Data Article

Dataset on audio records of animals from the northeast Andes of Colombia II: The vertebrate sounds of Santander department

Ana María Ospina-L a,*, Kevin G. Borja-Acosta b, Alexandra Buitrago-Cardona a, d, Mauricio Torres c, 1, Orlando Acevedo-Charry a, d

a Colección de Sonidos Ambientales, Colecciones Biológicas, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Carrera 8 # 15-08, Villa de Leyva, Boyacá, Colombia
b Sistematización y curación de datos, Colecciones Biológicas, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Claustro de San Agustín - Carrera 8 # 15-08, Villa de Leyva, Boyacá, Colombia
c Subdirección de Investigaciones, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Venado de Oro - Avenida Paseo Bolívar 16-20, Bogotá, Cundinamarca, Colombia
d Asociación Colombiana de Ornitología, Bogotá, Colombia

A R T I C L E  I N F O

Article history:
Received 29 November 2019
Received in revised form 5 February 2020
Accepted 10 February 2020
Available online 17 February 2020

Keywords:
Colombia
Natural history audio archives
Sound recordings
Sound collection

A B S T R A C T

Colombia holds one of the most spectacular biodiversity of the world. Yet, vast aspects of this biodiversity are still poorly inventoried. One of the least known aspects of Colombia’s biodiversity is the sound produced by its animals, even for the most conspicuous ones, the vertebrates. Here we reviewed and compiled the sound records available for the Department of Santander, a region in the North-East of Colombia, gathering the sound records of birds, anurans, mammals, and fishes. By conducting a detailed review in the environmental sound collection of the Humboldt Institute, the Macaulay Library of the Cornell Lab of Ornithology and Xeno-canto platform of the Naturalis Biodiversity Center, we present the first dataset of vertebrate sounds information from the Santander department. We selected recordings with a taxonomic resolution up to species and complete metadata information. Using latitude and longitude information, we assigned each recording to one of the six biotic units reported for Santander. We found a total of 1499 recordings, which belong to six biotic units: Guane-Yariguíes...
This dataset can have a wide scope of applications, from basic scientific questions, to analyses made by decision makers regarding conservation strategies, to support biodiversity-based economies such as ecotourism.

© 2020 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications Table

| Subject                  | Agricultural and Biological Sciences |
|--------------------------|-------------------------------------|
| Specific subject area    | Ecology, Evolution, Behaviour and Systematics. |
| Type of data             | Table |
| How data were acquired   | Field surveys, web searching in three repositories. |
| Data format              | Table with raw data |
| Parameters for data collection | Data consisted of records of sounds from the department of Santander that were identified to the taxonomical rank of species and that included geographic coordinates with an accuracy of decimal degrees. Additionally, we used records associated with an available catalogue number, to allow the unique identification of each record, avoiding duplicates. The search and compilation gathered information up to November 30th, 2019. |
| Description of data collection | The data was a combination of three datasets. First, we used Specify (www.speifysoftware.org), conducting a query of the IAvH-CSA database, searching for acoustic records collected in “Santander”. This query produced a DarwinCore file with catalogue number, taxonomical, and locality information. Secondly, we conducted a search in the Macaulay Library (www.macaulaylibrary.org), filtering localities from “Santander, Colombia (CO)”. Finally, we used the advanced search option in Xeno-canto (www.xeno-canto.org), sorting by country = “Colombia”, and Locality = “Santander”. Elevation data of the combined dataset were confirmed with the GPS Visualizer tool (http://www.gpsvisualizer.com). In addition, we adjusted the municipalities, and each record to biotic units and following the IDEAM and Laboratory of Applied Biogeography and Bioacoustics of the Humboldt Institute [1]. |
| Data source location     | Acoustic records were obtained from municipalities across the Department of Santander, which is located in the north-eastern Andes of Colombia and has an extent of 30,537 km². The coordinates systems in the tables are standardized in decimal degrees and the reference datum is WGS84. Latitude: 5.924 and 7.628 min and max, respectively. Longitude: −73.6589 and −73.779 min and max, respectively. Elevation: 51–4073.7 m. |
| Data accessibility       | With the article. |

Value of the Data

- To encourage data acquisition on acoustic signals of animals living on different biotic units in a hotspot of biodiversity. This is one of few efforts to compile and share the acoustic records available for different taxa in a selected location, encouraging other similar initiatives
- Provide acoustic information potentially useful for species recognition with passive acoustic monitoring methods, for delimiting species, understanding evolutionary patterns, monitoring populations for conservation strategies, and assessing questions from basic research on ecological or behavioural aspects of animals, to applied research on conservation exploring impacts on these basic aspects of animals.
- Provide baseline for extracting acoustic traits to test hypothesis, and for decision makers in the region to support biodiversity-based economies or other ecosystem services in their territory, as well as outreach activities.
1. Data description

We present the raw data of 1499 acoustic records of sounds obtained from organisms of four different taxa: birds, anurans, mammals, and fishes (Supplementary File 1). Bird records dominate the dataset (1440), followed by amphibians (48), fishes (9), and mammals (2). Most of the sound records (726) are included in XC, followed by IAvH-CSA (628), and ML (145). Six biotic units constitute Santander department: High Andes Eastern Cordillera, Catatumbo, Middle Magdalena Eastern Cordillera, Guane-Yarigüies, Middle Magdalena Valley and Mompox Depression, Nechí-San Lucas; while sound records were concentrated mainly in three of these biotic units: Guane-Yarigüies (597), Middle Magdalena Valley - Mompox Depression (484), and High Andes Eastern Cordillera (167); there is a lack of sound records in only one of the biotic units of Santander, the U’wa unit (Fig. 1). Our dataset includes a good representation of acoustic repertoires from vertebrate species, including electric fish signals, of the Department of Santander in different elevation gradients [2]. Species and number of recordings is showed in Table 1.

