ORIGINAL ARTICLE

Missing something? Importance of measurement criteria of acoustic parameters in the analysis of bats recordings

Cristian Kraker-Castañeda1,2,*, Wiliam Ku-Peralta3, Antonio Santos-Moreno4

ABSTRACT

Ultrasound detectors are becoming an increasingly used tool in bat research. Still, the lack of standardization of measurement criteria in the analysis of recordings is problematic when attempting identifications. For instance, the analytical procedures depend on the technique for obtaining the spectral content of an echolocation pulse (eg. zero-crossing, Fourier analysis)—which is directly related to the detection device and software—and they should be as consistent as possible among studies’ and observers’ to allow comparisons. Using full-spectrum recordings of a Myotis species, we measured the minimum frequency and maximum frequency on the spectrogram, as well as considering four thresholds on the power spectrum: at -55 and -50 decibels below peak frequency. Focusing on the minimum frequency, we found statistically significant differences of measurements obtained on the spectrogram vs. power spectrum, with higher mean values in the latter as product of pulse truncation, as well as statistically significant differences between two trained observers. The measurements relative to the peak frequency were less variable. Without proper considerations, these issues may represent confounding factors and derive data unsuitable for cross-studies comparisons and potential misidentifications. We make recommendations about measurement criteria and emphasize the importance of replication and more rigorous reports.

INTRODUCTION

Acoustic detection has become a useful tool for the study of bats. The possibility to transform sound to study its properties (eg. structure), and to measure acoustic parameters, bring clues about the identity of the recorded bats on different taxonomic levels. Some species vocalizations exhibit highly stable features, and are referred to as vocal signatures (see O’Farrell & Miller 1999, O’Farrell et al. 1999). Others, depending on the magnitude of intraspecific-to-interspecific variability, and as a product of phylogenetic relatedness and adaptive convergence, may exhibit similar structure and overlap of acoustic parameters (see Russo et al. 2018). In the Neotropics, many research groups have implemented ultrasonic detection as a sampling method for bats; however, they usually follow different analytical procedures that can prevent cross-studies comparisons. This situation can be attributed to many factors, but an obvious one is the availability of detection devices that vary in performance, with direct implications on the characteristics of the recording (see Fenton 2000, Adams et al. 2012). Moreover, the measurement criteria of acoustic parameters are not set by the recording device itself but during post-processing, so standardization is important.

For instance, for full-spectrum recordings, there is a criterion for the measurement of the minimum frequency and maximum frequency based on the spectrogram (hereafter FMINy and FMAXy, respectively). There are also different criteria, based on thresholds, for the measurement of the minimum frequency and maximum frequency on the power spectrum (hereafter FMINp and FMAXp, considered synonyms of the lowest frequency and highest frequency, respectively): 1) measurement of the frequency at -60, -55 or -50 decibels (dB sound pressure level SPL) from 0 dB (eg. Fenton 2002, Fenton et al. 2004, Biscardi et al. 2004), 2) measurement of the frequency at 18, 15, 12, 10, or 6 dB below peak frequency (= frequency of maximum energy or amplitude) (eg. Obrist 1988, Barclay et al. 1999, Surrykje & Moss 2000, MacSwiney et al. 2008). Some of the arguments found in the literature supporting the latter criteria are: that...
the thresholds set -50 to -60 dB from 0 dB are adequate for the elimination of background noise (Fenton 2002); that the threshold 6 dB below peak frequency marks half of the energy in the signal (Fenton 2002); that the threshold 15 dB below peak frequency controls for differences in signal level (Surlykke & Moss 2000).

Presumably, the different measurement criteria will produce different results, although it is not clear to what magnitude. Here, we aimed to test for differences between measurement criteria of frequency-based parameters, using full-spectrum recordings. Also, to test for differences of measurements between two trained observers. If there are differences, these might represent confounding factors, which might derive data unsuitable for cross-studies comparisons and potential misidentifications. Finally, we make recommendations about measurement criteria and emphasize the importance of replication and more rigorous reports.

