Taylor & Francis
Taylor & Francis Group

Check for updates

Passive acoustic surveys and the BirdNET algorithm reveal detailed spatiotemporal variation in the vocal activity of two anurans

Connor M. Wood [a], Stefan Kahl[a], Stephanie Barnes[b], Rachel Van Horne[c] and Cathy Brown[d]

[a]K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA; [b]USDA Forest Service, Sierra National Forest, Prather, CA, USA; [c]USDA Forest Service, Humboldt-Toiyabe National Forest, Sparks, NV, USA; [d]USDA Forest Service, Stanislaus National Forest, Sonora, CA, USA

ABSTRACT

Passive acoustic monitoring has proven effective for broad-scale population surveys of acoustically active species, making it a valuable tool for conserving threatened species. However, successful automated classification of anuran vocalisations in large audio datasets has been limited. We deployed five autonomous recording units at three known breeding areas of the Yosemite toad (*Anaxyrus canorus*), which is threatened and relatively uncommon, and the sympatric Pacific chorus frog (*Pseudacris regilla*), which is widespread and more common, to test the viability of bioacoustics as a means of supplementing ongoing, human survey efforts. We analysed the audio data with the BirdNET algorithm, which was originally developed for birds but has been expanded to include both species. We achieved efficient and accurate identification of both species in 2,756 h of audio, which yielded high-resolution phenological data about seasonal and daily vocal activity as well as daily detection counts. These findings demonstrate that a newly expanded machine learning detector, BirdNET, can effectively process passive acoustic surveys for these species. Further exploration of how passive acoustic monitoring may complement existing survey techniques for these and other Anurans is warranted.

Introduction

Amphibians are globally threatened, with many species declining precipitously (McCallum 2007; Sodhi et al. 2008; Grant et al. 2020). There is thus an urgent need to identify, understand, and address drivers of population decline. Yet the cryptic nature of many amphibians has posed a persistent challenge to population studies. Bioacoustics has emerged as a powerful survey tool for vocally active species, enabling researchers to leverage occupancy-oriented surveys of threatened species to study their behaviour and interspecific interactions, and even differentiate individuals (Campos-Cerqueira et al. 2016; Wood et al. 2020; Bolitho et al. 2021). Vocal activity is widespread among anurans and bioacoustic approaches are increasingly being used for amphibian conservation

---

research (Measey et al. 2017; Wijayathilaka et al. 2018; Bolitho et al. 2021). However, the development of tools to facilitate efficient processing of passive acoustic survey data via automated detectors has been limited but see (Noda et al. 2016; Lapp et al. 2021; Akbal et al. 2023), in part because the development of species-specific detectors is very time-consuming and performance is sensitive to detector settings (Crump and Houlahan 2017). The expansion of easy-to-use animal sound detectors could facilitate the growth of bioacoustics as a tool for anuran research and conservation.

A truly scalable acoustic survey approach low-cost autonomous recording units (ARUs) and a computer-assisted signal detection and classification tool to determine where and when a species of interest has (or has not) been vocally active in an audio dataset that could contain thousands or tens of thousands of hours. Aurans are typically vocally active when breeding, and ARUs can be deployed before the breeding period begins and programmed to record continuously or at designated intervals. Recording many thousands of hours of audio is relatively easy; efficiently determining whether the vocalisation of a species of interest has been recorded in a large quantity of audio is challenging. Manual review of audio is time-consuming and quickly becomes impossible to conduct at scale. Custom detectors can be developed using spectrogram cross-correlation (e.g. the template detector in Raven Pro 2.0 (Arvind et al. 2022)), Hidden Markov Models (e.g. Sound Scope (Crump and Houlahan 2017)), or with amplitude-based pulse rate detectors (Lapp et al. 2021). However, cross-correlation tools and Hidden Markov Models can be extremely sensitive to both recording settings (e.g. sample rate) and tool settings (Crump and Houlahan 2017), while pulse rate detectors rely on relative metrics that are unique to each ARU and thus limit the scalability of the tool. Machine learning detectors are gaining traction (e.g. (Noda et al. 2016; Akbal et al. 2023)), and one such tool, the BirdNET algorithm (Kahl et al. 2021), requires no computer science expertise, is freely available, and has recently been expanded to identify some amphibians by sound.

