Conservation Biology

Contributed Paper

Vocal Activity as a Low Cost and Scalable Index of Seabird Colony Size

ABRAHAM L. BORKER,*‡ MATTHEW W. MCKOWN,* JOSHUA T. ACKERMAN,†
COLLIN A. EAGLES-SMITH,† BERNIE R. TERSHY,* AND DONALD A. CROLL*

*Department of Ecology and Evolutionary Biology, Center for Ocean Health, University of California Santa Cruz, 100 Shaffer Road, Santa Cruz, CA 95060, U.S.A.

†U.S. Geological Survey, Western Ecological Research Center, Davis Field Station, One Shields Avenue, University of California, Davis, CA 95616, U.S.A.

Abstract: *Although wildlife conservation actions have increased globally in number and complexity, the lack of scalable, cost-effective monitoring methods limits adaptive management and the evaluation of conservation efficacy. Automated sensors and computer-aided analyses provide a scalable and increasingly cost-effective tool for conservation monitoring. A key assumption of automated acoustic monitoring of birds is that measures of acoustic activity at colony sites are correlated with the relative abundance of nesting birds. We tested this assumption for nesting Forster's terns (*Sterna forsteri*) in San Francisco Bay for 2 breeding seasons. Sensors recorded ambient sound at 7 colonies that had 15–111 nests in 2009 and 2010. Colonies were spaced at least 250 m apart and ranged from 36 to 2,571 m^2. We used spectrogram cross-correlation to automate the detection of tern calls from recordings. We calculated mean seasonal call rate and compared it with mean active nest count at each colony. Acoustic activity explained 71% of the variation in nest abundance between breeding sites and 88% of the change in colony size between years. These results validate a primary assumption of acoustic indices; that is, for terns, acoustic activity is correlated to relative abundance, a fundamental step toward designing rigorous and scalable acoustic monitoring programs to measure the effectiveness of conservation actions for colonial birds and other acoustically active wildlife.*

Keywords: adaptive management, bioacoustics, monitoring, population indices, seabirds

La Actividad Vocal como un Índice Escalable y de Bajo Costo del Tamaño de Colonia de las Aves Marinas

Resumen: *Aunque las acciones para la conservación de la fauna silvestre han incrementado globalmente en número y complejidad, la carencia de métodos de monitoreo escalables y rentables limitan el manejo adaptativo y la evaluación de la eficacia de la conservación. Los sensores automatizados y los análisis auxiliados por computadores proporcionan una herramienta escalable y rentable para el monitoreo de la conservación. Una suposición clave del monitoreo acústico automatizado de aves es que las medidas de la actividad acústica en los sitios de colonia están correlacionados con la abundancia relativa de aves anidando. Probamos esta suposición en colonias de golondrinas marinas (*Sterna forsteri*) en nidación en la Bahía de San Francisco durante dos temporadas de reproducción. Los sensores grabaron el sonido ambiente en siete colonias que tenían entre 15 y 111 nidos en 2009 y 2010. Las colonias estaban espaciadas al menos 250 m y abarcaban desde 36 hasta 2, 571 m^2. Usamos correlación cruzada de espectrograma para automatizar la detección del llamado de las aves a partir de las grabaciones. Calculamos la tasa promedio de llamados estacionales y la comparamos con la media del conteo de nidos activos en cada colonia. La actividad acústica explicó el 71% de la variación en la abundancia de nidos entre los sitios de reproducción y el 88% del cambio en el tamaño de la colonia entre los años. Estos resultados validan una suposición primaria de los índices acústicos; esto es que, para las golondrinas marinas, la actividad acústica está correlacionada con la abundancia relativa, un paso fundamental hacia el diseño riguroso y escalable de programas de monitoreo acústico para medir la efectividad de las acciones de conservación para aves coloniales y otros animales acústicamente activos.*

Palabras Clave: Aves marinas, bioacústica, índices de población, manejo adaptativo, monitoreo

---

‡*email aborker@ucsc.edu*

Paper submitted July 28, 2013; revised manuscript accepted November 14, 2013.

DOI: 10.1111/cobi.12264

Introduction

Adaptive management is widely advocated. In practice, however, it has not been widely adopted to a large extent because of the difficulty in designing and implementing cost-effective monitoring approaches (Keith et al. 2011); expensive monitoring can decrease resources available for conservation action (McDonald-Madden et al. 2010). As a result, the effectiveness of management or conservation often is not assessed or monitoring is underfunded, yielding data with limited statistical power and value (Legg & Nagy 2006). Either scenario can lead to a significant waste of limited conservation funding and no rigorous assessment of conservation actions.

