International Journal of application or innovation in engineering and management

ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

Classification of four wood-warbler species using different similarity-based methods

Bhanu Partap

Roorkee, India Bhanu Partap@gmail.com

Article Info

Received: 17-06-2019 Revised: 21 -07-2019 Accepted: 06-08-2019 Published:15/09/2019

A B S T R A C T

Although several techniques exist for analyzing bird sounds, very little has been done to evaluate current approaches to determining and rating call similarity, especially when field recordings are involved. Spectrographic cross-correlation, dynamic time warping, Euclidean distance between spectrogram-based feature measurements, and random forest distance are the four methods for calculating call similarity that are compared in this manuscript, which investigates a suite of methodologies for analyzing flight calls of New World warblers. Since these signals may include crucial demographic or ecological information, we put these techniques to the test on night cries, which are brief, structurally simple vocalizations often used during nighttime migration. We classified flight calls from three datasets—one with birds in captivity and two with birds in the field—using the four methods described above. The four warbler species that are often recorded during acoustic monitoring—American Redstart, Chestnut-sided Warbler, Hooded Warbler, and Ovenbird—had a same amount of sounds in each dataset. We developed four similarity-based classifiers using recordings from captives to train the classification models. These classifiers were then evaluated on both the captive and field datasets. Classification accuracy was lower on field recordings than captive recordings for each of the evaluated approaches, and we demonstrate that these methods are unable to completely characterize the sounds of these warbler species. With an accuracy of 67.6% in classifying field recordings, the random forest technique outperformed the other three approaches we tested. The most popular approach in flight call research, manual classification, was compared to the automated algorithms by having human specialists categorize calls from each dataset. Even if automated methods are quicker, they still can't compare to human classification when it comes to over 90% of field recordings that were accurately classified by the experts. Nevertheless, because to the difficulties of working with this data—for example, the fact that the field recordings include background noise and the fact that the f-light cries of several species are structurally similar—some of the automatic classification methods that were evaluated may be suitable for application in the actual world. Analysis, detection, and classification of signals of short durations might benefit from the information provided by this comparison of generally applicable approaches. Our findings suggest that, with human supervision, a mix of feature measurements and random forest classification may be used to assign flight sounds to species.

International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

1. Introduction

Bird vocalizations are used in numerous behavioral contexts and serve a variety of purposes, such as maintaining contact within social groups or between mates, issuing warnings about predators, and eliciting parental care from adults, and are often involved in mating dis- plays and territorial defense (reviewed in Catchpole and Slater, 1995; Marler, 2004a). Whereas singing is mostly associated with breeding be- haviors, calling can accompany a range of behaviors and represents a more versatile and comprehensive method of communication. Howev- er, calls, especially those of passerine songbirds, have often been neglected in the study of bird communication, only recently receiving attention .Lanzone et al., 2009). Although references to flight calls date to the turn of the 20th century (Libby, 1899), and despite more recent research, some basic features remain poorly known, including their function, their evolutionary origins, and the extent of within-individual and within-species variation (Farnsworth, 2005). Furthermore, applications of flight calls for conservation goals have not advanced much beyond applied natural history studies. Improved knowledge of these vocaliza- tions could be useful in a variety of applications, including efforts to mit- igate potential impacts of wind energy, track species movements during seasonal migration, and estimate bird density using vocalization counts (Farnsworth et al., 2004; Gagnon et al., 2010). A critical component for realizing such applications is the automated classification of flight calls to species.

Numerous methods have been developed for the automatic classifi- cation of avian vocalizations (reviewed in Blumstein et al., 2011). Such methods are typically employed to obtain information about migration patterns (Evans and Mellinger, 1999), monitor areas of human interest, such as wind farms (Evans, 1998; Kunz et al., 2007), facilitate the con- servation of protected areas (Brandes, 2008), and studying soundscape ecology (Kasten et al., 2012). For flight call analyses, however, automa- tion has been only partially realized. To date, studies have combined manual and automatic processes, yielding species-specific migration data (e.g., Larkin et al., 2002) and estimates of species richness (Wimmer et al., 2013), and permitting comparative analyses among spe- cies (Farnsworth and Lovette, 2005, 2008). However, more efficient flight call analysis is essential to monitor species across larger ecological scales. The automatic classification of flight calls could greatly increase analysis efficiency, thereby enhancing knowledge of avian ecology and facilitating improved conservation and management of wild birds.

Many of the methods commonly applied to the classification of bird vocalizations are based on traditional speech recognition techniques (Rabiner and Juang, 1993). These algorithms fall into three general categories, and all remain widely used. The first includes spectrogram- based template matching techniques, such as spectrogram cross corre- lation (Clark et al., 1987), which strictly compares corresponding spec- trogram values, and dynamic time warping (e.g., Anderson et al., 1996; Damoulas et al., 2010), which allows for some compression or expan- sion in time to permit better matching. The second category includes feature-based classifiers that define each call by a set of spectro- temporal measurements. These measurements are then fed into auto- matic classifiers,

which range from simple clustering techniques such as nearest neighbor (e.g., Fagerlund and Härmä, 2005) or Euclidean dis-tance between features (e.g., Tyagi et al., 2006), to more complex algo-rithms, including Gaussian mixture models (e.g., Marcarini et al., 2008), autonomous neural networks (e.g., Cai et al., 2007; Ranjard and Ross, 2008), and support vector machines (e.g., Fagerlund, 2007). The third is advanced pattern recognition, which has been used to classify entire bird song sequences, using algorithms such as Hidden Markov Models (Kogan and Margoliash, 1998; Somervuo et al., 2006; Trifa et al., 2008) or techniques such as ensemble processing using distributed pipelines (Kasten et al., 2010). Although several classification techniques have been developed for avian vocalizations, there is no clear consensus as to which method is most effective.

Here, we compare the ability of four similarity-based classifiers to automatically assign flight calls to the correct warbler species using flight calls recorded from wild birds in several locations. We investigate the effectiveness of the following four methodologies for calculating call similarity: (1) spectrographic cross-correlation, (2) dynamic time warping, (3) Euclidean distance between spectro-temporal measure- ments, and (4) random forest distance between spectro-temporal mea- surements. To compare the ability of each method to correctly group similar (i.e., conspecific) calls, we apply non-metric multidimensional scaling to the four similarity matrices for extraction of latent acoustic measures used in a linear discriminant analysis (e.g., Baker and Logue, 2003; Cortopassi and Bradbury, 2000, 2006). Taking advantage of recent studies of New World warbler (*Parulidae*) flight calls (Farnsworth, 2007b; Lanzone et al., 2009), we use calls from the American Redstart (*Setophaga ruticilla*), Chestnut-sided Warbler (*Setophaga pensylvanica*), Hooded Warbler (*Cardellina citrina*), and Ovenbird (*Seiurus aurocapillus*) to compare the methods listed above. These species were selected be- cause they are frequently recorded in North American nocturnal acoustic monitoring studies, and because they may be challenging to classify due to the structural similarity of their calls, therefore making our study more relevant for real-world applications of these classification tech- niques. Lastly, to compare performance between these automated tech- niques and manual classification, we contrast the correct classification rates of the four automated methods to those of human experts.

