TESTING THE EFFECTIVENESS OF AUTOMATED ACOUSTIC SENSORS FOR MONITORING VOCAL ACTIVITY OF MARBLED MURRELETS *BRACHYRAMPHUS MARMORATUS*

JENNA L. CRAGG[1,2], ALAN E. BURGER[1] & JOHN F. PIATT[3]

[1]*Department of Biology, University of Victoria, Victoria, BC V8W 3N5, Canada (jenna.cragg@gmail.com)*
[2]*Current address: 2876 Philip Street, Duncan, BC V9L 2A7, Canada*
[3]*US Geological Survey, Alaska Science Center, Anchorage, AK 99503, USA*

Received 26 June 2014; accepted 17 April 2015

SUMMARY

Cragg, J.L., Burger, A.E. & Piatt, J.F. 2015. Testing the effectiveness of automated acoustic sensors for monitoring vocal activity of Marbled Murrelets *Brachyramphus marmoratus*. *Marine Ornithology* 43: 151–160.

Cryptic nest sites and secretive breeding behavior make population estimates and monitoring of Marbled Murrelets *Brachyramphus marmoratus* difficult and expensive. Standard audio-visual and radar protocols have been refined but require intensive field time by trained personnel. We examined the detection range of automated sound recorders (Song Meters; Wildlife Acoustics Inc.) and the reliability of automated recognition models ("recognizers") for identifying and quantifying Marbled Murrelet vocalizations during the 2011 and 2012 breeding seasons at Kodiak Island, Alaska. The detection range of murrelet calls by Song Meters was estimated to be 60 m. Recognizers detected 20 632 murrelet calls (*keer* and *keheer*) from a sample of 268 h of recordings, yielding 5 870 call series, which compared favorably with human scanning of spectrograms (on average detecting 95% of the number of call series identified by a human observer, but not necessarily the same call series). The false-negative rate (percentage of murrelet call series that the recognizers failed to detect) was 32%, mainly involving weak calls and short call series. False-positives (other sounds included by recognizers as murrelet calls) were primarily due to complex songs of other bird species, wind and rain. False-positives were lower in forest nesting habitat (48%) and highest in shrubby vegetation where calls of other birds were common (97%–99%). Acoustic recorders tracked spatial and seasonal trends in vocal activity, with higher call detections in high-quality forested habitat and during late July/early August. Automated acoustic monitoring of Marbled Murrelet calls could provide cost-effective, valuable information for assessing habitat use and temporal and spatial trends in nesting activity; reliability is dependent on careful placement of sensors to minimize false-positives and on prudent application of digital recognizers with visual checking of spectrograms.

Keywords: Marbled Murrelet, *Brachyramphus marmoratus*, automated acoustic recording, population monitoring, habitat use, vocalization, Kodiak Island

INTRODUCTION

Marbled Murrelets *Brachyramphus marmoratus* nest at low density in cryptic, dispersed and often inaccessible locations, usually high in old-growth trees (Nelson 1997, Piatt *et al.* 2007). Flights to and from nests typically occur in dark pre-dawn or dusk hours. The secretive breeding behavior and cryptic nest sites of murrelets make it difficult to census local populations, study behavior and determine habitat use; however, this information is needed to monitor threatened populations and their nesting habitats and to set conservation priorities (McShane *et al.* 2004, Piatt *et al.* 2007, Miller *et al*. 2012). Logging of old-growth nesting habitat has been identified as a main factor in population declines, and consequently a major focus of recovery planning is identifying high-quality forest nesting habitat for protection (CMMRT 2003, Raphael 2006). Monitoring the occurrence and behavior of murrelets in forest stands can provide important information for managers attempting to identify, rank and map nesting habitat used by murrelets (Burger & Bahn 2004, Meyer *et al.* 2004, Stauffer *et al.* 2004, Bigger *et al.* 2006).

Automated acoustic recording systems have recently been developed as a cost-effective alternative to deploying personnel for monitoring remote, nocturnal and elusive populations of seabirds (Buxton & Jones 2012, McKown *et al*. 2012, Buxton *et al.* 2013, Borker *et al.* 2014, Oppel *et al.* 2014). In this study we tested whether automated acoustic methods could be used to monitor vocal behavior and relative abundance of Marbled Murrelets. Currently, murrelet presence and habitat use are evaluated at the forest stand-level using standardized audio-visual or radar surveys conducted by human observers (Evans Mack *et al.* 2003, Burger 1997, Cooper *et al.* 2001); however, these methods have several limitations (Evans Mack *et al.* 2003, Bigger *et al.* 2006). Costs are high to support field crews, especially in remote areas. Observers can survey only one site at a time; therefore, spatial and temporal replication is reduced. Finally, audio-visual observers are subject to observer bias and varying viewing conditions when monitoring murrelet behavior.

Marbled Murrelets (hereafter, murrelet(s), unless noted) are a suitable candidate for automated acoustic monitoring because of their conspicuous pre-dawn vocalizations while flying above or near nesting habitat during the breeding season (Nelson 1997, Dechesne 1998). We used automated acoustic sensors to record murrelet vocalizations across a range of habitat types in the Kodiak Archipelago, Alaska, to address three questions: What is the detection range of automated sensors for murrelet vocalizations? Do recognizer algorithms developed for identifying murrelet calls in digital recordings reliably detect the occurrence and frequency of these calls? How can automated acoustic methods best be applied for studying and monitoring murrelets? To help address the last question, we analyzed spatial and seasonal variations in acoustic detections.

STUDY AREA AND METHODS

Field methods

We tested automated acoustic sensors in the Kodiak Archipelago during the murrelet breeding season (June–August) in 2011 and 2012. Two types of Song Meter acoustic sensors (Wildlife Acoustics, Inc., Concord, MA) were used: general purpose SM2 Terrestrial sensors (mounted 1 m above ground), and SM2 Night Flight sensors designed to detect distant calls while attenuating noise from below (mounted 3–4 m above ground). The performance of two types of Song Meter simultaneously deployed at the same tree in 2011 were compared: the SM2 Night Flight model had lower mean detections than the SM2 Terrestrial model (Table 1). Although this difference was statistically significant (Wilcoxon signed rank test: $N = 15$, $V = 96$, $P = 0.043$), the numerical difference was small (4%), and we therefore used data from both sensors in analyses. Song Meters were programmed to record for 2 h each day, starting 2 h before sunrise (sample rate of 16 000 Hz, gain 0.0 dB on both microphones). Forests were considered to be potential nesting habitat if they provided large trees with suitable nest-site platforms, usually in the form of thick epiphytic moss mats (Burger 2004), but we did not attempt to confirm nesting (actual egg laying) in any forest sites. In forested habitat, sensors were mounted on trees. In unforested habitat, sensors were mounted on 1 m posts and the surrounding vegetation was cleared within a 1.5 m radius to reduce noise interference.

