2015

The Problem of Low Agreement among Automated Identification Programs for Acoustical Surveys of Bats

Cliff Lemen
University of Nebraska–Lincoln, clemen2@unl.edu

Patricia W. Freeman
University of Nebraska-Lincoln, pfreeman1@unl.edu

Jeremy A. White
University of Nebraska at Omaha, jeremywhite@unomaha.edu

Brett R. Andersen
University of Nebraska at Omaha

Western North American Naturalist 75(2), © 2015, pp. 218–225

THE PROBLEM OF LOW AGREEMENT AMONG AUTOMATED IDENTIFICATION PROGRAMS FOR ACOUSTICAL SURVEYS OF BATS

Cliff Lemen[1,3], Patricia W. Freeman[1], Jeremy A. White[2], and Brett R. Andersen[2]

ABSTRACT.—We compared 4 programs designed to identify species of bats from their echolocation calls (Bat Call ID, EchoClass, Kaleidoscope Pro, and SonoBat) using field data collected in Nebraska, USA (29,782 files). Although we did not know the true identity of these bats, we could still compare the pairwise agreement between software packages when identifying the same call sequences. If accuracy is high in these software packages, there should be high agreement in identification. Agreement in identification by species averaged approximately 40% and varied by software package, species, and data set. Our results are not consistent with the high accuracy often claimed by some software packages and may be a warning about the importance of understanding accuracy of acoustical identification in designing ecological experiments and interpreting results.

RESUMEN.—Comparamos 4 programas diseñados para identificar especies de murciélagos a partir de sus llamados de ecolocalización (Bat Call ID, EchoClass, Caleidoscopio Pro, y SonoBat) utilizando datos de campo obtenidos en Nebraska (29,782 archivos). No sabemos la verdadera identidad de estos murciélagos, pero comparamos el acuerdo por pares entre los paquetes de software en la identificación de las mismas secuencias de llamados. Si la precisión entre estos paquetes de software es alta, debe haber un acuerdo en la identificación. El promedio en la identificación por especies fue en promedio el 40% y varió por paquete de software, por especies, y por conjunto de datos. Nuestros resultados no son consistentes con la alta precisión aclamada por algunos paquetes de software y pueden ser una advertencia acerca de la importancia de entender la exactitud de la identificación acústica en el diseño de experimentos ecológicos y en la interpretación de los resultados.

With the rapid expansion of acoustical surveys of bats, there is a need for accurate, automated identification of bat calls by computer software. Large-scale studies may play an important role in understanding bat distribution and abundance in the face of emerging management problems such as white-nose syndrome and wind turbine installations. Further, if more bats are classified as endangered, there will be important management and economic decisions based on scientists' best knowledge of the distribution of endangered species. These software packages make possible computer identification of hundreds of thousands, or even millions, of bat call sequences. In disclaimers, the developers of these programs typically call for caution, noting the importance of making good recordings by avoiding cluttered areas and elevating microphones above the ground. We concur with these guidelines for making the best recordings possible. Some automated packages also supply performance data with rates of correct identification of approximately 95% or better for most species (white paper for Sonobat 3.1, or white paper for Kaleidoscope Pro). These high estimated accuracies create a situation where vast amounts of data can be collected and analyzed with apparent confidence, even for researchers with little experience in the identification of bat calls. Armed with these results, researchers can focus on important issues of conservation or ecology of bats.

