The potential for passive acoustic monitoring and automated detection to improve conservation efforts of tarsiers in Sulawesi, Indonesia

Dena J. Clink[1*] and Johny S. Tasirin[2]
[1]K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA.
[2]Faculty of Agriculture, Sam Ratulangi University, Manado, Indonesia

Corresponding author email: dena.clink@cornell.edu

Abstract

Tarsiers are small, haplorrhine primates that occur in Southeast Asia. Tarsiers on the island of Sulawesi range from Vulnerable to Critically Endangered, and many are data deficient, which means there is a great need for improved monitoring approaches. Sulawesi tarsiers are pair-living, territorial, and engage in duets within human hearing range, which makes them ideal candidates for passive acoustic monitoring (PAM), an approach that relies on autonomous acoustic recording units. Here, we provide a proof-of-concept workflow that combines PAM, automated detection, and a simple occupancy modeling example to monitor Gursky's spectral tarsier (*Tarsius spectrumgurskyae*) in Tangkoko National Park, North Sulawesi, Indonesia. We used a custom trained BirdNET model deployed over ~520 hours of PAM data and manually verified all detections. Similar to previous work, we found that the majority of tarsier duet vocalizations occurred around sunrise, with a few vocalizations emitted in a non-duet context other times during the night. Using the true positive detections, we were able to create a detection history for occupancy modeling. Our simple occupancy modeling yielded occupancy and detection estimates consistent with expectations for the well-studied population at this site. We advocate that future work includes occupancy modeling across land use gradients, different forest types, and under different management regimes, to improve tarsier conservation efforts across Sulawesi. We provide the labeled training data and the trained model to facilitate future work.

Keywords: bioacoustics, transfer learning, autonomous recording units, nonhuman primates

Introduction

Passive acoustic monitoring (PAM) is an approach that relies on autonomous acoustic recording units (ARUs). PAM can provide data on vocal animals across larger spatial and temporal scales (e.g. ecologically relevant scales) than relying on human observers alone. PAM can be used for many conservation-relevant applications, including monitoring of spatial and temporal distribution of animals (e.g. forest elephants, *Loxodonta cyclotis*, (Wrege et al., 2017)) and human activities (Katsis et al., 2022), multi-species occupancy modeling across entire biomes (Brunk et al., 2025), and understanding ecological impacts of restoration activities on biodiversity (Ramesh et al., 2023). The use of PAM in terrestrial, tropical environments has seen an increase in use in recent years (Sugai et al., 2019), due to the combined factors of reduced costs of the ARUs, improved battery life, and increased data storage capabilities. ARUs can be deployed to record continuously for long time periods, which leads to the accumulation of large amounts of acoustic data that are time- and cost-prohibitive to analyze manually.

Like other fields, the deep learning revolution has led to improved efficiency of automated detection and classification of acoustic signals from large acoustic datasets (Stowell, 2022). Many of the early advances focused on birds, for example BirdNET was initially trained on 984 North American and European bird species, but has since been extended to over 6000 species globally (Kahl et al., 2021). Perch 2.0 was trained on over 14,500 different avian, mammalian, insect, and amphibian species (van Merriënboer et al., 2025). Embeddings, or learned feature representations, from these models can be used as a feature extractor for signals not in the original training data, and these embeddings can be used to train a new classifier; this is an approach called transfer learning (Pan and Yang, 2009). BirdNET transfer learning was shown to be highly effective for training new classifiers for marine mammals, anurans, and bats (Ghani et al., 2023). BirdNET transfer learning has also been used successfully in two species of gibbons, the northern grey gibbon (*Hylobates funereus*) (Clink et al. 2025) and the southern yellow-cheeked crested gibbon (*Nomascus gabriellae*) (Clink et al. 2024, 2025).

Applications of PAM for studying sound-producing primates has seen increased interest in recent years. One of the foundational applications of PAM and automated detection in primates was used to model occupancy of diana monkeys (*Cercopithecus diana*), king colobus monkeys (*Colobus polykomos*), and chimpanzees (*Pan troglodytes*) (Kalan et al., 2015) in Taï National Park, Côte d'Ivoire. Since, PAM has been used to investigate ranging behavior (Kalan et al., 2016) of chimpanzees, vocal behavior of black and gold howler monkeys (*Alouatta caraya*) (Pérez-Granados and Schuchmann, 2021), northern grey gibbons (*Hylobates funereus*) (Clink et al. 2020), and black-and-white ruffed lemurs (*Varecia variegata*) (Batist et al., 2024). PAM has also been used for occupancy modeling in northern buffed-cheeked gibbons (*Nomascus annamensis*) (Vu et al., 2023) and Yucatán black howler monkeys (*Alouatta pigra*) (Wood et al., 2023).

Tarsiers are small haplorrhine primates that are found in Southeast Asia, specifically in the Philippines, and the islands of Borneo and Sulawesi. Out of the 14 recognized tarsier species, 12 of the species are found on Sulawesi and surrounding islands (Syahrullah et al., 2023). Tarsiers on Sulawesi are unique, as they engage in coordinated duets or choruses that are audible to the human ear. The most well-studied tarsier, Gursky's spectral tarsiers (*Tarsius spectrumgurskyae*) (Gursky, 2003, 2000a, 1998; MacKinnon and MacKinnon, 1980), is currently classified as Vulnerable by the International Union for Conservation of Nature (UCN), whereas other species are considered Critically Endangered (e.g. Siau island tarsier; *T. tumpara*) Endangered (e.g. Sangihe tarsier, *T. sangirensis*) or Data Deficient (Lariang tarsier; *T. lariang*) (IUCN, 2025). PAM-based studies of tarsiers on Sulawesi have been limited and have generally focused on using PAM to study behavioral ecology. For example, PAM was used to investigate ultrasonic vocalizations in Gursky's spectral tarsiers (Gursky, 2015) and montane pygmy tarsiers (*T. pumilus*) (Grow, 2019). Two other studies used PAM to investigate vocal coordination (Clink et al. 2020) and acoustic tradeoffs (Comella et al., 2022) in Gursky's spectral tarsier duets.