In 2011, sensors were placed at six locations across a variety of habitat types at one site, Monashka Bay (57°50′N, 152°26′W) on Kodiak Island (Table 1), from 15 June to 3 September (Cragg 2013). In 2012, three sensors were deployed at Monashka Bay throughout the breeding season, from 1 June to 27 August. Two locations used in 2011 were re-used in 2012, based on preliminary results (relatively low noise interference and consistent presence of murrelets). These included sensor “Forest 1,” located in a forest stand with potential nesting habitat, and sensor “Flight path 5,” located along a flight path regularly used by murrelets commuting to and from inland nesting areas, as identified by radar surveys (Cragg 2013). A third sensor (“Forest 3”) was added in 2012 to sample a separate patch of potential forest nesting habitat in the Monashka Bay study area. Also in 2012 we collected one 2 h pre-sunrise recording in potential forest nesting habitat or flight paths at each of five new sites in the northern Kodiak Archipelago (see Cragg 2013 for details). The single recordings from these additional sites were used to compare potential effects of habitat types on the performance of the Song Meters.

We tested the maximum distance at which murrelet calls could be detected by Song Meters in different acoustic environments. A 20 s series of 35 murrelet calls, including all four call types detected in the field (*keer*, *keheer*, *ay*, *quack*; Dechesne 1998, Cragg 2013) was played from an iPhone speaker at increasing 20 m intervals from the SM2 Terrestrial Song Meter. A loud call series (maximum amplitude 92 dB) and soft call series (maximum amplitude 83 dB) were played at each distance up to 80 m. Call amplitude was measured using a separate iPhone microphone (Decibel Meter). Because the amplitude of murrelet calls is not known, we estimated a conservative minimum decibel level from field recordings. The loudest murrelet calls recorded by Song Meters in field recordings from this study reached 90 dB (amplitude measured by the Song Meter microphone), and were most likely produced by murrelets flying at least 10 m from the Song Meter, based on probable flight height relative to the sensor. Thus, the loud call used in our tests was likely to produce conservative estimates of the detection range of Song Meters. Two trials were conducted in both forested and open habitat: one with +12 dB gain adjustments to both microphones on the Song Meter (to amplify weaker calls) and one without gain adjustment.

TABLE 1
Summary of Song Meter locations, sensor types and habitat descriptions at Monashka Bay, Kodiak Island, 2011–2012

Year	Site name[a]	Sensor type[b]	Habitat	No. of 2 h recordings in subsample	Mean detections per morning ± SE	Mean false-positives, % ± SE
2011	Forest 1[c]	SM2-T	Old-growth forest	15	57.4 ± 8.3	66.1 ± 7.0
	Forest 2[c]	SM2-NF	Old-growth forest	28	55.2 ± 5.8	57.0 ± 6.3
	Flight path 1	SM2-T	Grass and shrub	4	4.7 ± 2.8	98.7 ± 0.7
	Flight path 2	SM2-T	Grass	4	3.0 ± 0.9	98.9 ± 0.4
	Flight path 3	SM2-T	Alder shrub	4	19.3 ± 5.4	95.8 ± 0.9
	Flight path 4	SM2-T	Grass and shrub	4	3.5 ± 1.8	99.1 ± 0.7
	Flight path 5	SM2-NF	Sparse conifers	11	17.5 ± 4.3	64.9 ± 10.0
2012	Forest 1	SM2-T	Old-growth forest	16	90.5 ± 8.7	27.2 ± 6.0
	Forest 3	SM2-T	Old-growth forest	21	62.4 ± 5.5	35.0 ± 5.7
	Flight path 5	SM2-NF	Sparse conifers	21	11.7 ± 2.8	65.0 ± 6.6

[a] See Cragg (2013) for site locations.
[b] SM2-T: Song Meter SM2 Terrestrial; SM2-NF: Song Meter SM2 Night Flight.
[c] These two sensors were deployed in the same tree for comparison (see text).

Analysis of recordings

We subsampled a random selection of field recordings throughout the 2011 and 2012 breeding seasons from each sensor to develop automated recognition models called "recognizers" in the program Song Scope 4.1.3A (Wildlife Acoustics 2011; Buxton & Jones 2012). Details of recognizer model development are given in the Appendix 1 (available on the website; see also Cragg 2013). Briefly, recognizer models were developed by an iterative process to identify murrelet calls in the recorded spectrograms and distinguish them from the calls of other birds and other sounds, such as rain or leaves rustling. The final recognizers identified *keer* and *keheer* calls (Dechesne 1998), which made up >90% of all murrelet sounds in our recordings, and we ran the recognizers for both calls simultaneously to analyze recordings.

The spectrograms of the recognizer selections of suspected murrelet calls were then visually reviewed to confirm correct identification of a murrelet call or identify false-positive detections (other species or background noise). Correctly identified murrelet calls were labelled by call type and tallied for each recording. These murrelet sounds were then grouped together into "call series detections" (hereafter "detections"). Based on audio-visual survey standards (Evans Mack *et al.* 2003), a call series consists of repeated calls, given by the same bird or group of birds and separated by < 5 s. We focused on call series detections rather than individual calls because this measure was most appropriate for comparisons with radar and audio-visual surveys (Cragg 2013).

We determined false-positive proportions as the number of non-murrelet sounds divided by the total number of automated detections. False-positive proportions were plotted throughout the season to detect seasonal trends in noise interference, showing the effects of competing vocalizations of other bird species. Mean false-positive rates for each sensor were calculated for all subsampled recordings per season to show differences between sensor locations and the potential influence of different acoustic environments on the effects of noise interference on call recognition.

To determine the proportion of false-negatives (i.e. call series missed by the recognizers), we selected a random subsample of six 2 h recordings from different Song Meter units and acoustic environments for visual review (total sample 12 h of recordings, 2117 recognizer detections, 330 call series). The false-negative proportion was the number of murrelet call series missed by the automated recognizers relative to the total number found from visual inspections of the subsampled recordings. The false-negative proportion was calculated for missed call series (rather than individual calls) because this was the relevant unit of comparison between acoustic sensors and audio-visual surveys. By contrast, false-positive proportions were calculated for individual calls to describe how efficiently the recognizer could distinguish murrelet sounds from other noise. However, we also calculated the proportion of individual calls missed by the recognizers within each call series to understand how this source of error contributed to a single call series being counted as multiple detections; when multiple calls were missed within a series, this created time gaps of > 5 s between detected calls, so that a single call series was erroneously recorded as multiple detections.