Analysis of bat calls has a history of controversy over methods and equipment (Fenton 2000, Corben and Fellers 2001, Allen et al. 2011). Several efforts have explored approaches to optimize bat species identification (Parsons and Jones 2000, Skowronski and Harris 2006, Britzke et al. 2011). There are also warnings that identifications may prove problematic because bat calls are not as distinctive as bird calls (Barclay 1999). On one hand is a history of at least some controversy and skepticism about acoustical identification of bats, and on the other is the ease-of-use and apparent high levels of accuracy in available automated identification packages. As the use of detectors and automated

[1]School of Natural Resources and University of Nebraska State Museum, University of Nebraska–Lincoln, 428 Hardin Hall, Lincoln, NE 68583-0974.
[2]Department of Biology, University of Nebraska at Omaha, Omaha, NE 68182.
[3]E-mail: clemen2@unl.edu

identification spreads rapidly, not every user will be an expert in the statistical underpinnings of these packages nor in the identification of bat calls. Further, as more bat detectors are used over longer periods of time, many call sequences can be recorded. Even if an expert is involved, there may be too many files to be reviewed by eye. Indeed, these are the needs that the software packages are designed to accommodate.

We do not know the identity of the bats producing our call sequences; therefore, the ideal approach of quantifying accuracy by comparing the software identification to known identities is not possible. However, we reasoned that if the rate of correct identifications averages approximately 95% in all packages, this would lead to a level of agreement between any 2 packages, assuming errors are random, of 0.95^2 or approximately 90%. So our question is: do software packages agree in bat identification at a level consistent with high accuracy? We evaluated 4 of these programs (Bat Call ID, EchoClass, Kaleidoscope Pro, and SonoBat) by comparing the level of agreement of species identifications between programs.

There are certainly weaknesses to using agreement to gain insight into accuracy. First, low levels of agreement among the 4 identification programs cannot exclude the possibility that one program is near perfect and the other 3 programs are flawed. Second, high levels of agreement could result from the programs all making the same mistakes.

It should be mentioned that some software packages also supply probability-of-presence estimates at a site for each species identified as an alternative to call-level identifications. The idea is to determine how likely it is that a species is actually present given the rate of false positives from other species. This may prove a useful approach. However, we do not address this method here; rather, we concentrate on assessing the accuracy of the identification of call sequences by looking at the agreement among software packages.

METHODS

We compared 4 software packages: Bat Call ID 2.6 (BCID; Bat Call Identification, Inc., Kansas City, MO, USA; www.batcallid.com), EchoClass 2.0 (Dr. Eric Britzke; www.fws.gov/midwest/Endangered/mammals/inba/surveys/inbaAcousticSoftware.html), Kaleidoscope Pro 2.2.2 with North American 2.2.2 classifier (Wildlife Acoustics, Concord, MA, USA; www.wildlifeacoustics.com), and SonoBat 3.1 (Arcata, CA, USA; www.sonobat.com). Although acoustical classification of bats is an active area of research (Skowronski and Harris 2006, Skowronski and Fenton 2008, Britzke et al. 2011), we restricted our analysis to these widely available packages.

To make our recordings, we used Wildlife Acoustics SM2Bat+ detectors and SMX-US microphones. We restricted our recordings to one detector and microphone brand, but it is clear that hardware used to make bat recordings can have an important effects on results (Fenton 2000, Adams et al. 2012). We set up our recorders to manufacturer's recommendations. The recorders were set to sample at 192,000 Hz. In theory this would allow the recorder to detect sounds up to 96 kHz. However, because the SMX-US microphone suffers a drop in sensitivity of approximately 20 dB above 60 kHz, the effective upper limit of frequency is probably closer to approximately 60–70 kHz, rather than 96 kHz. The gain of the detector was set to the recommended level of 48 dB. The high-pass filter was set to one-twelfth of the sampling rate, or approximately 16 kHz, to remove some of the low-frequency noise files. Compression mode was set to WAC0, the lossless compression mode. The trigger level was set to 12SNR. This means that a recording will be triggered if a sound is detected ≥12 dB above the background level. All sequences recorded as full-spectrum in WAC format were later converted to WAV and Zero Crossing (ZC) formats using Kaleidoscope Pro software from Wildlife Acoustics. Within Kaleidoscope Pro, we specified the maximum duration of the WAV files as 5 s. The filter settings specified a signal of interest between 10 and 120 kHz and 2 to 500 ms. Further we used BCID to analyze each of these sound files and only used files that contained ≥5 pulses within the call sequence. We then analyzed the WAV files by using Sonobat and Kaleidoscope Pro and the ZC files by using BCID and EchoClass.