Cost-effective, reliable methods for measuring change in biological communities are needed to realize the potential of adaptive management and drive iterative improvements in management. Automated sensors offer 1 approach, and their use has increased recently (e.g., Karanth & Nichols 1998; Acevedo & Villanueva-Rivera 2006; Sherley et al. 2010). Advantages of automated sensors include increasing the spatial and temporal scale of wildlife surveys, lowering the cost of field surveys, decreasing interobserver and temporal biases in data collection (Brandes 2008; Blumstein et al. 2011), and reducing impacts of human observers on sensitive wildlife (Carey 2009).

Passive acoustic sensors and automated bio-acoustic analyses are powerful tools for monitoring sound-producing wildlife. Acoustic monitoring has been used to search for rare species such as North Pacific right whales (*Eubalaena japonica*) and Ivory Billed Woodpeckers (*Campephilus principalis*) (Fitzpatrick et al. 2005; Wade et al. 2006), monitor activity patterns of sympatric petrels in remote areas (Mckown 2008), and estimate species richness, abundance, or density of terrestrial birds (Dawson & Efford 2009; Celis-Murillo et al. 2009), marine mammals (Barlow & Taylor 2005), and terrestrial mammals (Thompson et al. 2010).

Colonial or aggregated populations present unique challenges and advantages for monitoring programs. Aggregated populations reduce the spatial coverage needed to monitor abundance; and the sensitivity to disturbance and remoteness of many colonial species is a challenge suited to automated acoustic sensors. To effectively monitor the relative abundance of colonial species, we must know the relationship between the number of individuals and acoustic activity.

Colonial seabirds are particularly suited for exploring the utility of acoustic monitoring. Most seabirds produce sounds at breeding colonies that are often logistically difficult, expensive, and even dangerous to visit. Colonies are sensitive to human disturbance; visits can cause temporary nest abandonment, high rates of depredation, reduced chick provisioning rates, or permanent abandonment (Carey 2009). This combination of remoteness and sensitivity makes it especially difficult to assess the effectiveness of conservation actions at seabird colonies, despite great need (28% of seabirds are threatened [Croxall et al. 2012]).

A number of studies have quantified seabird acoustic activity to examine colony attendance patterns (Mckown 2008), examine large-scale responses to island restoration (Buxton & Jones 2012), or detect elusive species at remote colonies (Wood et al. 2002). A previously untested assumption of these and other studies of colonial species is that changes in vocal activity at colonies are correlated with changes in the relative abundance of breeding individuals.

Forster's terns (*Sterna forsteri*) make an ideal system to test the relationship between acoustic activity and relative abundance for a colonial bird. They exhibit variability in colony size across sites and between years, and their conspicuous surface nesting habits make it relatively easy to directly measure nesting abundance. We tested the hypothesis that interannual and intercolony differences in Forster's tern nesting abundance are correlated with their acoustic calling activity. We measured calling rates while conducting weekly nest counts during 2 breeding seasons at 5 and then 7 tern colonies in San Francisco Bay (CA, U.S.A.).

Methods

Study Species, Site, and Design

Forster's terns nest in Don Edwards National Wildlife Refuge, San Francisco Bay, on small (<1 ha), human-made islands in shallow salt ponds delineated by levies (Strong et al. 2004) that are predominantly covered by bare ground and *Salicornia virginica*. Terns arrive at breeding colonies in late May and remain through August. Some American Avocets (*Recurvirostra americana*) and Black-necked Stilts (*Himanotopus mexicanus*; mean 36 nests/season summed across all sites) also breed on these small islands.

We selected 7 islands (5 in 2009) that represented a range of nest abundances (Ackerman et al. 2009; Bluso-Demers et al. 2010). We considered each island a distinct colony and considered colonies to be acoustically independent when at least 250 m from other islands occupied by breeding terns. Colony sizes and distance to neighboring breeding birds are summarized and mapped in Supporting Information.

