JOURNAL OF FIELD ORNITHOLOGY

Published by the
Association of Field Ornithologists

VOL. 76, NO. 1 | Winter 2005 | PAGES 1–108

J. Field Ornithol. 76(1):1–11, 2005

Bioacoustic monitoring of forest songbirds: interpreter variability and effects of configuration and digital processing methods in the laboratory

Robert S. Rempel,[1,4] Keith A. Hobson,[2] George Holborn,[3] Steve L. Van Wilgenburg,[2] and Julie Elliott[1]

[1] *Ontario Ministry of Natural Resources, 955 Oliver Road, Thunder Bay, Ontario P7B 5E1, Canada*
[2] *Canadian Wildlife Service, 115 Perimeter Road, Saskatoon, Saskatchewan S7N 0X4, Canada*
[3] *Ontario Ministry of Natural Resources, Natural Resource Centre, 25th Side-Road, Thunder Bay, Ontario P7C 4T9, Canada*

Received 21 January 2004; accepted 21 April 2004

ABSTRACT. Omnidirectional bioacoustic recording systems offer the ability to record forest songbirds in the field by technical staff, and then interpret the recordings later in the laboratory by skilled interpreters. Among several advantages to this approach are the ability to estimate variance among interpreters, obtain a permanent archival record of the point count, reduce costs by using regular field crews to collect data vs. those skilled in bird identification, and remove impediments to breeding bird surveys due to lack of available skilled birders. In this study we first evaluated the effects of microphone configuration and digital processing methods on the quality and effectiveness of the recordings, and then evaluated how consistently skilled birders interpreted the same songbird recordings collected under a mix of environmental conditions, and related this to the commonness of the species. At the time of this evaluation, the most cost-effective configuration of the bioacoustic monitoring system included use of a 180°/180° microphone combination, a minidisc digital recording system, analog transfer of the sound data via a digital soundcard, post-processing amplification of the signal, and data storage in an .MP3 format. This combination maintains high sound fidelity while minimizing both expense and data storage requirements. As recording device technology improves, the direct storage and digital transfer of .WAV format files will be the preferred and most effective recording option. Despite noisy conditions due to wind and other ambient sounds for many of the recordings, interpreters showed a high level of similarity in species identification and enumeration for the 34 most abundant species. Standardized coefficient of variance increased sharply when species had fewer than 10 occurrences, suggesting that birders are more variable in their identification of rare or uncommon species. Desktop identification systems that include type specimens of spectrographic signature and sound clips to aid interpreters could improve identification accuracy of rarer species.

SINOPSIS. **Monitoreo bioacústico de aves de bosques: variabilidad en el interpretador y efecto de la configuración del método de procedimiento digital en el laboratorio**

Los sistemas de grabación bioacústicos omnidireccionales permiten grabar aves de bosque por personal técnico y luego la interpretación de lo grabado por personal versado o con buena experiencia en el laboratorio. Estre las ventajas de este método se encuentran la habilidad para estimar la variación entre interpretadores, obtener y archivar un record permanente en el lugar de grabación, reducir los costos del trabajo utilizando un pequeño grupo para tomar los datos vs. personal experimentado para la identificación de aves y minimizar los impedimentos de censos de aves no-reproductivas debido a la limitación de observadores experimentados. En este estudio evaluamos, en primer lugar, el efecto de la configuración del micrófono y el método de procesamiento digital en la calidad y efectividad de la grabación. También se evaluó la consistencia del personal experimentado en la interpretación de la grabación del canto bajo condiciones ambientales mixtas, y el relacionar esto con el número de individuos de la especie. Al momento de la evaluación, la configuración más costo-efectiva del sistema de monitoreo bioacústico

[4] Corresponding author. Email: Rob.rempel@mnr.gov.on.ca

incluyó una combinación de micrófonos de 180/180 grados, un sistema de grabación digital con minidiscos, un sistema ánalogo de transferencia vía una tarjeta digital de sonidos, amplificación post proceso digital y el almacenamiento de los datos en formato .MP3. Dicha combinación mantiene un sonido de alta fidelidad y minimiza los gastos y requerimientos de almacenar los datos. Según vaya mejorando la tecnología, también mejorará, el almacenamiento directo y transferencia digital a archivos .WAV lo que será la opción más efectiva y perferida de grabación. No empece a condiciones en donde el viento y otros elementos produjeron ruido, los interpretadores mostraron un alto nivel de similaridad en la identificación de las especies y en el número de las 34 más abundantes. El coeficiente de varianza estandarizado, aumentó marcadamente cuando las especies fueron detectadas en menos de 10 casos, lo que sugiere que los observadores son más variables en la identificacion de especies poco común o raras. Sistemas de identificación que incluyan espectografias del canto típico de aves y muestras de sonido, muy bien pudieran ayudar a los interpretadores a identificar correctamente las especies raras.