We made recordings of bats during 2012–2013 in southeastern Nebraska. We used 2 data sets in this study. The first group of recordings consists of 6431 WAV files from 35 sites. These files were the ones left after Kaleidoscope Pro filtered out noise files and BCID tallied ≥5 pulses in the files. We made recordings for a single night at each of these sites in the spring and summer of 2012. We refer to this group of recordings as the single-night data set. We made the second group of recordings at one locality near Nebraska City, and it contained 23,351 files after processing by Kaleidoscope Pro and BCID. The Nebraska City recordings were made over 201 nights from fall 2012 to fall 2013. We analyzed all files for identifications with Kaleidoscope Pro, SonoBat, BCID, and EchoClass.

We made the single-night recordings under a variety of conditions, with some sites more open and some less. Although we attempted to put the microphone in the best location possible at a site, at times even the best spot may have been in a cluttered situation. Visual inspection of the call sequences indicated that the recordings from a few of the single-night sites did contain a high proportion of clutter calls. We are sensitive to the fact that some recording spots were in fairly tight situations, and clutter calls could compromise some results. Our initial intention was to eliminate these sites from the analysis. On further thought, we decided to leave them in to give a broad range of results that might be possible when clutter is at least an occasional problem. For comparison, we also used the Nebraska City data set, where all recordings were made from a single spot in a forest clearing. The microphone was on a recording post set 4 m high in a grassy field well away from trees and large bushes (approximately 35 m away from and in front of the microphone and 10 m away from and behind the microphone). This microphone was in a more open situation than the majority of sites from the single-night data set. Thus, if having the microphone in the open results in a greater number of identifiable recordings, then we might see at least some higher agreement among the results from this data set.

The bat fauna of southeastern Nebraska is not large; we expected to record calls from 8 species at our sites. In our study area, 5 species have high-frequency calls: little brown myotis (*Myotis lucifugus*), northern myotis (*Myotis septentrionalis*), tri-colored bat (*Perimyotis subflavus*), eastern red bat (*Lasiurus borealis*), and evening bat (*Nycticeius humeralis*); 2 species have midrange calls: big brown bat (*Eptesicus fuscus*) and silver-haired bat (*Lasionycteris noctivagans*); and one species has a low-frequency call: hoary bat (*Lasiurus cinereus*). This is the set of bats used in our analysis, and we refer to it as the SE Nebraska set of species. Different sets of species in different areas likely alter results, because calls of some species are more difficult to distinguish from others.

Determination of the bats to include in the analysis could have an effect on agreement of the software. If species that are extremely rare in our area (Mexican free tailed bat [*Tadarida brasiliensis*]) or would represent a range extension (the gray bat [*Myotis grisescens*] or the Indiana bat [*Myotis sodalis*]) are included, they might decrease agreement. Of course, there are places where these bats are present and would have to be included. We suspect there would be some false positives with species having similar call characteristics. We decided to use a conservative approach and exclude rare species and species that represent a range extension in our area. We understand that it can be argued that these limits will have the effect of increasing agreement.

Ideally we could set the possible species to be identified as our Nebraska set, however it was not possible because not all software packages give that level of control in including or excluding species in the analysis. SonoBat has a fixed set of species by region, and for identification of calls from southeastern Nebraska, we chose to use the Ozark region. This package includes all 8 species we expected at our sites, as well as 4 species that are unlikely to be encountered in the region: Townsend's big-eared bat (*Corynorhinus townsendii;* not known from SE Nebraska) Brazilian free-tailed bat (*Tadarida brasiliensis;* present but extremely rare), Indiana bat (possibly present, but although there are no records from Nebraska, it is known from sites in Iowa and northwestern Missouri), and gray bat (unlikely to be present; no records from Nebraska but known from Kansas and Missouri). The other 3 programs allow more control of the species included in the analysis. In Kaleidoscope Pro and Bat Call ID we were able to select just the SE Nebraska set of species. In EchoClass, we could only select from 3 species lists. To include all SE Nebraska species, we had to include *Myotis austroriparius, M. grisescens, Myotis leibii,* and *M. sodalis.* Fortunately, very few identifications were attributed to these species, so little harm was done by their inclusion. However, if SonoBat or EchoClass identified a species not in the SE Nebraska list of species, we scored it as invalid and did not use that identification to calculate agreement.