2. Materials and methods

2.1. *Data collection*

Three datasets were used in this study; one to train and test the clas- sification models, and two for testing only. The “captive” dataset, which was used to train and test the classifier models, contains flight call re- cordings taken from temporarily captured wild birds. The remaining two datasets contain calls recorded from wild birds in flight. The “diur- nal” dataset was recorded during daylight hours in Northeastern North America, and the “nocturnal” dataset was recorded during evening hours (i.e., after civil twilight at dusk) in the Gulf of Mexico. Because the diurnal and nocturnal datasets are field recordings, rather than re- cordings made in a controlled environment, these recordings have high amounts of wind noise and a much lower signal to noise ratio, making them realistic test cases for classification (Lanzone et al., 2009).

International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

2.1.1. *Captive recordings*

The captive flight call dataset was recorded at Powdermill Avian Research Center near Pittsburgh, PA during April–May and September– October 2005. Birds were captured using mist nets, individuals were banded with a United States Geological Survey band, and the date and time of recording, as well as the sex (male, female, unknown) and ap- proximate age (hatching-year, after hatching-year) of each individual was recorded. The birds were placed in an enclosed recording unit equipped with a microphone, and all vocalization produced by the focal bird were recorded (Lanzone et al., 2009). Playbacks of flight calls from conspecifics and heterospecifics, which were five minutes in duration, were used to elicit flight calls from non-calling, captive birds. The micro- phone was connected to a computer running Raven Pro 1.4 64-bit (Charif et al., 2004) and recordings were saved as 44,100 kHz, 24-bit WAV files. The birds were released after a 10-minute period regardless of the number of calls produced.

2.1.2. *Diurnal field recordings*

The diurnal flight call dataset was collected in Pennsylvania and New York in May-July and September and October 2005. Calls were recorded from wild individuals during flight in daylight hours. Recordings were made using a Sennheiser MKH 70 microphone (Sennheiser Electronic, Old Lyme CT), either to an analog recording device (Sony TCM-5000 re- corder) or a digital recording device (Nagra Ares BB + flash memory re- corder). All flight calls were either recorded or digitized as 16-bit, 22,050 Hz wav files. For analyses, only flight calls recorded with a clear line of sight between the microphone and the bird are included. Calls recorded with interfering vegetation or structures, or overlap- ping flight calls were excluded (see Farnsworth, 2007b). To avoid collecting multiple samples from the same individuals, the captive and diurnal datasets were collected at different locations and on dif- ferent days.

2.1.3. *Nocturnal field recordings*

The nocturnal flight call dataset was collected from September 8th to November 5th, 1999 on the Viosca Knoll oil platform (VK 786), approximately 145 km southeast of the Alabama coast (29°1344N;

87°4655W), during local civil twilight. A portable pressure zone micro- phone with a Knowles Electret EK3132 microphone element was used to record nocturnal vocalizations (see Evans, 1994, Evans and Mellinger, 1999; Evans and Rosenberg, 2000; Farnsworth et al., 2004; Larkin et al., 2002). The microphone element has a relatively flat fre- quency response in the 1-10 kHz range. The recordings were made on a videocassette recorder (VCR; Sony SLV-675) in extended play mode. The VCR recorded audio from the microphone on 8-h 19-min video home system (VHS) tapes through a portable stereo cassette tape deck (Optimus SCT-86; Radio Shack, Fort Worth, TX) that amplified input signal strength (gain setting: +7 dB). Recordings were digitized at 22,050 Hz, 16-bit sample size (see Farnsworth and Russell, 2007). Species-identity labels for the flight calls in the diurnal and nocturnal datasets were established by A. Farnsworth.

2.2. *Call characteristics*

Spectrograms of exemplar flight calls collected from the focal species are shown in Fig. 1. The calls of the American Redstart (AMRE) and the Ovenbird (OVEN) (Fig. 1A and D, respectively) both exhibit a "check-mark" shape when viewed as a spectrogram and have frequency mod- ulation in the tail end of calls, though the AMRE flight call starts with a longer downward slope. The flight calls of the Chestnut-sided warbler (CSWA) and Hooded warbler (HOWA) (Fig. 1B and C, respectively) have numerous inflections and exhibit high frequency modulation throughout the entire signal. The structural characteristics of calls col- lected from each species are summarized in Tables B.2, B.3, and B.4.

2.3. *Data extraction*

Flight calls were manually clipped from the captive and diurnal re- cordings by using Raven Pro 1.4 (Charif et al., 2004) to browse recording spectrograms and extract portions of audio files containing calls. A 256-sample Fast Fourier Transform (FFT) with 256-sample Hann windows and an advance of 38 samples was used to make spectrograms of cap- tive recordings, and a 128-sample FFT with 128-sample Hann windows and an advance of one sample was used for diurnal and nocturnal re- cordings. For captive recordings, flight calls from focal birds were differ-entiated from playback calls by their relatively higher signal-to-noise ratio (SNR). For the nocturnal recordings, calls were automatically detected using the Raven Pro. 1.4 Band Limited Energy Detector (Table B.1). Selections created by the Band Limited Energy Detector were reviewed by the authors for presence of the target species. Raven selection tables containing the selected calls were first consoli- dated using Google Refine (Huynh and Mazzocchi, 2012). Next, a buffer of 5 ms was added on either side of call selections, and calls were auto-matically clipped from recordings using SoX (Sound eXchange v. 14.3.1, http://sox.sourceforge.net/).

The number of calls collected for each species in the three datasets is shown in Table 1. To create our final three datasets, we randomly select-ed 400 calls (100 from each species) from the captive recordings, 360 calls (90 from each species) from the diurnal recordings, and 144 calls (36 from each species) from the nocturnal recordings, totaling 904 flight calls. The number of calls used in each dataset was dictated by the species with the fewest calls. The 400 selected captive calls were downsampled to 22,050 Hz, giving the final 904 calls the same sampling

International Journal of application or innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