Key words: acoustic surveys, digital recording techniques, microphones

Sustainable forest management programs are increasingly being scrutinized for the effectiveness of their conservation strategies. For example, as management agencies develop prescriptive indicators of their wildlife conservation strategies (e.g., amount and distribution of old growth forest), they are also finding that these strategies must be tested by actually monitoring the response of wildlife to changes in forest structure, extent, and pattern (Simberloff 1998; Carlson and Schmiegelow 2002). For forest birds, this task is easier said than done. For example, few resource management agencies have the skilled and experienced staff required to launch an intensive six-week field program to identify and count forest songbirds through song recognition. In addition to cost, the few highly qualified birders that do exist are in high demand during this short period (Hobson et al. 2002).

One possible solution to this lack of skilled individuals in the field is to record songbirds and use automated voice recognition software to identify them. Although promising in theory, the reality of noisy forest environments, bird songs that overlap in frequency and occur simultaneously, regional dialects of songs, and the mixture of loud and quiet birds makes the automated acoustic wave composition analysis of digital sound files very difficult (Anderson et al. 1996; Larkin et al. 1996; Kogan and Margoliash 1998). An alternative approach is to record songbirds using a high quality recording system that attempts to mimic what a birder would hear in the field, and then send these recordings to experienced birders for interpretation later in the year.

Bioacoustic recording systems utilizing microphones such as the Compression Zone Microphone (CZM) attempt to mimic what the birder would hear in the field, and an initial evaluation of such a system was positive (Hobson et al. 2002). These systems can be oriented in a number of configurations involving more than one microphone system. Monitoring schemes such as these quickly accumulate large volumes of digital audio information, necessitating efficient methods of data backup and storage that maintain sufficient audio fidelity to enable interpretations of the data. Many formats exist for performing these tasks, and in particular, recent advances in audio compression (e.g., .MP3) and sound card technology hold great promise. Several questions remain, however. How effective are different microphone configurations in terms of identifying and counting birds, what are the effects of different digital processing methods on the quality of the recordings, and how variable are the songbird interpretations among analysts? We attempted to answer these questions, and to assess the financial costs of alternative methods for a cost-effective configuration of a bioacoustic recording system to monitor the diversity and abundance of forest songbirds.

METHODS

Digital recording and processing procedures. We used the CZM microphone as described in detail by Hobson et al. (2002) to make digital recordings of songbirds. The nature and quality of digital recordings are very much affected by the algorithms and parameters used to encode, compress, store, and transfer the data. The information presented in this section is necessary and sufficient to reproduce our results. It is useful to note that the audio recording options used here are necessary to achieve audio quality similar to CD audio. CD audio consists of 16-bit resolution (16 bits of data to store audio) and a sampling frequency of 44.1

kHz (the number of times per second the data is sampled) in stereo format. A rule of thumb for selection of sampling rate is to double the highest audible frequency. The highest audible frequency for human hearing is about 22 kHz, so a sampling rate of 44.1 kHz should accurately reproduce any audible tone.

Forest songbirds were recorded on the Sony MiniDisc MZ-N707 (MD), which makes use of the Adaptive Transform Acoustic Coding (ATRAC) compression algorithm. This algorithm compresses sound files by omitting psychoacoustically inaudible data, i.e., frequencies that are so similar that the human ear "lumps" them together. This allows large recordings to be stored on relatively small media. ATRAC nomenclature is confusing, and this version is not the MiniDisc Long Play (MDLP) that uses a higher compression ratio (sometimes called ATRAC3), but is rather generation "DSP Type-R" of the ATRAC algorithm, which includes improvements to version 4.5 in the high frequency range, and is significantly improved over versions 1 and 2. This ATRAC perceptual coding system is similar in performance to the ISO-MPEG Audio Layer-3 (.MP3) coding system, popularized by Internet music file sharing (Sablatash and Cooklev 1996). To evaluate the effects of compression, we made simultaneous test recordings on the Tascam DA-P1 Digital Audio Tape (DAT) recorder, which stores recordings in an uncompressed format. A "y-cord" was specially constructed to allow simultaneous recording on both devices from a single microphone combination.

Digital data were transferred to the computer by digital sampling (44.1 or 48 kHz, 16-bit format) using the Soundblaster Live© soundcard. We used the Sony/Philips Digital Interconnect Format (S/PDIF) fibre-optic interface with Pulse Code Modulation (PCM) and saved data in an uncompressed file format, Waveform Audio (.WAV). Data were downloaded using the software provided with the soundcard. To assess quality of analog versus digital recordings, 28 of the recordings were also downloaded directly from the MD stereo headphone-out port (5 mW) to the analog line-in port on the soundcard; this is the typical line-in available on most soundcards (Fig. 1). The DAT recordings were transferred to the computer using the coaxial interface (Fig. 1).

Seven of 84 DAT recordings were recorded and downloaded at the 44.1 kHz sampling frequency; all others were recorded and downloaded at the 48 kHz rate. All the MD recordings were recorded and downloaded with a sampling frequency of 44.1 kHz. The Sony MD Walkmans do not have a digital output port, so the MD disks were placed in a Sony JB940 MiniDisc deck unit, and digital MD recordings were then transferred using the digital optical Toslink (fibre-optic) interface between the MiniDisc deck and the soundcard. Note that the transfer of data from the MiniDisc media requires sampling of the digital datastream via the soundcard by real time playback and recording (on PC using sound editing software), and is not simply a direct file transfer from the MD media to the computer. This is necessitated by proprietary software created by Sony for ATRAC audio compression, making the files unrecognizable to the PC except as an audio signal via the soundcard.

