

Identification of individual Eastern Screech-Owls *Megascops asio* via vocalization analysis

Christopher M. Nagy^{a,b,c,*} and Robert F. Rockwell^{a,b}

^aDivision of Vertebrate Zoology, American Museum of Natural History, Central Park West, New York, NY 10024, USA; ^bBiology Department, City College, City University of New York, 160 Convent Ave, New York, NY, 10031 USA; ^cMianus River Gorge Preserve, 167 Mianus River Rd, Bedford, NY, 10506, USA

(Received 26 July 2011; final version received 11 October 2011)

To more easily and non-invasively monitor urban Eastern Screech-Owl populations, we developed a method of distinguishing individual owls using their calls. A set of seven variables derived from recordings of ‘bounce’ calls taken from 10 known (either free-ranging birds recorded at a single site on a single night or identifiable captive owls) owls was tested using a model-based clustering analysis (Mclust) as a method of discriminating individual owls. The cluster analysis correctly classified these calls with 98% accuracy. A second set of calls from nine owls was used to further test the method and correctly classified 84% of the calls using the same variables. Four owls were recorded repeatedly from 2008 to 2010 to determine the extent to which calls changed over time; the cluster analysis correctly assigned 89% of the calls to the correct owl regardless of the year the recordings were made. Based on these results, we are confident that the Mclust analysis can be used to reliably and safely estimate abundance and survival of Eastern Screech-Owls within the time frame of a few years and of population sizes <15 owls.

Keywords: cluster analysis; Eastern Screech-Owl; *Megascops asio*; spectrograms; vocalization analysis

Introduction

The Eastern Screech-Owl (*Megascops asio*) is a small raptor that inhabits mixed-hardwood forests in the United States and southern Canada east of the Rocky Mountains. This species can tolerate some human development and can be found in suburban and urban parks, golf courses, and other semi-developed greenspaces (Gehlbach 1995). Screech-Owls are one of the few raptors that can persist in small urban parks (Lynch and Smith 1984; Smith and Gilbert 1984; Gehlbach 1995), and managers would benefit from knowledge about their survival rates and small-scale habitat use in these areas. As part of a larger study on urban Screech-Owls in the New York City (NYC) metropolitan area, we sought to develop a non-invasive method to identify and monitor individual Screech-Owls to estimate abundance and adult survival.

Screech-Owls are difficult to trap and there are concerns regarding the behavioural effects of telemetry (Gehlbach 1994). They have been monitored successfully via nest boxes for long-term studies in Texas (Gehlbach 1994) and Ohio (VanCamp and Henny 1975), but daily sampling of such boxes can be time-consuming and when we attempted such a survey

*Corresponding author. Email: cnagy@gc.cuny.edu

in NYC we did not capture a sufficient number of owls. Screech-Owls defend territories throughout the year, especially from spring through late summer (Ritchison et al. 1988), and like many other owls announce their presence to rivals via vocalizations. They also attract mates and communicate to mates and offspring via frequent vocalizations, and will readily respond to call-playback broadcasts (Lynch and Smith 1984; Dorn and Dorn 1994; Bosakowski and Smith 1997). If a method of identifying individual Screech-Owls via vocalization analysis could be developed, then current call surveys could yield mark-recapture data as well as site occupancy information.

Attempts at developing a method to discriminate individuals based on vocalizations has been successful in numerous species of birds (Corncrakes *Crex crex*: Peake et al., 1998; Barred Owls *Strix varia*: Freeman 2000; Wood Owls *Strix woodfordii*: Delpont et al. 2002; Great Bitterns *Botaurus stellaris*: Gilbert et al. 2002; Western Screech-Owls *Megascops kennicottii*: Tripp and Otter 2006; Woodcock *Scolopax rusticola*: Hoodless et al. 2008; Willow Flycatcher *Empidonax trailii extimus*: Fernandez-Juricic et al. 2009; summarized by Terry and MacGregor 2002) as well as a few mammals (male Fallow Deer *Dama dama*: Reby et al. 1998; Swift Fox *Vulpes velox*: Darden et al. 2003; Wild Dog *Lycaon pictus*: Hartig 2005). If a reliable method of discerning individuals based on their vocalizations can be found, researchers can non-invasively monitor otherwise cryptic or difficult-to-sample species, often for a fraction of the cost, effort, and negative effects associated with other methods (Terry et al. 2005; Hoodless et al. 2008; Fernandez-Juricic et al. 2009). To be truly effective, however, the vocalizations must be of consistent form so that a set of variables can be repeatedly measured from them. These variables should exhibit low within- relative to among-individual variation (Terry et al. 2005). In addition, an animal's vocalizations (and derived variables) should ideally stay consistent over time so that individuals can be tracked over many years (Delpont et al. 2002; Terry and MacGregor 2002; Terry et al. 2005).

To develop and test a method to census and monitor Screech-Owls with no previous knowledge regarding abundance, we used calls recorded from captive owls housed at rehabilitation clinics and free-living (i.e. wild) owls sampled in disparate locations to build a large set of recordings of 'known' individual owls. We measured a number of variables from the 'bounce' call (Cavanagh and Ritchison 1987; Gehlbach 1995) and assessed their usefulness as individual markers. Using a model-based cluster analysis, we classified the recordings from half of the owls, and then re-tested the analysis on the other half using the same variables. We then clustered the recordings from a subset of owls that were recorded over the course of 2 or 3 years to determine the extent of change in calls over time.

