Bird Census News 2025, 38/1–2: 37–42

Maximizing call detections of Boreal Owl by combining Kaleidoscope PRO and BirdNET

Marina Klaas-Fábregas[a*], Alba Estrada[a], Iván Afonso[b], Irene Figueroa[c], Martí Franch[d], Rafael López del Río[e], José María Martínez[f], Blas Molina[g], Juan Carlos Del Moral[g], Diego Villanúa[h] and Jesús Martínez-Padilla[a]

[a] Instituto Pirenaico de Ecología (CSIC). Avenida Nuestra Señora de la Victoria, S/N. 22700, Jaca, Spain
[b] Conselh Generau d'Aran. Plaça d'Aran, 1-2. 25530, Vielha, Spain
[c] Associació Paisatges Vius. Carrer Padró, 10. 08511 L'Esquirol, Spain
[d] Servei de Fauna i Flora, Generalitat de Catalunya. Carrer del Foc, 57. 08038, Barcelona, Spain
[e] Servicio de Biodiversidad, Dirección General de Medio Natural, Caza y Pesca. Departamento de Medio Ambiente y Turismo. Gobierno de Aragón. Plaza de San Pedro Nolasco, 7. 50071, Zaragoza, Spain
[f] Sección de Biodiversidad. Servicio Provincial de Medio Ambiente y Turismo. Avenida de la Paz, 5. 22004, Huesca, Spain
[g] SEO/BirdLife. Calle de Melquíades Biencinto, 34. 28053, Madrid, Spain
[h] Gestión Ambiental de Navarra (Orekan). Calle del Padre Adoáin, 219 bajo. 31015, Pamplona, Spain

*Corresponding author: m.klaas@ipe.csic.es; mklaasfabregas@gmail.com

Abstract. Ecoacoustics is an effective technique for monitoring elusive species, but generates large datasets that require processing and validation. In 2024, we deployed passive audio recordes to study the distribution of Boreal Owl (*Aegolius funereus*) in the Spanish Pyrenees. We assessed the performance of Kaleidoscope PRO and BirdNET against 25 manually-reviewed audio recorders. By combining both software, we detected 95% of positive audio recorders identified by manual revision. To maximize the detections of Boreal Owl calls, we recommend a two-step protocol based first on Kaleidoscope PRO with an advanced classifier, followed by BirdNET. We recommend developing a species-specific approach when aiming to study distribution or behavioural traits of different taxa.

Introduction

Ecoacoustics is a relatively new field of research that is constantly evolving (Stowell & Sueur 2020). In this discipline, Passive Acoustic Monitoring (PAM) has proven to be a useful tool for wildlife monitoring, and can help answer questions within the fields of biogeography, population dynamics, or animal behaviour (Ross et al. 2023; Stowell & Sueur 2020). More precisely, PAM has become a widely used technology to monitor uncommon or cryptic species (Campos-Cerqueira & Aide 2016; Freitas et al. 2023; Gibb et al. 2019).

The Boreal Owl (*Aegolius funereus*) is an elusive nocturnal raptor, and the Spanish Pyrenees is the meridional limit of its distribution (Keller et al. 2020). The low detectability and scattered distribution constrain its monitoring at population and individual level. As a consequence, knowledge about its distribution, abundance and ecology in Spain is very limited, affecting its conservation strategies.

During 2024, we designed and carried out the first national survey of Boreal Owl in Spain. Passive audio recorders were set along the Spanish Pyrenees, in order to update its distribution and abundance (Martínez-Padilla et al. 2024). PAM has proven to be a very useful technique to monitor owl species and to increase their detection rates in other regions (Freitas et al. 2023, Shonfield et al. 2018). Nevertheless, PAM has some constraints concerning the management of large amounts of data (Gibb et al. 2019). Audio recorders can record for several months with high frequency intervals accumulating large quantity of audio data, which requires a considerable amount of time and resources for data processing (Nieto-Mora et al. 2023). The use of artificial intelligence (AI) has considerably increased in the last few years in order to maximize the detection of targeted species, optimizing the time spent reviewing and improving data processing (Nieto-Mora et al. 2023). However, the reliability of AI differs between species (Ruff et al. 2021, Nieto-Mora et al. 2023, Manzano-Rubio et al. 2022, Kahl et al. 2021), and enhancing the specificity and sensibility of species detection is essential to improve the robustness of the analysis outcomes. Kaleidoscope PRO is a commercial software designed for bioacoustics analysis. One of its functionalities, cluster analysis, can be used for targeted searches of acoustic patterns using Hidden Markov Models, detecting and sorting signals of given characteristics within the audio files. Cluster analysis can also be combined with the use of "bait files", improving the detection rate by augmenting the statistical density of the target sounds. In addition, cluster analysis allows advanced classifiers to be built from training data, enabling a better isolation of the target species and labelling the detections with custom tags (Wildlife Acoustics, Inc. 2023). Alternatively, BirdNET is a deep neural network trained for the sound detection of more than 900 bird species (Kahl et al. 2021). One of its multiple features include the detection of target species by providing a specific species list, with a differing performance depending on the species to detect, its distance to the call source and the recorder itself (Manzano-Rubio et al. 2022, Pérez-Granados 2025, Wood & Kahl 2024). Once analysed, BirdNET provides detections in three seconds splits, with a confidence score for each one (Kahl et al. 2021, Wood & Kahl 2024). Despite its usefulness, both Kaleidoscope PRO and BirdNET require human validation of the results, since it is usual to find false-positive detections (Pérez-Granados et al. 2024, Tseng et al. 2025, Wood & Kahl 2024). Therefore, the reliability of results of both methods demands human validation.