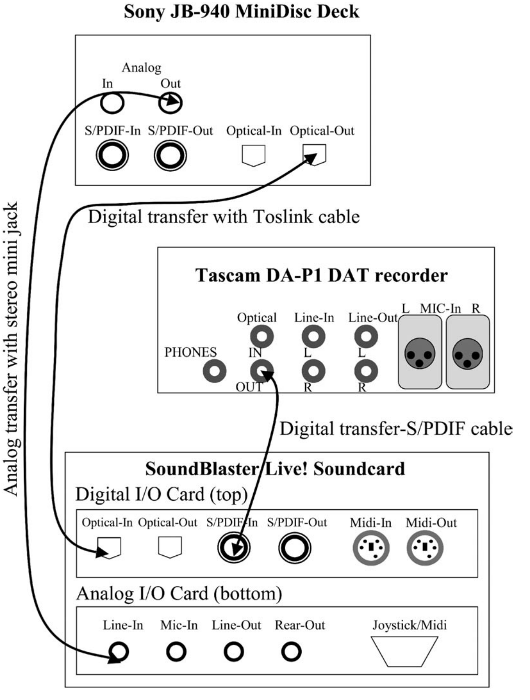

Fig. 1. Schematic of both digital and analog connections used for transfer of audio to the Soundblaster Live© soundcard from MiniDisc and Digital Audio Tape media.

Each recording was categorized according to level and type of environmental noise. Noise included wind effects (principally rattling of

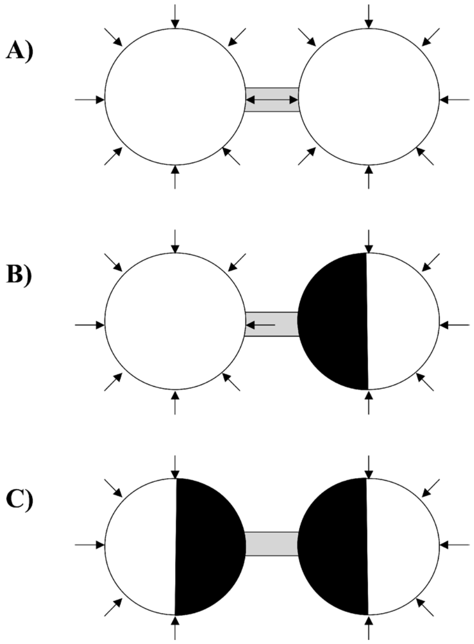

Fig. 2. Stereo configurations of CZM microphones used in experiments: (A) 360°/360°, (B) 360°/180°, (C) 180°/180°. Arrows indicate the directions from which the microphone collects sound waves. Black semi-circles indicate baffles preventing sound from entering the microphone waveguide.

leaves by trembling aspen), mosquitoes, and machinery operating in the area. For tests of perceived storage file format quality, data were re-sampled in Cool Edit 2000 (Syntrillium Software Corporation, Phoenix, AZ) to .MP3 format using the Fraunhofer-Thompson compression scheme, 44.1 kHz sampling rate, and a constant bit rate (256 kbps). The final compression ratio from .WAV file to .MP3 was 5.5: 1. This is a less aggressive compression than many default .MP3 schemes that result in a 10: 1 level of compression.

Analysis of the hardware/software configuration. Three microphone configurations were evaluated in terms of perceived sound quality, and the ability of interpreters to spatially discriminate the location of birds. The microphones are available as either fully omnidirectional (360°) or semi-omnidirectional (180°), where baffles block the sound for half the microphone opening. To achieve stereo recordings that allow estimation of bird abundances, two microphones separated by approximately 30 cm in configurations of 180°/180°, 360°/180°, and 360°/360° were used (Fig. 2). In each instance, the microphones create separate left and right audio channels, and the human ear perceives time delays between channels, creating the stereo effect. All three configurations were tested by playing recordings of White-throated Sparrows (*Zonotrichia albicolis*), Least Flycatchers (*Empidonax minimus*), and Hermit Thrushes (*Catharus guttatus*) along an arc at distances of 10 and 20 m. These songbird species were selected for the test because they range in frequency and quality of call, and hence their potential effect on listeners' expected ability to discern location.

Table 1. Assessment of recording quality by one interpreter (JE).

Recording quality	Number of recordings
Good	13
Hiss	5
Wind minimal	4
Wind moderate	4
Mosquitoes	2
Machinery	1
Raindrops	1

Sound-clips (10 s) of songbird recordings made as above were sent to four interpreters, which included three of us (SVW, RSR, GH) and one other experienced songbird interpreter. For each test, recordings were randomly mixed, and assigned an index number referencing the recording to test conditions by one of us that did not participate in the interpretation (JE). Hence, the test was "blind" because none of the interpreters had access to this information.

For perceived sound quality, interpreters were asked to assign each recording to one of four quality categories: poor (1), acceptable (2), high (3), and exceptional (4). For spatial discrimination, interpreters were asked which of three recordings provided better spatial discrimination (A, B or C? N = no obvious or easily discernable difference). Results were analyzed with ANOVA, with Student-Neuman-Keuls (SNK) post-hoc hypothesis testing.