In general, there have been a limited number of population surveys of mammals in central Indonesia, and a paucity of long-term population studies across the country (Ardiantiono et al., 2024). Importantly, there is a lack of clear understanding how land use change impacts tarsiers, and more data are needed to effectively assess the impact of human activities (Supriatna, 2022). Despite the potential of PAM to improve monitoring efforts and conservation interventions of tarsiers, to our knowledge there have been very few conservation-focused PAM studies of tarsiers on Sulawesi.

Our goal is to provide a case study from Tangkoko National Park to show how PAM combined with automated detection can be used to monitor tarsiers on Sulawesi. We had three main goals: 1) evaluate the performance of BirdNET transfer learning for automated detection and classification of duets from longer PAM recordings; 2) use manually verified true positive detections to explore the temporal patterns of tarsier vocal behavior; and 3) provide a simple example to demonstrate how PAM data can be used for occupancy modeling of tarsiers. Due to the considerable conservation concern for many species of tarsier, and current lack of data across Sulawesi, we propose that PAM combined with automated detection can be an effective way to fill these data gaps.

Methods

Acoustic data collection

All data collection occurred in August 2018 in Tangkoko National Park, North Sulawesi, Indonesia and focused on Gursky's spectral tarsier (hereafter tarsier). This area has marked seasonality, with distinct wet and dry seasons (Gursky, 2000b), and many different forest types, including lowland forests, submontane forests, and mossy cloud forests (MacKinnon and MacKinnon, 1980). We used two different ARUs to collect acoustic data. For the training data, we used four SwiftOne units (Koch et al., 2016) that recorded at 48 kHz and 16 bits, and were placed next to known sleep trees of tarsiers, spaced between 100 m- 250 m apart. The evaluation dataset came from a three-unit ARBIMON array (Aide et al., 2013) deployed in a triangular configuration with ~50 m spacing between August 20-28, 2018, with the closest ARU to the training dataset being ~ 300 m. The rolling 3-ARU array was intentionally deployed with ~50 m spacing to increase the probability of detecting tarsier duets on multiple ARUs. The array was moved to a new location at least 100-m from the initial location each day over nine consecutive days (see Figure 1), with all units recording at 44.1 kHz and 16-bit resolution. We decided on the evaluation dataset array design based on an older estimate of ~1 ha home range size for these tarsier (MacKinnon and MacKinnon, 1980). A circular 1-ha home range would correspond to a radius of ~56 m, suggesting that 50-m spacing would sample within a single group’s territory, and 100-m spacing would sample in neighboring tarsier territories. However, later work using radio telemetry showed that tarsier home ranges are substantially larger (1.6 –

4.1 ha) (Gursky, 1998), meaning our within-array spacing was too small to be considered independent sites. We therefore used only a subset of widely separated ARU locations (>150–200 m apart; shown in blue in Figure 1) for occupancy modeling (see below).

Figure 1. A map of the rolling ARU locations used for evaluation dataset and occupancy modeling. Three ARUs were deployed each day in a triangular shaped array, with approximately 50-m distance in between each unit, and units were moved each day from August 20-28, 2018. The points indicated in blue were included in the occupancy modeling. The inset shows the location of Tangkoko National Park in North Sulawesi, Indonesia. The map was made using QGIS v 3.34.1-Prizren (QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.org).

Model training

To create a training dataset, we used a band-limited energy detector implemented in the ‘gibbonR’ package v1.0.1 (Clink and Klinck 2019). We extracted sound events between 5 and 15 kHz from the training data ARU recordings; see Figure 2 for representative spectrograms. We set the noise quantile threshold to 0.75 and limited signal durations to between 6-s to 24-s. This function exports sound events to waveform audio files (.wav). We used the review tab in the BirdNET GUI v2.4.0 to manually assign sound events to one of two classes: ‘tarsier’ or ‘noise’. This resulted in 30 tarsier clips and 217 noise clips for training. We then used the command line interface (CLI) to train a custom classifier using BirdNET v2.4, setting fmin = 5000 Hz, fmax = 15000 Hz, and crop_mode = 'segments', which divides longer clips into 3-s clips for training (Kahl et al., 2021). All other settings were default.