We conducted a small-scale test of whether passive acoustic surveys and BirdNET sound analysis could be combined to efficiently identify two vocally active anurans, the US federally threatened Yosemite toad (*Anaxyrus canorus*) and the more common Pacific chorus frog (*Pseudacris regilla* or *P. sierra*; see (Recuero et al. 2006; Barrow et al. 2014)). The Yosemite toad is endemic to high elevations (>1,980 m asl) in the central Sierra Nevada, USA, whereas the sympatric Pacific chorus frog is more widely distributed throughout the western USA and southwestern Canada. In the Sierra Nevada, both species are explosive breeders that emerge at snowmelt when adult males form breeding choruses for relatively short periods (~2 weeks). Both species can readily be identified by sound despite nearly simultaneous breeding (Figure 1(a)). After breeding, adults disperse into the surrounding uplands (Liang 2013). Landscape genetic surveys have provided data about effective population size (e.g. (Wang 2012; Maier et al. 2022)), and the inaccessibility of most breeding sites when adults are present has forced bioregional monitoring efforts to rely on post-breeding occupancy surveys for tadpoles and metamorphs conducted over a 6–8-week period when breeding sites are accessible (Brown et al. 2012; Brown and Olsen 2013). Thus, little information is known about annual changes in adult populations across the toad's range. Passive acoustic surveys could provide information about breeding phenology and potentially adult male population sizes (only males are vocally active), but the viability of bioacoustic monitoring has not been evaluated. Therefore, we deployed
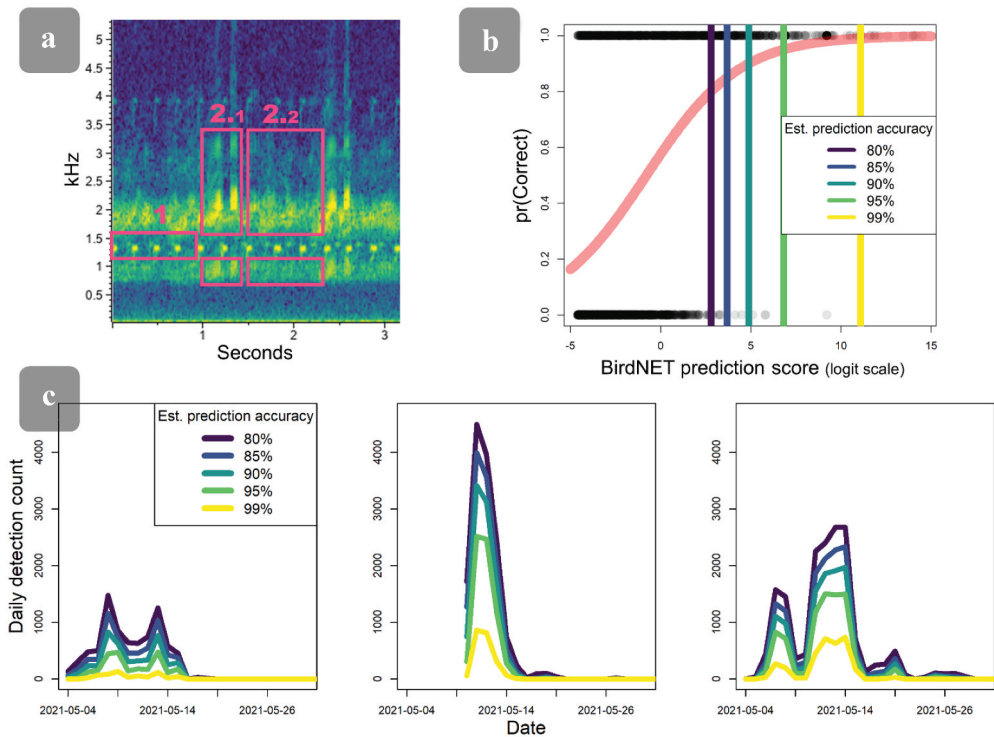
Figure 1. After scanning 2,756 hours of passively recorded audio with the BirdNET algorithm, we manually validated randomly selected BirdNET predictions of Yosemite toads and Pacific chorus frogs (a); related prediction outcome (correct/incorrect) to prediction score via logistic regression to set probabilistic score thresholds (b); and tested the sensitivity to patterns in daily vocalisation counts to different score thresholds (c). In the spectrogram (a), a foreground Yosemite toad is shown in box 1 and a second, more distant individual begins calling at about 1.6 seconds; these exemplify the 1–5-second trills which adult males produce during the breeding season. A foreground Pacific chorus frog is shown in box 2.1; adult males will spend several hours producing these brief calls, which collectively form the background chorus of many individuals is shown in box 2.2. BirdNET correctly identified both species in this audio clip. In the logistic regression plot (b), the logistic curve relating BirdNET score to prediction outcome is shown in red. Each panel in the bottom row (c) represents data from a recording unit deployed at a 9.9 ha breeding area (Site 1).

five ARUs across three sites known to be occupied by Yosemite toads and Pacific chorus frogs to evaluate: 1) the efficacy of passive acoustic surveys and the newly expanded BirdNET algorithm as a survey method for sympatric amphibian species, and 2) the fine-scale spatiotemporal variation in both species' vocal activity. If passive acoustic monitoring and emerging machine learning tools yield reliable estimates of vocal activity, they could support anuran research and conservation for both at-risk and common species.