Software packages vary in how they deal with confidence scores (such as a discriminant score) as a measure of confidence in identification. Discriminant scores by convention usually range from 0 to 1, with 1 being the highest level of confidence. Different software packages have different means of quantifying the level of confidence in a particular identification. In this paper we refer to such indices generically as confidence scores.

Sonobat, Kaleidoscope Pro, and BCID all have a means to control "accuracy" of the analysis. By raising the accuracy setting the programs will only identify call sequences that have higher confidence scores within the mathematics of the software. And although fewer call sequences are identified, the hope is that a higher proportion of identifications are correct. If true, the level of agreement among software packages should increase at higher accuracy settings. To test this hypothesis, we ran our analysis twice using lower and higher accuracy settings for SonoBat, Kaleidoscope Pro, and BCID. EchoClass supplies baseline identifications without the ability to alter an accuracy setting, with the result that we only ran one analysis of this program. SonoBat allows the user to input a decision threshold and acceptable call quality. In SonoBat, we set the decision threshold to 0.9 and acceptable call quality to 0.8 for the lower accuracy setting. We raised the decision threshold to 0.98 for higher accuracy. BCID confidence score was left at the default level for the baseline accuracy. For higher accuracy we raised the decision threshold to 0.4 out of the 0 to 1.0 scale of the confidence score. The value of 0.4 resulted in a similar number of calls being identified as with the 0.98 confidence score used in SonoBat. For the lower accuracy analysis in Kaleidoscope Pro we used the intermediate setting of "Sensitive/Accurate," and for the higher accuracy analysis we used the setting of "More Accurate." We then calculated the level of agreement within the lower and within the higher accuracy results separately to determine whether the accuracy setting would affect agreement.

To calculate pairwise agreement between methods, we only made comparisons where both software programs attempted identification. Our approach raises the question of how to treat sequences identified by one program but not the other. It could be argued that if Kaleidoscope Pro attempted a classification and SonoBat did not, then that should be scored as a disagreement. Such a method would increase the disagreement among software packages. We chose the more conservative method of demanding mutual classification, which avoids the problem of the proportion of bats identified by each program. As an example, because EchoClass identifies fewer calls, there would be many disagreements with SonoBat simply because EchoClass did not attempt an identification. Because we restrict our comparisons to mutual classifications, this will increase the level of agreement among software packages and, in some sense, may be viewed as biased toward increasing agreement.

RESULTS

Kaleidoscope Pro extracted 24,648 and 94,770 WAV sound files from the WAC files recorded at the single-night and Nebraska City data sets, respectively. Filtering the files to only those with 5 pulses reduced the number of files to 6431 for the single-night data set and 29,862 for the Nebraska City data set (Table 1). At the lower accuracy setting, SonoBat identified 88% of these WAV files to species. At the higher accuracy setting, SonoBat identified 58% of the files. For Kaleidoscope Pro, the identified percentages for lower and higher accuracy were 95% and 81%, respectively. For BCID, the percentages were 98% and 58%, and for EchoClass, the percentage was 18%. EchoClass identified the fewest files. Thus, the standard accuracy setting and only one available for EchoClass may have the strictest requirements for identification among the programs we tested here.