We evaluated the effect of recording methods on sound quality with the following three tests. (1) Similarity and quality of recordings made on MD versus DAT recording devices was tested by simultaneously recording to both devices.

To remove the possible effect of different microphone acoustics, a y-cord was constructed to simultaneously connect cables from a single set of microphones to both the MD and DAT recorder. This also ensured that the interpreter could interpret the same recording, and that the similarity (correlation) between spectrograms could be calculated. (2) Similarity and quality of recordings made by digital transfer versus analog transfer was tested by transferring the same MD recording using both digital and analog methods. Analog transfer, unlike digital transfer, results in a loss of acoustic gain (i.e., volume). As a result, analog recordings were initially of lower amplitude than the digital recording. To remedy this, all analog recordings were amplified by 750% using sound analysis software (Cool Edit) to create recordings of virtually identical volume. (3) Similarity and quality of recordings replayed using uncompressed file format (.WAV) versus compressed file format (.MP3) was tested by saving files in both formats. The file sizes of the two formats differ, so to hide the identity of the format, the .MP3 files were first compressed from .WAV to .MP3, and then saved again as .WAV. This resulted in files of identical size and the same extension name, but with data resolutions inherent to the originating format; data omitted in the initial compression to .MP3 files is permanently lost. This test was based on recordings downloaded from the MD (ATRAC internal format).

For the above three questions, interpreters were asked which recording sounded better (A, B, or N), with results later recoded to -1, 1, or 0, respectively. For example, the quality of MD versus DAT recordings was scored for 84 pairs of recordings by selecting the best recording (randomly labeled as either A or B), and later we assigned either -1 (MD), 1 (DAT), or 0 (no detectable difference) to that response. If the interpreters selected no detectable difference, and/or if DAT and MD were selected with no evident preference, then the test score approached 0. If one or the other device was preferred, then the score was expected to deviate significantly from 0. Results were analyzed using a one-sample t-test, where the test value was set to 0 (i.e., the null hypothesis is that there is no evident preference), and with a two-tailed test.

These three questions were also studied using numerical spectrographic cross-correlation (SPCC) techniques. For each pair of recordings, a spectrogram was created in AVISOFT, and saved in binary format. The correlation coefficient of the two spectrograms was calculated by a routine that incrementally shifts the position of 1 spectrogram relative to the other, and then determines the maximum correlation coefficient (r). Because we were testing the similarity of two theoretically identical sounds, we did not specify any cut-off frequencies, or frequency tolerances for the correlation. The sound clips were of identical length, and this eliminates potential issues with SPCC that result from comparisons of sounds of different lengths. If used correctly, the SPCC approach is an effective tool for examining bioacoustic hypotheses (Cortopassi and Bradbury 2000).

The effect of recording level on sound quality was tested by recording at low, medium, or high, or with Automatic Volume Limiter System (AVLS) engaged on the Sony MiniDisc systems.

Analysis of variability among songbird interpreters. Thirty 10-min field recordings were copied and transferred to CDs, and sent to each of six experienced songbird interpreters. Each interpreter had experience in songbird identification for programs such as the Ontario Breeding Bird Atlas (Cadman et al. 1987), North American Breeding Bird Survey (Bystrak 1981), Forest Bird Monitoring Program (Welsh 1995) or for studies where results were published in peer reviewed journals. Songs were recorded using the bioacoustic monitoring system based on the CZM (formerly CVX) microphone (Hobson et al. 2002) following standard point count protocols. Songbird recordings were made between 5:00 and 10:10, for the period from 30 May to 8 July 2002. Recordings were not made under high wind or rainy conditions, and followed the environmental condition protocol of the Forest Bird Monitoring Program (Welsh 1995). All data were collected in the mesic boreal mixedwood region of northwest Ontario, with a mid-point location of approximately 89.0°W, 49.6°N. No identifying information concerning site or habitat was included. The recording quality label (high or poor) was not revealed, and the names of other interpreters were not divulged. Interpreters were requested to identify birds heard during the first five minutes, and then to continue recording new individuals for the full 10 min of the re-

Table 2. Total number of individuals of each bird species heard across all sites, for interpreters A–F.