MATERIALS AND METHODS

The recordings came from 14 individuals of Myotis nigricans (Schinz, 1821) (Chiroptera, Vespertilionidae) (for details see Kraker-Castañeda et al. 2019), obtained with a high-speed recording device (Pettersson Ultrasound Detector D1000X, Pettersson Elektronik AB, Sweden). The echolocation pulses were displayed with the software BatSound Standard – Sound Analysis v. 3.31 (Pettersson Elektronik AB, Sweden): 1) on the spectrogram (frequency vs. time; Fig. 1a), 2) on the oscillogram (amplitude vs. time; Fig. 1b), 3) on the power spectrum (amplitude vs. frequency; Fig. 1c). We obtained the frequency-based parameters on the fundamental harmonic, specifically from one pulse with a high signal-to-noise ratio (see Fig. 1b), per sequence and file.

First, we displayed the entire recording on the spectrogram and once we zoomed in the pulse of interest, we configured the milliseconds (ms) per plot to 100 ms, maximum frequency depending on the pulse bandwidth, threshold to 10, amplitude contrast to 3, Fast Fourier Transformation (FFT) size to 1024 samples to enhance frequency resolution (see Parsons et al. 2000), and FFT window to Hanning. We obtained FMIN and FMAX in kHz (Fig. 1a). Then, we marked the pulse and displayed it on the power spectrum, which we configured with a relative intensity minimum level at -60 dB, and FFT size to 1024 samples. We obtained FMIN and FMAX in kHz, considering four thresholds (Fig. 1c): at -55 and -50 dB from 0 dB, and at 18 and 6 dB below peak frequency. To allow two observers to measure the same pulse with the same criterion, we relied on references on the time axis (pulse location in ms).

We plotted the data for a graphic comparison of measurement criteria and observer’s measurements, and calculated the mean and coefficient of variation. We chose to limit the statistical analyses to minimum frequency because it is an acoustic parameter proposed to be key for differentiating vespertilionid species (see O’Farrell & Miller 1999, Arias-Aguilar et al. 2018). We corroborated normality in the data sets with the Kolmogorov-Smirnov test to proceed with parametric analyses. Then, we performed the following tests to explore for differences between means: 1) standard

---

**Fig. 1** - Echolocation pulses of Myotis nigricans recorded with a Pettersson Ultrasound Detector D1000X (Pettersson Elektronik AB, Sweden), displayed as: a) spectrogram (frequency vs. time), b) oscillogram (amplitude vs. time), c) power spectrum (amplitude vs. frequency). In the spectrogram, we highlight the pulse of interest along with the cursor position for obtaining the minimum frequency (FMIN) and maximum frequency (FMAX), which is subsequently displayed on the oscillogram and power spectrum. On the oscillogram, it can be appreciated that the pulse exhibits a good signal-to-noise ratio (above 20%). On the power spectrum, the arrows are differentiated for minimum frequency (FMIN) and maximum frequency (FMAX), and indicate four thresholds: 1) at -55 and -50 dB from 0 dB, and at 18 and 6 dB below peak frequency.
one-way Analysis of Variance (ANOVA) for measurements on the same pulses with different criteria by each observer, 2) paired t-tests for measurements on the same pulse with the same criterion between observers. Depending on the probability values, we proceeded with paired comparisons based on the Tukey-Kramer (TK) post hoc test. The box plots were elaborated in Statistica v. 8.0 (Statsoft, Inc.), and the calculations were performed in PAST v. 2.17c (Hammer et al. 2001).

RESULTS

The data can be visualized graphically and comparatively between measurement criteria and observer’s measurements in Fig. 2 (FMIN) and 3 (FMAX). For FMIN$_s$ and FMIN$_p$ based on the coefficient of variation (CF), we obtained the following results in descending order of variation: for observer 1, FMIN$_s$ (mean = 47,4 kH, $CF = 6,6$), FMIN$_p$ at -50 dB from 0 dB (mean = 49,4, $CF = 4,4$), FMIN$_p$ at -55 dB from 0 dB (mean = 49,2, $CF = 4,3$), FMIN$_p$ at 18 dB below peak frequency (mean = 50,0, $CF = 3,8$), and FMIN$_p$ at 6 dB below peak frequency (mean = 51,1, $CF = 3,2$); for observer 2, FMIN$_s$ (mean = 45,4, $CF = 6,8$), FMIN$_p$ at -55 dB from 0 dB (mean = 48,3, $CF = 5,4$), FMIN$_p$ at -50 dB from 0 dB (mean = 48,8, $CF = 5,0$), FMIN$_p$ at 18 dB below peak frequency (mean = 49,4, $CF = 4,4$), and FMIN$_p$ at 6 dB below peak frequency (mean = 50,3, $CF = 3,7$).