Methods

Eastern Screech-Owls are typically thought to have two calls that are used as broadcast vocalizations. The 'whinny' call is a territorial call and general alarm call (Cavanagh and Ritchison 1987; Gehlbach 1994) and can be quite variable even within a single bout (Figure 1A). The 'bounce' call generally consists of a series of quickly repeated notes on a steady pitch. It is also used as a territorial call as well as communication between mates and between parents and offspring (Cavanagh and Ritchison 1987; Gehlbach 1994). Thus, we thought the bounce call would be the most likely to contain information that was individual-specific, as also recommended by Cavanagh and Ritchison (1987). While gathering and analysing our recordings, we found that there appeared to be two forms of the bounce call: a 'long bounce' which could range from approximately 5 seconds to as long as 45 seconds and was delivered at a steady note rate and frequency (Figure 1B); and a 'short bounce' (Figure 1C) which was approximately 2–4 seconds long and had three distinct phases where

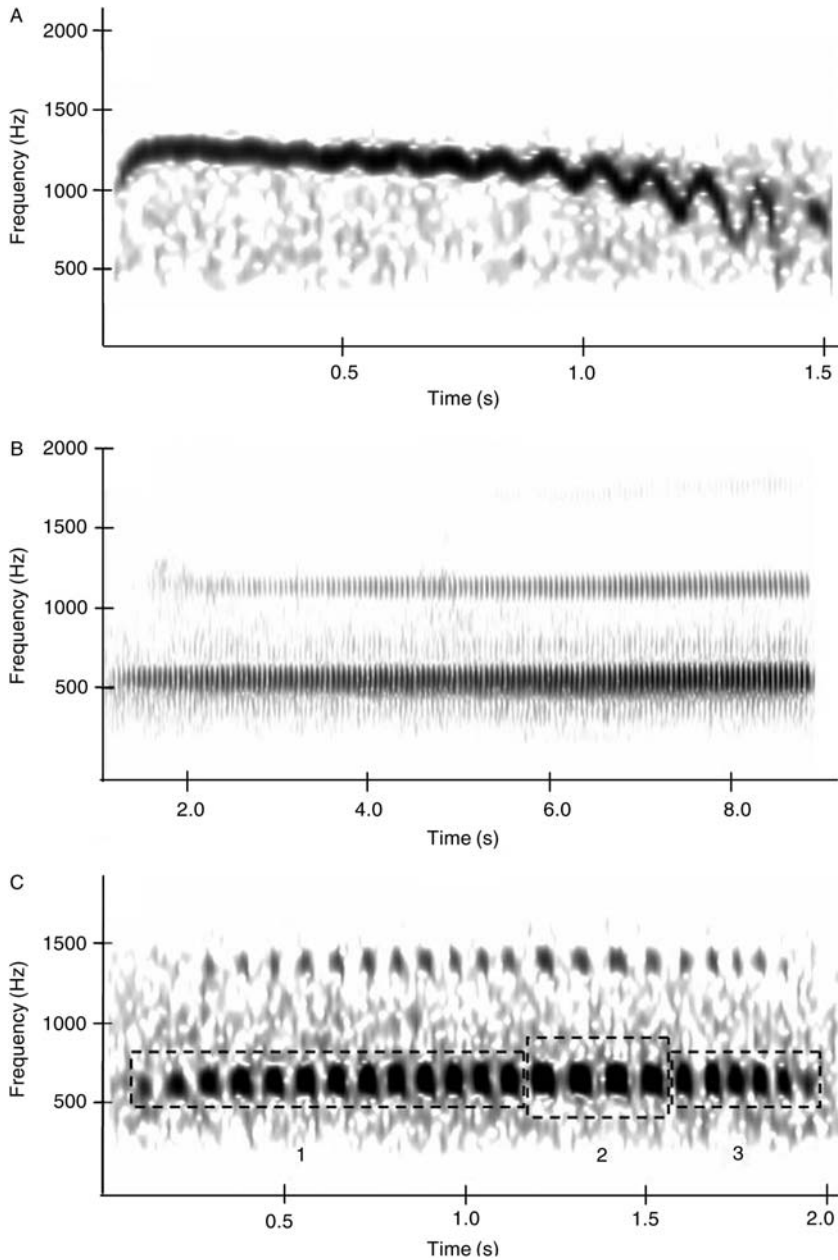


Figure 1. Sample spectrograms of whinny (A), long bounce (B), and short bounce (C) calls of Eastern Screech-Owl. Selections 1, 2, and 3 in C represent the three phases of the call.

the note length and the time between each note changed. The first and third phases had substantially faster note rates than the middle (second) phase, and the phases could be easily identified visually on a spectrogram and/or by listening to the call at 0.4 speed. Occasionally the frequency of the short bounce changed slightly from phase to phase.

In our experience, the long bounce was used less often than the short bounce in response to broadcast surveys. If an owl was heard calling independently (i.e. not in

response to our broadcasts) it always vocalized long bounces. When an owl used long bounces in response to our broadcasts, in most cases they eventually switched to a short bounce. Also, when we observed owls calling to fledglings in the late spring, only short bounces were used. Thus, because the different phases of the short bounce would allow more variables and more variability to be measured, and because of the use of the short bounce in parent-offspring communication and in response to our broadcasts, we thought the short bounce had the best potential as an individual identifier and as a tool to monitor Screech-Owls via call-playback surveys.

Three sets of Eastern Screech-Owls were recorded during this study. The first set of ten owls was recorded between April and December 2008 and was used for the initial model development ('build set'). Six of these were free-living owls recorded at parks and preserves in New York State: Mianus River Gorge Preserve, Bedford; Ward Pound Ridge, Pound Ridge; Harriman State Park, Rockland and Orange Counties; and Saxon Woods Park, White Plains. In these cases we only used recordings from one owl at each site, or recordings from two owls that were recorded simultaneously, to ensure that each free-living owl was in fact a separate individual. The remaining four owls were permanent captive birds at rehabilitation centres in New York, New Jersey, and Pennsylvania and thus could be identified. The second set of nine owls ('test set') was recorded in 2008 and 2009 to test the method on independent recordings. Two of the captive owls in the build set were re-recorded 3 months later in the same season and used in the test set. The six free-living owls in the test set were recorded at the Mianus River Gorge Preserve and Riverdale, Van Cortland, and Inwood Parks, NYC. In addition, we were able to record four owls repeatedly from 2008–2010 to determine if owls' calls changed across years ('multi-year set'). One of these, a free-living owl, was recorded in June 2008, April 2009, and June 2010. We were reasonably certain that this owl was a single individual because of the reliability with which we could find it and elicit calls at the same location and the distinctive timbre of its calls (this owl was actually the inspiration for investigating this method). Two captive owls were recorded in May 2008, December 2008, December 2009, and one from this pair was recorded again in October 2010 (the second died in early 2010). A final captive owl was recorded in December 2009 and November 2010. Overall, we recorded 265 calls from 17 owls: 10 unique owls were used in the build set, seven unique owls plus new calls from two build set owls were used in the test set, and two owls from the build set and two owls from the test set were used in the multi-year set.