Materials and methods

Field methods

We deployed five ARUs (SwiftOne Recorder, rugged edition, K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, New York, USA) at three known

breeding areas of Yosemite toads on lands managed by the Sierra, Humboldt-Toiyabe, and Stanislaus National Forests (Sites 1, 2, and 3 respectively; Table 1). The breeding areas were 25–150 km apart; previous surveys suggest Site 1 had 20–25 adult male Yosemite toads, while the other two sites had 5–15. At each breeding area, we chose the deployment locations based on our knowledge of where Yosemite toad breeding had occurred in the previous year; Site 1 had three ARUs, which were 190–360 m apart (Table 2). Initial deployments were made on March 30 – 2 April 2021, and follow-up visits were made in early May to resolve a minor glitch in the ARU recording schedules.

The Yosemite toad was the priority species, rather than the Pacific chorus frog, so we designed our recording schedule to focus on that species. The adult male Yosemite toad calls diurnally during their two-week breeding season so we programmed our ARUs to record daytime hours (06:00–17:00 PDT) and two hours in the evening (21:00–23:00 PDT) to record the more nocturnally active chorus frog (similarly, adult males are vocally active during their 2–3-week breeding season) and because we have documented Yosemite toad calling during evening hours (C. Brown unpubl data). All ARUs recorded daily at a sample rate of 32 kHz (microphone gain+32 dB, bit depth 16). Battery power constraints precluded continuous recording, although this would have allowed for more detailed temporal analyses.

Qualitative range tests of these ARUS suggested that if either species was audible to members of the field team, it could also be recorded with sufficient clarity to be reliably identified. For the Yosemite toad, this was 10–50 m, but for both human and machine observers, the effective range is a function of the signal to noise ratio, which will vary with vegetation, weather, and other factors, such that it varies continuously over time.

Bioacoustic data preparation

We analysed the audio using the BirdNET algorithm, a deep artificial network initially developed to identify and classify bird sounds (see (Kahl et al. 2021) and https://github.com/kahst/BirdNET-Analyzer for further detail). The algorithm was recently expanded from 984 bird species to >3,000 species, including the Yosemite toad and Pacific chorus frog. BirdNET analyses input audio in 3-s chunks (e.g. Figure 1(a)) and generates a unitless

Table 1. Breeding site and ARU deployment metadata, and vocal activity data for both species.

| | | | | | | Vocal Activity | | | |
| | | | | ARU Deployment | | Yosemite toad | | Pacific chorus frog | |
Site	Area (ha)	Elevation (m)	n	Start	End	Peak	End	Peak	End
1	9.9	2774	3	14 May	16 June		30 May		8 June
2	0.5	2768	1	11 May	25 June	13 May	31 May	13 May	31 May
3	1.3	2621	1	18 May	1 July	31 May	9 June	30 May	26 June

Table 2. Pairwise distance (m) and Pearson's correlation (r) between daily detection counts among recording units at Site 1.

| | Distance (m) | | r (Yosemite Toad) | | r (Pacific Chorus Frog) | |
	Unit 3	Unit 2	Unit 3	Unit 2	Unit 3	Unit 2
Unit 1	346	361	0.895	0.892	0.894	0.903
Unit 2	193		0.852		0.970	

numeric prediction score for each target species. We converted BirdNET's 'confidence' score to its original logit scale via:

$$Logit\ score = {}^{1}\!/(1 - confidence\ score) \tag{1}$$

To determine a BirdNET score threshold that could be consistent for both species, we generated two sets of validation data: we randomly selected 20 h of audio to manually review and we used the 'segments.py' script to randomly select individual BirdNET predictions; collectively this yielded almost 4,600 selected BirdNET predictions spanning the range of dates, times, and locations within our study ($n_{\text{Yosemite toad}} = 1{,}995$; $n_{\text{Pacific chorus frog}} = 2{,}603$), and manually validated each one. For each species, we then used logistic regression to relate the BirdNET prediction score to the binary outcome of the validation process (correct/incorrect), yielding an equation that allowed us to convert BirdNET scores to the probability that any given prediction is correct (Figure 1(b)). For Yosemite toads and Pacific chorus frogs, those equations were:

$$logit(p_{Yosemite\ toad}) = 0.307 + \beta_{score} * 0.387 \tag{2}$$

and

$$logit(p_{Pacific\ chorus\ frog}) = 2.187 + \beta_{score} * 0.621 \tag{3}$$

respectively.

We solved each species' logistic regression equation for prediction score thresholds that would yield a pr(true positive) rates of 80%, 85%, 90%, 95%, and 99%, and applied those thresholds to daily detection count data to test the sensitivity of overall patterns in vocal activity (i.e. relative peaks and troughs) to different score thresholds (Figure 1(c)). Finally, we used the validation dataset to estimate precision (the proportion of predictions above a given threshold that are correct) and recall (the proportion of observed vocalisations that are correctly identified at a given threshold) for the probabilistic score threshold we selected for our analyses.