Data analysis

Data analysis was performed in R (v. 2.11.0). To test whether the number of detections increased as a linear function of all murrelet sounds detected, we plotted the number of call series as function of total confirmed murrelet calls using local polynomial regression fitting. We used pooled data from daily confirmed murrelet detections for all sensors at forested locations at Monashka Bay in 2011 and 2012 to plot the seasonal trend in vocal activity using non-parametric smoothing (generalized additive models), since each of the sensors covered slightly different periods of the breeding season. Only the Forest 2 sensor recorded continuously throughout the breeding season in 2011, and therefore this sensor alone was used to compare murrelet detections to false-positives across the breeding season. In 2012, sensors were deployed on 1 June but did not record daily until 14 July because of a programming bug. We used the Friedman test with Julian date as the blocking factor to compare seasonal trends in counts between years. This test was performed on sensors that provided the longest series of daily counts in two sampling locations: 1) in forest habitat Forest 2 for 2011 was compared with Forest 1 for 2012, as these sensors were deployed at the same tree (although different types of sensors were used: an SM2 Night Flight sensor in 2011 and an SM2 Terrestrial sensor in 2012); and 2) at a commuting flight path sensor Flight Path 5 was compared between years 2011 and 2012. Finally, to test whether the two sensors in nearby patches of potential forested nesting habitat had different daily counts, we used a paired *t*-test to compare sensors Forest 1 and Forest 3 (250 m apart) in 2012.

RESULTS

Summary of calling behavior

Recognizers detected 20632 murrelet sounds, yielding 5870 detections (Table 2). The most frequent call types identified by recognizers were *keheer* (73% of calls) and *keer* (25%). Other vocalizations (*ay* and *quack*) made up the remaining 2% of calls detected. There were 14 detections of non-vocal murrelet sounds (wing beats and jet sounds) incidentally detected while visually reviewing recognizer detections of calling bouts.

TABLE 2
Murrelet calls, non-vocal sounds ("jet" sounds and wing beats), and call series detected by Song Meters in 2011 and 2012

Year	Number of mornings sampled (2 h each)	Number of calls				Wing beats and jet sounds	Total number of call series detections
		Keheer	*Keer*	*Quack*	*Ay*		
2011	70	6285	2221	117	39	13	2725
2012	64	8721	2995	72	168	1	3145
Total	134	15006	5216	189	207	14	5870

Effect of detection distance

In tests using the SM2 Terrestrial Song Meters, visual scanning of spectrograms detected murrelet calls at greater distances (approximately 20 m farther) than call recognizers, which performed poorly at detecting faint calls (Fig. 1). The +12 dB gain adjustment did not consistently affect the number of calls detectable by visual scan or by recognizers. Acoustic environment had a strong effect: calls were detectable by both recognizers and visual scans at least 20 m farther from the Song Meter in open habitat than in forest (Fig. 1). In forest, loud (92 dB) calls could be detected by visual scan of the spectrograms up to 60 m, while recognizers could detect loud calls only up to 40 m from the Song Meter. In open habitat, 47% of loud calls could be visually detected on the spectrogram at 80 m, yet only one call (3%) was detected by the recognizer at this distance. Soft calls were more easily lost in background noise in both environments, especially beyond 20 m. The call types recognized at the greatest distance from the Song Meter were *keer*, *keheer* and *ay*, while *quack* calls were rarely detected beyond 20 m.

Audit of recognizer detections

Song Scope recognizers detected similar numbers of detections (call series) compared with the audit by visual scan in a review of six randomly selected recordings totaling 12 h (296 and 323 detections, respectively; Table 3: Rows C and D). Excluding two mornings with significant noise interference (25 June 2011 and 12 August 2012), the number of detections identified by recognition models (70.8 ± 6.7 detections; n = 4) was on average 5% less than the number of call series detected visually (74.8 ± 8.4 detections; n = 4), and there was no significant difference in total detections across all sampled recordings (Wilcoxon signed rank test n = 6, $V = 3$, $P = 0.1411$). However, this apparent match in identification is misleading, since automated recognition missed many short call series (consisting of 1–3 calls) and, conversely, counted many long call series as multiple detections when faint calls were missed. Although the numbers of detections were similar between the two methods, this was not because the recognizer identified the same detections as the visual scan. The true proportion of call series detected by the recognizers was only 68% of the call series identified by visual scans (n = 299 call series; Table 3, Row F). In other words, 32% of call series seen during visual scans were missed by the recognizer; visual inspection showed that those missed were either soft calls and short call series.

Longer call series could be double-counted by the recognizers (“multiple detection”) if they missed softer or more distant calls during a longer calling bout (such as when murrelets were circling while calling over a patch of forest, with call amplitudes decreasing as the birds moved away from the Song Meter, and increasing as they returned). The mean proportion of all calls detected by the recognizers within a call series was 30% (299 visually detected call series, 3 632 total calls, n = 4 mornings). Thus, in long call series