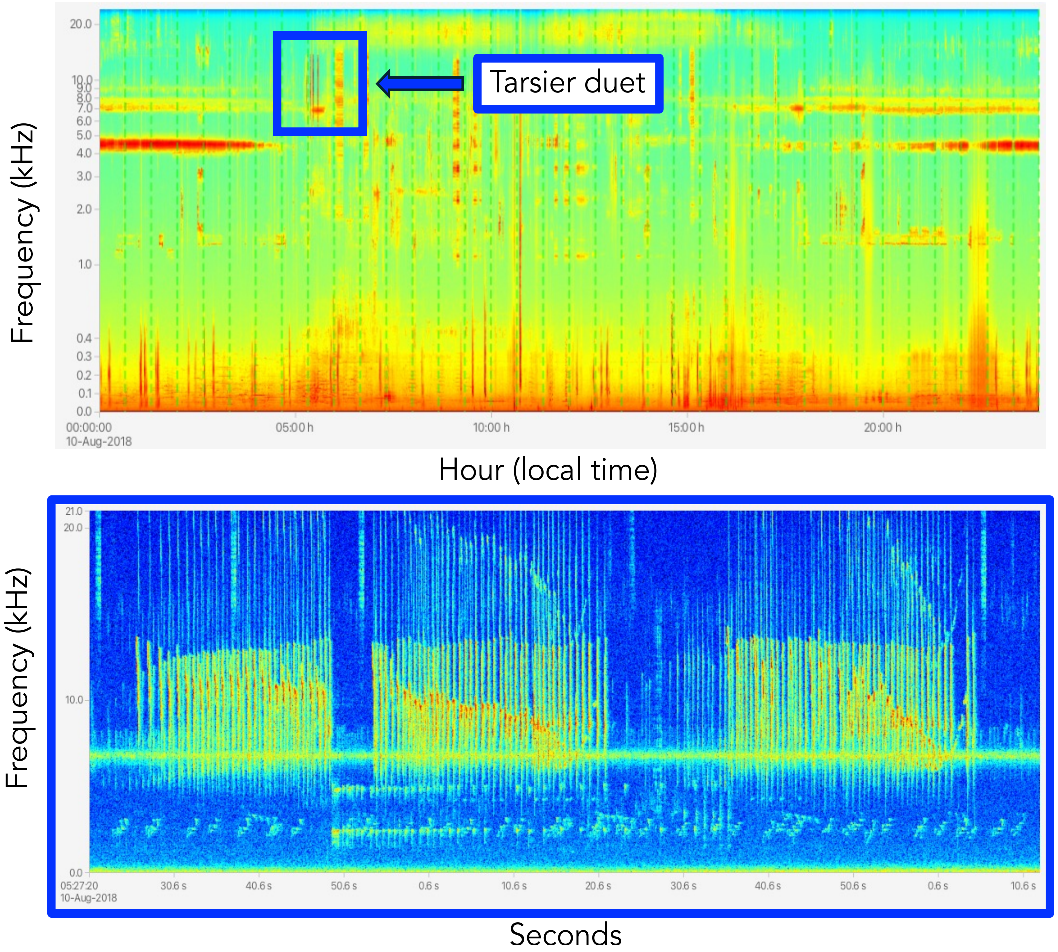

Figure 2. Long-term spectral average (LTSA) of 24-hrs of recording indicating the presence of a tarsier duet (top) and a representative spectrogram of a tarsier duet (bottom). Both made using Raven Expedition v1.2 (K. Lisa Yang Center for Conservation Bioacoustics, 2025).

To evaluate the performance of BirdNET and determine the optimal confidence threshold, we deployed the model over the full evaluation dataset (~ 523 hours of data) and used the 'segments' function in BirdNET to export 100 segments randomly selected over a range of confidence scores (0.1 – 1.0). We then used the review tab in the BirdNET GUI to manually assign detections into two categories: true or false positive. The review tab also calculates a logistic regression, with the outcome variable being whether the detection was a true or false positive, and the predictor variable being BirdNET confidence score. Confidence scores that are output by models such as BirdNET are not probabilities, however the use of a logistic regression can convert the scores to probability that the detection is a true positive (Wood and Kahl, 2024). This approach can provide insight into the precision of the model and also help decide on an optimal confidence score to use. Based on the results of this analysis, we found that precision at higher confidence scores was very high, we used a confidence score of 0.65 which translates to ~ 95% probability that the detection is a true positive (Figure 3).

To further estimate performance of the automated detection system, we manually annotated six hours of recordings taken between 05:00-06:00 WITA from six different ARU locations using RavenPro v1.6 (K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology, 2023). We annotated every instance of a tarsier duet bout. We then used the 'review' tab in the BirdNET GUI to calculate area under the receiver operating characteristic curve (AUC-ROC; which summarizes the tradeoff between true positive and false positive rates across all confidence thresholds), precision (proportion of detected clips that were true positives), recall (proportion of true positive clips detected), and F1 score (harmonic mean of precision and recall). Precision, recall and F1 score all depend on a confidence score, so we used 0.65. These metrics were calculated at the 3-s clip level, as this version of BirdNET makes predictions on 3-s clips.

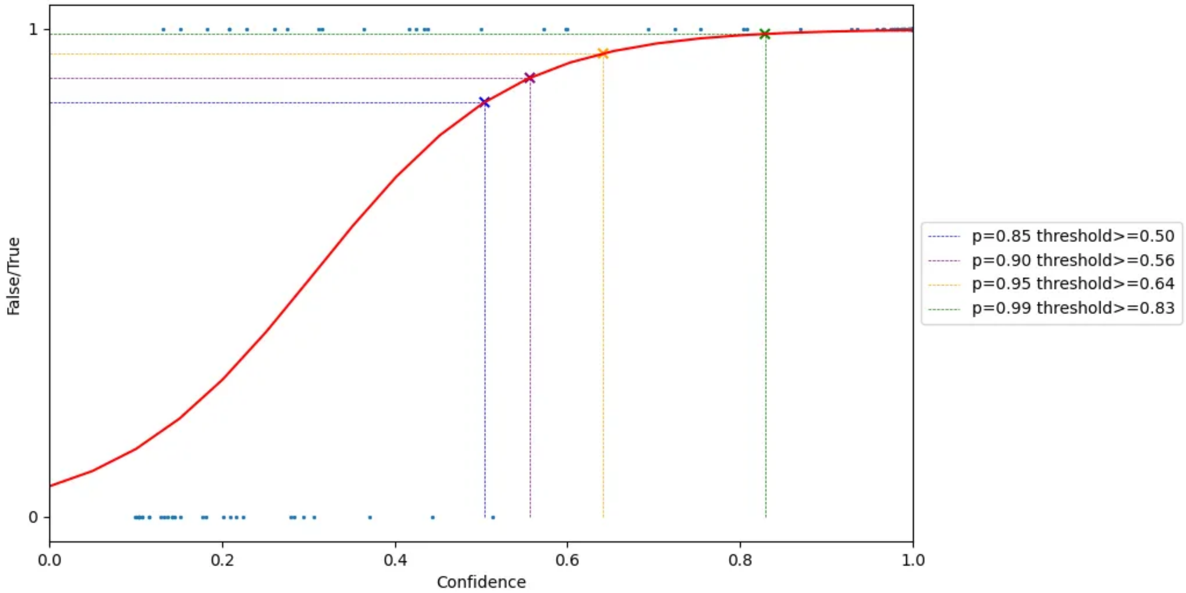

Figure 3. Logistic regression with true or false positive as the outcome, and confidence score as the predictor. The colored vertical lines indicate the confidence score (or threshold) associated with 0.85 (blue), 0.90 (purple), 0.95 (yellow), and 0.99 (green) probability that the detection is a true positive. Figure was created in the BirdNET GUI v2.4.0.