Common name	Scientific name	Total number of individuals per species after 10 min						CoV
		A	B	C	D	E	F	
Ruby-throated Hummingbird	*Archilochus colubris*	0	0	0	0	1	0	223.61
Yellow-bellied Flycatcher	*Empidonax flaviventris*	27	23	24	19	18	16	17.97
Alder Flycatcher	*E. alnorum*	12	12	8	8	12	8	20
Least Flycatcher	*E. minimus*	4	1	1	1	2	2	58.21
Solitary Vireo	*Vireo solitarius*	2	2	2	1	0	1	55.9
Philadelphia Vireo	*V. philadelphicus*	0	2	0	0	0	0	223.61
Red-eyed Vireo	*V. olivaceus*	11	7	8	10	12	8	19.23
Gray Jay	*Perisoreus canadensis*	3	2	2	0	1	1	63.83
Blue Jay	*Cyanocitta cristata*	4	3	4	1	5	3	37.42
Boreal Chickadee	*Poecile hudsonica*	4	3	5	6	3	4	25.61
Red-breasted Nuthatch	*Sitta canadensis*	4	1	4	4	5	2	41.23
Brown Creeper	*Certhia americana*	2	2	1	0	1	1	58.9
Winter Wren	*Troglodytes troglodytes*	16	17	18	19	17	18	5.47
Golden-crowned Kinglet	*Regulus satrapa*	21	21	20	10	20	16	21.99
Ruby-crowned Kinglet	*R. calendula*	15	11	12	16	14	11	14.82
Veery	*Catharus fuscescens*	0	1	0	0	0	0	223.61
Swainson's Thrush	*C. ustulatus*	16	17	22	17	18	11	19.22
Hermit Thrush	*C. guttatus*	17	13	17	11	10	16	20.2
American Robin	*Turdus migratorius*	9	11	14	5	7	8	32.08
Cedar Waxwing	*Bombycilla cedrorum*	4	6	3	1	4	1	55.95
Tennessee Warbler	*Vermivora peregrina*	10	4	5	5	6	4	36.26
Orange-crowned Warbler	*V. celata*	0	0	0	0	11	0	223.61
Nashville Warbler	*V. ruficapilla*	46	47	32	22	38	56	27.51
Northern Parula	*Parula americana*	1	1	0	6	0	0	160.08
Yellow Warbler	*Dendroica petechia*	0	0	1	0	0	0	223.61
Chestnut-sided Warbler	*D. pensylvanica*	6	4	1	4	1	2	60.86
Magnolia Warbler	*D. magnolia*	20	23	17	21	21	29	16.81
Cape May Warbler	*D. tigrina*	2	2	1	1	4	0	74.83
Black-throated Blue Warbler	*D. caerulescens*	1	1	1	1	3	1	55.9
Yellow-rumped Warbler	*D. coronata*	13	5	21	7	8	10	49.21
Black-throated Green Warbler	*D. virens*	3	2	3	1	3	1	41.42
Blackburnian Warbler	*D. fusca*	0	1	0	4	4	0	120.19
Palm Warbler	*D. palmarum*	9	6	3	10	4	4	44.1
Bay-breasted Warbler	*D. castanea*	11	12	8	2	1	0	85.65
Blackpoll Warbler	*D. striata*	0	0	1	0	0	0	223.61

Table 2. Continued.

Common name	Scientific name	Total number of individuals per species after 10 min						CoV
		A	B	C	D	E	F	
Black-and-white Warbler	*Mniotilta varia*	2	1	2	3	1	0	63.83
American Redstart	*Setophaga ruticilla*	0	0	8	2	0	12	127.6
Ovenbird	*Seiurus aurocapilla*	15	16	13	13	14	16	8.68
Northern Waterthrush	*S. noveboracensis*	3	0	4	3	2	7	66.78
Connecticut Warbler	*Oporornis agilis*	2	2	1	1	0	0	81.65
Mourning Warbler	*O. philadelphia*	5	7	3	3	1	4	48.61
Common Yellowthroat	*Geothlypis trichas*	4	2	1	2	3	5	47.43
Wilson's Warbler	*Wilsonia pusilla*	0	3	0	0	0	0	223.61
Canada Warbler	*W. canadensis*	1	0	0	3	0	0	165.83
Scarlet Tanager	*Piranga olivacea*	0	0	0	0	2	0	223.61
Chipping Sparrow	*Spizella passerina*	8	6	5	2	5	7	34.42
Clay-colored Sparrow	*S. pallida*	1	0	0	1	0	0	141.42
Vesper Sparrow	*Pooecetes gramineus*	0	0	1	0	0	0	223.61
Lincoln's Sparrow	*Melospiza lincolnii*	9	6	3	3	0	3	70.71
Swamp Sparrow	*M. georgiana*	1	0	0	1	0	0	141.42
White-throated Sparrow	*Zonotrichia albicollis*	45	42	55	50	47	46	8.66
Dark-eyed Junco	*Junco hyemalis*	4	5	9	1	0	2	85.32
Rose-breasted Grosbeak	*Pheucticus ludovicianus*	1	2	1	0	0	0	111.8
Indigo Bunting	*Passerina cyanea*	0	0	0	0	1	0	223.61
White-winged Crossbill	*Loxia leucoptera*	11	3	0	0	0	0	172.61
Pine Siskin	*Carduelis pinus*	4	2	4	0	0	0	107.7
American Goldfinch	*C. tristis*	0	0	0	1	0	0	223.61
Evening Grosbeak	*Coccothraustes vespertinus*	1	2	1	0	0	0	111.8
Total species		45	45	44	42	38	34	

Table 3. Community similarity measure (for all 58 identified songbirds) to evaluate similarity of the observed bird community among interpreters (A–F).

	Sorensen coefficient (CC)[a]				
	B	C	D	E	F
A	0.9111	0.8989	0.8966	0.7952	0.8354
B		0.8764	0.8276	0.7952	0.8101
C			0.8372	0.8049	0.8718
D				0.8000	0.8421
E					0.8333

[a] CC = $2c/s_1 + s_2$, where CC is the Sorensen community coefficient, c is the number of species in common between observers, and s_1 and s_2 are the number of species identified by interpreters 1 and 2, respectively.