Analyses of vocal activity

We applied the species-specific score threshold required to achieve 85% accuracy to the BirdNET outputs and tested for spatiotemporal patterns in vocal activity of both species. First, for all five ARUs across the three breeding sites, we summarised the daily and hourly variation in vocal activity of both species to identify phenological patterns, notably the seasonal peak and termination of vocal activity, and to identify diel patterns of vocal activity. Second, we tested for spatial variation in vocal activity across the three ARUs deployed at Site 1, a single large breeding area. We compared variation in daily detection counts among the different ARUs over time, and by computing the pairwise Pearson's correlation between daily detections at the three ARUs.

Results

Detector performance

In total, we collected 2,756 h of audio. BirdNET was highly effective at identifying both species of interest. We chose an 85% accuracy threshold because it minimised the possibility of false positives for both species: at that level, precision was 0.99 and 0.89 for Yosemite toads and Pacific chorus frogs, respectively. Recall varied much more between species at the 85% accuracy threshold, with an estimated recall of just 0.19 for the Yosemite toad but 0.98 for the Pacific chorus frog. Despite relatively low recall for the Yosemite toad, observed patterns in vocal activity were robust to the score threshold (Figure 1(c)): probabilistic thresholds from 80% to 95% accuracy yielded the same relative peaks and troughs in vocal activity, though, of course, observed detection counts decreased as the threshold increased.

Temporal variation

Across all sites, both species exhibited seasonal (Table 1; Figure 2) and diel (Figure 3) peaks in vocal activity. At Sites 1 and 2, both species were vocalising on the first day of recording, but substantial increases in daily detection counts in subsequent days suggested that vocal activity was just beginning for the season. Peak vocal activity dates different by more than 2 weeks across sites (Table 1), and daily vocal activity rates for both species had multiple peaks and troughs across sites (Figure 2). Across all five ARUs (three of which were at Site 1), Yosemite toad vocal activity peaked at midday and had a minor peak between 21:00–22:00; Pacific chorus frog vocal activity had minor peaks between 10:00–11:00 and 16:00–17:00, then tripled between 21:00–23:00 (Figure 3). The ARUs were not recording 17:00–21:00.

Spatial variation

Daily detection counts of Yosemite toads at Site 1 were highly correlated among the three ARUs ($r = 0.852$–0.895; Table 2; Figure 2, solid lines), though the variation in counts among units could be quite high (e.g. on 11 May, there were 5,777 detections at Unit 1, 4,419 more than at Unit 2). Daily detection counts of Pacific chorus frogs were even more correlated ($r = 0.894$–0.970; Table 1; Figure 2, solid lines). Although the absolute magnitude of the differences among units at Site 1 was similar to what we observed for Yosemite toads, the greater overall detection counts meant that the proportional differences in detection counts of Pacific chorus frog was smaller (e.g. twice as many detections at one unit vs another, as opposed to four or five times as many detections).

Discussion

Identifying target signals in many thousands of hours of audio has long been a limiting factor in passive acoustic monitoring. We have successfully expanded BirdNET, which was originally designed to identify birds, to reliably and efficiently detect Yosemite toads and Pacific chorus frogs, even when they vocalise simultaneously. Thus, the combination of passive acoustic monitoring and this new machine learning animal sound
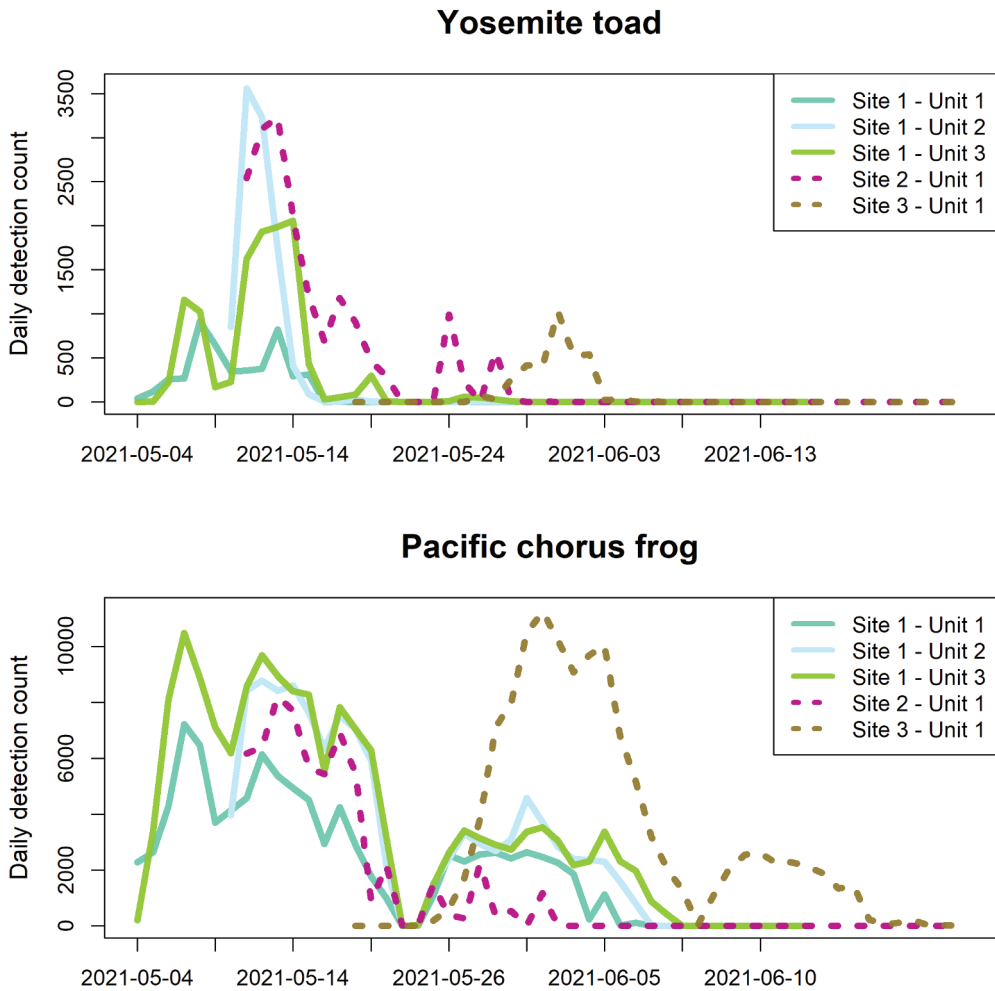
Yosemite toad