Occupancy modeling

Rolling 3-ARU arrays frequently detected the same events (see results below), confirming that ARUs in the triangular array were not independent sampling locations. However, each 3-ARU array was moved to a new location each day, and no two 3-ARU array recorded simultaneously. To obtain spatially independent sites for occupancy modeling, we selected one ARU from each of the nine daily 3-ARU array deployments, ensuring that units were >150 m apart. This follows recent guidance on occupancy study design from camera traps, where simulation studies suggest that distance between detectors should be at least 0.67 times the diameter of the largest reported home ranges (Fuller et al., 2022).

For tarsiers, the largest reported home range is 4.1 ha (Gursky, 1998), which corresponds to a circular home-range radius of approximately 114 m (r = $\sqrt{A}/\pi$ = $\sqrt{41{,}000\ m^2}/\pi$ = 114.2 m) and a diameter of ~228 m; therefore, 0.67 × diameter ≈ 152 m, supporting our use of ≥150–200 m spacing to approximate spatial independence. We acknowledge that our study design is not optimal for occupancy modeling due to the limited spatial and temporal coverage but provide this as a proof of concept to show how PAM combined with automated detection can be used for occupancy modeling of tarsiers in the future.

We considered between 05:00 - 06:00 WITA as our survey period. To create multiple survey occasions required by occupancy models, we divided the 1-hr period into shorter time intervals. Tarsier duet bouts tend to be relatively long with durations ranging from 60-s to 251-s (median 99-s; n=13; this study), so to avoid the potential for consecutive windows containing the same tarsier duet, we compared survey periods of 10, 15, and 20 minutes. We created detection histories based on the manually verified true positive tarsier detections for each of the three different survey period lengths (see example in Table 1).

Table 1. Example detection history for each 20-min survey period and ARU location. Detection history was created using the manually verified true positive tarsier detections.

	Survey 1	Survey 2	Survey 3
TPAM1	1	1	0
TPAM4	0	1	0
TPAM7	0	0	0
TPAM11	1	1	1
TPAM14	1	1	0
TPAM17	0	0	0
TPAM21	0	1	0
TPAM24	1	1	0
TPAM27	1	1	0

We used the ‘unmarked’ R package (Fiske and Chandler, 2011; Kellner et al., 2023) to create three single-season occupancy models with the ‘occu’ function using detection histories created using survey periods of 10, 15, and 20 minutes. Occupancy (ψ) is defined as the probability that a site is occupied, and detection probability (p) is the probability of detecting the species if it is present (MacKenzie et al., 2002). We specified intercept-only models for both occupancy (ψ) and detection probability (p) components, which treats detection probability and occupancy as constants across sites and surveys. We did not include environmental covariates as there was minimal variation across the ~ 122,500 m² (12.25 ha) study area. We could not use the ‘fitList’ function from the ‘unmarked’ R package for model selection, as this requires each model to have the same underlying data structure, and this was not the case since we used varying survey period lengths for each of our models. However, we were able to extract each Akaike Information Criterion (AIC) values manually and calculated the delta AIC by subtracting all AIC values by the minimum AIC value. We used the ‘plogis’ function in base R to convert the logit-scale values for occupancy (ψ) and detection (p) components into a probability between 0 and 1. All occupancy modeling analyses were done using R version 4.4.0 (R Core Team, 2024).

Data and code availability

All scripts needed to recreate the analyses are openly available on GitHub (https://github.com/DenaJGibbon/tarsier-automated-detection). Training data, manually verified true and false positive tarsier detections, and the trained BirdNET model are available on Zenodo (https://doi.org/10.5281/zenodo.18483496).

Results

Across the full evaluation dataset, we had 1825 tarsier detections that were above a confidence score of 0.65, of which 1809 were true positive and 16 were false positive, resulting in a precision of ~0.99. Based on six hours of manually annotated recordings, model performance at the 3-s clip level was satisfactory, with an AUC-ROC of 0.86, high precision (0.96), moderate recall (0.74), and a balanced F1 score of 0.81.

The majority of tarsier vocalizations occurred between 05:00-06:00 (Waktu Indonesia Tengah, WITA) time. Sunrise occurred at approximately 05:37 WITA during our study period. We also found some tarsier detections around 03:00 WITA and after sunset which occurred around 17:46 WITA (Figure 4). Visual inspection of spectrograms of these events that occurred outside of sunrise show that they were tarsier vocalizations but did not follow the typical duet structure.