Captive and free-living owls were recorded after dark using a Sennheiser ME67 shotgun microphone with a foam windscreen and a Marantz PMD 661 digital recorder at a 44.1 kHz sampling rate. Vocalizations were elicited via broadcasting a mixture of alternating bounce and whinny calls (Stokes et al. 1997) with a portable CD player. The entire bout was recorded and we used as many calls from each bout as possible. Some calls were censored if background noise (car traffic, airplanes, trains, police sirens, other wildlife, etc.) made it impossible to measure frequency or note variables. We converted all recordings to spectrograms and measured variables on usable short bounce calls using Raven 1.3 (Cornell Lab of Ornithology 2008). The discrete Fourier transform (DFT) used by Raven 1.3 to generate spectrograms from waveforms must be parameterized by a DFT size that determines the number of discrete frequency-amplitude measurements plotted on the spectrogram from the waveform. This value was held constant at the highest value of 65,536 samples (0.732 Hz grid size). Spectrogram transformation also requires a parameter called window size that determines how precisely the spectrogram will measure frequency, i.e. the bandwidth of the frequency filters. Frequency changes less than the chosen bandwidth will not be discernible by the DFT. There is a trade-off between

frequency and time resolution: a small window size will provide high resolution on the temporal scale and low frequency resolution, while a large window will provide high frequency but poor temporal resolution (Charif et al. 2008). We tested three common window sizes with our build dataset to determine which was optimal for discriminating owls. To do this we measured all frequency-based variables from spectrograms built with windows of 256, 512, and 1024 samples. Temporal variables were measured directly on the waveform when possible or on the smallest window size.

We measured the number of notes, duration, centre frequency (CF; the frequency that divides the call into two frequency intervals of equal energy), first quartile frequency (1QF; the frequency that divides the call into two intervals that contain 25% and 75% of the energy in the call), third quartile frequency (3QF; the frequency that divides the call into two intervals that contain 75% and 25% of the energy in the call), the interquartile range (the frequency difference between 3QF and 1QF), and the note rate (NR). These measurements were taken on the entire call and each of the three phases of the short bounce. We also calculated the proportion of total notes and the proportion of total duration in each phase of the short bounces (28 variables in total). As an initial index of the amount of variation within individuals compared to the variation amongst individuals, we calculated the proportion for individuality coding (PIC; Sokal and Rohlf 1995), which, for a given variable, is the coefficient of variation for the total set of measurements divided by the average of the coefficients of variation for each individual. If the ratio of these CVs is greater than 1, then there is likely more variation amongst individuals than there is within them, and the variable can potentially be used as a predictor of individual identity (Robisson et al. 1993). Within-owl CV was calculated by $(SD_i/\bar{x}_i) \times (1 + 1/4n_i) \times (100)$, where SD_i , \bar{x}_i , and n_i are the standard deviation, means and number of calls for owl i , respectively. Total CV for the entire sample was calculated by $(SD/\bar{x}) \times 100$. PIC was calculated by total CV divided by the average within-owl CV (Sokal and Rohlf 1995; Charrier et al. 2004).

Owls called at approximately 650 Hz (centre frequencies ranged from 516.4 to 1051.8 Hz across all owls) and 14.5 notes/second. After eliminating uninformative variables first with PIC and then iteratively to maximize cluster accuracy with the build set, the final clustering variables (regardless of choice of window size) were centre frequency (CF), first- and third-quartile frequencies (1QF and 3QF), the note rate of the entire call (NRall), and the note rates for the each call phase (NR1, NR2, and NR3). These variables were standardized and entered into a model-based cluster analysis using the Mclust package (Fraley and Raftery 2007) for R. This agglomerative clustering method considers clusters (in this case, individual owls) as multivariate normally distributed components in a mixture, and can estimate the total number of clusters (G) by finding the maximum likelihood estimate for G given a range of possible clusters. Models are then ranked with Bayesian Information Criterion (BIC) to determine which model best fits the data without overfitting. In addition, models can be parameterized to allow for varying volumes, shapes, and orientations among clusters (Fraley and Raftery 2007; Xu and Wunsch 2009), although this adds additional parameters to the model and thus penalizes the model's BIC. We did not use priors for modelling (Fraley and Raftery 1998; Fraley and Raftery 2007). Since we were interested in a method that would estimate the number of animals from a set of calls without any prior knowledge of G , an approach that provided estimates of G and associated likelihoods as well as assigning all observations to clusters was necessary. Group membership likelihoods for each observation were also calculated to assign observations (calls) to clusters (owls).

The final set of spectrogram variables was reduced to the set that gave the highest accuracy via Mclust in estimating G and assigning individual calls to the correct cluster. When these variables were determined using the build set, the same variables were measured from calls of the test set (using the optimal spectrogram window size in Raven 1.3) and entered into the cluster analysis to determine if the variables performed well on an independent set of recordings (i.e. was our variable set generally applicable, or specific only to data in the build set?). Next, to informally examine the maximum number of owls that could be discriminated, we ran a pooled dataset of the entire build and test sets together (17 owls in total). In particular, we were interested in whether the calls from the two owls found in both the build and tests sets would cluster together despite being recorded at different times in the season in this large dataset.