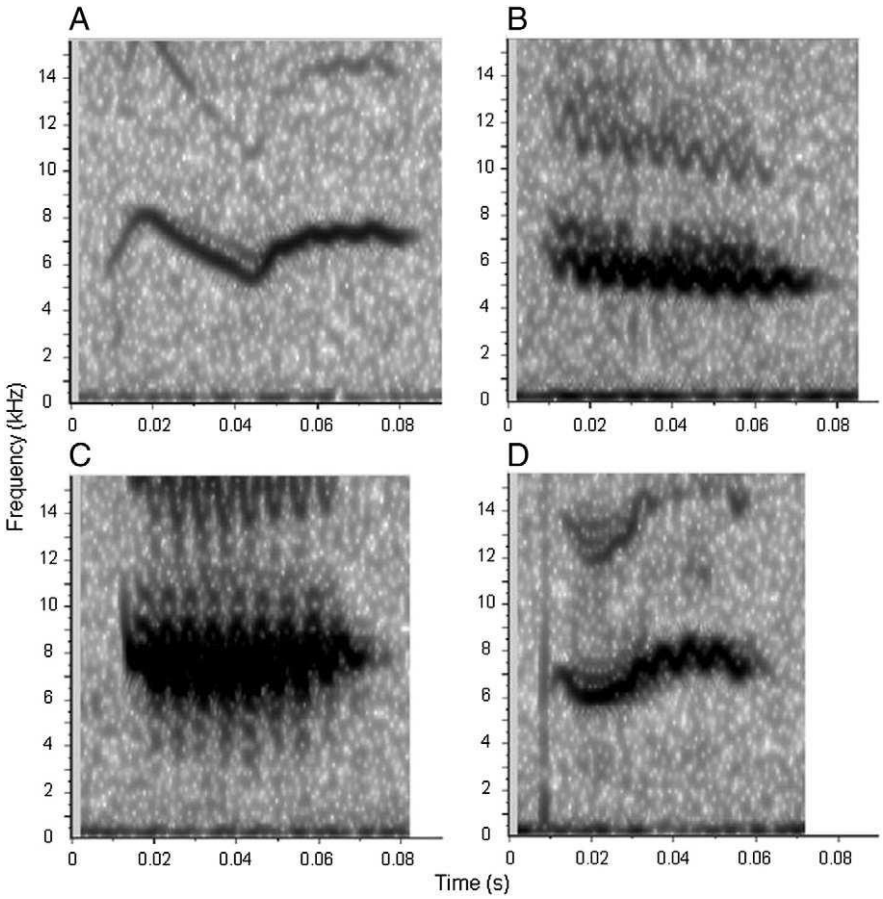


Fig. 1. Time-frequency spectrogram of typical warbler flight calls. Shown here are the four focal species used in this study: (A) American Redstart, (B) Chestnut-sided warbler, (C) Hooded warbler, and (D) Ovenbird.

Table 1
Summary of flight calls used in this study. This table contains the number of flight calls collected from each species and location, the datasets to which collected calls belong, and the number of calls used for training and testing of classification models. For the captive calls, we collected between 1 and 63 (mean ± SD: 16.35 ± 13.82) calls from each AMRE in- dividual, 1 and 98 (25.24 ± 24.11) calls from CSWA, 1 and 39 (11.53 ± 10.60) calls from HOWA, and 1 and 37 (12.31 ± 9.49) calls from OVEN.

Species	Dataset	Total calls collected	Calls used in analysis
AMRE	Captive	790	100
AMRE	Diurnal	266	90
AMRE	Nocturnal	241	36
CSWA	Captive	385	100
CSWA	Diurnal	118	90
CSWA	Nocturnal	66	36
HOWA	Captive	192	100
HOWA	Diurnal	91	90
HOWA	Nocturnal	36	36

Ellis (2008), and executed using Matlab, 2010a (The Mathworks, 2010).

International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

rate. Spectrograms were then created for each call using a 1024-point FFT with 256-point frame length using Hann windows and an advance of one point between frames. This small advance between frames re- sulted in a time resolution of 0.045 ms and a frequency resolution of 21.5 Hz in all call spectrograms.

2.4. *Call similarity analyses*

2.4.1. *Spectrogram cross correlation*

Spectrogram cross-correlation (SPCC) typically measures the simi- larity between pairs of spectrograms by calculating similarity as a function of time (Clark et al., 1987; Cortopassi and Bradbury, 2000; Farnsworth, 2007b). We used the SPCC-PCO tool (SoundXT; Cortopassi *unpublished data*; used previously in Cortopassi and Bradbury, 2000, 2006) to calculate the peak correlation values for all pairs of flight calls across both time and frequency. Although most previous applications of SPCC involved correlation along time only, it was necessary for our study to also allow sliding along the frequency axis as flight calls may exhibit frequency shifting both within and among individuals and species. We used a maximum time lag of 127 ms (the duration of the longest flight call in the dataset) and a maximum frequency lag of 3 kHz. The frequency lag was determined using visual inspection of call spectrograms, which confirmed that calls did not vary in center frequency or bandwidth by more than this amount. All spectrogram matrix entries represented power in decibel scale (dB/Hz), and spectrogram matrices were normalized to contain values between 0 and 1, ensuring that peak correlation values would fall within this range. To minimize correlation between background noise in pairs of recordings we used only spectrogram values within the 3-10.5 kHz range, in effect applying a bandpass filter to all call recordings. Frequency shifting, matrix normalization, and bandlimiting employed here are all built-in features and configuration options of the SoundXT tool itself. Applying this tech- nique, we calculated pair-wise peak correlation values for all pairwise combinations of flight calls.

2.4.2. *Dynamic time warping*

Dynamic time warping (DTW; Vintsyuk, 1968) calculates pairwise similarity between vectors or matrices while permitting some expan- sion or compression in time in order to maximize similarity. DTW has been most commonly applied in automatic speech recognition (Deller et al., 1993; Rabiner et al., 1978; Sakoe and Chiba, 1978), and, more recently, in the detection and classification of avian

vocalizations (e.g., Brown et al., 2006; Damoulas et al., 2010; Kogan and Margoliash, 1998). DTW effectively stretches or shortens calls in time, allowing for calls with similar contour shapes but different durations to be scored

International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

2.4.3. *Euclidean distance between feature measurements*

Direct call measurement was performed on acoustic features (sum-marized in Table B.5) of all flight calls using an adapted version of the Acoustat sound measurement tool (Fristrup and Watkins, 1992, 1993). Acoustat measurements were applied to spectrogram values within "event" boxes manually created in Raven Pro 1.5 (Charif et al., 2004), which entirely enclosed the power within the fundamental frequency of each flight call. The measurement process involves collapsing the signal's time-frequency spectrogram into an aggregate power envelope as a function of time and an aggregate power spectrum as a function of frequency. From these, robust measures of central tendency and disper- sion are extracted using order statistics (Cortopassi, 2006; Fristrup and Watkins, 1992, 1993, www.birds.cornell.edu/brp/research/algorithms/RSM.html, Cortopassi and Fristrup, personal communication). We cal-culated Euclidean distances between the Acoustat energy distribution measurements using the ecodist package (Goslee and Urban, 2007) in R (R Core Team, 2013). This method is hereafter referred to as Feature-ED.

2.4.4. *Random forest distance between feature measurements*

Using the feature measurements described above, we created a sec- ond set of pairwise similarity measurements by applying a random for-est (Breiman, 2001) decision tree to these values and calculating the proximity metric between every flight call pair in the dataset. Random forests are combinations of decision trees created using a shared feature space, where each tree is built on a subsample of the given dataset in a technique called "bagging". Individual decision trees perform classifica-tions based on decisions at nodes within each tree, using logical or arith-metic comparisons of a subset of feature measurements which is itself chosen randomly for each node. A random forest is typically used in a supervised manner for classification tasks. However, it is also possible to use the algorithm in an unsupervised manner producing a measure of similarity between data points without considering their class mem-bership (see Liaw and Wiener, 2002). We used the randomForest (Liaw and Wiener, 2002) package in R version 3.0.0 (R Core Team, 2013) to generate a pairwise proximity matrix between the 904 calls in the dataset, based on the 81 feature measurements. We used 500 trees, with 9 features tried at each node. We converted this proximity matrix to a distance matrix using the transformation one minus proximity. This method is abbreviated as Feature-RF.