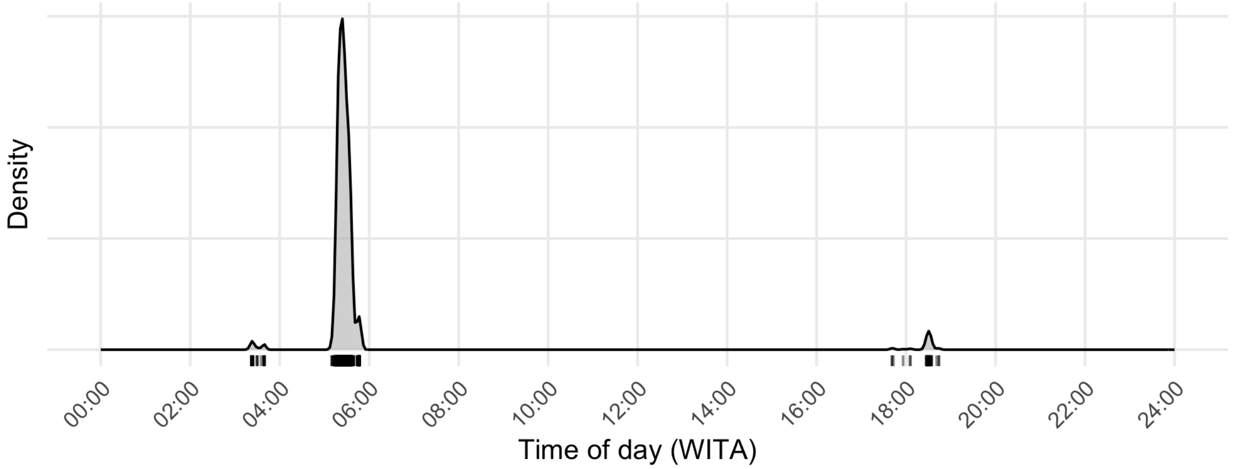

Figure 4. Density of manually verified true positive tarsier detections in Tangkoko National Park as a function of time of day (Waktu Indonesia Tengah, WITA).

We also found that it was highly variable whether tarsiers were detected at the same date and time across the rolling 3-ARU arrays or not. In some cases, all three ARUs had tarsier detections at the same time, however in some cases only one ARU had detections. We found that at least one ARU had tarsier detections each day of deployment (see Figure 5).

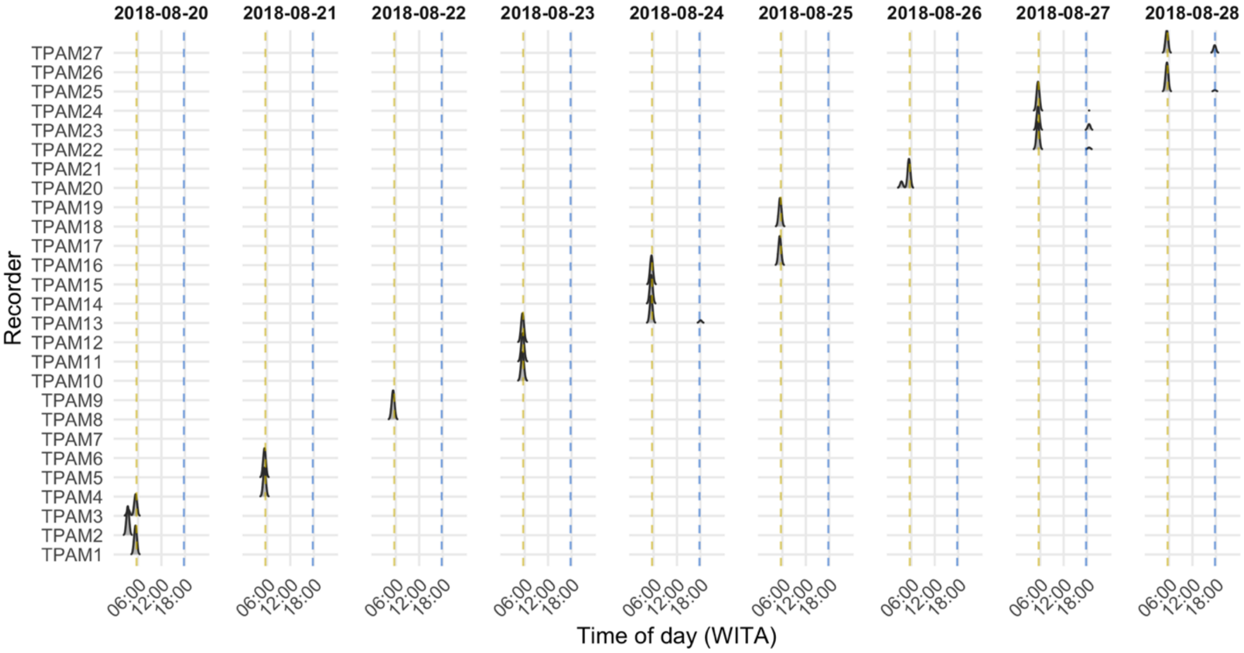

Figure 5. Density plots showing tarsier detections by date and time for each 3-unit rolling deployment array. The dashed gold line indicates approximates sunrise time and the dashed blue line indicates approximate sunset time for each survey day.

For occupancy modeling, we found that the model with the longest survey period (20-min) had the lowest AIC, indicating it was the best supported model (Table 2). We found that, based on the top supported model, occupancy (ψ)was 0.85 (95% CI: 0.32–0.98) and detection probability (p) was 0.57 (95% CI: 0.32–0.79). Note that the estimates are similar across all survey duration periods, with very wide confidence intervals.

Table 2. Estimated occupancy (ψ), detection probability (p), AIC, and ΔAIC for three survey durations. Values in parentheses are 95% confidence intervals.

Survey duration	Occupancy (ψ)	Detection (p)	AIC	ΔAIC
10-min	0.84 (0.34–0.98)	0.35 (0.21–0.53)	68.4	27.9
15-min	0.94 (0.01–1.00)	0.35 (0.17–0.59)	49.8	9.3
20-min	0.85 (0.32–0.98)	0.57 (0.32–0.79)	40.5	0

Discussion

Here, we illustrate how PAM can be combined with automated detection to provide insights into temporal and spatial patterns of calling in tarsiers and generate detection histories that can be used for occupancy modeling. Similar to previous work (MacKinnon and

MacKinnon, 1980), we found that the majority of tarsier vocalizations occurred around sunrise, however there were a few that occurred around 03:00 and 18:00 WITA. Our rolling 3-ARU arrays frequently detected the same tarsier vocalization events, with differences across detections and ARUs presumably due to tarsier calling location within the 3-ARU array. Our simple example of occupancy indicated relatively high occupancy (~0.85) and moderate detection probability (~0.57), however these results should be interpreted with caution due to the small number of ARU locations used for this analysis. Given the data deficient status of many tarsier species across Sulawesi, and the effectiveness of this PAM and automated detection workflow, we advocate for an island-wide acoustic census of tarsiers on Sulawesi. Importantly, incorporating studies across habitat- and land use-types will provide insights into how human activities impact tarsier populations (Supriatna, 2022).