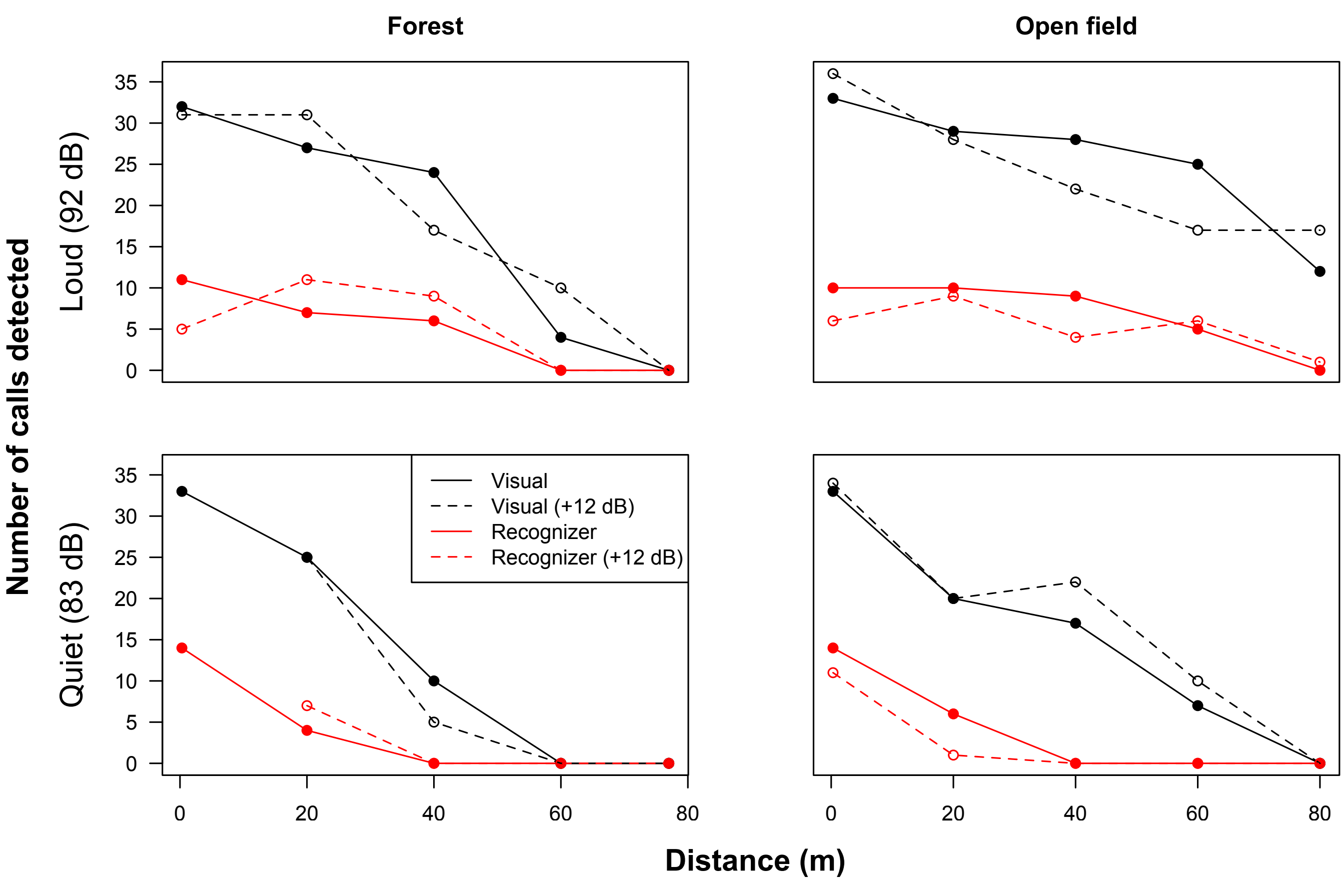

Fig. 1. Song Meter field test results showing the number of calls (loud, 92 dB; quiet, 83 dB) detected by visual scans of spectrograms and by automated recognizers in Song Scope at increasing distance from the Song Meter in two acoustic environments (forest and open habitat), with 12 dB gain amplification (dashed lines) or without (solid lines).

(many series exceeded 50 calls), the recognizer was likely to miss enough calls to result in gaps between recognized calls exceeding 5 s and, therefore, produce multiple detections from the same series.

When detections (call series) identified by recognizers were compared with the raw number of calls, the relationship was linear when the number of calls within a recording was <150, but non-linear when the number of calls exceeded 150 (Fig. 2). The curve of the smoothed trend line was not affected by the removal of outliers. The variance in numbers of detections relative to calls also increased as the number of calls increased.

Noise interference and false-positive detections

The proportion of false-positives (out of all automated detections) varied seasonally and between sites as a function of two factors: the relative abundance of murrelet calls (more calls meant a lower proportion of false-positive detections) and noise interference. Analysis focused on the relative influences of two types of noise interference: weather (strong wind, heavy rain) and vocalizations of other birds. Forested sites had higher rates of murrelet detections than unforested sites, while competing bird vocalizations were greater at sites with brushy habitat supporting high densities of songbirds (Table 4). Forest rated as potential nesting habitat had the lowest mean false-positive detection rate (48%), while lower-ranked habitat had higher false-positive detections (65%–98%). In general, grass and shrub sites had few murrelet detections and a high proportion of false-positive detections (97%–99%) because of high densities of songbirds (Table 4).

Noise interference created by vocalizations of other bird species constituted the majority of false-positive detections and occurred in nearly all recordings. Songbird interference, indicated by the false-positives, was most intense in June and early July when these birds were maintaining territories (Fig. 3; e.g. Table 3: 25 June 2011). During the height of the songbird dawn chorus in June and early July, competing vocalizations occasionally became so intense that murrelet calls could not be detected even by visually reviewing the spectrogram. At the same time, recognizers often falsely detected loud calls made by other bird species at the expense of fainter murrelet calls, resulting in a higher proportion of false-negatives (e.g. 50% on 25 June 2011; Table 3) compared with other times in the season when songbird calls were less frequent (32% averaged over 8 h of recordings between 10 July and 16 August when there was no noise interference; Table 3).

Heavy rain caused high proportions of false-positives and false-negatives (e.g. 99% and 86%, respectively, of all recognizer detections on 12 August 2012; Table 3), and generated broad-spectrum background "white noise" that masked most murrelet

TABLE 3
Sample of spectrogram audits comparing detections of murrelet calls and call series by visual spectrogram review vs. recognizer detections

Date	Samples with low noise interference					Samples with high noise interference	
	10 Jul 2011	12 Aug 2011	16 Jul 2012	16 Aug 2012	Mean ± SE	25 Jun 2011	12 Aug 2012
Sensor	Forest 1	Forest 2	Forest 3	Forest 1		Forest 1	Flight path 5
Noise interference	None	None	None	None		Bird vocalizations	Heavy rain/wind
(A) No. of automated detections	443	186	382	440	362.8 ± 60.6	575	91
(B) No. of visually confirmed murrelet calls detected by recognizer	234	179	158	340	227.8 ± 40.7	33	1
False-positives (%)[a]	47	4	59	23	33.3 ± 12.3	94	99
(C) No. of call series detections (recognizer)[b]	81	52	70	80	70.8 ± 6.7	13	1
(D) No. of call series detections (visual)	72	53	81	93	74.8 ± 8.4	24	7
(E) No. of call series detections missed by recognizer	20	17	28	29	23.5 ± 3	12	6
(F) False-negatives (%)[c]	28	32	35	31	31.5 ± 1.4	50	86

[a] The false-positive rate was calculated as 100*(A-B)/A (proportion of calls to describe the efficiency of recognizer in distinguishing murrelet sounds from other noise).

[b] The number of call series detected by the recognizer includes call series counted multiple times due to missed calls in the middle of the series. Therefore, the number of call series detections is sometimes larger than the number detected visually, while some call series were still missed.

[c] The false-negative rate indicates the proportion of call series identified by visual review that were missed by the recognizer, per recording; it is calculated as 100*E/D (proportion of detections).

calls; strong wind produced a similar effect. Inspection of the spectrograms showed that these severe weather conditions were relatively rare, occurring on only three survey days in both 2011 and 2012 seasons.

Seasonal and spatial trends

Analyzing seasonal and spatial variations in detectability was an important test of the acoustic method and allowed us to determine optimal periods for deployment. Vocal activity of murrelets detected by Song Meters in forested habitat at Monashka Bay increased from June through early July, with a peak in activity between 15 July and 5 August, followed by a decline in activity through the remainder of August (Fig. 4). The seasonal patterns of detections were not significantly different between years either in forested habitat (Friedman test; $S = 0.2857$, $P = 0.593$) or along a commuter flight path ($S = 0.0909$, $P = 0.763$) at Monashka Bay. When daily counts in 2012 were compared between two simultaneously operating forest sensors, separated by 250 m at Monashka Bay 2012, there were significantly more detections at Forest 1 than at Forest 3 (Table 1; $t = 4.06$, $P = 0.001$).