Lastly, the calls from the four owls that were recorded repeatedly across years were entered as a third dataset. In this analysis we constrained the model structure to components with equal shape and volume. Calls that were quite different from each other yet occupied otherwise ‘sparse’ areas of the dataspace (e.g. the high or low extremes) might cluster together if clusters were allowed to be very large or take alternate shapes. Restricting volume and shape ensured that the Mclust analysis would not group calls together that were in reality quite different from each other. The trade-off for this constraint was to risk over-estimating the number of owls (clusters) by assuming owls have similar variation in calls. If the respective calls from each of these owls clustered together across years, then we could have some confidence that the method could be used to track owls from year to year (at least to a maximum of three years). If this was not the case, the method might still be useful in obtaining a ‘snapshot’ abundance from year to year but could not be used to monitor individual owls (e.g. for annual survival estimation) over long time periods.

Results

The cluster analysis, using data measured from spectrograms with a 256 sample window size, correctly assigned all but two of the 88 calls (98%) in the build set to the correct owl using a model with ellipsoidal components of equal volume, shape, and orientation (‘EEE’ Table 1, Figure 2). The other two spectrogram window sizes did not perform as well: data measured with a window size of 512 samples yielded a BIC-selected best model with the correct number of clusters (10) but used a more complicated cluster structure with variable volume and orientation (‘VEV’). Measuring call variables with a window of 1024 samples yielded a BIC-selected best model with 20 clusters (twice as many owls as there actually were). Subsequent measurements were therefore derived from spectrograms with window sizes of 256 samples.

The analysis using the test set selected the same model form of ‘EEE’ but was slightly less accurate (Table 1), with a BIC-selected best model of 10 components (Figure 3). The correct number of owls was nine, not 10; however, the extra cluster was made up of only two observations from two different owls and thus could easily be identified and removed by looking at the classifications of individual calls. The model correctly classified 53 out of 63 calls (84%). When all calls from both sets were pooled together, they were classified correctly with 84% accuracy, again with an ‘EEE’ model structure. The model properly classified the two re-sampled owls that were present in the original build and test sets. However, the method predicted one extra cluster (18 owls) than was truly present, similarly to the test set alone.

Of the 114 calls from the four owls recorded in 2008–2010, 102 were clustered with the correct owl (89%; Tables 1 and 2) using a model with ellipsoidal components of

Table 1. BIC scores for initial (build set) model-based clustering of Eastern Screech-Owl bounce calls, New York City, Long Island, and Westchester, NY, and Millington, NJ, 2008–2010.

Dataset	Model ^a	d ^b	G ^c	BIC ^d	Log-likelihood
Build	EEE, 10 clusters	7	10	336.89	403.55
	EEE, 11 clusters	7	11	316.22	410.79
	EEE, 12 clusters	7	12	302.11	421.31
Test	EEE, 10 clusters	7	10	-84.81	179.25
	EEE, 11 clusters	7	11	-109.77	183.34
	EEE, 12 clusters	7	12	-116.99	196.31
Multi-year	EEV, 4 clusters	7	4	-611.91	-17.05
	EEV, 3 clusters	7	14	-677.98	-118.76
	EEV, 6 clusters	7	11	-734.86	-9.84

^aEEE indicates a model with ellipsoidal components of equal shape, volume, and orientation; EEV indicates a model with ellipsoidal components of equal shape and volume and variable orientation;

^bnumber of dimensions (variables);

^cnumber of components (clusters) estimated by the model;

^dBIC = (2 * Log-Likelihood) - (k * log(n)); see Fraley and Raftery 2007:

for EEE models, $k = (G * d) + (G - 1) + [(d * (d + 1))/2]$;

for EEV models, $k = 1 + (d - 1) + G * ((d * (d - 1))/2)$

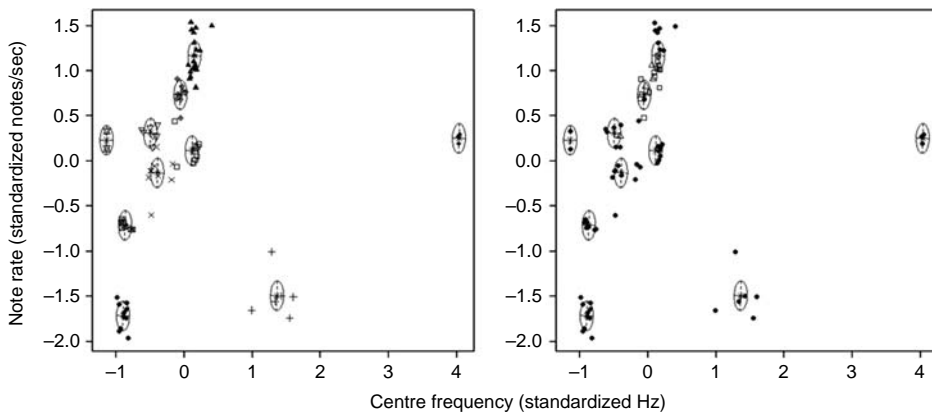


Figure 2. Final clusters and covariance ellipses of the build set of 10 Eastern Screech-Owl short bounces calls plotted along two of the seven variables used in the analysis. Cluster assignments (left) are based on the BIC-selected best model with 10 components. Black dots in uncertainty plot (right) represent observations in the 0–75% quantile of uncertainty; square dots represent observations in the 75–95% quantile, and triangular dots are observations in the 95% quantile of uncertainty.

consistent shape and volume but variable orientation (‘EEV’; Figure 4). The calls from each owl seemed to change only slightly from year to year, with some indication that calls get lower in pitch and somewhat slower with time.