2.5. *Creation of similarity-based classifiers*

Our aim was to determine which analysis method could provide the best relative similarity measurements of calls in our dataset, and thus could best facilitate classification of call recordings by species. The anal-yses described above yielded four similarity matrices representing pairwise relationships between the 904 calls in the three datasets. To determine which method best calculates similarity between call record- ings, we first used non-metric multidimensional scaling (NMDS) to model pairwise relationships between calls in 5-dimensional space using the ecodist package (Goslee and Urban, 2007) in R (R Core Team, 2013), creating four unique NMDS ordinations each computed using five iterations (minimum stress and R^2 for each ordination: SPCC: 0.107, 0.903; DTW: 0.105, 0.924; Features-ED: 0.001, 0.99;

Features-RF: 0.21, 0.465). Although the four pairwise similarity matrices were created as an intermediate step and were ultimately used to train

classification models for each method, we illustrate the effectiveness of each similarity analysis to cluster calls of the same species by creating 2-D ordination plots showing the distribution of calls in NMDS space for each set of similarity measurements (Fig. B.1).

To classify calls by species, we used the MASS R package (Venables and Ripley, 2002) to create linear discriminant analysis (LDA) models for each of the four similarity-based methods, using the captive calls as training data. To evaluate LDA performance, we used 100-fold, leave-one out cross-validation of the training data, and calculated the percentage of calls correctly classified for each method (mean ± sd: SPCC 0.715 ± 0.024, DTW: 0.658 ± 0.024; Features-ED: 0.833 ± 0.019; Features-RF: 0.888 ± 0.25). After characterizing the perfor- mance of each classifier using the training dataset, calls from the diurnal and nocturnal datasets were then classified using the four LDA models. By using individuals as output variables and latent measures as input variables, it was possible to generate classification rates for correctly identifying calls to known species. A summary of the steps involved in each method and the motivation for implementation can be found in Table A.1. A flowchart illustrating the order in which each method was applied is shown in Fig. A.1.

2.6. *Performance measures*

2.6.1. *Comparisons among automated methods*

To determine how well each of the models described above per- formed relative to one another, we summarized the classification results from each model in several confusion matrices. By comparing the distri- bution of values across confusion matrices we were able to compare the abilities of each method to correctly separate species and identify com- mon sources of error. Additionally, we calculated the sensitivity and specificity of each LDA model when classifying calls from each species. Sensitivity was calculated as the percentage of calls known to be from a certain species that were correctly classified as such. Specificity was found by summing all calls that were neither known to be from a certain species or predicted to be from that species and dividing that by the total number of calls known to be from all other species. These values are reported in the confusion matrices created for each model. To eval- uate overall performance of each LDA model, we summed the number of correct classifications for each species (the numbers along the di- agonal of the confusion matrix), and divided by the total number of calls being classified. The resulting value is referred to as the "correct classification rate", and this measurement represents a common metric for assessing the abilities of an LDA to discriminate among species.

2.6.2. *Expert human reviewers*

To compare performance of the four similarity-based methods to human classification, we asked three expert human reviewers to manu- ally classify a random subset of 36 calls per species from the three datasets, totaling 432 calls. Each of the expert reviewers has extensive knowledge of flight calls and years of experience studying avian vocali- zations, but had not previously seen any call in this dataset. The sound files were distributed to the experts without accompanying metadata and in random order, ensuring that calls from the same dataset and/or species wouldn't

be viewed consecutively, potentially biasing the ex- perts' classification. Using Raven Pro 1.5 (Charif et al., 2004), the experts visually inspected spectrograms and listened to recordings for each flight call in order to classify each call by species to the best of their ability. Spectrogram parameters were determined by the personal preference of the experts to optimize their ability to accurately clas- sify the flight calls, which is common practice amongst experts. To compare human performance to the automated classification methods, we calculated correct classification rates for each expert reviewer as well.

International Journal of application or Innovation in Engineering and Management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

3. Results

3.1. *Classifier performance*

The four automated classification methods performed relatively well when tested on the captive dataset, which was also used to train each model (correct classification rate: SPCC: 71.5%; DTW: 65.75%; Feature- ED: 83.25%; Feature-RF: 88.75%). Despite the relatively high correct classification rates of captive calls by the automated techniques, the classification accuracy of all techniques dropped significantly when tested on the diurnal and nocturnal datasets (Fig. 2). The classification accuracy of three of the automated techniques, SPCC, DTW and Feature-ED, decreased when tested on the diurnal dataset (correct clas- sification rates were 55.83%, 52.22%, and 26.94%, respectively) as well as the nocturnal dataset (correct classification rates were 59.03%, 54.17%, and 25%). However, the classification accuracy of Feature-RF did not de- crease as severely when tested on field recordings; this method had cor- rect classification rates of 67.78% for the diurnal dataset and 67.36% for the nocturnal dataset. Whereas Feature-ED had the lowest correct clas- sification rates on field recordings, at least 25% lower than SPCC and DTW on both the diurnal and nocturnal datasets, Feature-RF performed better than all other automated methods when tested on field record- ings, with correct classification rates approximately 10% higher than SPCC and DTW (Fig. 2). As expected, the human experts had a high clas- sification accuracy, with each person having correct classification rates over 90% for the captive calls, over 92% for diurnal calls, and over 88% for nocturnal calls.

3.2. *Sources of classification error*

Each of the tested methods exhibited some degree of classification error when discriminating between calls from CSWA and HOWA, as well as calls from AMRE and OVEN. When tested with the field record- ings, SPCC and Feature-RF often misclassified HOWA calls as CSWA calls, and DTW often had equal amounts of misclassification between these two species. This is evidenced by the relatively low sensitivity scores found for AMRE and HOWA, and low specificity scores found for CSWA and, in some cases, OVEN (Tables 2 and 3). Feature-ED classified all HOWA calls as CSWA calls in both field datasets, and incorrectly clas- sified nearly all calls as CSWA when tested with the nocturnal dataset. SPCC, DTW, and Feature-RF often misclassified AMRE calls as OVEN calls when tested with both field datasets, and also misclassified OVEN as AMRE, to a lesser extent (Tables 2 and 3). Although the human ex- perts had relatively high correct classification rates overall, most mis- classification errors arose from confusion between AMRE and OVEN calls, and CSWA and HOWA calls (Tables B.6, B.7, and B.8).

