puede verse afectado por valores de la tasa de muestreo mucho más alta.

El alto número de selecciones de cantos en archivos donde no había ningún canto o sonido parecido fue un resultado inesperado, para el cual aún no tenemos una clara explicación. Muchas de estas instancias y las selecciones generadas por el algoritmo no aparentaban estar relacionadas con presencia de otras señales que pudieran considerarse falsos positivos, ni con ruido ambiental. Debido a esto, creemos que la introducción del umbral de selecciones ( $Th_{seleccion}$ ), como medida de post-procesamiento, puede ser una buena aproximación para filtrar estos audios que presentan un número alto de selecciones sin razón aparente. De una u otra manera, si se buscaran señales similares a las de *M. centralis*, con condiciones que sugieran señales de calidades medias y buenas, es recomendable emplear una tasa de muestreo de 22 kHz, un umbral de amplitud de 10 % y considerar un archivo de audio como positivo al presentar un máximo de 26 detecciones. Esto considerando que, aunque no fue evaluado, durante la realización de las pruebas se evidenció un menor tiempo de procesamiento para los archivos con una tasa de muestreo de 22 kHz.

En resumen, es claro que existe una relación directa entre la calidad de la señal y el desempeño de autodetec,

lo cual es de esperar en este tipo de algoritmos. Lo complicado es saber si la causa de la baja calidad de la señal se debe a interferencia por ruido ambiental, o a lejanía del emisor. Si la causa de la mala calidad es la lejanía del emisor, puede que las inferencias que se realicen sean consistentes con el hecho de que el organismo de interés no se encuentra dentro del rango de operación de la grabadora. Si la causa es interferencia por ruido, el problema al interpretar el audio como una ausencia se haría evidente. Es por ello que el criterio de inclusión de las señales debería darse, en parte, por la distancia a la que los emisores se encuentren de las unidades de recepción y por ello, se hacen necesarios estudios que calculen la relación entre dichas distancias y las amplitudes registradas ([Darras \*et al.\*, 2018](#)), en especial, bajo las diferentes condiciones que ofrece el Neotrópico.

## Conclusiones

Nuestros resultados sugieren que el desempeño de este algoritmo no es recomendable bajo las condiciones acústicas que implican sensores remotos, asociadas a la naturaleza del canto de una especie como *M. centralis*, que se caracteriza por sus bajas frecuencias, viéndose altamente afectado por el ruido. Sin embargo, es de esperarse que algoritmos sencillos, como el evaluado en este estudio, presenten mejores resultados en señales de otra naturaleza, como murciélagos, donde la relación señal-ruido sea mejor. Es probable que aproximaciones desde técnicas de aprendizaje de máquina o inteligencia artificial permitan obtener mejores resultados para escenarios como el presentado. No obstante, implicando un conocimiento detallado del algoritmo que se emplee. Una vez más, hacemos énfasis en la necesidad de contar con un conjunto de datos anotado y lo suficientemente grande, que permita realizar las respectivas evaluaciones

## Agradecimientos

Agradecemos al Grupo Herpetológico de Antioquia de la Universidad de Antioquia (GHA) por la proporción de los datos empleados, que fueron obtenidos gracias al convenio Universidad de Antioquia- ISAGEN, y al Grupo de Ecología y Evolución de Vertebrados de la Universidad de Antioquia (GEEV), por las licencias y el recurso humano proporcionados.