Discussion

The Mclust clustering algorithm performed well, clustering two independent sets of data and discriminating individual owls over a few years. Vocal individuality had been found for a number of owl species (Galeotti and Pavan 1991; Galeotti et al. 1993; Freeman 2000; Delport et al. 2002; Tripp and Otter 2006) and this is not surprising as aural communication and identification would likely be important for nocturnal birds. The classification

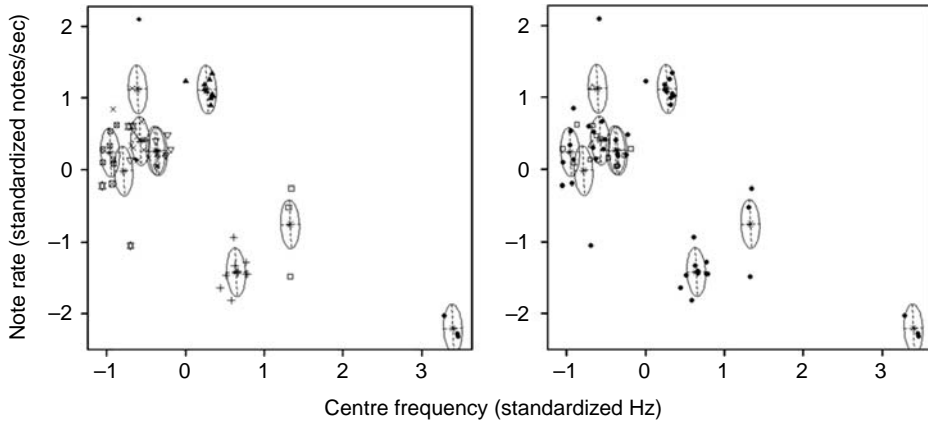


Figure 3. Final clusters and covariance ellipses of the test set of nine Eastern Screech-Owl short bounces calls plotted along two of the seven variables used in the analysis. Cluster assignments (left) are based on the BIC-selected best model with 10 components. Black dots in uncertainty plot (right) represent observations in the 0–75% quantile of uncertainty; square dots represent observations in the 75–95% quantile, and triangular dots are observations in the 95% quantile of uncertainty.

Table 2. Cluster designations, number of calls per owl/season, and number of calls misclassified per owl/season using a 4-cluster model with ellipsoid clusters of consistent shape and volume and variable orientation (EEV), New York and New Jersey, 2008–2010.

Owl ^a	Spring '08 ^b	Winter '08 ^b	Spring '09 ^b	Winter '09 ^b	Spring '10 ^b	Winter '10 ^b
RT1 (C)	A(2, 0)	A(9, 0)	NS	A(17, 0)	NS	NS
RT2 (C)	B(26, 0)	B(10, 0)	NS	B(5, 0)	NS	B(3, 0)
SC1 (C)	NS	NS	NS	C(10, 4)	NS	C(6, 2)
VC1 (F)	D(4, 0)	NS	D(16, 1)	NS	D(6, 3)	NS

^a 'C' indicates a captive owl; 'F' indicates a free-living owl;

^b Cluster designation (total number of calls, number of misclassified calls); NS, the owl was not sampled in that season.

accuracies of this method (85–98%) are comparable to other vocalization-based methods for other species (Freeman 2000; Delpont et al. 2002; Gilbert et al. 2002; Tripp and Otter 2006; Hoodless et al. 2008; Fernandez-Juricic et al. 2009). Traditionally, discriminant function analysis has been used to categorize observations into groups, but the groups must be known and fixed (e.g. male/female, known species or subspecies, age class, etc.). For wildlife monitoring or abundance estimation, discriminant function analysis is thus of limited use because, first, the number of groups is often unknown and, second, the groups (individual animals) being measured disappear and appear over time as individuals die or emigrate and are born or immigrate. Using discriminant function analysis, observations from new individuals would be assigned to the most similar starting cluster, not assigned to a new group. Model-based clustering allows for classification of observations as well as maximum likelihood estimation of the number of groups.

We were initially surprised that the smallest spectrogram window size provided the most useful data, as greater window size should provide more precise frequency measurements. However, when we compared measurements taken across the three window sizes, only Q1F and Q3F appeared to vary substantially (Table 3). As window size increased, the quartile frequency measurements moved closer to the centre frequency. This may have caused data

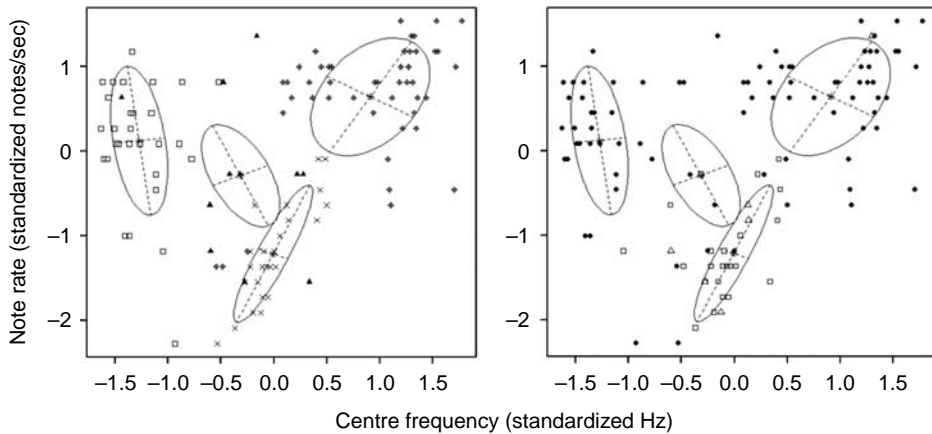


Figure 4. Final clusters and covariance ellipses of the multi-year set of four Eastern Screech-Owl short bounces calls plotted along two of the seven variables used in the analysis. Cluster assignments (left) are based on the BIC-selected best model with four components. Black dots in uncertainty plot (right) represent observations in the 0–75% quantile of uncertainty; square dots represent observations in the 75–95% quantile, and triangular dots are observations in the 95% quantile of uncertainty.

points to ‘constrict’ across these two frequency variables and thus pull away from otherwise similar points. This could lead to a greater number of clusters being predicted by the Mclust routine since only the most similar points remained close to each other in multidimensional space. Indeed, the top two models using the largest window size had 20 and 19 clusters (although the 10 cluster ‘EEE’ model had the third best BIC). When the classifications given by the 20-cluster model were examined, the extra clusters were in fact wholly contained within individual owls’ bouts.