4. Discussion

4.1. *Challenges inherent to the dataset*

International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

Despite much interest in the classification of avian vocalizations and evidence that flight calls are highly useful in comparative analyses and monitoring migration patterns (e.g., Farnsworth, 2005), relatively few studies have focused specifically on automated flight call classification (e.g., Mills, 1995). Flight calls are extremely short (typically less than 100 ms), and thus sequence-based classifiers that have been shown to successfully classify song by species are not applicable to these signals. Therefore, classifiers that rely upon template matching and feature ex-traction are more common in previous studies of flight and contact calls, and have been shown to achieve high accuracy rates (Anderson et al., 1996; Bradbury et al., 2001; Damoulas et al., 2010; Schrama et al., 2008; Vehrencamp et al., 2003). However, to our knowledge, no previous studies have explicitly examined the classification of highly similar flight calls from species that exhibit similar spatiotemporal mi-gration phenology (but see Farnsworth and Lovette, 2008), although

International Journal of application or innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

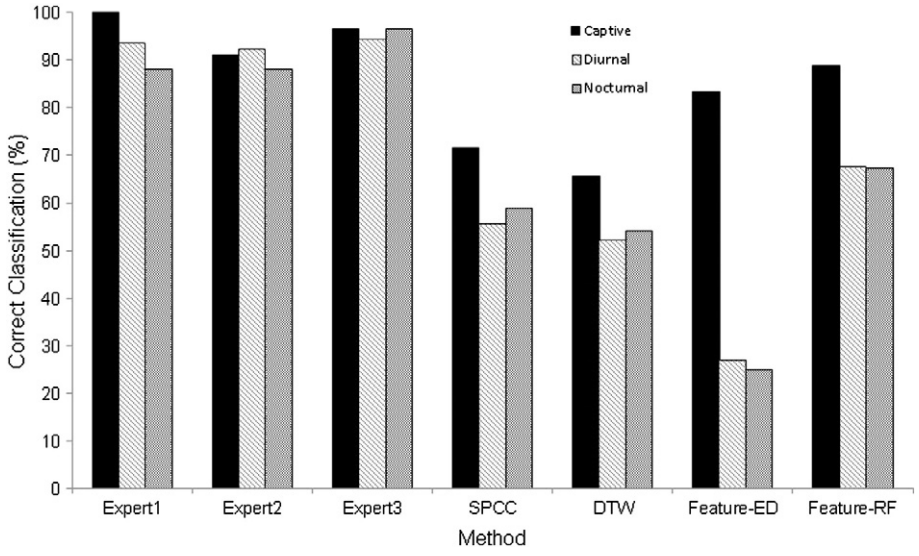


Fig. 2. Percentage of flight calls correctly classified by human experts and automated methods. The correct classification rates are shown for the three datasets used in this study: captive recordings, diurnal recordings, and nocturnal recordings. Human experts classified 144 flight calls from each dataset, and automated methods were used to classify 400 captive calls, 360 diurnal calls, and 144 nocturnal calls. The automated methods were trained using the captive call recordings, thus leave-one-out cross validation was used to obtain classification results for this dataset.

this is an increasingly important challenge for biologists studying inter-specific migration patterns.

The low correct classification rates observed in our study, rangingfrom 25 to 68% in field recordings, can partially be attributed to our de-liberate choice of some species that are similar to one another and theuse of field recordings as test data. Previous studies have often tested classification techniques only on captive recordings, or those taken at extremely close range. Here, we intentionally included calls recordedunder typical field conditions to more accurately test methods that arecommonly used on data characteristic of passive acoustic monitoring

Results of classification of diurnal flight calls. Field recordings of flight calls collected during daylight hours were classified by models based on (A) SPCC, (B) DTW, (C) Feature-ED, and
(D) Feature-RF. The models were tested on a dataset of 360 calls, comprising 90 from each focal species.

International Journal of application or Innovation in engineering and management

ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

Observed	Predicted AMRE	Predicted CSWA	Predicted HOWA	Predicted OVEN	Sensitivity
A					
AMRE	19	7	6	58	0.21
CSWA	0	80	10	0	0.89
HOWA	4	62	24	0	0.27
OVEN	3	9	0	78	0.87
Specificity	0.97	0.72	0.94	0.79	
B					
AMRE	14	2	8	66	0.16
CSWA	0	48	42	0	0.53
HOWA	0	47	42	1	0.47
OVEN	0	0	6	84	0.93
Specificity	1.00	0.82	0.79	0.75	
C					
AMRE	24	65	1	0	0.27
CSWA	17	73	0	0	0.81
HOWA	30	60	0	0	0.00
OVEN	43	45	2	0	0.00
Specificity	0.67	0.37	0.99	1.00	
D					
AMRE	33	2	8	47	0.37
CSWA	4	73	11	2	0.81
HOWA	1	19	70	0	0.78
OVEN	5	0	3	82	0.91
Specificity	0.96	0.92	0.92	0.82	

projects. The diurnal and nocturnal datasets exhibit high noise levels due to wind and other ambient noise, making "clean" signal analysis impossible. Thus, despite relatively poor performance of automated methods in field conditions, the use of field recorded datasets in this study enhances the relevance of our results for real-world applications.

4.2. *Evaluation of classification techniques*

Each of the tested methods has drawbacks or limitations, particular-ly when applied to field recordings. SPCC and DTW are both computa-tionally expensive and are susceptible to classification error due to matching background noise in spectrograms rather than the signal of in-terest, a common shortcoming of template matching algorithms. For our field recordings, we recognize and accept this as an integral source of error in such data. SPCC and DTW are also much more computationally expensive, and without the use of a kernel (e.g., Damoulas et al., 2010), they may be impractical to implement. Both feature-based techniques are computationally inexpensive, but are ultimately limited by the qual-ity of features used; a universal challenge in classification of acoustic data. The distance metric used to measure similarity between record-ings of calls can also have a large impact on correct classification rates, as shown in Fig. 2. We suspect that using Euclidean distance to estimate call similarity was less effective, as background noise may have resulted in highly inaccurate feature measurements and may have introduced errors into distance calculations. The Feature-RF method is likely more resilient to background noise because pairwise distance is not a function of the difference of pairs of feature measurements, but instead the distance between decision trees. With additional features that are less affected by background noise, it may be possible to improve per-formance of both feature-based methods. Other feature sets have been developed for the purpose of classifying avian vocalizations (e.g., Tchernichovski et al., 2000), and it would be interesting to test classification based on these.

4.3. *Similarity among flight call structures*

The common classification errors among all models were confusion between AMRE and OVEN calls, and between CSWA and HOWA calls. Confusion between these species occurred in human classification as well, though to a lesser extent (Tables B.6–B.8). AMRE and OVEN flight

International Journal of application or innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

Table 3
Results of classification of nocturnal flight calls. Field recordings of flight calls collected during nighttime hours were classified by models based on (A) SPCC, (B) DTW,
(C) Feature-ED, and (D) Feature-RF. The models were tested on a dataset of 144 calls, com-prising 36 from each focal species.

Observed	Predicted AMRE	Predicted CSWA	Predicted HOWA	Predicted OVEN	Sensitivity
A					
AMRE	17	2	3	14	0.47
CSWA	0	31	0	5	0.86
HOWA	4	24	7	1	0.19
OVEN	1	5	0	30	0.86
Specificity	0.95	0.71	0.97	0.81	
B					
AMRE	7	6	7	16	0.19
CSWA	0	25	7	4	0.69
HOWA	0	7	28	1	0.78
OVEN	0	4	14	18	0.50
Specificity	1.00	0.84	0.73	0.81	
C					
AMRE	0	36	0	0	0.00
CSWA	0	36	0	0	1.00
HOWA	1	35	0	0	0.00
OVEN	0	36	0	0	0.00
Specificity	0.99	0.01	1.00	1.00	
D					
AMRE	24	1	3	8	0.67
CSWA	3	24	3	6	0.67
HOWA	0	15	19	2	0.53
OVEN	6	0	0	30	0.83
Specificity	0.92	0.85	0.94	0.85	

calls resemble a “check-mark” (Fig. 1A and D), but can be distinguished most often by the longer downsweep at the beginning of AMRE calls, the overall longer duration of AMRE calls, and the higher level of mod-ulation in the second half of OVEN calls (Evans and O'Brien, 2002).