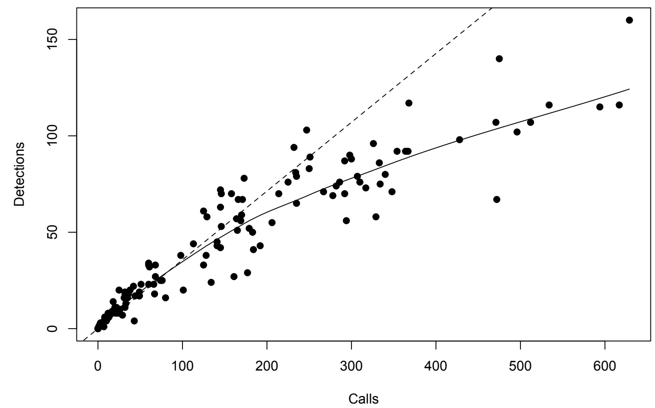

Fig. 2. Comparison of the number of detections (call series) with calls detected by recognizers in all recordings with smoothed trend line (local polynomial regression fitting), and linear regression line (dashed line). The number of detections relative to calls decreased as the number of calls detected in the recording increased, producing a non-linear relationship.

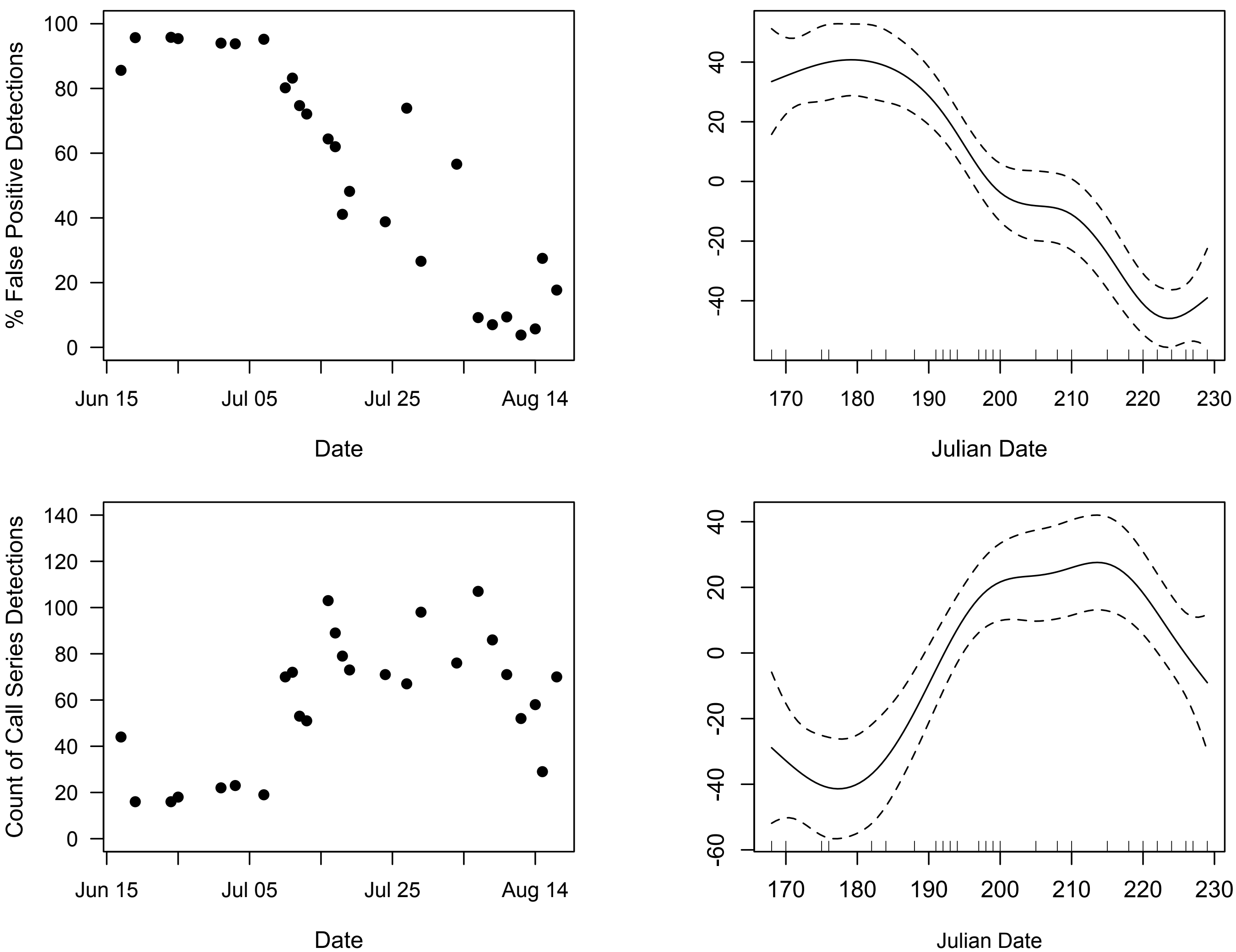

Fig. 3. Seasonal trends from acoustic sensor Forest 2 at Monashka Bay in 2011, showing false-positive detections resulting from competing vocalizations of other bird species (top), and murrelet call series detections (bottom). Generalized additive models are shown on the right-hand panels, with dashed lines indicating two standard errors or approximately 95% confidence limits of the prediction. The y-axis is the centered smooth value of the number of detections in logits; the tick marks just above the x-axis indicate the sampling distribution.

DISCUSSION

Detection range and factors affecting automated recorder sensitivity

In our tests, Song Meters failed to detect most murrelet calls beyond 40–60 m, depending on habitat and call amplitude. The amplitude of calls used in our field tests was quieter than most real murrelet calls, based on our subjective impression after multiple seasons of auditory field observations in a wide range of habitats. The functional detection range of murrelet calls by Song Meters in forested habitat was therefore likely to be at least 60 m. Other variables not tested that would also affect Song Meter detection range are wind speed and direction, flight altitude and direction of the birds, forest foliage density, and noise generated by rain, wind or human activities. Tests of Song Meters with boreal songbirds showed that detection probability declined below 50% beyond 50 m (Venier *et al*. 2012). A conservative estimated range of 60 m gives a sampling area of approximately 1.1 ha. In standard murrelet audio-visual surveys, detection rates decrease substantially beyond 200 m, (sampling area of about 12.6 ha; Cooper & Blaha 2002).