Table 4. Community similarity measure (for the 34 most abundant songbirds) to evaluate similarity of the observed bird community among interpreters (A–F).

	Sorensen coefficient (CC)				
	B	C	D	E	F
A	0.9851	0.9851	0.9538	0.9375	0.9375
B	0	0.9697	0.9375	0.9206	0.9206
C	0	0	0.9688	0.9524	0.9524
D	0	0	0	0.9508	0.9508
E	0	0	0	0	0.9333

cording. Interpreters were asked to listen to the recordings in the same manner as they would in an operational setting, with no specific advice on listening to recordings more than once. Interpreters were not informed a priori that the results would be part of an evaluation of the CZM microphone; hence these results are not tests of interpreter ability, but simply an evaluation of consistency among interpreters in an operational setting.

The analysis of how similar songbird identification and counts were among interpreters was conducted using the Sørensen community similarity coefficient (Sørensen 1948). This metric considers both the commonness of species identification, and the similarity of species counts. A computer program, based on Basic code in Brower et al. (1997), was written by RSR in Visual Basic to calculate similarity between all pairs of observation sets and produce the similarity matrix.

RESULTS AND DISCUSSION

Analysis of the hardware/software configuration. *Effect of microphone configuration on sound quality.* Interpreters rated the 180°/180° combination best in terms of spatial discrimination. The average scores for the 28 recordings of 180°/180°, 360°/180°, and 360°/360° combinations were 2.96, 2.39, and 2.32, respectively. The combinations were significantly different (among combinations $F_2 = 6.4$, $P = 0.003$), with the 180°/180° combination significantly higher than the 360°/180° and 360°/360° group (SNK $\alpha = 0.05$). Most recordings scored either a 2 (acceptable quality) or 3 (high quality), so the 180°/180° combination was scored as high quality on average and the other two combinations as acceptable quality.

Effect of microphone configuration on spatial discrimination. Interpreters rated the 180°/180° combination best in terms of sound quality. The average scores for 90 recordings of 180°/180°, 360°/180°, and 360°/360° combinations were 0.408, 0.317, and 0.117, respectively, where the recording was scored 1 if it provided the best spatial discrimination among the three recordings. The combinations were significantly different (among combinations $F_2 = 14.2$, $P < 0.0001$), with the 180°/180° and 360°/180° combinations being significantly different from the 360°/360° combination in terms of spatial discrimination (SNK $\alpha = 0.05$). The 180°/180° combination provided the best ability to spatially discriminate the location of birds, but the 360°/180° combination also performed well. There was no effect of species used in the playback ($F_2 = 0.10$, $P = 0.90$)

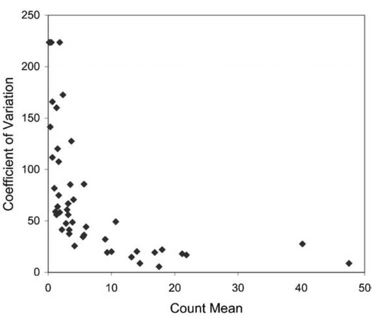

Fig. 3. Coefficient of variation plotted against mean count (among the six interpreters), for each forest songbird species.

nor an interaction of microphone combination and playback species ($P = 0.11$).

MD versus DAT. Interpreters could not discern the difference in quality between recording formats. The average score of 84 pairs was 0.143 (SE = 0.078), and the score did not differ significantly from 0 (one-sample *t*-test; df = 83, $P = 0.07$). The recordings had almost identical spectrographic signatures, with a mean correlation (r) of 0.968 (SD = 0.032). Both the MD and DAT datastream are digitally sampled in the soundcard at the same rate (44.1 kHz/s), and the resulting digital spectrographic signature is similar between the two sources. Although this test found no evidence that the two recording formats differed, these results must also be viewed with some caution because the variance resulted in a relatively weak test with a power of only 0.56. There is considerable variance in anecdotal opinion of the quality of the ATRAC internal compression algorithm used by the MD, with some Internet reports saying ATRAC is much better than .MP3 and others reporting that ATRAC should never be used for recording birds. Essentially these results compare the pure, uncompressed .WAV format (as download from the Tascam DAT), with the compressed MD ATRAC format, and found no evidence of a strong difference.

Recording level. The effect of recording level on sound quality was evaluated by listening to recordings made at either low, medium, high, or set to AVLS, and selecting the best quality recording. Of eight pairs, only recordings made at high or moderate levels were selected as best, with AVLS and low levels never selected. AVLS is responsive to the closest (or loudest) bird singing, and hence will decrease the volume if a nearby bird is calling loudly, making it more difficult to hear the quieter birds.

Digital versus analog download. Interpreters could not detect a difference between the amplified analog and digital downloaded versions of the recordings (one-sample *t*-test, mean = -0.071, df = 27, $P = 0.16$). The spectrographic similarity between pairs was moderate (average $r = 0.85$, SD = 0.091). Thus downloading data from a recorder that does not have a digital output port (using the headphone output port) did not have a detectable effect on interpretability of the sounds. This value was lower than the similarity of recording pairs made using the MD versus DAT devices.