We calculated species-level agreement in the single-night data set and the Nebraska City data set for both the higher accuracy and lower accuracy settings (Table 2). The sample sizes for these comparisons are shown in Table 1. We did not calculate agreement for species with fewer than 100 identifications at the lower accuracy setting. Note that the percentage agreement is highly variable between software packages, species, and data sets. However, a clear result is that there are not uniformly high levels of agreement by species for any software comparisons at either the higher or lower

TABLE 1. Number of call sequences identified simultaneously by both software programs used for calculation of agreement in identification. Header columns indicate the programs being compared to calculate percentage agreement at both the lower and higher accuracy levels. Sono = SonoBat, Kal = Kaleidoscope Pro, BCID = Bat Call ID, and Echo = EchoClass. At low setting both programs are set to the lower accuracy setting, and at high setting both programs are set to higher accuracy setting. The exception is EchoClass, which could only be run at one accuracy setting.

	Sono/Kal		Sono/BCID		Sono/Echo		Kal/BCID		Kal/Echo		BCID/Echo	
Species	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Single night												
E. fuscus	2208	1391	2359	816	684	517	3635	1020	738	602	847	249
L. borealis	832	429	691	70	912	599	1773	271	1115	879	1233	381
L. cinereus	133	65	234	97	150	127	574	159	188	138	189	85
L. noctivagans	632	446	590	208	87	61	1896	664	149	134	136	77
M. lucifugus	349	182	319	86	86	56	744	162	121	86	94	32
N. humeralis	496	148	573	91	147	54	1484	424	166	90	264	133
P. subflavus	1014	886	1143	957	144	138	1475	1056	148	128	215	162
Nebraska City												
E. fuscus	10,471	6333	10,805	4247	1044	528	5129	1448	480	386	542	118
L. borealis	5047	1495	5223	574	2288	699	4387	1239	2252	1883	2306	1005
L. cinereus	1760	1115	4480	2009	1510	1097	6497	3851	1676	1488	1679	1230
L. noctivagans	8597	5157	6515	3426	395	207	13,321	8599	999	972	877	659
N. humeralis	2075	326	3785	604	199	20	3744	291	431	246	949	508
P. subflavus	6281	4773	6467	4219	820	561	6157	2385	684	591	752	518

TABLE 2. Percentage species-level agreement of identification of bat sequences by pairwise comparisons of software packages. Header columns indicate the programs being compared to calculate percentage agreement at both the lower and higher accuracy levels. Abbreviations and settings as in Table 1.

	Sono/Kal		Sono/BCID		Sono/Echo		Kal/BCID		Kal/Echo		BCID/Echo	
Species	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Single night												
E. fuscus	78	76	79	82	26	26	73	81	30	34	24	14
L. borealis	37	39	39	17	25	19	35	7	31	39	21	3
L. cinereus	83	86	39	42	54	58	32	42	47	39	42	62
L. noctivagans	28	23	14	19	10	7	33	68	19	21	19	31
M. lucifugus	65	76	71	88	16	16	59	91	12	17	16	19
N. humeralis	44	76	50	66	26	9	42	61	34	52	18	30
P. subflavus	86	87	86	95	60	62	75	85	65	70	46	59
Nebraska City												
E. fuscus	29	29	23	7	12	16	30	9	24	28	9	6
L. borealis	53	58	23	5	49	19	23	5	40	47	16	3
L. cinereus	74	92	25	45	64	76	26	26	62	60	68	76
L. noctivagans	10	10	9	10	13	10	42	55	32	32	37	46
N. humeralis	22	12	18	16	3	0	39	41	3	3	1	1
P. subflavus	88	92	84	98	14	18	82	97	16	16	15	21

accuracy settings. The average species-level agreement across both data sets between SonoBat and Kaleidoscope Pro was 54% and 58% for the lower and higher accuracy settings, respectively. This comparison was the highest average agreement between any of the software packages. The agreement percentages for other comparisons were as follows: SonoBat and BCID were 43% and 45% (for lower and higher accuracy settings, respectively); SonoBat and EchoClass were 29% and 26%; Kaleidoscope Pro and BCID were 45% and 51%; Kaleidoscope Pro and EchoClass were 32% and 35%; and BCID and EchoClass were 26% and 29%. Therefore, although increasing the accuracy settings threshold improves agreement for some species, it does not alter the pattern of low average agreement among software packages (Table 2).