![Fig. 1](image_url).

**Fig. 1.** A. Map representing audio records of vertebrates in Santander, Colombia. Kernel density estimation for acoustic records in the study area with a radius of 10 km, shows the zones with highest concentration of sound records. Symbols indicate the repository source of records. B. Biotic units represented in the Department of Santander.
Table 1
Taxa representation of acoustic data in Santander department.

| Class          | Scientific name | Number of audio recordings | Percentage of recordings |
|----------------|-----------------|----------------------------|--------------------------|
| Aves           | 350             | 1440                       | 96.06                    |
| Amphibia       | 13              | 48                         | 3.20                     |
| Actinopterygi  | 2               | 9                          | 0.60                     |
| Mammalia       | 1               | 2                          | 0.13                     |

2. Experimental design, materials, and methods

2.1. Study area

The acoustic data was recorded in rural areas within 38 municipalities of the Department of Santander, which is located at the North-East of Colombia (Fig. 1). The sampling sites cover an altitudinal gradient, from 51 to 4073.7 m a.s.l., involving different types of ecosystems, such as open areas in agricultural systems, gallery forests, high Andean forests, mixed and oak forests, and paramos.

2.2. Data collection

We conducted a detailed review of the acoustic records of birds, anurans, mammals, and fishes in Santander. We used three repositories, the Environmental Sound Collection (Colección de Sonidos del Instituto Humboldt (IAvH-CSA) [3], the Macaulay Library of the Cornell Lab of Ornithology (ML: https://www.macaulaylibrary.org), and the Xeno-canto platform of the Naturalis Biodiversity Center (XC: https://www.xeno-canto.org). The first one (IAvH-CSA) combined acoustic data from historical records and data gathered from a recent explorative project in Santander performed by the authors. IAvH-CSA uses high quality recorders (SoundDevice MixPre3-I, Marantz MPD660-MKII), recording in wav format 48 kHz - 24-bit/sec, and coupled with unidirectional shotgun microphones (Sennheiser ME67) or a parabolic system (Telinga 22' with a cardioid microphone). Fishes were collected and set in a plastic cage and left for 5 minutes to adaptation after disturbance. Then, electric signals produced by fishes were recorded with a system of a bipolar electrode attached to the plastic cage and connected to a digital recorder (MPD660-MKII) [4]. The remaining repositories, ML and XC, included citizen science multimedia, which could include high quality recorders but also records made using commercial cell phones or other not professional recording equipment. Audio files in ML and XC are submitted mainly in wav and mp3 format, respectively.

We only included the sounds identified up to species category from the IAvH-CSA, ML or XC. After exploring the repositories, we combined the data (up to November 2019) and excluded duplicates using their catalogue numbers and remarks, as we reviewed this information in each database. Elevation for each record was confirmed by the GPS visualizer tool (http://www.gpsvisualizer.com), municipalities were adjusted and each record was assigned to a region and a biotic unit [1] using ArcGis 10.2. Biotic units were: High Andes Eastern Cordillera, Catatumbo, Middle Magdalena Eastern Cordillera, Guane-Yariguíes, Middle Magdalena Valley and Mompox Depression, Nechí-San Lucas. Finally, we performed a Kernel density analysis to visualize the zones within Santander with the highest concentration of sound records.

Acknowledgments

This study was supported in part by the project Santander Bio, which was funded by the Sistema General de Regalías de Colombia, administered by the Departamento Nacional de Planeación (BPIN 2017000100046), executed by the Gobernación de Santander, and operated by the Instituto de Investigación de Recursos Biológicos Alexander von Humboldt and the Universidad Industrial de Santander (Inter—administrative Agreement 2243, Gobernación de Santander). We thank to Alejandro Méndez-López and Fernando Alda for their help in the field taking electric signals of fishes.
Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2020.105298.

References

[1] M.C. Londorío, M.H. Olaya, C. Bello, I. González, C. Gutiérrez, D. López, J. Velásquez, Cuatro (4) mapas con regiones bióticas delimitadas utilizando como unidad de análisis los polígonos resultados del proceso de delimitación por la unidad ejecutora. Un mapa para cada grupo taxonómico: aves, mamíferos y herpetos y un mapa consenso., Bogotá D.C., Colombia, 2015, https://doi.org/10.1017/CBO9781107415324.004.

[2] M. Campos-Cerqueira, T.M. Aide, Changes in the acoustic structure and composition along a tropical elevational gradient, J. Ecoacoustics 1 (2017), https://doi.org/10.22261/JEA.PNCO71 #PNCO71.

[3] O. Acevedo-Charry, A.M. Ospina-Larrea, A. Buitrago-Cardona, K.G. Borja-Acosta, D. Murillo-Bedoya, J. Tovar-García, M. Alvarez-Rebolledo, M. Medler, E.T. Miller, A.M. Cuervo, C. Gómez-Pozada, H. García, A repository to the environmental sounds of Colombia: the Colección de Sonidos Ambientales at the Humboldt Institute, To Be Submitt. (n.d.).

[4] J. Sullivan, How to Record EODs in Mormyridae - African Weakly Electric Fishes, 2019, 15:05:20 +0000, http://cite.scratchpads.eu/mormyrids.myspecies.info/2019-11-29/How-to-record-EODs.pdf, (Accessed 29 November 2019).