Pacific chorus frog

Figure 2. Daily vocalisation counts of Yosemite toads and Pacific chorus frogs at known breeding sites of Yosemite toads in the central Sierra Nevada, USA. Data were collected between 06:00–17:00 and 21:00–23:00 every day.

identification tool can be a scalable means of providing precise phenological data about vocal activity and detection counts for either or both species. Furthermore, Wood et al. (2023) outline several possibilities for scaling up passive acoustic survey coverage despite the challenges associated with accessing remote breeding areas in difficult terrain.

Despite ARU deployments that did not span the entire breeding season, we observed relative peaks in vocal activity for both species at all three sites, which may help provide insights into breeding phenology and its relationship to climate – a connection that is considered a global priority in amphibian conservation (Grant et al. 2020). Although Sites 1 and 2 are nearly 150 km apart, their snow depth was similar in 2021 suggesting similar snowmelt timing, which may explain the similarity of their peaks of vocal activity (Table 1). With greater spatial coverage of acoustic surveys, including pre-breeding ARU deployments (Wood et al. 2023),
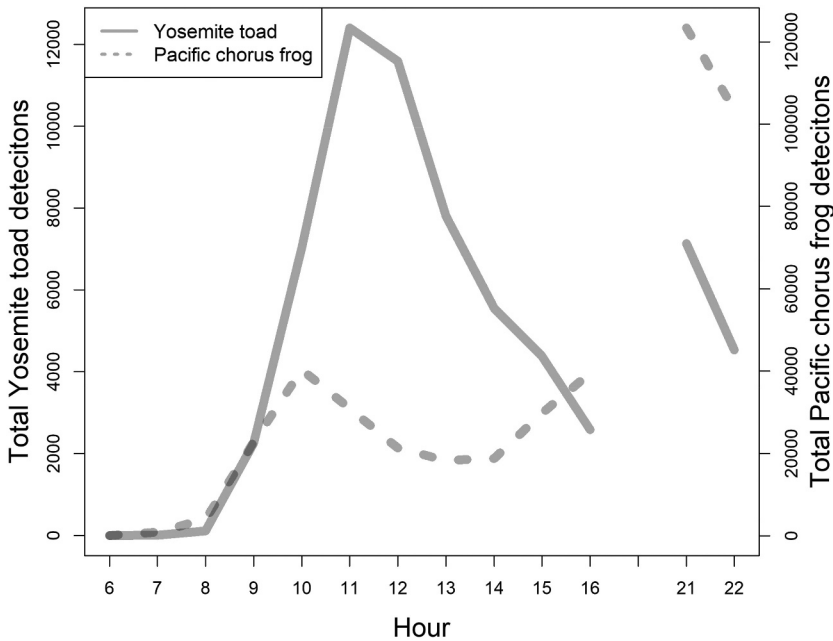
Figure 3. Per-hour counts of Yosemite toad (solid line, left axis) and Pacific chorus frog (dashed line, right axis) vocalisations summed across all days and all units (2,756 hours).

a comprehensive understanding of range-wide breeding phenology could be obtained. Understanding changes in breeding phenology, as well as diel shifts in vocal activity, could improve our understanding of how climate-linked environmental factors like drought and temperature are affecting Yosemite toad and Pacific chorus frog populations (Blaustein et al. 2001; Walpole et al. 2012). For example, there were 2 days with an unexpected absence of vocal activity at Sites 1 and 2 (20, 21 May; Figure 2). Manual review of this audio confirmed that there was indeed no anuran vocal activity and that high winds were present. Regional weather archives indicate that temperature lows dropped from ~2.8–6.0°C to ~ −6.5° during those days, suggesting that the animals – like most other amphibians – display seasonal and daily responses to environmental conditions (Vitt and Caldwell 2014).