.MP3 versus .WAV format. Interpreters could not discern the difference in quality between MD recordings stored using the .MP3 versus .WAV formats. The mean response value (-0.107) did not differ significantly from 0 (one-sample *t*-test, df = 27, $P = 0.18$). Spectral similarity of each pair of recordings was again moderate ($r = 0.86$; SD = 0.021). However, these tests were based on data originating from the MD, which uses the real time, internal ATRAC compression algorithm. Thus this test simply determines if archiving the .WAV format download files as .MP3 has a negative effect on interpretation quality. Because the MD files were already compressed internally, it is not unexpected that there was no discernable difference. Note that this result does not apply if the source of the data was the Tascam DAT recorder, in which case the difference in spectrographic signatures between .MP3 and .WAV would likely be much greater.

Variability among songbird interpreters. To represent the range of conditions that occurs in an "operational" setting, test recordings of variable quality were selected. Verbal and written response from the interpreters confirmed that the interpreters felt many of the recordings were of poor quality, with only 13 of the 30 recordings classified as "good" (Table 1). Wind and mosquitoes were the two most disturbing sounds reported.

A cumulative total of 58 songbird species were identified (Table 2) by all observers. The number of species identified by individual interpreters varied between 45 and 34 species over the 30 recordings, with interpreters A–C identifying 44–45 species. None of the identifications were subjected to further scrutiny or verification, as the emphasis of the study is on variability among interpreters, not absolute accuracy.

The full list of species counts (Table 2) was analyzed using community similarity measures. These indices compare the similarity of "communities" based on the combined similarity of species presence, and the relative abundance of individuals within each species (Table 3). The bird communities, as identified by interpreters A–D, had high similarity with each other, whereas the communities identified by interpreters E–F had lower similarity. This is in part because interpreters E–F identified fewer and different species, and their counts were more variable.

Substantially higher similarity values were achieved among all interpreters when only the 34 most abundant species were included in the analysis (Table 4). For these species most observers identified at least one or two individuals. Visual inspection of the species counts (Table 2) confirms the finding that there is a high degree of similarity among interpreters for the more abundant species. Variance in estimated abundance (expressed as coefficient of variance [COV], where COV = SD/mean) decreased with increasing abundance (Fig. 3). For total abundance of 10 or more individuals, variance stabilized, and there was little variability among birders in abundance counts.

These results indicate that even with some distracting noise, interpreters consistently identified the same common species, and produced similar counts of individuals. Each interpreter made sound clips of the species they identified, so confirmation of rare species could be achieved by comparison of the sound clips with archived recordings.

CONCLUSIONS

Considering cost, complexity of setup, and quality of recordings, the most effective recording configuration is the 180°/180° microphone pair, using the MD digital recorder. The MD recorder is substantially less expensive than a DAT recorder, is much lighter, and requires fewer and lighter batteries. There was no listener discernable difference in quality of recordings using either method. The 180°/180° microphone configuration provides the best spatial discrimination, and the best quality of sound in the recordings, although the 360°/180° configuration is also effective.

The most cost-effective data transfer approach of those tested is to download recordings via the headphone port to an analog line-in port on the computer, and post-process this data by amplifying 750% and saving the data as .MP3 using a single batch processing script. There was no difference in quality between digital and analog downloads, nor between .WAV and .MP3 file formats. This approach avoids the purchase of the expensive MD deck unit (e.g., JB940), and the complex setup of an optical digital I/O firmware system in the computer. It also allows considerable (~80%) savings of disk storage space by using the compressed .MP3 data format. The only additional cost is the purchase of digital sound editing software (for amplification and conversion of files to .MP3 format), but at this time good software can be purchased for < $80 US.

Sound recording equipment will undoubtedly change over the next few years. Recording media will change and portable memory (e.g., microdrives and flash memory cards) will replace disks as the favored media upon which recordings are stored and transferred. The results from this study will continue to apply in terms of sound quality assessment if similar compression algorithms are used by the recording device (e.g., .MP3). The cost-effectiveness of alternative strategies, however, will likely shift if uncompressed data (e.g., linear PCM .WAV format) can be easily transferred to computers using inexpensive microdrives. The preferred recording configuration will have no moving parts (reducing mechanical failure), have sufficient memory to store multiple recordings (avoiding costly storage disks), allow direct recording of sounds in .WAV format, allow direct file transfer of songs to the computer (speeding data transfer rates and maintaining original recording quality), and have low power requirements (reducing battery weight and expense).

Our research suggests that interpreter variability is likely most significant for rare species, as has been found by previous work (Kepler and Scott 1981). Interpreter variability declined and stabilized at a frequency of about 10 individuals. This is likely no different than what would occur with normal field sampling, and suggests that for statistical analysis of trends, or establishing relationships with resource management practices, interpreter error potentially becomes an important component for analysis of species that occur on less than 10 occasions in a survey. Analysis of these bioacoustic recording data demonstrates that at these lower levels, observations of rare or uncommon songbirds are more likely to be in error.