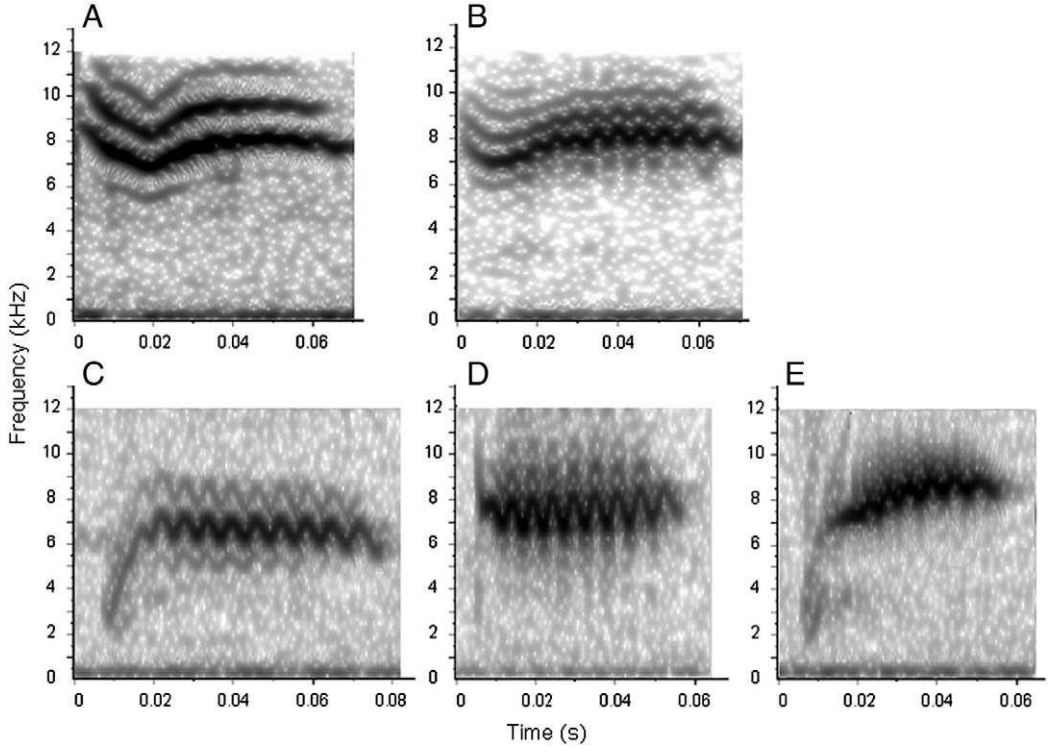
International Journal of application or Innovation in engineering and management

ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

However, we observed high variation in call structure found both within and among individuals in each species in our dataset, and certain call variants can appear more similar to those of heterospecifics than con-specifics, as shown in Fig. 3. CSWA flight calls (Fig. 1B) and HOWA flight calls (Fig. 1C) both have high frequency modulation, though the CSWA call is longer and maintains a constant average frequency over time, whereas HOWA calls have a "swooping" quality (Evans and O'Brien, 2002), with fluctuating rates of modulation and increasing or decreas-ing average frequency. Within-individual and within-species variation play a large role in confusion between these species as well (Fig. 3). Fur-thermore, when calls are somewhat masked by background noise, even calls with relatively different spectral structure may become indistin-guishable, which appears to be the primary cause of misclassification of field recordings.

4.4. *Comparison to previous studies*

Of the four similarity-based methods, the Feature-RF proved to be the most accurate when classifying calls from each of the three datasets. Other bioacoustics studies have had similar success using random for-ests in the classification of bird calls (Briggs et al., 2009), as well as calls of bats (Armitage and Ober, 2010) and cetaceans (Barkely et al., 2011; Henderson et al., 2011). Briggs et al. (2009) showed that a 100-tree random forest had a classification accuracy of up to 48.6%, compa-rable with adaboost and support vector machines, when classifying calls from 20 bird species found in the western United States. Armitage and Ober (2010) obtained correct classification rates of 84-96% when classifying calls from 11 different bats species using a random forest with 1000 trees. Taking feature measurements from eight species of delphinids, Barkely et al. (2011) had an overall correct classification score of 65.0% using a 500 tree random forest design. Using a random forest model with 5000 trees, Henderson et al. (2011) had a correct clas-sification rate of 64.8% for two types of whale vocalizations.



International Journal of application or Innovation in engineering and management
ijaiem.com/Sep 2019/ Volume 8/Issue 1/Article No-1/36-61
ISSN: 2319-4847

Fig. 3. Heterospecific flight calls that exhibit similar call shape and structure. Flight calls collected from the (A) American Redstart and (B) Ovenbird often exhibit similar “check mark” structures with fundamental frequencies between 7 and 9 kHz and durations of approximately 70 ms.

4.5. *Recommendations*

Based on our results and those mentioned above, we recommend using feature-based random forest classification, as both feature extrac-tion and computation of random forest distance are fast and compu-tationally inexpensive. To illustrate the ability of the Feature-RF methods to group calls within species, a two-dimensional NMDS or- dination plot of captive calls is shown in Fig. B.1. Although only an unsupervised random forest model is appropriate for this study, as it permits direct comparisons to the three other similarity-based methods, we also investigated the use of a supervised random forest model. In a supervised random forest, trees are constructed with knowledge of the class of each sample. The supervised random forest resulted in an overall accuracy rate of 96.25% on the captive record-ings, 76.67% on the diurnal recordings, and 75.69% on the nocturnal recordings. These results are not directly comparable to the other techniques tested here, but we recommend that this method be fur- ther explored in future studies. Additionally, we recommend the ex-ploration of the option to have an additional class to which calls may be assigned if they are not well suited to any existing classes, as might arise when working with "real world" datasets. Presently, the classifiers presented here cannot assign calls that do not belong to any class to an unknown category, though there is value in adding this option.

Though there is still much unknown about the extent of variation in flight calls, and best practices for applying the few similarity-based clas-sification methods practical for analysis of field recordings are unclear, the different techniques presented here offer insight into potential solu-tions for this challenging problem. If there is any hope of using acoustic monitoring of flight calls in real-time or near-real-time applications, classifications methods that take advantage of feature sets similar to the one used here will be key. Knowing that feature-based methods, rather than template-based methods, are more appropriate for flight call classification, sets the stage for others to apply more advanced clas- sification techniques, such as support vector machines or artificial neu-ral networks, to these challenging signals.