## Referencias

- Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ*, 2013(1), 1-19.  
<https://doi.org/10.7717/peerj.103>
- Almeira, J., & Guecha, S. (2019). Dominant power spectrums as a tool to establish an ecoacoustic baseline in a premontane moist forest. *Landscape and Ecological Engineering*, 15(1), 121-130.  
<https://doi.org/10.1007/s11355-018-0355-0>
- Analytics, S. (2015). Arbimon II: Bio-acoustics analysis platform. Retrieved from <https://arbimon.sieve-analytics.com/>
- Araya-Salas, M., & Smith-Vidaurre, G. (2017). warbleR: an R package to streamline analysis of animal acoustic signals. *Methods in Ecology and Evolution*, 8(2), 184-191.  
<https://doi.org/10.1111/2041-210X.12624>
- Araya-Salas, M., & Smith-Vidaurre, G. (2018). Package "warbleR."
- Baldo, S., & Mennill, D. J. (2011). Vocal behavior of Great Curassows, a vulnerable Neotropical bird. *Journal of Field Ornithology*, 82(3), 249-258.  
<https://doi.org/10.1111/j.1557-9263.2011.00328.x>
- Bardeli, R., Wolff, D., Kurth, F., Koch, M., Tauchert, K. H., & Frommolt, K. H. (2010). Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters*, 31(12), 1524-1534.  
<https://doi.org/10.1016/j.patrec.2009.09.014>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Rhrhristensen Bojesen, R. H., Singmann, H., ... Scheipl, F. (2020). Package "lme4." R topics documented.
- Blumstein, D. T., Mennill, D. J., Clemens, P., Girod, L., Yao, K., Patricelli, G., ... Hanser, S. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48(3), 758-767.  
<https://doi.org/10.1111/j.1365-2664.2011.01993.x>
- Byrnes, R. M. (2013). Assessing bioacoustic techniques for inventory and monitoring of forest owls in the Central Sierra Nevada, California. (Tesis). Humboldt State University, Science department. 50 pp.
- Campos-Cerqueira, M., & Aide, T. M. (2016). Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. *Methods in Ecology and Evolution*, 7(11), 1340-1348.  
<https://doi.org/10.1111/2041-210X.12599>
- Claudino, R. M., Carlos, M.-J. J., & Antonini, Y. (2018). Owl assemblages in fragments of atlantic forest in Brazil. *Ornitología Neotropical*, (29), 281-288.
- Dantas, S. M., Weckstein, J. D., Bates, J. M., Krabbe, N. K., Cadena, C. D., Robbins, M. B., ... Aleixo, A. (2016). Molecular systematics of the new world screech-owls (Megascops: Aves, Strigidae): Biogeographic and taxonomic implications. *Molecular Phylogenetics and Evolution*, 94, 626-634.  
<https://doi.org/10.1016/j.ympev.2015.09.025>
- Darras, K., Furnas, B., Fitriawan, I., Mulyani, Y., & Tschardtke, T. (2018). Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling. *Methods in Ecology and Evolution*, 9(9), 1928-1938.  
<https://doi.org/10.1111/2041-210X.13031>
- Deichmann, J. L., Acevedo-Charry, O., Barclay, L., Burivalova, Z., Campos-Cerqueira, M., d'Horta, F., Game, E. T., Gottesman, B. L., Hart, P. J., Kalan, A. K., Linke, S., Nascimento, L. Do, Pijanowski, B., Staatterman, E., & Mitchell Aide, T. (2018). It's time to listen: there is much to be learned from the sounds of tropical ecosystems. *Biotropica*, 50(5), 713-718.  
<https://doi.org/10.1111/btp.12593>
- Ferraz, G., Sberze, M., & Cohn-Haft, M. (2010). Using occupancy estimates to fine-tune conservation concerns. *Animal Conservation*, 13(1), 19-20.  
<https://doi.org/10.1111/j.1469-1795.2010.00350.x>
- Goyette, J. L., Howe, R. W., Wolf, A. T., & Robinson, W. D. (2011). Detecting tropical nocturnal birds using automated audio recordings. *Journal of Field Ornithology*, 82(3), 279-287.  
<https://doi.org/10.1111/j.1557-9263.2011.00331.x>
- Heinicke, S., Kalan, A. K., Wagner, O. J. J., Mundry, R., Lukashevich, H., & Kühl, H. S. (2015). Assessing the performance of a semi-automated acoustic monitoring system for primates. *Methods in Ecology and Evolution*, 6(7), 753-763.  
<https://doi.org/10.1111/2041-210X.12384>
- Kalan, A. K., Mundry, R., Wagner, O. J. J., Heinicke, S., Boesch, C., & Kühl, H. S. (2015). Towards the automated detection and occupancy estimation of primates using passive acoustic monitoring. *Ecological Indicators*, 54, 217-226.  
<https://doi.org/10.1016/j.ecolind.2015.02.023>
- Katz, J., Hafner, S. D., & Donovan, T. (2016). Tools for automated acoustic monitoring within the R package monitoR. *Bioacoustics*, 25(2), 197-210.
- Keen, S. C., Shiu, Y., Wrege, P. H., & Rowland, E. D. (2017). Automated detection of low-frequency rum-