The SM2 Night Flight Song Meter (cost approximately US$950, including microphones and mounting devices) detected a number of murrelet calls similar to that of the cheaper general-purpose SM2 Terrestrial model (US$700; sound recognition software used for either model was US$500). Although the Night Flight model is designed to increase sensitivity of the microphone to calls from above, it appeared that this increased sensitivity may have increased interference from background noise, which slightly reduced overall counts of murrelet detections. Alternatively, since the Night Flight sensor was mounted higher in the tree among branches, the microphone may have been closer to potential perch sites for songbirds that generated noise interference during the dawn chorus. The SM2 Terrestrial model (or equivalents) should therefore suffice for murrelet monitoring, although we encourage further testing of these and similar sensors.

In general, conditions that lead to poor sampling with automated sensors (loud songbird singing, wind and heavy rain) were also those that affect audio-visual surveys by human observers. Noise interference from these sources affected the performance of recognizers applied to Song Meter recordings, in some situations leading to high proportions of false-positives and false-negatives. Noise interference could be reduced by avoiding complex understory habitats that provide perch sites for songbirds and leafy vegetation. These habitats are noisier in wind or rain than moss-dominated sites. Forests rated as potential nesting habitat for murrelets in the Kodiak Archipelago often lacked dense subcanopy vegetation (J.L.C., personal observation), and false-positives were lower in this habitat than in more open shrubby habitat. Automated analyses of murrelet calls could also be improved by omitting periods of heavy rain and high winds. Analysis of seasonal trends in false-positives, caused mainly by bird song (Fig. 3), could guide researchers on when to forgo sampling during the peak of territorial singing by songbirds.

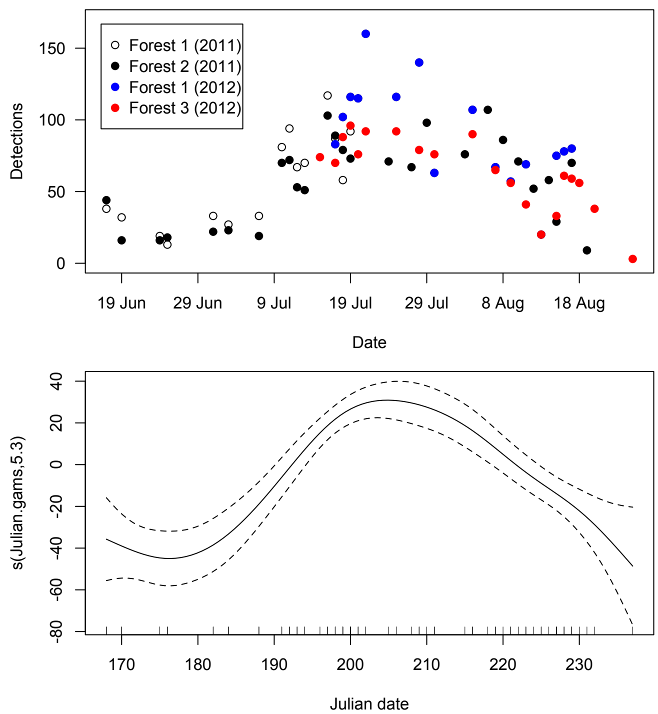

Fig. 4. Seasonal trends in Song Meter detections from all forest sensors at Monashka Bay in 2011 and 2012 (top), and non-parametric smoothed model (generalized additive model) of detections from sensor Forest 2 from both years of data (bottom). The dashed lines indicate two standard errors above and below the mean, or approximately 95% confidence interval. The y-axis is the centered smooth value of the number of detections in logits; the tick marks just above the x-axis indicate the sampling distribution.

TABLE 4
Proportions of false-positive detections (non-murrelet sounds selected by Song Meter recognizers) by habitat type or forest habitat quality (ranked according to nesting habitat potential)

Habitat	Vegetation subclass or habitat quality rank[a]	Number of 2 h recordings	Total recognizer detections	Total visually confirmed recognizer detections	Mean false-positive rate ± SE (%)	Mean detections per 2 h recording (call series) ± SE
Forested	High quality	82	40 277	18 656	48 ± 4	64 ± 4
	Low/marginal	3	1 830	201	89 ± 6	12 ± 6
Unforested	Sparse conifer	32	6 820	1 480	65 ± 5	14 ± 2
	Grass and shrub	8	9 089	215	97 ± 2	13 ± 4
	Grass	8	6 504	63	99 ±1	3 ± 1

[a] Habitat quality based on availability of large trees, potential nest platforms and canopy access (Burger 2004).

Application of automated call recognizers

Automated call-recognition software (recognizers in Song Scope) did not perfectly sample recordings (missing some calls and counting some call series as multiple detections), but their overall performance provided an approximation of the detection frequencies derived from visual inspections of the spectrograms. Automated recognizers greatly improved the efficiency of processing field recordings. It took approximately 1 h to review and group automated call detections (eliminating false-positives and grouping call series into detections) for every 2 h of field recording, compared with approximately 3 h to visually scan the same spectrogram to search for murrelet calls. Using automated recognizers alone, however, resulted in some missed call series that were detected only through visual inspection. Visual inspection of the spectrograms was more likely to detect soft or distant calls (reducing false-negatives) and was less likely to split a detection into two or more detections erroneously.

We grouped calls into call series detections in order to match existing metrics used in standard audio-visual and radar surveys (Evans Mack *et al*. 2003) and in studies comparing murrelet occurrence and behavior with habitat and temporal parameters (e.g. Rodway *et al*. 1993, Burger & Bahn 2004, Stauffer *et al.* 2004, Bigger *et al.* 2006). Recognizer-based detections had a non-linear relationship to the raw count of calls, especially when the number of calls per recording exceeded 150. This non-linear trend could be due to a saturation effect, in which calling bouts overlap more frequently during high murrelet vocal activity; alternatively, when vocal activity is greatest, each bird might be stimulated to give more prolonged call series. Most studies using audio-visual methods do not report the number of calls per detection, but Rodway *et al*. (1993) found mean values of 25.6 and 24.0 calls per detection at two heavily used sites in British Columbia. Rodway *et al.* (1993) showed that frequencies of detections and raw counts of calls showed similar seasonal trends.

Vocal behavior of Marbled Murrelets and implications for acoustic monitoring

During the breeding season, murrelets engage in conspicuous vocal and flight displays at dawn, in and above forests. The reasons for these conspicuous behaviors are not well understood but likely involve a combination of pair-bond maintenance and perhaps spacing (deterrence advertising) behaviors (Nelson 1997, Dechesne 1998). Standardized audio-visual surveys record these behaviors to assess stand occupancy, seasonal trends and spatial distributions of murrelets (Ralph *et al.* 1995, Evans Mack *et al*. 2003). Although the exact numbers of birds involved are never known, strong positive correlations have been shown between rates of audio-visual detections, especially behavior indicating stand occupancy, and various metrics of murrelet habitat quality (e.g. Meyer *et al*. 2004, Burger & Bahn 2004, Stauffer *et al*. 2004). In these contexts, automated sound recording might fill a role similar to audio-visual surveys, but the limitations of both methods need to be examined relative to other survey techniques.