Long-duration recordings provide an opportunity to learn more about species' natural history and interactions. Consistent with the acoustic niche hypothesis, sympatric anurans tend to vocalise simultaneously but in different frequency bands (Villanueva-Rivera 2014), as do the Yosemite toad and Pacific chorus frog (Figure 1(a)). Although Yosemite toads are generally diurnal, we recorded substantial night-time vocal activity; while Pacific chorus frogs are generally nocturnal, we recorded them calling during the day (Figure 2). Neither species exhibited vocal activity from 6:00 to 9:00, when temperatures were likely too cold (Vitt and Caldwell 2014), but recording the dawn chorus for a week or more could enable reliable bioacoustic assessments of avian diversity (Wood et al. 2021) as an ancillary benefit of anuran-oriented survey efforts. Further research into temporal patterns of vocal activity of both species is warranted and would be supported by continuous recording data, as opposed to the periodic schedule we used.

However, applying BirdNET (and likely other machine learning tools) to amphibian research may be challenged by simultaneous vocal activity. Convolutional neural networks like BirdNET are not sensitive to sound outside the frequency band of the target signal, but they can be highly sensitive to poorly labelled training data. The vocalisations of highly abundant chorusing frogs (e.g. Pacific chorus frog in our case and spring peeper [*P. crucifer*] in eastern North America) are likely to be in the background of recordings of less common sympatric species. If training data for species A contains the vocalisations of B but is only labelled as species B, the detector will begin to identify the vocalisations of species B as species A. High-scoring false positives (misclassifications of B as A) will necessitate higher score thresholds for species A, meaning that high precision will result in low recall. In our case, achieving pr(true positive) ≥0.85 for the Yosemite toad required a substantially higher BirdNET score than it did for the Pacific chorus frog (see the intercepts in Equations 2 and 3), which almost certainly drove the lower recall achieved by the former species at the same probabilistic score threshold. Thus, species with low population densities may be difficult to detect by virtue of their scarcity – an added challenge for threatened species. Therefore, we recommend that researchers developing machine learning tools for amphibians, or those improving existing tools like BirdNET, take the following steps for training data: tightly crop training data in time to focus on target signals, and use multi-label data when multiple species are present. For common species, simultaneous conspecific vocalisations may also be a challenge because the sound of an individual and the emergent sound of a group chorus can differ substantially, though BirdNET was able to identify both individual foreground Pacific chorus frogs and their background chorusing with high confidence (e.g. Figure 1(a) 2.1 and 2.2, respectively). Therefore, we offer a third recommendation for training data: include both individual (foreground) and chorus (background) examples where applicable.

Differential recall among species, as we observed (0.19 and 0.98 for the Yosemite toad and Pacific chorus frog, respectively) means that Yosemite toad detection counts were more substantially underestimated than those of the Pacific chorus frog. However, the importance of low recall depends on research goals. In the context of a broad-scale, occupancy-based monitoring program targeting species who are likely to vocalise many times during an ARU deployment, low call-based recall is not problematic as long as the species is correctly identified at a given site (Wood et al. 2019). If detection counts are important (e.g. (Pérez-granados and Traba 2021)), low recall may be more problematic. As we demonstrated, a consistent threshold could be applied and count-based analyses could be conducted with appropriate limitations, such as not comparing species with different recall rates. In this case, developing a probabilistic threshold is highly recommended because it provides a concrete, quantitative assessment of the likely error in the counts. We recommend our validation approach to other researchers (manually validate random samples, use logistic regression to relate BirdNET prediction score to the probability that a prediction is correct); however, the particular coefficients of our regression models (Equations (2) and (3)) are unique to our audio dataset. Alternatively, an automated detector could be used to identify the approximate periods of vocal activity, and manual review and annotation could then be used to identify all vocalisations.

Spatial variation in both species' vocal activity was evident at Site 1, where we had three ARUs. Daily detection counts were highly correlated among our three ARUs

($r = 0.852–0.970$; Table 2) but the counts themselves varied substantially (Figure 2). The Site 1 breeding area is relatively large and melts out in stages, and Yosemite toad breeding, and thus calling, follows the melt-out. Nonetheless, even daily detection counts varied substantially, suggesting that the ARUs were not sampling all the same individuals, though overlap is certainly possible (Figure 2). The potential sensitivity of acoustic detection count to ARU placement has important implications. First, whenever possible, placement of ARUs should be informed by existing information about within-meadow breeding locations. The distance between an ARU and known breeding pools may be an important covariate in count-based models. Second, detailed, species-specific assessments of the ARU's listening range will be essential to estimating the sampling coverage of each unit and, thus, determining how many ARUs may be necessary to obtain comprehensive acoustic survey coverage of a breeding site. Third, large breeding meadows, such as Site 1, may need multiple ARUs to ensure comprehensive acoustic survey coverage, necessitating some form of integration across ARUs at a given meadow.