DISCUSSION

Because we use agreement of classification, we cannot exclude the possibility that one program is near-perfect, and the other 3 programs are flawed. We can conclude that, at best, no more than one program has uniformly high accuracy because of the low levels of agreement among packages. Although beyond the scope of the data presented here, reviewing many calls by eye and comparing them with the identifications of all 4 programs has not led us to any conclusion about which program is the best.

Although average agreement among programs is disappointing, there are areas of high agreement in our data (Table 2). The highest average agreement was for the tri-colored bat (63%), but one comparison was 98%. The tri-colored bat's calls average the highest in frequency in this set of species. Next highest in agreement (at 55%) is the hoary bat, with one comparison at 92%. Calls of the hoary bat average the lowest frequency among these species. Thus, the 2 bats with relatively distinctive calls also have the highest average agreements.

We do not know what causes the discrepancy between the level of accuracy found by software developers and our levels of agreement. But we can point to some obvious possibilities—one is differences in calls recorded with different recording equipment. We used Wildlife Acoustics SM2BAT+ detectors with SMX-US microphones for our recordings. EchoClass and BCID were developed to be used with Anabat recorders and microphones specifically. Recordings made with different hardware will differ (Fenton 2000, Adams et al. 2012). However, Janos (2013) conducted a similar study to ours, but used Anabat equipment and compared Echoclass and BCID. He found overall low levels of agreement (average of 50%). Although somewhat higher than the agreement levels we found at our sites with Wildlife Acoustics equipment, it is still a disappointingly low average agreement.

Another possible reason for the often low levels of agreement we found is that software developers use libraries composed of call sequences of known identifications, whereas we used field recordings. There may be a tendency to select the best and most diagnostic calls (i.e., those with both high technical recording quality and low ambiguity [good search calls and not clutter calls]) to put into a library. This is exactly what has been suggested as essential for making call libraries (Britzke et al. 2013). However, there is no such selection process in our field data beyond what is done by the software. Before placing any confidence in the reported high levels of accuracy, researchers must ask the question whether their data sets are more similar to a carefully selected library groomed by experts or to data such as ours that were collected in the field.

We included the Nebraska City data set because all the recordings were made from a single spot that was more in the open than the average site from the single-night data set. The average agreement by species is higher in the single-night data set (45%) than in the Nebraska City data set (33%). Therefore we find no evidence that a more open microphone placement improved the agreement among software packages. It could be argued that the Nebraska City site was not open enough, indicating that even more isolation from cover might be needed to obtain truly high-quality results. We certainly concur with the idea that clutter calls are a problem in some situations and care should be taken in the positioning of microphones. Unfortunately, we do not have data sets from even more open sites and cannot address this important problem here. More research on this question is needed, particularly with replication of open sites, to help answer this question.

Conclusions

We do not know the true identity of bats producing the calls we analyzed; therefore, we cannot calculate the accuracy of software packages directly. However, by looking at the level of agreement among the programs, we can draw some conclusions. First, the level of agreement across all species, programs, and data sets is not consistent with overall high accuracy for all software packages and species. Second, increasing the confidence score threshold has little effect on the average agreement among packages in this study. Third, our more open site did not produce higher levels of agreement.

The promise of automated identification software was that mathematical quantification of the calls coupled with statistical analysis by the software packages would clarify identification and offer repeatability free of subjective biases of human identification. Given the low levels of agreement and the relatively high percentage of improbable identifications we found, this may not yet be the case. Naturally, these programs may be revised in the future and new programs may be created. Hopefully, new methods will achieve much higher accuracy rates than our study suggests. At this point, we cannot recommend any software package as superior. If data sets are prohibitively large for identification by experts, then we suggest it is prudent to use multiple software packages to analyze data and compare the results to test for agreement before placing too much confidence in the results. Naturally, high levels of agreement do not prove accuracy, but careful comparisons across packages can help researchers to understand the uncertainties in identification that may be present. There is also another approach where presence–absence is based on an expected rate of false positives. This approach may hold promise, but it could suffer from an incorrect assessment of the rate of false positives. The rate of false positives from call libraries might be much lower than from field-collected data such as ours. Our study certainly raises this possibility.