Sound-clips for unidentified birds can be extracted and compared to validated type specimens of songbird songs and calls, and this approach may be of particular value for the monitoring and documentation of rare and uncommon species. An audio identification system, comprised of desktop software to display observed and type specimen spectrographic signatures and audio clips would undoubtedly aid the identification of the rarer species, and result in a higher similarity among all observers and possibly reduced processing time. Analogous recording approaches have been used by field observers, with recordings of unidentifiable calls being made by miniparabolic microphones, with identification against type specimens proving useful in improving data quality (e.g., Hobson and Schieck 1999). Furthermore, sharing of sound clips between interpreters may serve as a training tool, allowing for further reduction in observer or interpreter variability. Robbins and Stallcup (1981) identified numerous sources of error in species identification, including for example observers' lack of familiarity with local dialect and dependence on bird song tapes recorded mostly in eastern North America as training aids. Comparison of sound-clips would be particularly useful in reducing observer variability and/or errors in these circumstances.

This omnidirectional recording system described here is not designed for recording single individuals, and hence is not necessarily suited for the creation of an archive of type specimen songs and calls (although we have obtained excellent recordings using this system). The twin plate design of the microphone may filter sounds in the 200–400 Hz band, thus limiting it usefulness for recording certain owls, columbids, and grouse (A. McCallum, pers. comm.). Directional parabolic or shotgun microphones are more suited for recording type specimen sound clips, and the appropriate hardware/software configuration would probably differ, with no compromises in quality acceptable (Wickstrom 1982). A general recommendation that we make is that directional microphone recordings may be improved if they are used in a stereo configuration, with recordings made with a stereo recording device.

The bioacoustic configuration described here, however, is appropriate for cost-effective songbird monitoring, and for the collection and documentation of sounds arising from entire bird communities occurring in different forest environments.

ACKNOWLEDGMENTS

Brian Turnbull developed the CZM microphone system and has been instrumental in assisting with its further refinement and development. We thank the individuals who participated in our testing of the audio recordings. We also thank Arch McCallum for his very helpful comments on the manuscript.

LITERATURE CITED

ANDERSON, S. E., A. S. DAVE, AND D. MARGOLIASH. 1996. Template-based automatic recognition of birdsong syllables from continuous recordings. Journal of the Acoustical Society of America 100: 1209–1219.

BROWER, J. E., J. H. ZAR, AND C. N. VON ENDE. 1997. Field and laboratory methods for general ecology, 4th ed. Willian C. Brown, Dubuque, IA.

BYSTRAK, D. 1981. The North American Breeding Bird Survey. Studies in Avian Biology 6: 34–41.

CADMAN, M. D., P. F. J. EALGES, AND F. M. HELLEINER. 1987. Atlas of the breeding birds of Ontario. University of Waterloo Press, Waterloo, ON.

CARLSON, M., AND F. SCHMIEGELOW. 2002. Cost-effective sampling design applied to large-scale monitoring of boreal birds. Conservation Ecology 6: 11. http://www.consecol.org/vol6/iss2/art11.

CORTOPASSI, K. A., AND J. W. BRADBURY. 2000. The comparison of harmonically rich sounds using spectrographic cross-correlation and principal coordinates analysis. Bioacoustics 11: 89–127.

HOBSON, K. A., AND J. SCHIECK. 1999. Changes in bird communities in boreal mixedwood forest: harvest and wildfire effects over 30 years. Ecological Applications 9: 849–863.

———, R. S. REMPEL, H. GREENWOOD, B. TURNBULL, AND S. L. VAN WILGENBURG. 2002. Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. Wildife Society Bulletin 30: 709–720.

KEPLER, C. B., AND J. M. SCOTT. 1981. Reducing count variability by training observers. Studies in Avian Biology 6: 366–371.

KOGAN, J. A., AND D. MARGOLIASH. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. Journal of the Acoustical Society of America 103: 2185–2196.

LARKIN, R. P., D. MARGOLIASH, J. A. KOGAN, AND L. L. PATER. 1996. Recognition of the utterances of terrestrial wildlife: a new approach. Journal of the Acoustical Society of America 99: 2532.

ROBBINS, C. S., AND R. W. STALCUP. 1981. Problems in separating species with similar habitats and vocalizations. Studies in Avian Biology 6: 360–365.

SABLATASH, M., AND T. COOKLEV. 1996. Compression of high-quality audio signals, including recent methods using wavelet packets. Digital Signal Processing 6: 96–107.

SIMBERLOFF, D. 1998. Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? Biological Conservation 83: 247–257.

SØRENSEN, T. 1948. A method for establishing groups of equal amplitude in plant sociology based on similarity of species content. K. Dansk. Vidensk. Selsk. Biol. Skrift 5: 2–16.

WELSH, D. A. 1995. An overview of the forest bird monitoring program in Ontario, Canada. In: Monitoring bird populations by point counts (C. J. Ralph, J. R. Sauer, and S. Droege, eds.), pp. 93–97. U.S. Department of Agriculture Forest Service General Technical Report PSW-GTR-149, Pacific Southwest Research Station, Albany, CA.

WICKSTROM, D. C. 1982. Factors to consider in recording avian sounds. In: Acoustic communication in birds, vol. 1 (D. E. Kroodsma, E. H. Miller, and H. Ouellet, eds.), pp. 2–52. Academic Press, New York, NY.