The aim of this study is assessing the performance of Kaleidoscope PRO and BirdNET in detecting Boreal Owl calls compared to human-expert detection, in order to outline a species-specific protocol for its monitorization using passive audio recorders (PAM). Specifically, we aimed to i) evaluate the performance of several cluster analyses in Kaleidoscope PRO; ii) develop an advanced classifier with Kaleidoscope PRO; iii) assess the performance of BirdNET at different confidence thresholds; iv) evaluate the performance of BirdNET in recorders not detected by Kaleidoscope PRO; and v) quantify the performance of both software together.

Material and methods

Collection of acoustic data

We used two types of passive audio recorders, Audiomoth and Song Meter Micros, in order to record the Boreal Owl male calls in the Spanish Pyrenees. The devices were deployed in subalpine forests during the breeding season (March – April) for a mean of 19 days (range: 6 to 42 days). They were programmed to record 90 and 270 minutes before and after sunset, respectively, and 150 and 45 minutes before and after sunrise, respectively. From all the recorders installed, we selected 25 and reviewed them manually (visualization and earing) in order to look for Boreal Owl calls. After human validation, 20 of them were positive and five of them negative to the presence of the species. The amount of manual validated files was equivalent to 1 TB.

Kaleidoscope PRO

We assessed the performance of three cluster analysis methods of Kaleidoscope PRO 5.6.6 (Wildlife Acoustics, 2024) in the 25 manually reviewed recorders. The methods evaluated were: i) signal extraction and clustering; ii) signal extraction and clustering with baits; and iii) advanced classifier. Each analysis took a rough of 12–15 minutes (not including the validation process). The signal parameters settled were: 400 and 900 Hz (minimum and maximum frequency range, respectively); 0.1 and 7.5 seconds (minimum and maximum length of detection, respectively); a maximum inter-syllable gap of 0.35 seconds; a maximum distance from cluster centre of 1.0; a Fast Fourier Transform Window of 21.33 ms; and default settlings for: maximum states, maximum distance to cluster centre for building clusters, and maximum clusters. Likewise, the advanced classifier was built with the same parameters and 553 audio files with calls from the same study area. The bait files were selected from a dataset of recordings collected in 2023 from the same study area.

For each method and each recorder, we checked the detections of the first four clusters, and com-

Table 1. Comparison of performance of cluster analysis tools in Kaleidoscope PRO for positive audio recorders. Method 1: signal extraction and clustering; Method 2: signal extraction and clustering with baits; Method 3: advanced classifier.

Kaleidoscope PRO methods	Method 1	Method 2	Method 3
Positive audio recorders detected	6	4	12
Positive audio recorders not detected	14	16	8
Success rate	30	20	60

pared the results with the results of the manual review. If a Boreal Owl call was found in at least one detection of the first four clusters, the recorder was considered positive to the presence of the species. If not, the recorder was considered negative to the presence of the species.

BirdNET

Using the same 25 audio recorders, we analysed the accuracy of the detections of Boreal Owl in BirdNET in a range from 0 to 0.99 confidence. We manually reviewed the three-second detections delivered by the programme, and noted: 1) whether the detection was a *true positive* (a Boreal Owl present in the audio fragment and correctly identified by BirdNET as a Boreal Owl) or a *false positive* (any sound different to a Boreal Owl but incorrectly identified by BirdNET as a Boreal Owl), and 2) the confidence score given by BirdNET for both true and false positive detections. We then evaluated the relationship between the confidence score of the detections and their classification as true positive or false positive, by applying a logistic regression model using a generalized linear model (GLM) with a binomial distribution in R 4.2.2 (R Core Team 2024). In such models, "detection" (1: true positive; 0: false positive) was the dependent variable and confidence the explanatory term.