If our conclusions are generally found to be correct, then it affects how these software programs should be used in research projects and environmental inventories, and how the results of such studies should be viewed by the wider research community. It is also important to point out that we use and rely on some of the software used in this study (SonoBat and Kaleidoscope Pro) for our own research. These programs provide us with a first sort of call sequences to organize the recordings for manual review. The software provides an absolutely critical step in our analysis of the recordings. However, we are pointing out the problem of uncritically accepting software identifications. This can have serious consequences in the larger acoustical studies that are proceeding as recording equipment becomes more common.

ACKNOWLEDGMENTS

We thank our respective academic homes for their support: School of Natural Resources and University of Nebraska State Museum, University of Nebraska–Lincoln; Department of Biology, University of Nebraska at Omaha. This research was possible with funding from the U.S. Fish and Wildlife Service in Grand Island; and Mike Fritz, Nebraska Game and Parks Commission, who loaned much of the equipment. Hans Otto assisted in collection of data; and Keith Geluso, University of Nebraska at Kearney, had much to do with getting this project started. Without the help of these institutions and colleagues, this project could not have been done.

LITERATURE CITED

ADAMS, A.M., M.K. JANTZEN, R.M. HAMILTON, AND M.B. FENTON. 2012. Do you hear what I hear? Implications of detector selection for acoustic monitoring of bats. Methods in Ecology and Evolution 3:992–998.

ALLEN, C.R., S.C. ROMELING, AND L. ROBBINS. 2011. Acoustic monitoring and sampling technology. Pages 173–188 *in* K.C. Vories, H.C. Caswell, and T.M. Price, editors, Protecting threatened bats at coal mines: a technical interactive forum. Southern Illinois University, Carbondale, IL.

BARCLAY, R.M.R. 1999. Bats are not birds—a cautionary note on using echolocation calls to identify bats: a comment. Journal of Mammalogy 80:290–296.

BRITZKE, E.R., J.D. DUCHAMP, K.L. MURRAY, R.K. SWIHART, AND L.W. ROBBINS. 2011. Acoustic identification of bats in the eastern United States: a comparison of parametric and nonparametric methods. Journal of Wildlife Management 75:660–667.

BRITZKE, E.R., E.H. GILLAM, AND K.L. MURRAY. 2013. Current state of understanding of ultrasonic detectors for the study of bat ecology. Acta Theriologica 58: 109–117.

CORBEN, C., AND G.M. FELLERS. 2001. Choosing the 'correct' bat detector—a reply. Acta Chiropterologica 3:253–256.

FENTON, M.B. 2000. Choosing the 'correct' bat detector. Acta Chiropterologica 2:215–224.

JANOS, G. 2013. Utilizing acoustic methods to classify bat species and to assess their habitat use and perception of owls. Master's thesis, Bowling Green State University, Bowling Green, OH.

PARSONS, S., AND G. JONES. 2000. Acoustic identification of twelve species of echolocating bats by discriminant function analysis and artificial neural networks. Journal of Experimental Biology 203:2641–2656.

SKOWRONSKI, M.D., AND M.B. FENTON. 2008. Model-based detection of synthetic bat echolocation calls using an energy threshold detector for initialization. Journal of the Acoustical Society of America 123: 2643–2650.

SKOWRONSKI, M.D., AND J.G. HARRIS. 2006. Acoustic detection and classification of microchiroptera using machine learning: lessons learned from automatic speech recognition. Journal of the Acoustical Society of America 119:1817–1833.

Received 15 August 2014
Accepted 11 February 2015