We were particularly concerned with developing a method that uses variables that can be reliably measured even in sub-optimal recording conditions, namely near roads, major highways and busy NYC flyways. In a field setting, where background noise and inconsistent recording conditions are a reality, measures such as call duration, raw number of notes, or upper harmonics are often unreliable because they can be recorded poorly. A few notes at the beginning or end of a call may not be sufficiently recorded, so call duration, numbers of notes, and measures taken on a specific start or ending note can vary not by individual animal but by recording conditions. Often, animals may not be close enough for the recording equipment to pick up harmonics and other faint characteristics. Note rate, however, requires only a few notes in each component of the call, and a frequency measured across the total duration of the call can be calculated reliably with only the middle and the loudest portion of the call – provided the call generally remains at a steady frequency, as is

Table 3. Frequency measurements (mean and SD) using three different window sizes for spectrogram production of calls from 10 eastern screech owls, New York and New Jersey, 2008– 2010.

Window size (samples)	Centre frequency (Hz)	First quartile frequency (Hz)	Third quartile frequency (Hz)
256	644.1 ± 100.6	569.2 ± 99.8	719.9 ± 101.7
512	644.8 ± 100.5	603.7 ± 99.6	685.3 ± 101.8
1024	645.4 ± 100.9	619.7 ± 101.2	670.1 ± 100.5

the case with Screech-Owls' bounce calls. Additionally, the use of energy-based frequency measures available in the Raven software was more accurate – compared to measuring the maximum, middle, and minimum by hand – because small discrepancies in the spectrogram selections of call boundaries do not substantially affect the resulting frequency calculations.

We were able to distinguish 10 and nine individual owls with >85% accuracy. As more individuals are added to an analysis, one would expect greater and greater amounts of overlap between clusters, and eventually discrimination among individuals would become difficult. When the two datasets were pooled (17 owls total), the method did cluster the two identical owls together across separate bouts, but the overall classification rate dropped below 85% and the number of clusters was over-estimated by one owl. This suggests that clusters may begin to overlap excessively around 15 or 16 owls. However, any identification method need only have the capacity to discriminate up to the maximum number of individuals that would reasonably be expected to inhabit an area of interest. For urban parks in NYC with rather small, fragmented woodlands (e.g. the wooded areas of Inwood, Riverdale, and Van Cortland are 55 ha, 45 ha, and 362 ha, respectively), one could expect to encounter more than 10–15 owls in only the largest sites. Proper classification will also depend on the particular individuals that are sampled. Most centre frequencies hovered around 570–620 Hz, with two individuals calling above 950 Hz on average. Owls that call at very high frequencies will be more distinguishable than those who call within the 'average' range of 570–620 Hz. Cavanagh and Ritchison (1987) observed that female eastern screech owls generally call at higher frequencies and at slower rates than males. Unfortunately, we did not know the sexes of the owls we recorded. However, our frequencies tended to be lower on average (654 Hz) than both males (721 Hz) and females (823 Hz) in central Kentucky, and those owls with higher frequency calls tended to sing faster (in contrast to Cavanagh's and Ritchison's [1987] findings). Some species have been known to modify their calls to sound more (MacGregor and Krebs 1989) or less like their neighbours (Walcott et al. 2006) and/or to stand out from background noise in urban areas (Warren et al. 2006; Wood and Yezerinac 2006). Future research can determine if Eastern Screech-Owls that live close to one another exhibit more or fewer differences than would be expected by chance or if urban owls seem to shift their calls relative to their rural counterparts. While centre frequency appeared to be the most important single factor to determine individuality (PIC = 3.17, see Appendix), the specific combinations of phase-specific note rates was important, particularly between NR1 and NR3.

In within-season analyses, 'EEE' models (with ellipsoidal shapes and equal volumes and orientations) were always selected. The selection of an 'EEV' model – with varying component orientations – in the multi-year analysis suggests that vocalizations were beginning to diverge from their initial measurements. While the analysis was able to account for this variation, researchers should be aware that (not surprisingly) owls' calls do not stay completely consistent for their entire lives. Screech-Owls lived on average for 4.1 ± 2.8 (SD; median 2.6) years in Texas (Gehlbach 1994) and 3.1 ± 2.6 years (SD; median 2.0) in Ohio (VanCamp and Henny 1975) so a limit of approximately 3 years in terms of call consistency is adequate for most owls. Still, since those owls that live longer play a large role in recruitment rates over their lifetimes, monitoring long-lived individuals would be important for population studies. We also caution others that our multi-year dataset was rather limited, owing to the difficulty of finding captive owls that can be recorded and identified for many years, and thus large populations may not be discernible over time. However, one could potentially perform multiple analyses of data from consecutive years and then link clusters across years. Any new owls that establish

themselves in the study area should appear as new clusters, unlinked to any cluster in the previous year. Alternatively, researchers could simply estimate total abundance each year and compare these census counts to abundances derived from multi-year analyses. While the latter is able to provide individual survival rates, annual census counts may be analytically simpler and can still provide information on population trends and status.