We demonstrated that passive acoustic surveys and a newly expanded machine learning tool, the BirdNET algorithm (Kahl et al. 2021), can be used to effectively survey adult male Yosemite toad and Pacific chorus frog populations. High temporal resolution phenological data can be obtained, with options ranging from seasonal patterns to hourly or even minute-by-minute analyses. Patterns in detection counts may also be quite valuable. Already, high-density ARU arrays have been used to estimate amphibian density (e.g. (Measey et al. 2017)), and detection counts from moderate- and low-density deployments, in which ARUs do not overlap in their recording coverage, have been used to estimate density and abundance for birds (Pérez-granados and Traba 2021). If metrics of acoustic activity (e.g. maximum daily or hourly detection count) could be calibrated with mark-recapture abundance estimates, the possibility of landscape-scale bioacoustic estimates of adult male abundance could be explored as a complement to ongoing occupancy-based surveys and genetic assessments (Brown et al. 2012; Brown and Olsen 2013; Maier et al. 2022). Our study underscored the multi-species capabilities of passive acoustic surveys and machine learning tools, as well as the potential for passive acoustic surveys to yield a wide variety of data, such as phenology, interspecific interactions, and potentially abundance (Wood et al. 2019, 2020). More importantly, we demonstrate the viability of bioacoustics as a research tool to aid in the conservation of a declining species, and the monitoring of many other threatened amphibians.

Disclosure statement

Funding

ORCID

Connor M. Wood (iD) http://orcid.org/0000-0002-0235-5214

Data accessibility

It is challenging to make large audio datasets available, particularly when they contain sensitive information about threatened species. However, the authors welcome inquiries about data availability.

References

Akbal E, Barua PD, Dogan S, Tuncer T, Acharya UR. 2023. Explainable automated anuran sound classification using improved one-dimensional local binary pattern and Tunable Q Wavelet Transform techniques. Expert Syst Appl. 225:120089. doi:10.1016/j.eswa.2023.120089.

Arvind C, Joshi V, Charif R, Jeganathan P, Robin VV. 2022. Species detection framework using automated recording units: a case study of the Critically Endangered Jerdon's courser. Oryx. 57:1–8. doi:10.1017/S0030605321000995.

Barrow LN, Ralicki HF, Emme SA, Lemmon EM. 2014. Species tree estimation of North American chorus frogs (Hylidae: Pseudacris) with parallel tagged amplicon sequencing. Mol Phylogenet Evol. 75:78–90. doi:10.1016/j.ympev.2014.02.007.

Blaustein AR, Belden LK, Olson DH, Green DM, Root TL, Kiesecker JM. 2001. Amphibian breeding and climate change. Conserv Biol. 15:1804–1809. doi:10.1046/j.1523-1739.2001.00307.x.

Bolitho LJ, Rowley JJL, Hines HB, Newell D, Bolitho LJ, Rowley JJL, Hines HB, Newell D. 2021. Occupancy modelling reveals a highly restricted and fragmented distribution in a threatened montane frog (Philoria kundagungan) in subtropical Australian rainforests. Aust J Zool. 67:231–240. doi:10.1071/ZO20037.

Brown C, Kiehl K, Wilkinson L. 2012. Advantages of long-term, multi-scale monitoring: assessing the current status of the Yosemite toad (*Anaxyrus [Bufo] canorus*) in the Sierra Nevada, California, USA. Herpetol Conserv Biol. 7:115–131.

Brown C, Olsen AR. 2013. Bioregional monitoring design and occupancy estimation for two Sierra Nevadan amphibian taxa. Freshw Sci. 32:675–691. doi:10.1899/11-168.1.

Campos-Cerqueira M, Aide TM, Jones K. 2016. Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. Methods Ecol Evol. 7:1340–1348. doi:10.1111/2041-210X.12599.

Crump PS, Houlahan J. 2017. Designing better frog call recognition models. Ecol Evol. 7:3087–3099. doi:10.1002/ece3.2730.

Grant EHC, Miller DAW, Muths E. 2020. A synthesis of evidence of drivers of amphibian declines. Herpetologica. 76:101–107. doi:10.1655/0018-0831-76.2.101.

Kahl S, Wood CM, Eibl M, Klinck H. 2021. BirdNET: a deep learning solution for avian diversity monitoring. Ecol Inform. 61:101236. doi:10.1016/j.ecoinf.2021.101236.