The performance of automated detectors depends on many different aspects including the amount of background noise, presence of other signals that are similar to the target signal, stereotypy of the target signal, quality and diversity of training data, and parameters of the underlying model (Kershenbaum et al., 2025). Performance also depends heavily on the evaluation or test dataset (Clink et al. 2024). Previous work in southern yellow-cheeked crested gibbons found that BirdNET performance stabilized when the number of positive training samples was > 25 (Clink et al. 2024). Here, we found that the BirdNET model trained on 30 high signal-to-noise ratio positive samples obtained using a band-limited energy detector led to satisfactory performance, with very high precision over the evaluation dataset (~0.99). We also found that at the 3-s clip level recall was satisfactory, however for occupancy modeling, detection histories are often created at the survey level, which means that only a single detected clip would be necessary to verify presence within the survey window.

When deploying an automated detector there is an inherent tradeoff between precision and recall, that is often mediated by the choice of a confidence score. We used a logistic regression to determine the confidence score (0.65) that was associated with a 99% probability that the detection was a true positive (Wood and Kahl, 2024). In our case, the full evaluation data set consisted of approximately 520 hours, and we only had sixteen false positive detections. For larger deployments it is possible that a higher confidence score would be desirable if deploying over many hundreds or thousands of hours of recordings resulted in a very large number of false positives. Importantly, the ideal confidence score will vary for deployments at different sites, or using different recording units, which would require confidence score tuning similar to our approach.

Our preliminary analyses of occupancy and detection probability are consistent with the fact that tarsier populations are known to occur at high density in Tangkoko National Park (Gursky, 1998) and that they call consistently most mornings around sunrise (MacKinnon and MacKinnon, 1980). Even within Tangkoko National Park there are many different forest types (MacKinnon and MacKinnon, 1980); and it is unclear how tarsier populations vary across these forest types. Tangkoko National Park has been impacted heavily by selective logging and human encroachment on the borders, and is an active ecotourism site, with tarsiers in heavily touristed areas showing modified behaviors, including emitting more audible and ultrasonic alarm calls (Gursky, 2022). Therefore, future work comparing tarsier occupancy estimates in different landuse types, forest types, and gradients of ecotourism will be highly informative.

Sulawesi tarsiers are ideal candidates for future PAM + occupancy studies for this approach as they are highly vocal and territorial, and we provide the following suggestions for future work. To determine the optimal number of ARUs to answer the question of interest, e.g. how does occupancy vary across landuse types, or to detect a change in populations over time due to shifts in management practices, we suggest simulation-based studies to determine optimal number of ARUs (Fuller et al., 2022; Wood et al., 2019). Regarding duration or number of surveys, this has varied substantially across different PAM studies, for example 5- to 7- night deployments were used for owl species in the Sierra Nevada, California, USA (Wood et al., 2019), whereas three consecutive mornings were used for occupancy of northern yellow-cheeked crested gibbon (*N. annamensis*) using human observers (Vu et al., 2020) and ARUs (Vu and Tran, 2019). One of the benefits of using PAM over human observers is the ability to collect data more easily across larger spatial and temporal scales, so we advocate for longer surveys, if possible, and empirical tests on the impact of survey length on occupancy results. It is highly possible that detection probability will be lower during the rainy season, due to either behavioral modifications (Gursky, 2000b) and/or reduced performance of the automated detector under noisy conditions, so seasonal differences must be taken into account. It is also unclear if the diel patterns of vocalizations, with the majority occurring around sunrise, is generalizable across all species of Sulawesi tarsiers. Therefore, pilot studies investigating temporal patterns of vocal behavior will be needed for understudied species, but if vocalizations do occur mostly around sunrise, limiting acoustic data collection to a few hours around sunrise could help with data storage and processing limitations.

Given the data deficient status of many tarsier species on Sulawesi, a need for a better understanding of how human activity impacts tarsiers (Supriatna, 2022), and a general lack of long-term surveys across Indonesia (Ardiantiono et al., 2024), we propose that occupancy modeling using PAM and automated detection can be an effective way to fill some of these knowledge gaps. Advances in technology for automated detection of acoustic data, like BirdNET and the associated code-free GUI (Kahl et al., 2021), have made these approaches more accessible to conservation practitioners. Importantly, some barriers still remain to wide scale adoption of PAM in Sulawesi and beyond, including access to sufficient computing power to be able to process large amounts of acoustic data, and the need to build local capacity for using these tools (Speaker et al., 2022). We hope that our work inspires practitioners to work towards the conservation of Sulawesi's tarsiers through coordinated, long-term, and locally led acoustic monitoring efforts.

Acknowledgements

We thank Holger Klinck for his unwavering support and never-ending enthusiasm for tarsiers. We thank Diana Parker for loaning us her ARBIMON recorders to use for this project. We thank Vandem Tundu for his assistance with field data collection. Funding for this project came from the Fulbright ASEAN Research Award for U.S. Scholars (no award number given). We also acknowledge Astrid Lim and other staff at the American Indonesian Exchange Foundation for their assistance in obtaining necessary permits and permissions.

Compliance with Ethical Standards

All data were collected noninvasively; however institutional approval was provided by the Cornell University (IACUC 2017-0098). Research was conducted under an Indonesian foreign research permit issued by the Ministry of Research, Technology, and Higher Education (permit no. 2881/FRP/E5/Dit.KI/VII/2018).