For the observer 1 parameter’s measurements, we found statistically significant differences ($F = 5,24, p = 0,001$). Based on the TK test paired comparisons, we found statistically significant differences between FMIN$_s$ and FMIN$_p$ at 18 dB below peak frequency ($q = 4,441, p = 0,02$); FMIN$_s$ and FMIN$_p$ at 6 dB below peak frequency ($q = 6,267, p = 0,004$). In both cases, with higher mean values in the power spectrum. For all the other comparisons the probability values indicated no statistically significant differences ($p > 0,05$).

For the observer 2 parameter’s measurements, we found statistically significant differences ($F = 7,932, p = 2,857E-05$). Based on the TK test paired comparisons, we found statistically significant differences between FMIN$_s$ and FMIN$_p$ at -55 dB from 0 dB ($q = 4,412, p = 0,02$); FMIN$_s$ and FMIN$_p$ at -50 dB from 0 dB ($q = 5,169, p = 0,004$); FMIN$_s$ and FMIN$_p$ at 18 dB below peak frequency ($q = 6,066, p = 0,0006$); FMIN$_s$ and FMIN$_p$ at 6 dB below peak frequency ($q = 7,439, p = 0,0001$). In all cases, with higher mean values in the power spectrum. For all the other comparisons the probability values indicated no statistically significant differences ($p > 0,05$).

Between observers 1 and 2, we found statistically significant differences for FMIN$_p$ (mean = 47,4 kH and 45,4 kH, respectively, $difference = 2,0$ kH, $t = 9,274, d.f. = 13, p = 4,281E-07$); FMIN$_s$ at -55 dB from 0 dB (mean = 49,2 kH and 48,3 kH, respectively, $difference = 0,9$ kH, $t = 3,455, d.f. = 13, p = 0,004$); FMIN$_s$ at -50 dB from 0 dB (mean = 49,4 kH and 48,8 kH, respectively, $difference = 0,6$ kH, $t = 3,176, d.f. = 13, p = 0,007$); FMIN$_s$ at 18 dB below peak frequency (mean = 50,0 kH and 49,4 kH, respectively, $difference = 0,6$ kH, $t = 3,518, d.f. = 13, p = 0,003$); FMIN$_s$ at 6 dB below peak frequency (mean = 51,1 kH and 50,3 kH, respectively, $difference = 0,8$ kH, $t = 4,067, d.f. = 13, p = 0,001$).

DISCUSSION

By analyzing the minimum frequency, we provide evidence of significant differences between measurement criteria (spectrogram vs. power spectrum), and no significant differences between thresholds on the power spectrum, so these might be considered equally useful. Between observers, we found the highest difference of measurements (mean’s difference) on the spectrogram, and significant differences in all paired comparisons. It is important to note that, although the differences of measurements between observers were not of great magnitude, the test applied (paired t-test) is highly sensible. The measurements on the power spectrum were less variable compared to those on the spectrogram, as well as the ones relative to the peak frequency which were less variable as they approach the peak compared to those relative to 0 dB.
To our knowledge, there is not much information on this subject in the bat literature, specifically as a source of confusion. Fenton (2002) reports that, for a small sample of pulses (n = 10) of Otomops martienssensi, the difference for minimum frequency and maximum frequency between -55 dB from 0 dB vs. 6 dB below peak frequency, was less than 0.5 kHz. The latter difference seems to be insignificant and nothing to worry about; however, we argue that it can represent a potential confounding factor when attempting identifications, and include some examples below.