The use of the short bounce for this method was based on a few factors. As stated above, bounce calls seem to be used in behavioural contexts that would be expected to require individually distinguishable characters (e.g. mate-mate and parent-offspring communications). Second, they were the most common call recorded. However, since the vast majority of our calls were elicited as a result of an artificial broadcast, Screech-Owls may in fact use other calls more often, or call characters may change substantially in other contexts. For example, occasionally an owl was heard calling before we began our artificial broadcasts and these bouts consisted almost exclusively of long bounces only. Thus, it is possible that short bounces are primarily used in aggressive territorial disputes or courtship displays, and currently our method is untested using calls from other behavioural contexts.

It would also be useful to find variables that could determine individuality using long bounce and whinny calls, something we did not have the time or resources to pursue for this study. Long bounces may be used in pair or parent-offspring communication as in the short bounces but lack the phase-based variation of the short bounce. To develop a method based on long bounces, reliable variables other than centre and quartile frequencies and note rates will have to be found. Preliminary analyses of long and short bounces using only the three frequencies and the note rate of the entire call (the only note rate measurable on long bounces since they have no discernable phases) had poor discriminating power. Development of a method based on whinny calls would probably be even more difficult: to our own ears, we noted substantial variability in whinny durations, frequencies, and general forms even within a single bout. However, if possible, using the three types of call would allow additional verification of cluster classifications and should allow more individuals to be discriminated.

Using vocalization analysis to monitor individuals has many advantages. Non-invasive techniques such as this minimize the danger to study animals and can be used in urban areas where project visibility is often high and public opinion regarding trapping and handling local wildlife may be quite unfavourable (C.M. Nagy, pers. obs.). The cost of recording equipment and analysis software (~ US\$2,000 total) is less than the cost of a telemetry-based study. The latter also has constant costs involved in replacing or refurbishing transmitters, while a vocalization study has only the initial cost. Telemetry will still be necessary if the study objectives require precise and numerous locations, however, especially if one uses call-playback surveys – which draw owls to the researcher – to obtain recordings. Mist netting and banding may be less expensive but can be more labour-intensive and, without telemetry, usually cannot yield repeated samples of individuals within a single season. For biologists and land managers with limited time and budgets – as is usually the case among researchers studying urban wildlife and common, non-game species – methods that can be performed on a small budget are often the only options.

Acknowledgements

We would like to thank the several wildlife rehabilitators and preserve managers that allowed us to record owls/on their property: L. Butcher, K. Hundgen, M. Smith, E. Smithies, R. Schubert, and S. Tatem. We also thank J. Groth for his advice on recording equipment, making field recordings, and manuscript comments. This study was funded by the Mianus River Gorge Preserve, a

Frank M. Chapman Memorial Grant from the American Museum of Natural History, and a generous donation from M. Winn. We thank two anonymous reviewers for their thoughtful and helpful comments.

References

- Bosakowski T, Smith DG. 1997. Distribution and species richness of a forest raptor community in relation to urbanization. *Journal of Raptor Research* 31(1):26–33.
- Cavanagh PM, Ritchison G. 1987. Variation in the bounce and whinny songs of the eastern screech-owl. *Wilson Bulletin* 99:620–627.
- Charif RA, Waack AM, Strickman LM. 2008. Raven Pro 1.3 User's Manual. Ithaca, NY: Cornell Laboratory of Ornithology.
- Charrier I, Bloomfield LL, Sturdy CB. 2004. Syllable types and coding in parid vocalizations. I: The chick-a-dee call of the black-capped chickadee (*Poecile atricapillus*). *Canadian Journal of Zoology* 82:769–779.
- Cornell Lab of Ornithology. 2008. Raven Pro 1.3. Bioacoustics Research Program. Ithaca, NY: Cornell Laboratory of Ornithology.
- Darden SK, Dablesteen T, Pedersen SB. 2003. A potential tool for swift fox *Vulpes velox* conservation: individuality of long-range barking sequences. *Journal of Mammalogy* 84:1417–1427.
- Delpont W, Kemp AC, Ferguson WH. 2002. Vocal identification of individual African wood owls *Strix woodfordii*: A technique to monitor long-term adult turnover and residency. *Ibis* 144:30–39.
- Dorn RD, Dorn JL. 1994. Further data on screech-owl distribution and habitat use in Wyoming. *Western Birds* 25:35–42.
- Fernandez-Juricic E, Delnevo AJ, Poston R. 2009. Identification of individual and population-level variation in vocalizations of the endangered southwestern willow flycatcher (*Empidonax trailii extimus*). *Auk* 126:89–99.
- Fraley C, Raftery AE. 1998. How many clusters? Which clustering method? Answers via model-based cluster analysis University of Washington Dept of Statistics Technical Report No. 329.
- Fraley C, Raftery AE. 2007. Mclust version 3 for R: Normal mixture modeling and model-based clustering. University of Washington Dept of Statistics Technical Report No. 504.
- Freeman PL. 2000. Identification of individual barred owls using spectrogram analysis and auditory cues. *Journal of Raptor Research* 43:85–92.
- Galeotti P, Paladin M, Pavan G. 1993. Individually distinct hooting in male pygmy owls *Glaucidium passerinum*: a multivariate approach. *Ornis Scandinavica* 24:15–20.
- Galeotti P, Pavan G. 1991. Individual recognition of male tawny owls (*Strix aluco*) using spectrograms of their territorial calls. *Ethology Ecology & Evolution* 3:113–126.
- Gehlbach FR. 1994. The Eastern screech-owl: Life history, ecology, and behavior in suburbia and the countryside. College Station, TX, USA: Texas A&M Univ. Press.
- Gehlbach FR. Eastern screech-owl (*Megascops asio*). In: A Poole, editor. 1995. The birds of North America online. Ithaca, NY: Cornell Laboratory of Ornithology, Available online: <http://bna.birds.cornell.edu/bna/species/165> (accessed 24 January 2011) .
- Gilbert G, Tyler GA, Smith KW. 1990. Local annual survival of booming male great bittern *Botaurus stellaris* in Britain, in the period 1990–1999. *Ibis* 144:51–61.
- Hartig S. 2005. Individual acoustic identification as a non-invasive conservation tool: an approach to the conservation of the African wild dog *Lycaon pictus*. *Bioacoustics* 15:35–50.
- Hoodless AN, Inglis JG, Doucet JP, Aebischer NJ. 2008. Vocal individuality in the roding calls of woodcock *Scolopax rusticola* and their use to validate a survey method. *Ibis* 150:80–89.
- Lynch PJ, Smith DG. 1984. Census of Eastern Screech-Owls in urban open space areas using tape-recorded song. *American Birds* 38:388–391.
- Macgregor PK, Krebs JR. 1989. Song learning in adult great tits (*Parus major*): effects of neighbours. *Behaviour* 108:139–150.
- Peake TM, Mcgregor PK, Smith KW, Tyler G, Gilbert G, Green RE. 1998. Individuality in corncrake *Crex crex* vocalizations. *Ibis* 140:120–127.
- Reby D, Joachim J, Lauga J, Lek S, Aulagnier S. 1998. Individuality in the groans of fallow deer (*Dama dama*) bucks. *Journal of Zoology* 245:79–84.
- Ritchison G, Cavanagh PM, Belthoff JR, Sparks EJ. 1988. The singing behavior of Eastern Screech-Owls: Seasonal timing and response to playback of conspecific song. *Condor* 90:648–652.