Acoustic surveys cover limited spatial areas (approximately 1.1 ha with automated sensors and 12.6 ha with human observers) compared with areas scanned by radar (typically 1–1.5 km radius, 300–700 ha covered; Burger 2001). Other limitations of audio monitoring include high variability in vocalization rates; inclusion of calls from passing murrelets that are not associated with the habitat being sampled; and inclusion of non-breeding murrelets in dawn flights (Burger 1997, Cooper & Blaha 2002, Jodice & Collopy 2000). Overall, these sources of error in estimating the relative abundance of murrelets at a given forest site result in high variation in audio-visual counts relative to radar counts, and low power of audio-visual data to detect population trends (Jodice *et al.* 2001). Radar surveys show that audio-visual surveys miss many passing murrelets and fail to sample peak pre-sunrise periods of flight, during which murrelets seldom vocalize (Burger 1997, Cooper & Blaha 2002, Bigger *et al.* 2006). Radar surveys can differentiate commuting flight paths not associated with nearby habitat and circling behavior over potential nesting habitat (Cragg 2013). By contrast, in this study, Song Meters recorded far fewer detections on known commuting flight paths than in potential forest nesting habitat, indicating that their application would be more successful in the latter location.

Because murrelet calls are relatively simple and often monosyllabic (e.g. *keer* and *ay* calls), recognizers often misidentified components of other, more complex bird calls as murrelet calls, generating large numbers of false-positives. Many species contributed false-positives, and their relative contributions depended on song amplitude (obviously affected by proximity to the sensor) and on the length and complexity of calls. For example, two species with prolonged and complex calls generated the majority of false-positives in our study: out of a sample of 967 false positive detections of other birds, Fox Sparrows *Passerella iliaca* accounted for 36% of false-positives while Pacific Wrens *Troglodytes pacificus* accounted for 26%.

Although limited as a method for estimating population size, automated acoustic sensors could improve monitoring of murrelets in forest stands in several ways: by supplementing audio-visual surveys to increase spatial and temporal replication; by eliminating variation due to observer bias; and by reducing variation due to site characteristics that affect viewing conditions for human observers. The main strength of auditory detections is in evaluating murrelet activity that is relevant to a given forest stand, as an indication of nesting habitat use or suitability. In this context, the limited spatial coverage of automated sensors can be beneficial, by reducing the probability of including vocalizations associated with other nearby stands. Our study showed consistent differences in vocal activity of murrelets between nearby forest patches, a pattern that could guide more focused research into nest sites, stand occupancy and habitat suitability. Although the method fails to provide estimates of actual murrelet numbers, long-term acoustic sampling at key locations could reveal changes in relative abundance or habitat use by murrelets in response to environmental changes such as climate change, habitat fragmentation, or changes in predator abundance.

Strengths and limitations of automated acoustic monitoring

Automated acoustic sensors have several practical advantages over other types of surveys: the relatively low cost of units, automation (greatly reducing costs of field personnel), portability, and efficient use of power (battery life generally lasts >2 months for daily 2 h recordings). Acoustic monitoring also provides a permanent recording of each survey that can be reviewed multiple times by the same observer or multiple observers to check for errors. Recordings also allow the use of spectrograms to distinguish between murrelet call types that are virtually indistinguishable by ear (such as *keer* and *keheer*), providing a more detailed record of vocal activity. Archived recordings provide opportunities for detailed future analyses of vocal behavior (e.g. a subsample of 268 h of our recordings yielded 20632 calls) and opportunities to investigate sounds made by other species of management interest.

The main strength of acoustic monitoring is in collecting large volumes of data at low cost, but processing such data remains time-consuming, even using automated recognition software (in this study, processing time was approximately half the duration of recordings). Processing could be streamlined for murrelet recordings by simply reviewing results of automated scan results to eliminate false-positives, without grouping call series. This would produce an index of vocal activity, rather than a number of detections that could be directly compared with audio-visual survey detections. Another option is to subsample smaller time intervals each morning, once representative times have been identified (e.g. Wimmer *et al.* 2013). Processing efficiency could also be improved by avoiding sampling areas with high frequencies of songbird vocalizations, which greatly increased false-positives, false-negatives and processing time.

In conclusion, our study reveals some of the advantages and limitations of using automated sound recorders and recognition software for studying Marbled Murrelets. At the least, the recordings provide evidence of murrelet presence. With careful application of methods, automated acoustic recording and recognition could provide quantitative measures of vocal activity and relative measures of habitat use. The key methodological considerations we recommend are the placement of sensors to minimize false-positives, attention to seasonal trends, and prudent application of automated recognizers coupled with visual inspections of detections to remove false-positives. The advances being made in acoustic sensor systems and digital sound recognition (McKown *et al.* 2012, Wildlife Acoustics 2013) should help overcome some of the limitations we have identified.

ACKNOWLEDGEMENTS

The study was supported by a grant from the North Pacific Research Board. Additional financial, logistical and equipment support was provided by the US Geological Survey (Alaska Science Center), the Kodiak National Wildlife Refuge, US Fish and Wildlife Service and the University of Victoria. J.L.C. was supported by grants from the Natural Sciences and Engineering Research Council of Canada and University of Victoria Graduate Fellowships and awards. For major contributions to this study we thank A. Borker, R. Buxton, R. Corcoran, S. Hrushowy, J. Lewis, B. Pyle and S. Snyder. Comments from K. Kuletz and anonymous reviewers greatly improved the paper. Any use of trade names is for descriptive purposes only and does not imply endorsement by the US federal government. This is North Pacific Research Board publication number 546.

REFERENCES

BIGGER, D., PEERY, Z.M., CHINNICI, S. & COURTNEY, S.P. 2006. Efficacy of audiovisual and radar surveys for studying Marbled Murrelets in inland habitats. *Journal of Wildlife Management* 70: 505–516.

BORKER, A.L., MCKOWN, M.L., ACKERMAN, J.T., EAGLES-SMITH, C.A., TERSHY, B.R. & CROLL, D.A. 2014. Vocal activity as a low cost and scalable index of seabird colony size. *Conservation Biology* 28: 1100–1108.

BURGER, A.E. 1997. Behavior and numbers of Marbled Murrelets measured with radar. *Journal of Field Ornithology* 68: 208–223.

BURGER, A.E. 2001. Using radar to estimate populations and assess habitat associations of Marbled Murrelets. *Journal of Wildlife Management* 65: 696–715.