We obtained the following minimum and maximum values for observers 1 and 2, respectively: 1) for FMIN: 38.9-51.6 kHz and 37.0-49.2 kHz, 2) for FMIN, including all thresholds: 44.6-54.6 kHz and 42.4-53.6 kHz. We compared our data with measurements reported by MacSwiney et al. (2008; supplementary material), as well as with measurements from recordings provided by C. MacSwiney (pers. comm.), of small vespertilionid bats. MacSwiney et al. (2008) extracted FMIN from full-spectrum recordings with a threshold at 10 dB below peak frequency. We determined that the FMIN of M. nigricans (37.0-51.6) overlapped with the FMIN of Rhogeessa aeneus (free-flying bats, n = 13 pulses, 42.7-51.9) and Eptesicus furinalis (free-flying bats, n = 22 pulses, 33.0-39.5), so the variation is large enough to represent a confounding factor. On the other hand, the FMIN of M. nigricans (42.4-54.6) did not overlap with E. furinalis. Rhogeessa aeneus is a lowland species and E. furinalis has been reported along with M. nigricans (see Williams-Guillén & Perfetto 2011); a similar situation could happen with potential sympatric species of Rhogeessa and E. brasiliensis.

Regarding Myotis species, the FMIN of M. pilosatibialis (8 captured bats, n = 8 pulses, 52.9-59.1) overlapped with the FMIN of M. nigricans, while the FMIN of M. pilosatibialis (free-flying bats, n = 48 pulses, 55.4-61.8) did not overlap with the FMIN of M. nigricans. We also compared our data to measurements of M. velifer obtained by Rizo-Aguilar (2008) (6 captured bats, n = unreported pulses, mean = 38.6, s.d. = 2.7), which is considered a large size representative in this genus, specifically FMIN, with a threshold at 25 dB below peak frequency. We determined that the FMIN of M. nigricans encompassed the mean FMIN, of M. velifer, while the FMIN of M. nigricans did not.

It is important to recognize that the samples are small, and this could have an effect on the calculations and the possibility of more robust data (B. Miller, pers. comm.); however, we consider it important evidence that should be taken into account. There is also the possibility that the hand-release technique for bat recordings might affect pulse features vs. free-flying bat recordings (B. Miller, pers. comm.). Unfortunately, we had difficulties finding published comparable data of other potential sympatric Myotis species distributed in the area.

To decide what measurement criterion is adequate, it is important to define the level of precision required depending on the objectives of the research. For instance, the acoustic differentiation of very similar species relies on high-resolution data. In other situations, this is not critical and even potential differences between measurement criteria and/or observers will not impede identifications. Another consideration is long-term and large-scale vs. short-term and local acoustic surveys, which produce different amounts of data. The processing of large amounts of data has to be time-efficient, and manual measurements might not be practical. Anecdotally, we experienced a variation on the time spent with the different measurement criteria, which was higher on the power spectrum because it involves marking the pulse on the spectrogram and subsequent change of display. However, it should be noticed that as an alternative to manual analysis, the classifiers incorporated in many software reduce workload by filtering noise, and also assign identities to recordings and provide acoustic parameters (Skowronska & Fenton 2009, Lemen et al. 2015, Rydell et al. 2017). Nevertheless, the automated methods should be used with caution and it is important to validate the output (performance level), as well as determine confidence levels (Rydell et al. 2017).

Some limitations should be acknowledged based on the characteristics of the pulse and recording to proceed with analyses on the power spectrum. For instance, if the intensity level of the pulse of interest is too low it may not reach the desired threshold, and if the recording exhibits a low signal-to-noise ratio then background noise might mask the pulse of interest (Surlykke & Moss 2000, Fenton 2002). Also, it is important to note that the measurements on the power spectrum are not faithful to the “real” frequency emitted by the bat, as these are truncated. The measurements on the spectrogram can be closer to the real frequency, as well as less impacted by the signal-to-noise ratio. This is because the high-amplitude noise often exhibits lower spectral density and a lower frequency than the pulse of interest (see Parsons et al. 2000); however, it is also expected the measurement criteria will have attributes such as inter-observer bias, therefore representing a trade-off.