- Robisson P, Aubin T, Brémond J. 1993. Individuality in the voice of the emperor penguin *Aptenodytes forsteri*: adaptation to a noisy environment. *Ethology* 94:279–290.
- Smith DG, Gilbert R. 1984. Eastern Screech-Owl home range and use of suburban habitats in southern Connecticut. *Journal of Field Ornithology* 55:322–329.
- Sokal RR, Rohlf FJ. 1995. *Biometry*. 3rd ed. New York, NY, USA: W.H. Freeman.
- Stokes D, Stokes L, Eliot L. 1997. *Stokes field guide to bird songs: eastern region*. New York, NY, USA: Time Warner.
- Terry AMR, McGregor PK. 2002. Census and monitoring based on individually identifiable vocalizations: the role of neural networks. *Animal Conservation* 5:103–111.
- Terry AMR, Peake TM, McGregor PK. 2005. The role of vocal individuality in conservation. *Frontiers in Zoology* 2:1–16.
- Tripp TM, Otter KA. 2006. Vocal individuality as a potential long-term monitoring tool for western screech-owls, *Megascops kennicottii*. *Canadian Journal of Zoology* 84:744–753.
- Vancamp LF, Henny CJ. 1975. The screech owl: Its life history and population ecology in northern Ohio. *North American Fauna* 71:1–65.
- Walcott C, Mager JN, Piper W. 2006. Changing territories, changing tunes: male loons, *Gavia immer*, change their vocalizations when they change territories. *Animal Behaviour* 71:673–683.
- Warren P, Katti M, Ermann M, Brazel A. 2006. Urban bioacoustics: It's not just noise. *Animal Behaviour* 71:491–502.
- Wood WE, Yezerinac SM. 2006. Song sparrow (*Melospiza melodia*) song varies with urban noise. *Auk* 123:650–659.
- Xu R, Wunsch Ii DC. 2009. *Clustering*. Piscataway, NJ, USA: IEEE.

Appendix. Descriptive statistics and PIC (proportion for individuality coding) values for all variables measured on calls from 10 Eastern Screech-Owls, New York and New Jersey, 2008–2010. Variables with PIC > 1 were considered for discriminatory value. After iterative testing with these variables, whole-call centre frequency (CFall), first quartile frequency (Q1Fall), third quartile frequency (Q3Fall), and note rate (NRall), and note rates of each call phase (NR1, NR2, and NR3) had the best discriminatory power.

	Notes_all ^a	Notes1 ^a	Notes2 ^a	Notes3 ^a	Dur1 ^b	CF1	Q1F1	Q3F1	Dur2 ^b	CF2	Q1F2	Q3F2	Dur3 ^b
Mean	31.9	13.1	5.1	13.7	0.95	645.2	570.8	720.8	0.45	645.3	571.9	719.9	0.81
SD	4.4	2.3	1.2	3.5	0.16	101.0	101.5	103.5	0.11	101.4	102.5	103.1	0.22
PIC	0.82	0.68	0.49	0.64	0.59	2.93	1.86	2.15	0.54	2.98	1.98	2.18	0.66
	CF3	Q1F3	Q3F3	Dur_all ^b	CFall	Q1Fall	Q3Fall	NRall	NR1	NR2	NR3	PrNotes1 ^c	
Mean	645.4	571.8	720.3	2.22	644.1	569.2	719.9	14.39	13.66	11.48	16.96	0.41	
SD	99.3	100.1	101.5	0.28	100.6	101.4	102.8	0.81	0.81	1.02	0.97	0.08	
PIC	3.05	2.03	2.36	0.82	3.17	2.00	2.28	1.68	1.23	1.86	1.05	0.65	
	PrNotes2 ^c	PrNotes3 ^c	PrDur1 ^d	PrDur2	PrDur3	IQR1	IQR2	IQR3	IQRall				
Mean	0.16	0.43	0.43	0.20	0.36	150.0	148.0	148.5	150.7				
SD	0.03	0.07	0.08	0.05	0.07	33.7	32.9	33.2	33.6				
PIC	0.42	0.59	0.62	0.47	0.58	0.60	0.64	0.64	0.63				

^aNotes, number of notes in whole call (all) or phase 1, 2 or 3; ^bDur, duration in seconds of whole call (all) or phase 1, 2, or 3; ^cPrNotes, the proportion of total notes in the call in that phase (1, 2, or 3); ^dPrDur, the proportion of the total call duration in that phase (1, 2, or 3).