5. Conclusions

We provide strong evidence that similarity-based approaches have great potential for correct species-level classifications of four species of wood-warblers. Although human reviewers are still substantially bet-ter at classification of these signals, human classification is impractical even at the present-day scales of analysis, as tens of thousands of flight calls can be recorded from a single monitoring station in a single migration season. We recommend a continuing line of research using feature-based classification and the random forest distance metric. For the present, we recommend a combination of human review and unsu- pervised random forest methods for the most efficient and accurate analysis of flight call signals. In future work, assigning labels to flight calls automatically could permit human reviewers to continue to over-see classification while greatly improving efficiency and allowing more tedious, low-level classification to be performed by detection-classification algorithms.

The need for automation in acoustic monitoring is critical at multiple steps. The methods that we compared provide a solid foundation for au- tomating the classification process and for future research on additional feature sets needed to

improve classification. Recent work on automatic detection is part of the solution to improving efficiency (Ross and Allen, in press).

References:

A.S. Dave, S.E. Anderson, and D. Margoliash, 1996. Automatic syllable identification from continuous recordings of bird songs using a template-based approach. Publication: Journal of the Acoustical Society of America, Volume 100, Pages 1209-1219 (2010).
In 2010, Armitage and Ober published a work. Bat echolocation call categorization using various supervised learning approaches. Ecological Informatics 5, 465-473.
Published in 2003 by Baker and Logue. Three bioacoustical analysis methods compared for the detection of population difference in a complicated bird song. Published in the Journal of Ethology, volume 109, pages 223-242.
This is from the year 1994 by Baptista and Gaunt. Research on the language of birds has progressed. Article 96, pages 818–830.
In 2011, Barkely et al. published a study. A comparison of ROCCA (real-time odontocete call classification algorithm) species identification of dolphin whistles during and after a voyage was conducted. Request for Proposal (RFP): NOAA-TM-NMFS-SWFSC-473.
By 2011, Blumstein, Mennill, Clemins, Girod, Yao, Patricelli, Deppe, Clark, Cortopassi, Hanser, Ali, McCowan, Kirschel, and Krakauer had made significant contributions to the field. Using micro-phone arrays for acoustic monitoring in terrestrial environments: potential uses, technical issues, and a proposal. Research in Applied Ecology, 48, 758–767.
Published in 2001 by Bradbury, J.W., Cortopassi, K.A., and Clemmons, J.R. The contrast cries of orange-fronted parakeets vary depending on their location. See Auk 118, 958-972.
T.S. Brandes (2008). Instruments for the automated recording and analysis of bird calls for use in conservation efforts. Volume 18, Sections 163–173, of Bird Conservation International.
Breiman first published in 2001. Forests that use random variables. Learning Machines 45, 5-32.
In 2009, Briggs, Fern, and Raich published.... A research on the efficacy of using syllables for the purpose of auditory species identification in birds. See the technical report on bioacoustics at http://eecs.oregonstate.edu/research/.
In 2006, Brown, J.C., Hodgins-Davis, A., and Miller, P.J.O. Published. Analysis of killer whale call types by use of dynamic temporal warping. Volume 119, Issues 34–40, Journal of the Acoustical Society of America.
Jinglan, Z., Binh, P., Roe, P., and Cai, J.(2007). Identifying bird species using a sensor network for environmental monitoring. page 293-298. Catchpole, C.K., and Slater, P.J.B., 1995. Third International Conference on Intelligent Sensors, Sensor Networks, and Information. Themes and Variations in Bird Song: A Biological Perspective. Press of Cambridge University, Cambridge, United Kingdom.
Fristrup, K.M., Clark, C.W., and Charif, R.A. (2004). Raven 1.2 Guide for Users. Location: Ithaca, New York, Cornell Laboratory of Ornithology.
Marler, P., and Beeman, K. (eds.), 1987. Clark, C.W. An application to the song of the swamp sparrow using quantitative study of vocal phonology in animals. No. 76, Ethology, 101–115.
In 2006, Cortopassi was cited. Signal feature measurement that is both automated and robust. This URL is: http://www.birds.cornell.edu/brp/research/algorithms/RSM.html.
In 2000, Cortopassi and Bradbury published a paper. Spectrographic cross-correlation and main coordinates analysis were used to compare sounds that were rich in harmonics. Journal of Bioacoustics 11, pages 89-127.
K.A. Cortopassi and J.W. Bradbury published in 2006. Parrot pairings in the wild, Aratinga canicularis, exhibit a wide spectrum of contact calls. Behavior of animals, 71, 1141–1154.
This sentence is a citation for a 2010 study by Damoulas, Henry, Farnsworth, Lanzone, and Gomes. Aircraft flight call classification using a new dynamic temporal warping kernel developed by Bayes. Page numbers 424–429 from the IEEE Ninth Annual International Conference on Machine Learning and Some Applications.
In 1993, Deller, J.R., Proakis, J.G., and Hansen, J.H.L. Transient Speech Signal Processing.

New York, NY: MacMillian Publishing Co.
This is Ellis (2008). Coordinating musical notation with audio files. here: http://www.ee.columbia.edu/dpwe/resources/matlab/alignmidiwav/;.
[Evans, W.R.] since 1994. The Bicknell's Thrush takes flight at night. Page numbers 55–61 in Wilson Bulletin 106. Published in 1998 by Evans, W.R. Utilization of acoustic bird monitoring techniques in wind power applications. San Diego, California, Proceedings of the Third National Avian-Wind Power Planning Meeting, pages 141–152.
In 1999, Evans and Mellinger published a work. Tracking migrating grassland birds at night. Bird Biology Studies 19, 219–229.
This information is from a 2002 publication by Evans and O'Brien. Flight cries of landbirds in North America's Eastern states that are migrating. “CD-ROM” format. The Ithaca, New York, Old Bird.
In 2000, Evans and Rosenberg published a paper. Using sound to track migratory birds at night: a progress report. Methods for Preserving Birds: The Joint Flight Planning Process with Partners. Workshop on Third-Partner Flight; Cape May, New Jersey.
(Fagerlund, 2007). Use of support vector machines for the identification of bird species. Advanced Signal Processing: An EURASIP Journal, 2007, pp. 1–7.
Authors: Fagerlund and Härmä (2008). The optimization of non-harmonic bird calls for AI detection. Annual Conference on Signal Processing (EUSIPCO2005), 13th Annual, Antalya, Turkey.
Farnsworth, A. (2005) writes. Bird sounds in flight and their significance for ornithological research and preservation efforts. Volume 123, pages 733–746, Apologies.
The flight cries of wood warblers are not just linked to their migration habits (Farnsworth, 2007a). Dean Wilson Journal of Ornithol. 119, 334-341.
Light cries of wood warblers (Parulidae): ecological and evolutionary aspects (Farnsworth, 2007b). Cornell University, Ithaca, New York, USA (Ph.D. Dissertation).
Published in 2005 by Farnsworth and Lovette. Nighttime call evolution in migratory wood warblers: no morphological constraints found. Avian Biology Journal 36, 337-347.
With thanks to Lovette and Farnsworth (2008). Interspecific variation in vocalizations of structurally simple birds: phylogenetic and ecological impacts. Volume 94, Issue 17, pages 155–173. Published in the Biol. Journal of the Linnaeus Society.
In 2007, Farnsworth and Russell collaborated. From an oil rig in the northern Gulf of Mexico, we are listening to the flight sounds of birds that are migrating. (Journal of Field Ornithol., 78, 279-289).
Authors: Farnsworth, A., Gauthreaux Jr., S.A., and Van Blaricom, D., 2004. Doppler radar reflectivity data (WSR-88D) and the number of bird calls heard throughout the night by migratory birds are compared. The citation is from the Journal of Avian Biology, volume 35, pages 365-369.
Written by Fristrup and Watkins in 1992. Distinguishing the auditory characteristics of aquatic animal noises. Report No. WHOI-92-04, authored by the Woods Hole Oceanographic Institution.
Published in 1993 by Fristrup and Watkins. Animal sounds in the ocean categorization. Technical Report WHOI-94-13, Woods Hole Oceanographic Institution.
In 2010, Gagnon, F., Belisle, M., Ibarzabal, J., Vaillancourt, P., and Savard, J. A study comparing the radar reflectivity of a Canadian weather monitoring system with nighttime auditory counts of passerines. No. 127, pages 119–128.
In 2007, Goslee and Urban published a paper. For ecological data analysis based on dissimilarity, there is the ecodist package. Report on statistical software 22, 1–19.
E.E. Henderson, J.A. Hildebrand, and M.H. Smith (2011). Behavior categorization of Lagenorhynchus obliquidens, or Pacific white-sided dolphins, based on their vocalizations. Acoust. Soc. Am. Journal 130, 557-567.
In 2012, Huynh and Mazzocchi published a paper. Please visit OpenRefine at http://openrefine.org.. In 2010, Kasten, E.P., McKinley, P.K., and Gage, S.H. Using ensemble extraction, we can classify and detect different species of birds. Environmental Journal, 5, 153–166.
With contributions from Kasten, E.P., Gage, S.H., Fox, J., and Joo, W. of 2012. An repository for investigating soundscape ecology, the acoustic library is part of the distant environmental assessment