In bat acoustic research, it is common to make use of references that include spectrograms to aid species identification; however, it can be inadequate to do so for recordings obtained with detection devices that incorporate different transformation techniques for obtaining the spectral content of the pulse, and hence produce signals with differences in the information retained (see Parsons et al. 2000). These kinds of comparisons are only appropriate for species that can be identified with a high degree of confidence, such as those that exhibit vocal signatures that are characterized by low structure variability (O’Farrell & Miller 1999, O’Farrell et al. 1999). As mentioned before, even when considering the same detection device and software, the measurement criteria can be different during post-processing. Moreover, we have found published articles that omit details about the measurement criteria and measurements itself, and/or do not provide enough detail about the configuration settings. For instance, a basic consideration is that the appearance of the pulse is affected by the windowing on the spectrogram (see Parsons et al. 2000, Avila-Flores & Fenton 2005, MacSwiney et al. 2009), and this can have important implications on identifications.

Here, we found that the measurements relative to the peak frequency are less variable and this should be an important criterion depending on the research objectives.
as well as to reduce inter-observer bias. The criteria selection should also be based on supporting references. We recommend including detailed information on the software configuration (see methods section), measurement criteria, and basic descriptive statistics (mean, coefficient of variation, minimum-maximum), to allow cross-studies comparisons. The acoustic analysis should always be complemented with the qualitative approach (e.g. structure of the pulse and diagnostic features) and coupled with time-based parameters (e.g. duration of the pulse and interpulse interval). Finally, we include an important premise by Fenton (2002: 138): “clearly articulate the procedures you followed to collect data and the criteria you used in the analysis to ensure that others can replicate your study”.

ACKNOWLEDGEMENTS

We would like to thank Cristina MacSwiney for providing recordings, and Bruce Miller for valuable comments. We are very grateful to two anonymous reviewers for their suggestions, which helped improve this manuscript substantially.

REFERENCES

ADAMS, A. M., JANTZEN, M. K., HAMILTON, R. M. & FENTON, M. B. (2012). Do you hear what I hear? Implications of detector selection for acoustic monitoring of bats. *Methods in Ecolology and Evolution*, 3(6): 992-998. https://doi.org/10.1111/j.2041-210X.2012.00244.x

ARIA-AGUILAR, A., HINTZE, F., AGUIAR, L. M. S., RUFRAV, V., BERNARD, E. & RAMOS PEREIRA, M. J. (2018). Who’s calling? Acoustic identification of Brazilian bats. *Mammal Research*, 63: 231-253. https://doi.org/10.1007/s13364-018-0367-z

AVILA-FLORES, R. & FENTON, M. B. (2005). Use of Spatial Features by Foraging Insectivorous Bats in a Large Urban Landscape. *Journal of Mammalogy*, 86(6): 1193-1204. https://doi.org/10.1644/04-MAMM-A-08SR1.1

BARCLAY, R. M. R., FULLARD, J. H. & JACOBS, D. S. (1999). Variation in the echolocation calls of the hoary bat (*Lasiurus cinereus*): influence of body size, habitat structure, and geographic location. *Canadian Journal of Zoology*, 77(4): 530-534. https://doi.org/10.1139/z99-008

BISCARDI, S., ORPRCIEJO, J., FENTON, M. B., TSOAR, A. & RATCLIFFE, J. M. (2004). Data, Sample Sizes and Statistics Affect the Recognition of Species Of Bats by Their Echolocation Calls. *Acta Chiropterologica*, 6(2): 347-363. https://doi.org/10.3161/001.006.0212

FENTON, M. B. (2000). Choosing the ‘correct’ bat detector. *Acta Chiropterologica*, 2(2): 215-224.

FENTON, M. B. (2002). Reporting: essential information and analysis. In: Bat echolocation research: tools, techniques and analysis, ed.: Bat Conservation International. Austin, Texas, USA, p.133-149.

FENTON, M. B., JACOBS, D. S., RICHARDSON, E. J., TAYLOR, P. J. & WHITE, W. (2004). Individual signatures in the frequency-modulated sweep calls of African large-eared, free-tailed bats *Otomops martiensseni* (Chiroptera: Molossidae). *Journal of Zoology*, 262(1): 11-19. https://doi.org/10.1017/S095283690300431X

HAMMER, Ø., HARPER, D. A. T. & RYAN, P. D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 4(1): 1-9.