Secondly, we selected the *false negative* audio recorders (Boreal Owl presence confirmed manually, but considered negative by Kaleidoscope PRO) from the 25 sample. We analysed them with BirdNET and a custom species list to compare its performance in relation to Kaleidoscope PRO and manual reviewing. Each analysis took about three hours, and the inference settings were: 0.5 minimum confidence threshold, 1 sensitivity, 0 overlap, 0 Hz minimum bandpass frequency and 1500 Hz maximum bandpass frequency. If a true positive was found in at least one detection of BirdNET, the recorder was considered to be correctly identified as positive to the presence. To the contrary, a recorder was considered negative to the presence if there were no detections of Boreal Owl, or no true positive detections.

Results

Kaleidoscope PRO

The advanced classifier built with Kaleidoscope PRO was the cluster analysis method with the best performance, by detecting the 60% (n=12 out of 20 positive audio recorders; Table 1) of the positive audio recorders of the sample. The signal extraction and clustering tool was not efficient, as it only detected the 30% (n=6) of the positive audio recorders. The use of baits did not improve the detection rates, as it detected the 20% (n=4) of the positive audio recorders. As a consequence, eight out of 20 positive audio recorders were incorrectly considered by Kaleidoscope PRO as negatives to the presence of Boreal Owl. Furthermore, none of the three methods found Boreal Owl calls in the five audio recorders that were negative to the presence of the species by manually reviewing, validating the negative results.

BirdNET

Concerning the detections of BirdNET and its confidence, we found that detections were associated with confidence interval (estimate 10.108 ± 0.264, $z=38.34$, $p<0.001$). Specifically, both true positives and false positives could be found at both very low and very high confidence scores. Noteworthy, in our analysis, the number of true positive detections reached a maximum specially from a 0.5 confidence (Fig. 1).

On the other hand, BirdNET was able to detect the 87.5% (n=7 out of 8) of the positive audio audio recorders that were not detected by Kaleidoscope PRO.

The combination of the best method of Kaleidoscope PRO (in this case, the advanced classifier)

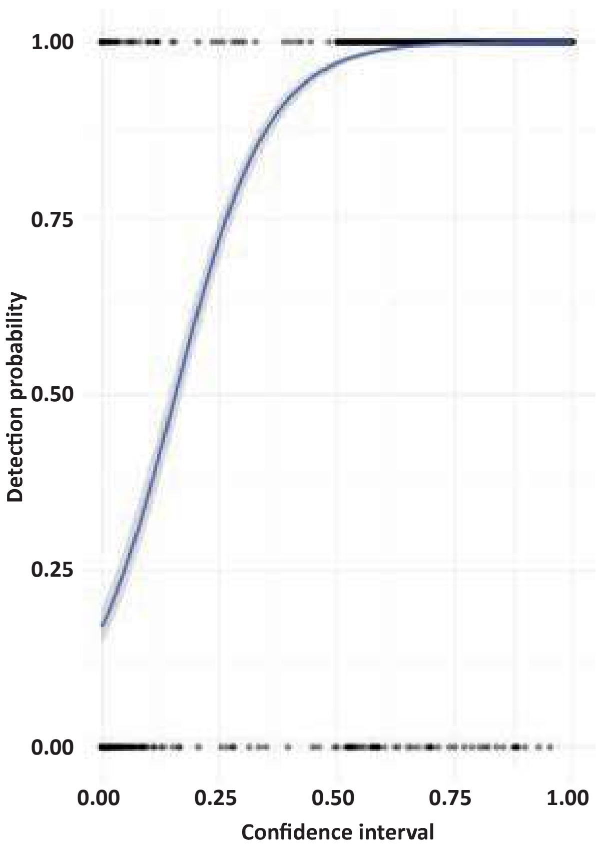

Figure 1. The positive relationship between BirdNET confidence scores and the probability of a true detection for Boreal Owl vocalisations. The blue line represents the predicted probability of detection as a function of confidence. The individual points represent observed detections classified as true positives (1) and false positives (0).

and BirdNET resulted in a correct classification of the positive audio recorders of 95% (n=19).