References

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 1, e103. https://doi.org/10.7717/peerj.103

Ardiantiono, Pinondang, I.M.R., Chandradewi, D.S., Semiadi, G., Pattiselanno, F., Supriatna, J., Tasirin, J.S., Winarni, N.L., Voigt, M., Bull, J.W., Humle, T., Deere, N.J., Struebig, M.J., 2024. Insights from 20 years of mammal population research in Indonesia. Oryx 58, 485–492. https://doi.org/10.1017/S0030605323001539

Batist, C.H., Dufourq, E., Jeantet, L., Razafindraibe, M.N., Randriamanantena, F., Baden, A.L., 2024. An integrated passive acoustic monitoring and deep learning pipeline for black-and-white ruffed lemurs (Varecia variegata) in Ranomafana National Park, Madagascar. American Journal of Primatology 86, e23599.

Brunk, K.M., Goldberg, J.F., Maxwell, C., Peery, M.Z., Jones, G.M., Gallagher, L.R., Kramer, H.A., Westerling, A.L., Keane, J.J., Kahl, S., Wood, C.M., 2025. Bioregional-scale acoustic monitoring can support fire-prone forest restoration planning. Frontiers in Ecology and the Environment n/a, e2843. https://doi.org/10.1002/fee.2843

Clink, Dena Jane, Cross-Jaya, H., Kim, J., Ahmad, A.H., Hong, M., Sala, R., Birot, H., Agger, C., Vu, T.T., Thi, H.N., Chi, T.N., Klinck, H., 2024. Benchmarking for the automated detection and classification of southern yellow-cheeked crested gibbon calls from passive acoustic monitoring data. bioRxiv. https://doi.org/10.1101/2024.08.17.608420

Clink, Dena J., Cross-Jaya, H., Kim, J., Ahmad, A.H., Hong, M., Sala, R., Birot, H., Agger, C., Vu, T.T., Thi, H.N., Chi, T.N., Klinck, H., 2024. Benchmarking automated detection and classification approaches for monitoring of endangered species: a case study on gibbons from Cambodia. https://doi.org/10.1101/2024.08.17.608420

Clink, Dena J., Hamid Ahmad, A., Klinck, H., 2020. Gibbons aren’t singing in the rain: presence and amount of rainfall influences ape calling behavior in Sabah, Malaysia. Scientific Reports 10, 1282. https://doi.org/10.1038/s41598-020-57976-x

Clink, D.J., Kim, J., Cross-Jaya, H., Ahmad, A.H., Hong, M., Sala, R., Birot, H., Agger, C., Vu, T.T., Thi, H.N., Chi, T.N., Klinck, H., 2025. Automated Detection of Gibbon Calls From Passive Acoustic Monitoring Data Using Convolutional Neural Networks in the “Torch for R” Ecosystem. Ecology and Evolution 15, e71678. https://doi.org/10.1002/ece3.71678

Clink, D.J., Klinck, H., 2019. GIBBONR: An R package for the detection and classification of acoustic signals using machine learning. arXiv preprint arXiv:1906.02572.

Clink, Dena J, Tasirin, J.S., Klinck, H., 2020. Vocal individuality and rhythm in male and female duet contributions of a nonhuman primate. Current zoology 66, 173–186.

Comella, I., Tasirin, J.S., Klinck, H., Johnson, L.M., Clink, D.J., 2022. Investigating note repertoires and acoustic tradeoffs in the duet contributions of a basal haplorrhine primate. Frontiers in Ecology and Evolution 10, 910121.

Fiske, I., Chandler, R., 2011. unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. Journal of Statistical Software 43, 1–23.

Fuller, A.K., Augustine, B.C., Morin, D.J., Pigeon, K., Boulanger, J., Lee, D.C., Bisi, F., Garshelis, D.L., 2022. The occupancy-abundance relationship and sampling designs using occupancy to monitor populations of Asian bears. Global Ecology and Conservation 35, e02075. https://doi.org/10.1016/j.gecco.2022.e02075

Ghani, B., Denton, T., Kahl, S., Klinck, H., 2023. Global birdsong embeddings enable superior transfer learning for bioacoustic classification. Scientific Reports 13, 22876.

Grow, N.B., 2019. Cryptic Communication in a Montane Nocturnal Haplorhine, Tarsius pumilus. Folia Primatologica 90, 404–421. https://doi.org/10.1159/000497427

Gursky, S., 2015. Ultrasonic vocalizations by the spectral Tarsier, *Tarsius spectrum*. Folia Primatologica 86, 153–163. https://doi.org/10.1159/000371885

Gursky, S., 2003. The behavioral ecology of the spectral tarsier, *Tarsius spectrum*. Evolutionary Anthropology: Issues, News, and Reviews 11, 226–234. https://doi.org/10.1002/evan.10035

Gursky, S., 2000a. Sociality in the spectral tarsier, *Tarsius spectrum*. American Journal of Primatology 51, 89–101. https://doi.org/10.1002/(SICI)1098-2345(200005)51:1%253C89::AID-AJP7%253E3.0.CO;2-7

Gursky, S., 2000b. Effect of Seasonality on the Behavior of an Insectivorous Primate, Tarsius spectrum. International Journal of Primatology 21, 477–495. https://doi.org/10.1023/A:1005444020059

Gursky, S., 1998. Conservation Status of the Spectral Tarsier *Tarsier spectrum*: Population Density and Home Range Size. Folia Primatologica 69, 191–203. https://doi.org/10.1159/000052713

Gursky, S.L., 2022. Primate, *Tarsius spectrum*, in Indonesia, in: Gursky, S.L., Supriatna, J., Achorn, A. (Eds.), Ecotourism and Indonesia’s Primates. Springer, Cham, Switzerland, pp. 81-?

IUCN, 2025. The IUCN Red List of Threatened Species.

K. Lisa Yang Center for Conservation Bioacoustics, 2025. Raven Expedition: Interactive Sound Analysis Software.

K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology, 2023. Raven Pro: Interactive Sound Analysis Software (Version 1.6.4) [Computer software].