Lapp S, Wu T, Richards-Zawacki C, Voyles J, Rodriguez KM, Shamon H, Kitzes J. 2021. Automated detection of frog calls and choruses by pulse repetition rate. Conserv Biol. 35:1659–1668. doi:10.1111/cobi.13718.

Liang CT. 2013. Movements and habitat use of Yosemite toads (*Anaxyrus* (formerly *Bufo*) *canorus*) in the Sierra National Forest, California. J Herpetol. 47:555–564. doi:10.1670/12-054.

Maier PA, Vandergast AG, Ostoja SM, Aguilar A, Bohonak AJ. 2022. Gene pool boundaries for the Yosemite toad (*Anaxyrus canorus*) reveal asymmetrical migration within meadow neighborhoods. Front Conserv Sci. 3:851676. doi:10.3389/fcosc.2022.851676.

McCallum ML. 2007. Amphibian decline or extinction? Current declines dwarf background extinction rate. J Herpetol. 41:483–491. doi:10.1670/0022-1511(2007)41[483:ADOECD]2.0.CO;2.

Measey GJ, Stevenson BC, Scott T, Altwegg R, Borchers DL, Bellard C. 2017. Counting chirps: acoustic monitoring of cryptic frogs. J Appl Ecol. 54:894–902. doi:10.1111/1365-2664.12810.

Noda JJ, Travieso CM, Sánchez-Rodríguez D. 2016. Methodology for automatic bioacoustic classification of anurans based on feature fusion. Expert Syst Appl. 50:100–106. doi:10.1016/j.eswa.2015.12.020.

Pérez-granados C, Traba J. 2021. Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. Ibis. 63:765–783. doi:10.1111/ibi.12944.

Recuero E, Martínez-Solano Í, Parra-Olea G, García-París M. 2006. Phylogeography of Pseudacris regilla (Anura: Hylidae) in western North America, with a proposal for a new taxonomic rearrangement. Mol Phylogenet Evol. 39:293–304. doi:10.1016/j.ympev.2005.10.011.

Sodhi NS, Bickford D, Diesmos AC, Lee TM, Koh LP, Brook BW, Sekercioglu CH, Bradshaw CJA, Freckleton R. 2008. Measuring the meltdown: drivers of global amphibian extinction and decline. PLos One. 3:e1636. doi:10.1371/journal.pone.0001636.

Villanueva-Rivera LJ. 2014. *Eleutherodactylus* frogs show frequency but no temporal partitioning: implications for the acoustic niche hypothesis. Peer J. 2:e496. doi:10.7717/peerj.496.

Vitt LJ, Caldwell JP. 2014. Thermoregulation, performance, and energetics. In: Herpetology. 4th ed. Elsevier; p. 203–227.

Walpole AA, Bowman J, Tozer DC, Badzinski DS. 2012. Community-level response to climate change: shifts in anuran calling phenology. Herpetol Conserv Biol. 7:249–257.

Wang IJ. 2012. Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. Divers Distrib. 18:1033–1041. doi:10.1111/j.1472-4642.2012.00897.x.

Wijayathilaka N, Senevirathne G, Bandara C, Rajapakse S, Pethiyagoda R, Meegaskumbura M. 2018. Integrating bioacoustics, DNA barcoding and niche modeling for frog conservation – the threatened balloon frogs of Sri Lanka. Global Ecol Conserv. 16:e00496. doi:10.1016/j.gecco.2018.e00496.

Wood CM, Champion J, Brown C, Brommelsiek W, Laredo I, Rogers R, Chaopricha P. 2023. Challenges and opportunities for bioacoustics in the study of rare species in remote environments. Conserv Sci Pract N/A. e12941. doi:10.1111/csp2.12941.

Wood CM, Gutiérrez RJ, Peery MZ. 2019. Acoustic monitoring reveals a diverse forest owl community, illustrating its potential for basic and applied ecology. Ecology. 100:1–3. doi:10.1002/ecy.2764.

Wood CM, Kahl S, Chaon P, Peery MZ, Klinck H, Desjonquères C. 2021. Survey coverage, recording duration and community composition affect observed species richness in passive acoustic surveys. Methods Ecol Evol. 12:885–896. doi:10.1111/2041-210X.13571.

Wood CM, Klinck H, Gustafson M, Keane JJ, Sawyer SC, Gutiérrez RJ, Peery MZ. 2020. Using the ecological significance of animal vocalizations to improve inference in acoustic monitoring programs. Conserv Biol. 35:336–345. doi:10.1111/cobi.13516.

Wood CM, Popescu VD, Klinck H, Keane JJ, Gutiérrez RJ, Sawyer SC, Peery MZ. 2019. Detecting small changes in populations at landscape scales: a bioacoustic site-occupancy framework. Ecol Indic. 98:492–507. doi:10.1016/j.ecolind.2018.11.018.