Discussion

The fast development of AI is providing researchers and managers an increasing number of efficient tools to work with. However, AI needs qualified personnel, big training datasets and a large amount of time to be trained (Manzano-Rubio et al. 2022; Pérez-Granados et al. 2024, Wood et al. 2023). Moreover, the large amount of data generated by PAM leads to a trade-off between precision and time investment. In this context, resources like Kaleidoscope PRO and BirdNET emerge as useful tools when facing constraints in time and/or specialised staff. Nevertheless, it is important to emphasize that the performance of each software varies depending on the species (Ruff et al. 2021, Nieto-Mora et al. 2023, Manzano-Rubio et al. 2022, Kahl et al. 2021), for example, between an owl that hoots or trills and a songbird that warbles. Therefore, our study stresses the need to properly characterise the calls of the target species to avoid false positive detections of the target species. Particularly, the length of the call, the pauses or gaps and the range of frequencies are crucial traits that need to be carefully determined (Wildlife Acoustics, Inc. 2023). In our study, the most frequent false positives were: 1) Tawny Owl detections (*Strix aluco*; similar range of frequencies), 2) environmental noise (wind, storms, screeching trees), 3) drumming of several species of woodpeckers (similar range of frequencies), and 4) anthropic sounds (church bells, machines, cars or humans). Regarding methodological aspects, the quality of an advanced classifier in Kaleidoscope PRO relies on the good characterisation of the sound, the performance of the signal extraction and the quality and source of origin of the audio files used for building it. Building an advanced classifier with sound files from the same study area can partially tackle these issues (Wildlife Acoustics, Inc. 2023). In the case of BirdNET, the training dataset is not publicly available, so the characteristics and quality of vocalisations used to train the algorithm remain unknown (Pérez-Granados et al. 2024). This is especially relevant when aiming to study ecological patterns, as it is necessary to determine whether the software detects the full range of vocal variation in the call of the species. In our study, BirdNET only detected the so-called 'staccato' song and the prolonged staccato song of male Boreal Owls. It remains unclear whether this is because the BirdNET algorithm has only been trained with these types of calls, or because other vocalisations were absent from our recordings. We did not find other types of vocalisations described in the literature, such as the delivery call, screech, peeping call, weak call, 'chuuk' call or hiss (Korpimäki & Hakkarainen 2012) by manual reviewing. Besides BirdNET not detecting all the calls of each positive audio recorders manually reviewed, it allowed us to correctly classify as positives the 87.5% of the positive audio recorders that had not been detected by Kaleidoscope PRO, resulting in a 95% of correct classification when using both software. The remaining underestimation of the detections of positive audio recorders by the two software may be due to the quality of the recordings; the distance between the sound source and the recorder; and the differences between the Boreal Owl calls of the study area and the calls used to train the advanced classifier in Kaleidoscope PRO, and the algorithm in BirdNET. However, we expect that the underestimation rate of Boreal Owl calls will decrease in the near future, with the improvement of AI and new technologies.

Given the differences in computing time between Kaleidoscope PRO and BirdNET, we recommend to design a workflow in which the audio recorders are first analysed with the advanced classifier in Kaleidoscope PRO, and secondly analysed with BirdNET if they are negative in the first stage. This combination of both tools was also recommended by Pérez-Granados et al. (2024) to detect calls of an amphibian species. Above all, when combining the use of PAMs and AIs, human validation is crucial to quantify the reliability of the detection outcome for each species, as these tools may produce a biased understanding of species distribution or individual behaviour (Pérez-Granados 2025). In conclusion, we support the use of advanced classifiers with Kaleidoscope PRO followed by BirdNET as an effective way of maximizing the detection of Boreal Owl calls in passive acoustic recorders.

Acknowledgements

We acknowledge the help of the forestry rangers of Aragón, Aran, Catalonia and Navarra, the technicians of protected areas in Catalonia, and those of Science for Forest Management, Biodiversity and Bioeconomy (CTFC) for the deployment of the passive audio recorders; Ramón Jato and Juan Carlos Albero for the bait files; Alberto Pastoriza for the configuration of the devices; and José Lahoz Monfort for help with BirdNET.

This study was funded by the European Union through an agreement between MITECO and TRAGSATEC. Job3088153_Ref. P2.C4I1.P1.S000.A2.E1, “Fauna Terrestre y Aves marinas, Especies autóctonas y Exóticas Invasoras. Mejora de Conocimiento del Estado de Conservación. TRAGSATEC agreed to SEO/BirdLife to carry out this study”. Such European funds are associated within the Recovery Plan, Transformation and Resilience (NEXTGENERATIONEU) under the following project: Conservation state of Spanish avian species included in Bird Directive, invasive species (article 24), and Spanish breeding bird atlas.