KRAKER-CASTAÑEDA, C., SANTOS-MORENO, A., LORENZO, C. & MACSWINEY, M. C. (2019). Effect of intrinsic and extrinsic factors on the variability of echolocation pulses of *Myotis nigricans* (Schinz, 1821) (Chiroptera: Vespertilionidae). *Bioacoustics*, 28(4): 366-380. https://doi.org/10.1080/09524622.2018.1461685

Lemen, C., Freeman, P. W., White, J. A. & Andersen, B. R. (2015). The Problem of Low Agreement among Automated Identification Programs for Acoustical Surveys of Bats. *Western North American Naturalist*, 75(2): 218-225. https://doi.org/10.3398/0067.05210

MACSWINEY, M. C., CLARKE, F. M. & RACEY, P. A. (2008). What you see is not what you get: the role of ultrasonic detectors in increasing inventory completeness in Neotropical bat assemblages. *Journal of Applied Ecology*, 45(5): 1364-1371. https://doi.org/10.1111/j.1365-2664.2008.01531.x

MACSWINEY, M. C., BOLÍVAR CIMÉ, B., CLARKE, F. M. & RACEY, P. A. (2009). Insectivorous Bat Activity at Cenotes in the Yucatan Peninsula, Mexico. *Acta Chiropterologica*, 11(1): 139-147. https://doi.org/10.3161/15081109X465758

O’FARRELL, M. J. & MILLER, B. W. (1999). Use of Vocal Signatures for the Inventory of Free-flying Neotropical Bats1. *Biotropica*, 31(3): 507-516. https://doi.org/10.1111/j.1744-7429.1999.tb00394.x

O’FARRELL, M. J., MILLER, B. W. & GANNON, W. L. (1999). Qualitative Identification of Free-Flying Bats Using the Anabat Detector. *Journal of Mammalogy*, 80(1): 11-23. https://doi.org/10.2307/1383203

OBRIST, M. K. (1988). Individually recognizable freeflying bats: a new method to record and analyze their echolocation calls. *Myotis*, 26: 87-95.

PARSONS, S., BOONMAN, A. M. & OBRIST, M. K. (2000). Advantages and Disadvantages of Techniques for Transforming and Analyzing Chiropteran Echolocation Calls. *Journal of Mammalogy*, 81(4): 927-938. https://doi.org/10.1644/1545-1542(2000)081<0927:AADOTF>2.0.CO;2

Rizo-Aguilar, A. (2008). Descripción y análisis de los pulsos de ecolocación de 14 especies de murciélagos insectívoros aéreos del estado de Morelos. MS thesis, Instituto de Ecología A. C. Xalapa, Veracruz, México.
RYDELL, J., NYMAN, S., EKLÖF, J., JONES, G. & RUSSO, D. (2017). Testing the performances of automated identification of bat echolocation calls: A request for prudence. Ecological Indicators, 78: 416-420. https://doi.org/10.1016/j.ecolind.2017.03.023

RUSSO, D., ANCILLOTTO, L. & JONES, G. (2018). Bats are still not birds in the digital era: echolocation call variation and why it matters for bat species identification. Canadian Journal of Zoology. 96(2): 63-78. https://doi.org/10.1139/cjz-2017-0089

SKOWRONSKI, M. D. & FENTON, M. B. (2009). Detecting Bat Calls: An Analysis of Automated Methods. Acta Chiropterologica, 11(1): 191-203. https://doi.org/10.3161/150811009X465811

SURLYKKE, A. & MOSS, C. F. (2000). Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory. The Journal of the Acoustical Society of America, 108(5): 2419-2429. https://doi.org/10.1121/1.1315295

WILLIAMS-GUILLÉN, K. & PERFECTO, I. (2011). Ensemble Composition and Activity Levels of Insectivorous Bats in Response to Management Intensification in Coffee Agroforestry Systems. PLoS ONE, 6(1): e16502. https://doi.org/10.1371/journal.pone.0016502