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

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

Kalan, A.K., Piel, A.K., Mundry, R., Wittig, R.M., Boesch, C., Kühl, H.S., 2016. Passive acoustic monitoring reveals group ranging and territory use: a case study of wild chimpanzees (*Pan troglodytes*). Frontiers in Zoology 13, 34. https://doi.org/10.1186/s12983-016-0167-8

Katsis, L.K.D., Hill, A.P., Piña-Covarrubias, E., Prince, P., Rogers, A., Patrick Doncaster, C., Snaddon, J.L., 2022. Automated detection of gunshots in tropical forests using convolutional neural networks. Ecological Indicators 141, 109128. https://doi.org/10.1016/j.ecolind.2022.109128

Kellner, K.F., Smith, A.D., Royle, J.A., Kery, M., Belant, J.L., Chandler, R.B., 2023. The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology. Methods in Ecology and Evolution 14, 1408–1415. https://doi.org/10.1111/2041-210X.14123

Kershenbaum, A., Akçay, Ç., Babu-Saheer, L., Barnhill, A., Best, P., Cauzinille, J., Clink, D., Dassow, A., Dufourq, E., Growcott, J., Markham, A., Marti-Domken, B., Marxer, R., Muir, J., Reynolds, S., Root-Gutteridge, H., Sadhukhan, S., Schindler, L., Smith, B.R., Stowell, D., Wascher, C.A.F., Dunn, J.C., 2025. Automatic detection for bioacoustic research: a practical guide from and for biologists and computer scientists. Biological Reviews 100, 620–646. https://doi.org/10.1111/brv.13155

Koch, R., Raymond, M., Wrege, P., Klinck, H., 2016. SWIFT: A small, low-cost acoustic recorder for terrestrial wildlife monitoring applications, in: North American Ornithological Conference. Washington, D.C., p. 619.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83, 2248–2255.

MacKinnon, J., MacKinnon, K., 1980. The behavior of wild spectral tarsiers. International Journal of Primatology 1, 361–379. https://doi.org/10.1007/BF02692280

Pan, S.J., Yang, Q., 2009. A survey on transfer learning. IEEE Transactions on knowledge and data engineering 22, 1345–1359.

Pérez-Granados, C., Schuchmann, K.-L., 2021. Passive acoustic monitoring of the diel and annual vocal behavior of the Black and Gold Howler Monkey. American Journal of Primatology 83, e23241.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramesh, V., Hariharan, P., Akshay, V., Choksi, P., Khanwilkar, S., DeFries, R., Robin, V., 2023. Using passive acoustic monitoring to examine the impacts of ecological restoration on faunal biodiversity in the Western Ghats. Biological Conservation 282, 110071.

Speaker, T., O'Donnell, S., Wittemyer, G., Bruyere, B., Loucks, C., Dancer, A., Carter, M., Fegraus, E., Palmer, J., Warren, E., Solomon, J., 2022. A global community-sourced assessment of the state of conservation technology. Conservation Biology 36, e13871. https://doi.org/10.1111/cobi.13871

Stowell, D., 2022. Computational bioacoustics with deep learning: a review and roadmap. PeerJ 10, e13152. https://doi.org/10.7717/peerj.13152

Sugai, L.S.M., Silva, T.S.F., Ribeiro, J.W., Llusia, D., 2019. Terrestrial Passive Acoustic Monitoring: Review and Perspectives. BioScience 69, 15–25. https://doi.org/10.1093/biosci/biy147

Supriatna, J., 2022. Tarsidae, in: Field Guide to the Primates of Indonesia. Springer Nature.

Syahrullah, F.N., Maddus, U., Mustari, A.H., Gursky, S., Indrawan, M., 2023. Distribution and abundance of Peleng Tarsier (Tarsius pelengensis) in Banggai Island group, Indonesia. Sci Rep 13, 11445. https://doi.org/10.1038/s41598-023-30049-5

van Merriënboer, B., Dumoulin, V., Hamer, J., Harrell, L., Burns, A., Denton, T., 2025. Perch 2.0: The bittern lesson for bioacoustics. arXiv preprint arXiv:2508.04665.

Vu, T.T., Doherty, P.F., T. Nguyen, H., Clink, D.J., Nguyen, M.D., Dong, H.T., Cheyne, S.M., Giang, T.T., Phan, D.V., Ta, N.T., Tran, D.V., 2023. Passive acoustic monitoring using smartphones reveals an alarming gibbon decline in a protected area in the central Annamite Mountains, Vietnam. American Journal of Primatology 85, e23544. https://doi.org/10.1002/ajp.23544

Vu, T.T., Hoa Anh, N.Q., Rawson, B.M., Tran, D.V., Nguyen, H.T., Van, T.N., 2020. Monitoring occurrence, extinction, and colonization probabilities for gibbon populations. American Journal of Primatology 82, e23171.

Vu, T.T., Tran, L.M., 2019. An Application of Autonomous Recorders for Gibbon Monitoring. International Journal of Primatology 40, 169–186. https://doi.org/10.1007/s10764-018-0073-3

Wood, C.M., Barceinas Cruz, A., Kahl, S., 2023. Pairing a user-friendly machine-learning animal sound detector with passive acoustic surveys for occupancy modeling of an endangered primate. American Journal of Primatology 85, e23507.

Wood, C.M., Kahl, S., 2024. Guidelines for appropriate use of BirdNET scores and other detector outputs. Journal of Ornithology 1–6.

Wood, C.M., Popescu, V.D., Klinck, H., Keane, J.J., Gutiérrez, R., Sawyer, S.C., Peery, M.Z., 2019. Detecting small changes in populations at landscape scales: A bioacoustic site-occupancy framework. Ecological Indicators 98, 492–507.

Wrege, P.H., Rowland, E.D., Keen, S., Shiu, Y., 2017. Acoustic monitoring for conservation in tropical forests: examples from forest elephants. Methods in Ecology and Evolution 8, 1292–1301. https://doi.org/10.1111/2041-210X.12730