References

Campos-Cerqueira, M. & T. M. Aide. 2016. Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. *Methods in Ecology and Evolution*, 7: 1340–1348.

Freitas, B., Bas, Y., Robert, A., Doutrelant, C. & M. Melo. 2023. Passive acoustic monitoring in difficult terrains: The case of the Principe Scops-Owl. *Biodiversity and Conservation*, 32: 3109–3132.

Gibb, R., Browning, E., Glover-Kapfer, P. & K. E. Jones. 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, 10: 169–185.

Kahl, S., Wood, C. M., Eibl, M. & H. Klinck. 2021. BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics*, 61: 101236.

Keller, V., Herrando, S., Voříšek, P., Franch, M., Kipson, M., Milanesi, P., Martí, D., Anton, M., Klvanová, A., Kalyakin, M. V., Bauer, H.-G. & R. P. B. Foppen. 2020. *European Breeding Bird Atlas 2: Distribution, Abundance and Change*. European Bird Census Council & Lynx Editions.

Korpimäki, E. & H. Hakkarainen. 2012. *The Boreal Owl. Ecology, Behaviour and Conservation of a Forest-Dwelling Predator*. Cambridge University Press.

Manzano-Rubio, R., Bota, G., Brotons, L., Soto-Largo, E. & C. Pérez-Granados. 2022. Low-cost open-source recorders and ready-to-use machine learning approaches provide effective monitoring of threatened species. *Ecological Informatics*, 72: 101910.

Martínez-Padilla, J., Estrada, A. & M. Klaas-Fábregas. 2024. Distribución y estima de abundancia poblacional de Mochuelo boreal (*Aegolius funereus*) en el Pirineo. *Report on the agreement between CSIC and SEO-BirdLife.*

Nieto-Mora, D. A., Rodríguez-Buritica, S., Rodríguez-Marín, P., Martínez-Vargaz, J. D. & C. Isaza-Narváez. 2023. Systematic review of machine learning methods applied to ecoacoustics and soundscape monitoring. *Heliyon*, 9: e20275.

Pérez-Granados, C. 2025. BirdNET confidence scores decrease with bird distance from the recorder: Revisiting Pérez-Granados 2023. *Ardeola*, 72: 149–159.
Pérez-Granados, C., Feldman, M. J. & M. J. Mazerolle. 2024. Combining two user-friendly machine learning tools increases species detection from acoustic recordings. *Canadian Journal of Zoology*, 102: 403–409.
R Core Team. 2024. *R: A Language and Environment for Statistical Computing* [Software]. R Foundation for Statistical Computing. https://www.R-project.org/
Ross, S. R. P.-J., O'Connell, D. P., Deichmann, J. L., Desjonquères, C., Gasc, A., Phillips, J. N., Sethi, S. S., Wood, C. M. & Z. Burivalova. 2023. Passive acoustic monitoring provides a fresh perspective on fundamental ecological questions. *Functional Ecology*, 37: 959–975.
Ruff, Z. J., Lesmeister, D. B., Appel, C. L. & C. M. Sullivan. 2021. Workflow and convolutional neural network for automated identification of animal sounds. *Ecological Indicators*, 124: 107419.
Shonfield, J., Heemskerk, S. & E. M. Bayne. 2018. Utility of automated species recognition for acoustic monitoring of owls. *Journal of Raptor Research*, *52*: 42–55.
Stowell, D. & J. Sueur. 2020. Ecoacoustics: Acoustic sensing for biodiversity monitoring at scale. *Remote Sensing in Ecology and Conservation*, 6: 217–219.
Tseng, S., Hodder, D. P. & K. A. Otter. 2025. Setting BirdNET confidence thresholds: Species-specific vs. universal approaches. *Journal of Ornithology,* 166: 1123–1135.
Wildlife Acoustics. 2024. *Kaleidoscope PRO* (Version 5.6.6) [Software]. Inc.: Maynard, MA, USA.
Wildlife Acoustics, Inc. 2023. *Kaleidoscope User Guide*.
Wood, C. M. & S. Kahl. 2024. Guidelines for appropriate use of BirdNET scores and other detector outputs. *Journal of Ornithology*, 165: 777–782.
Wood, C. M., Kahl, S., Barnes, S., Van Horne, R. & C. Brown. 2023. Passive acoustic surveys and the BirdNET algorithm reveal detailed spatiotemporal variation in the vocal activity of two anurans. *Bioacoustics*, 32: 532–543.

Received: 17th June 2025
Accepted: 28th October 2025