- bles of forest elephants: A critical tool for their conservation. *The Journal of the Acoustical Society of America*, 141(4), 2715-2726.  
<https://doi.org/10.1121/1.4979476>
- König, C., Weick, F., & Jan-Hendrik, B. (2008). *Owls of the World*. Helm Identification guides. London. 528 pp.
- Krabbe, N. K. (2017). A new species of *Megascops* (strigidae) from the Sierra Nevada de Santa Marta, Colombia, with notes on voices of new world screech-owls. *Ornitologia Colombiana*, 2017(16).
- Kuhn, M. (2011). *The caret Package* (pp. 1-27). Retrieved from <http://cran.r-project.org/web/packages/caret/vignettes/caretTrain.pdf%0Apapers2://publication/uuid/D8CA271E-F548-44F3-B081-B129A8765F04>
- Marsland, S., Priyadarshani, N., Juodakis, J., & Castro, I. (2019). AviaNZ: A future-proofed program for annotation and recognition of animal sounds in long-time field recordings. *Methods in Ecology and Evolution*, 10(8), 1189-1195.  
<https://doi.org/10.1111/2041-210X.13213>
- Marten, K., Quine, D., & Marler, P. (1977). Sound transmission and its significance for animal vocalization—II. Tropical forest habitats. *Behavioral Ecology and Sociobiology*, 2(3), 291-302.  
<https://doi.org/10.1007/BF00299741>
- Nagy, C. M., & Rockwell, R. F. (2012). Identification of individual Eastern Screech-Owls *Megascops asio* via vocalization analysis. *Bioacoustics—the International Journal of Animal Sound and Its Recording*, 21(2), 127-140.  
<https://doi.org/10.1080/09524622.2011.651829>
- Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transaction American Institute of Electrical Engineers*, 614-644.
- Ovaskainen, O., Moliterno de Camargo, U., & Somervuo, P. (2018). Animal Sound Identifier (ASI): software for automated identification of vocal animals. *Ecology Letters*, 21(8), 1244-1254.  
<https://doi.org/10.1111/ele.13092>
- Parker, T. (1991). On the use of tape recorders in avifaunal surveys. *The Auk: Ornithological Advances*, 108(02), 443-444.
- Program B. R. (2011). *Raven Pro: Interactive sound analysis software* (1.5). The Cornell Lab of Ornithology.  
<http://ravensoundssoftware.com/software/raven-pro/>
- Ralph, C. J., Geupel, G. R., Pyle, P., Martin, T. E., & DeSante, D. F. (1993). *Handbook of Field Methods for Monitoring Landbirds*. Pacific Southwest Research Station, Albany, California.
- Ranft, R. (2004). Natural sound archives: Past, present and future. *Anais da Academia Brasileira de Ciencias*, 76(2), 455-465.  
<https://doi.org/10.1590/S0001-37652004000200041>
- Rosado-Hidalgo, S. (2018). *To sing or not to sing: Effects of the moon cycles on the vocal activity of American owls*. [Pontificia Universidad Javeriana].  
<https://repository.javeriana.edu.co/handle/10554/35421?locale-attribute=frSACC>,
- SACC. (2018). Proposal (771) to South American Classification Committee Break up *Megascops guatemalae* into several species (II). Retrieved April 18, 2020, from <http://www.museum.lsu.edu/~Remsen/SACCprop771.htm>
- Sberze, M., Cohn-Haft, M., & Ferraz, G. (2010). Old growth and secondary forest site occupancy by nocturnal birds in a neotropical landscape. *Animal Conservation*, 13(1), 3-11.  
<https://doi.org/10.1111/j.1469-1795.2009.00312.x>
- Sethi, S. S., Ewers, R. M., Jones, N. S., Signorelli, A., Pincinali, L., & Orme, C. D. L. (2020). SAFE Acoustics: An open-source, real-time eco-acoustic monitoring network in the tropical rainforests of Borneo. *Methods in Ecology and Evolution*, 2020(February), 1-4.  
<https://doi.org/10.1111/2041-210X.13438>
- Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., & Glotin, H. (2019). Automatic acoustic detection of birds through deep learning: The first bird audio detection challenge. *Methods in Ecology and Evolution*, 10(3), 368-380.  
<https://doi.org/10.1111/2041-210X.13103>
- Ulloa, J. S., Gasc, A., Gaucher, P., Aubin, T., Réjou-Méchain, M., & Sueur, J. (2016). Screening large audio datasets to determine the time and space distribution of Screaming Piha birds in a tropical forest. *Ecological Informatics*, 31, 91-99.  
<https://doi.org/10.1016/j.ecoinf.2015.11.012>

**Laura Andrea Hoyos Cardona**

Universidad de Antioquia

Medellín, Colombia

[lauraa.hoyos@udea.edu.co](mailto:lauraa.hoyos@udea.edu.co)

<https://orcid.org/0000-0003-2089-215X>

**Juan Sebastián Ulloa**

Instituto de Investigación de Recursos Biológicos

Alexander von Humboldt

Bogotá, Colombia

[julloa@humboldt.org.co](mailto:julloa@humboldt.org.co)

<https://orcid.org/0000-0001-5666-6942>

**Juan Luis Parra Vergara**

Universidad de Antioquia

Medellín, Colombia

[juanl.parra@udea.edu.co](mailto:juanl.parra@udea.edu.co)

<https://orcid.org/0000-0002-5689-1872>

---

**Detección automatizada de cantos de aves continúa siendo un desafío: el caso de warbleR y *Megascops centralis* (búho del Chocó)**

**Citación del artículo:** Hoyos-Cardona, L.A., Ulloa, J.S. & Parra-Vergara, J.L. (2021). Detección automatizada de cantos de aves continúa siendo un desafío: el caso de warbleR y *Megascops centralis* (búho del Chocó). *Biota Colombiana*, 22(1), 149-163.

<https://doi.org/10.21068/c2021.v22n01a10>

**Recibido:** 30 de mayo 2020

**Aceptado:** 9 de noviembre 2020