laboratory. Ecological Informatics 12, 50-77.
Jovan A. Kogan and David Margoliash, 1998. Using dynamic temporal warping and hidden Markov models, this research compares automated detection of bird song segments from continuous recording. Published in the Journal of the Acoustical Society of America, volume 103, pages 2185-2196.
A group of authors including Kunz, Arnett, Cooper, Erickson, Larkin, Mabee, Morrison, Strickland, and Szewczak published a study in 2007. Evaluation of the effects of wind energy development on bats and birds that are active at night: a manual for researchers. (Journal of Wildlife Management, 71, 2449–2486).
In 2009, Lanzone, DeLeon, Grove, and Farnsworth published a study. Using a new way to record birds in captivity, we were able to uncover the previously unknown or very little understood flight cries of warblers (Parulidae). Pages 511–519 of Auk 126.
In 2002, Larkin, Evans, and Diehl published a work. During the spring in south Texas, you may hear the nocturnal flight cries of Dickcissels and the echoes of Doppler radar. Publication: Journal of Field Ornithol. 73, lines 2–8.
Liaw and Wiener (2002) produced this work. Use randomForest for classification and regression. The article is in R. News 2, pages 18-22.
Author: Libby, O.G., year: 1899. The flight of migratory birds at night. Pau 16, 140–145.
In 2008, Marcarini, Williamson, and Garcia published a paper. Evaluation of several approaches to the automatic detection of nighttime flight sounds made by birds. pages 2029–2032, 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP.
Apr. 2004 by Marler, P. Calls of birds and their possible applications in behavioral neuroscience. Article number: 1016 in the Annals of the New York Academy of Sciences, pages 31–44.
Apr. 2004 (Marler, P.). An abundance of bird calls: a language toolbox. Birdsong and Their Scientific Significance in the Natural World, edited by P. Marler and H. Slabbekoorn. The academic press Elsevier, New York, pp. 132-177.
Release 2010a of MATLAB. The MathWorks, Inc., located in Natick, Massachusetts, USA. This information is from Mills (1995). Nighttime migratory bird call recognition and categorization using an automated system.
Paper published in the Journal of the Acoustical Society of America (97, 3370).
Mundinger, a law firm, 1970. Perpetual identification of finch calls and vocal replication of such sounds. Publication: Science 168, 480.
In 2013, the R Foundation for Statistical Computing in Vienna, Austria published R: A Language and Environment for Statistical Computing (http://www.R-project.org).
Written by Rabiner and Juang in 1993. An Introduction to Voice Recognition. Prentice Hall, Englewood Cliffs, N.J.
A. Rosenberg, S. Levinson, and L. Rabiner 1978. Points to think about in discrete word recognition systems that dynamically warp time. This is the citation for the article: IEEE Trans. Acoust. Speech 26, 575-582.
In 2008, Ranjard and Ross collaborated. Using evolving neural networks for unsupervised syllable classification in bird songs. Article 123, pages 4358–4368, published in the Journal of the Acoustical Society of America.
James C. Ross and Peter E. Allen, 2014. Random Forest to enhance the effectiveness of passive acoustic monitoring analyses. Environmental Data Science. Retrieved from http://dx.doi.org/10.1016/j.ecoinf.2013.12.002 on December 12, 2013, this article is in press and has the ISSN number 1574-9541.
Sakoe and Chiba published in 1978. An optimization approach for spoken word recognition using dynamic programming. Acoust. Speech, 26(2019), 43–49, published by IEEE.
In 2008, Schrama, Poot, Robb, and Slabbekoorn published a paper. Tracking bird flight calls during night migration using automated systems. Computerized Bioacoustics for Biodiversity Assessment. Proceedings of the International. Meet the Experts on Bioacoustic Pattern Detection using Cyberspace. Parts 131–134, German island of Vilm.
With contributions by Somervuo, Härmä, and Fagerlund (2006). Automatic species identification using

parametric representations of avian noises. 14(252), 2252–263, IEEE Transactions on Acoustic Speech.
In 2000, Tchernichovski et al. worked with Nottebohm, Ho, Pesaran, and Mitra. A method for the automatic evaluation of musical similarity. Publication: Anim. Behav. 59, 1167-1176.
In 2008, Trifa, V.M., Kirshel, A.N.G., and Taylor, C. Using hidden Markov models, antbirds in a Mexican jungle may be automatically recognized. Published in the Journal of the Acoustical Society of America, volume 123, pages 2424–2431.
Authors: Tyagi, H., Hegde, R.M., Murthy, H.A., and Prabhakar, A., 2006. Automatic bird call recognition using averaged spectral voiceprints. Conf. on European Signal Processing, 13th Edition, Florence, Italy. Report.
S.L. Vehrencamp, A.F. Ritter, M. Keever, and J.W. Bradbury (2003). The orange-fronted conure (Aratinga canicularis) reacts differently to nearby and faraway contact calls played back. No. 109, pages 37–54, ethology.
The authors of this work are Venables and Ripley (2). Advanced Statistical Methods with S, 4th edition. New York: Springer. Published in 1968 by Vintsyuk, T.K. Dynamic programming for speech discrimination. Theriogenology 4, 81–88.