BURGER, A.E. 2004. Standard methods for identifying and ranking nesting habitat of Marbled Murrelets in British Columbia using air photo interpretation and low-level aerial surveys. Victoria, BC: Ministry of Water, Land and Air Protection, Biodiversity Branch.

BURGER, A.E. & BAHN, V. 2004. Inland habitat associations of Marbled Murrelets on southwest Vancouver Island, British Columbia. *Journal of Field Ornithology* 75: 53–66.

BUXTON, R.T. & JONES, I.L. 2012. Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. *Journal of Field Ornithology* 83: 47–60.

BUXTON, R.T., MAJOR H.L., JONES I.L. & WILLIAMS, J.C. 2013. Examining patterns in nocturnal seabird activity and recovery across the Western Aleutian Islands, Alaska, using automated acoustic recording. *Auk* 130: 331–341.

CANADIAN MARBLED MURRELET RECOVERY TEAM (CMMRT). 2003. Marbled murrelet conservation assessment 2003, Part B: Marbled Murrelet recovery team advisory document on conservation and management. Canadian Marbled Murrelet Recovery Team Working Document No. 1. Delta, BC: Canadian Wildlife Service, Pacific and Yukon Region. [Available online from: http://www.sfu.ca/biology/wildberg/bertram/mamurt/PartB.pdf; accessed 6 July 2015.]

COOPER, B.A. & BLAHA, R.T. 2002. Comparisons of radar and audio-visual counts of Marbled Murrelets during inland forest surveys. *Wildlife Society Bulletin* 30: 1182–1194.

COOPER, B.A., RAPHAEL, M.G. & EVANS MACK, D. 2001. Radar-based monitoring of Marbled Murrelets. *Condor* 103: 219–229.

CRAGG, J.L. 2013. *Brachyramphus* murrelets at high latitude: behavioural patterns and new methods for population monitoring. M.Sc. thesis. Victoria, BC: University of Victoria.

DECHESNE, S.B.C. 1998. Vocalizations of the Marbled Murrelet (*Brachyramphus marmoratus*): vocal repertoire and individuality. M.Sc. thesis. Victoria, BC: University of Victoria.

EVANS MACK, D., RITCHIE, W.P., NELSON, S.K., KUO-HARRISON, E., HARRISON, P. & HAMER, T.E. 2003. Methods for surveying Marbled Murrelet in forests: a revised protocol for land management and research. Technical Publication No. 2. Portland, OR: Pacific Seabird Group.

JODICE, P.G.R. & COLLOPY, M.W. 2000. Activity patterns of Marbled Murrelets in Douglas-fir old-growth forests in the Oregon Coast Range. *Condor* 102: 275–285.

JODICE, P.G.R., GARMAN, S.L. & COLLOPY, M.W. 2001. Using resampling to assess reliability of audio-visual survey strategies for Marbled Murrelets at inland forest sites. *Waterbirds* 24: 331–344.

McKOWN, M.W., LUKAC, M., BORKER, A., TERSHY, B. & CROLL, D. 2012. A wireless acoustic sensor network for monitoring wildlife in remote locations. *Journal of the Acoustical Society of America* 132: 2036.

McSHANE, C., HAMER, T., CARTER, H., SWARTZMAN, G., FRIESEN, V., AINLEY, D., TRESSLER, R., NELSON, K., BURGER, A., SPEAR, L., MONAGEN, T., MARTIN, R., HENKEL, L., PRINDLE, K., STRONG, C. & KEANY, J. 2004. Evaluation report for the 5-year status review of the Marbled Murrelet in Washington, Oregon, and California. Portland, OR: US Fish and Wildlife Service, Region 1.

MEYER, C.B., MILLER, S.L. & RALPH, C.J. 2004. Stand-scale habitat associations across a large geographic region of an old-growth specialist, the Marbled Murrelet. *Wilson Bulletin* 116: 197–210.

MILLER, S.L., RAPHAEL, M.G., FALXA, G.A., STRONG, C., BALDWIN, J., BLOXTON, T., GALLEHER, B.M., LANCE, M., LYNCH, D., PEARSON, S.F., RALPH, C.J. & YOUNG, R.D. 2012. Recent population decline of the Marbled Murrelet in the Pacific Northwest. *Condor* 114: 771–781.

NELSON, K. 1997. Marbled Murrelet (*Brachyramphus marmoratus*). In: Poole A. & Gill G. (Eds.) Birds of North America, No. 276. Washington, DC: Academy of Natural Sciences and American Ornithologists' Union.

OPPEL, S., HERVIAS, S., OLIVEIRA, N., PIPA, T., SILVA, C., GERALDES, P., GOH, M., IMMLER, E. & McKOWN, M. 2014. Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. *Nature Conservation* 7: 1–13.

PIATT, J.F., KULETZ, K.J., BURGER, A.E., HATCH, S.A., FRIESEN, V.L., BIRT, T.P., ARIMITSU, M.L., DREW, G.S., HARDING, A.M.A. & BIXLER, K.S. 2007. Status review of the Marbled Murrelet (*Brachyramphus marmoratus*) in Alaska and British Columbia. US Geological Survey Open-File Report 2006-1387. Reston, VA: US Geological Survey.

RALPH, C.J., HUNT, G.L. JR., RAPHAEL, M.G. & PIATT, J.F. (Eds.). 1995. Ecology and conservation of the Marbled Murrelet in North America. General Technical Report PSW-GTR-152. Albany, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture.

RAPHAEL, M.G. 2006. Conservation of the Marbled Murrelet under the Northwest Forest Plan. *Conservation Biology* 20: 297–305.

RODWAY, M.S., REGEHR, H.M. & SAVARD, J.-P.L. 1993. Activity patterns of Marbled Murrelets in old-growth forest in the Queen Charlotte Islands, British Columbia. *Condor* 95: 831–848.

STAUFFER, H.B., RALPH, C.J. & MILLER, S.L. 2004. Ranking habitat for Marbled Murrelets: new conservation approach for species with uncertain detection. *Ecological Applications* 14: 1374–1383.

VENIER, L.A., HOLMES, S.B., HOLBORN, G.W., MCILWRICK, K.A. & BROWN, G. 2012. Evaluation of an automated recording device for monitoring forest birds. *Wildlife Society Bulletin* 36: 30–39.

WILDLIFE ACOUSTICS. 2011. Song Scope Bioacoustics software version 4.0 documentation. Concord, MA: Wildlife Acoustics Inc.

WILDLIFE ACOUSTICS. 2013. Overview of the Song Stream Remote Access Module. Concord, MA: Wildlife Acoustics Inc.

WIMMER, J., TOWSEY, M., ROE, P. & WILLIAMSON, I. 2013. Sampling environmental acoustic recordings to determine bird species richness. *Ecological Applications* 23: 1419–1428.