Acoustic Sampling

Acoustic activity rates in seabird colonies can be highly variable (Buxton & Jones 2012) and are influenced by weather (Piatt et al. 1990), co-occurring species, non-breeder activity, synchrony of nesting phenology, and factors independent of breeding bird abundance. Thus, at each colony, we collected ambient acoustic recordings during the entire 2009 and 2010 nesting seasons to minimize the effects of this variance. We used automated recording units (ARU) (Cornell Lab of Ornithology, Ithaca, NY, U.S.A.) or SongMeter SM2 (Wildlife Acoustics, Concord, MA, U.S.A.) acoustic sensors. A single sensor was deployed within 35 m of the field-assessed center of each colony at the same location each year. In 2009, ARUs were attached to a 1 m high t-post with bird spikes to deter perching. In 2010, 3 of the ARUs were replaced with SongMeters (SM2s), and microphones were placed closer to the ground (0.2 m).

We tested for differences in both sensor types to detect tern calls with a simultaneous 120-min recording period at the same survey site. The ARU's had a 1.3% higher rate of call detections per unit time (ANCOVA, $F_{(\mathrm{df}=2088)} = 103700, p \leq 0.01$). This slight advantage of the ARU could be due to higher microphone placement. We attempted to correct for sensor hardware changes in the second year by increasing the mean rate of calls per minute by 1.3%; there were no significant effects on results.

Both sensors recorded with omni-directional microphones (ARU sensitivity: mean [manufacturer's tolerance] = -35 dB [4], SM2 sensitivity: mean = -36 dB [4], frequency response: 20 Hz to 20 kHz, signal to noise ratio (S:N): >62 dB). The ARUs recorded at a sampling rate of 24,000 Hz, and SM2s recorded at 22,050 Hz. All sensors recorded ambient acoustic activity during a 1-min period every 10 mins throughout the day (i.e., 144, 60-s sound files/d). We restricted analysis to recordings made between 0000 and 1200 Pacific Standard Time because high wind levels in the afternoon and evening masked acoustic activity by terns. Samples were binned by day when computing means.

Nest Monitoring

During 2009 and 2010, we visited colonies weekly from nest initiation (May) until the last chicks fledged (August). At each visit, we marked newly initiated nests with a uniquely numbered anodized aluminum tag and recorded their location with a handheld GPS (approximately 8 m accuracy). We followed the fate of all known nests throughout the season to determine whether the nest was depredated, abandoned, or active. A nest was considered active from the date the first egg was laid to when the chicks hatched or the nest failed. Initiation date was estimated based on the assumption that terns lay 1 egg/d and begin incubation on the day the last egg is laid. A nest was considered inactive after all eggs hatched or the nest was abandoned or depredated and did not include the chick rearing period. The total number of nests found during the season did not reflect large differences across colonies in nest survival; therefore, we calculated the number of active nests each day and report the mean abundance of active nests per day in addition to the total nest abundance. When reporting mean active nest abundance, we excluded daily active nest counts for days when the sensor was not recording due to occasional hardware failures.

Automated Acoustic Analysis

We used computer software (eXtensible BioAcoustic Tool – XBAT; Figueroa 2007) to automatically detect tern calls in recordings with spectrogram cross-correlation (Mellinger & Clark 2000; Goyette et al. 2011). We used the number of calls detected to quantify acoustic activity rates at each colony in an efficient and replicable manner.

Forster's terns have at least 9 adult vocalizations and 4 chick vocalizations (Hall 1998). We focused on the advertisement *kerr* call, which is used for communication between mated individuals and their chicks (Hall 1998). We used a clip of a call with high signal to noise ratio as an exemplar for the search template. The template focused on the stereotypic declining tonal frequency sweep of the call (peak frequency between 3.0 and 3.5 kHz), which helped differentiate the call from other tern vocalizations and other sounds (Hall 1998). We calculated the seasonal mean rate of calls per minute for all breeding sites and seasons.

To measure the effectiveness of automated detection, we compared manual (human) detections of kerr calls with XBAT detections on a subset of recordings. We had a human observer identify the first 10 kerr calls in 30 randomly selected minutes of recordings from each site. We compared these known call detections with the results obtained from automated analysis to measure false alarms (type I errors, detector accuracy) and missed detections (type II errors, detector sensitivity) across colony sites and years. Variation in detector sensitivity and accuracy from each site and year was compared with a two-way analysis of variance (ANOVA).

Data Analyses

To understand the relationship between acoustic activity and colony abundance, we used a linear mixed model to account for site, year, and sensor type as random factors when testing the variance in call rate explained by mean active nest abundance and total nest abundance. We used restricted maximum likelihood to estimate model parameters and tested fixed effects with a type III ANOVA. To assess model fit we calculated the marginal R^2, the proportion of variance explained by fixed effects (Nakagawa & Schielzeth 2013). A sample size of 12 colonies is near the lower limit of data sets that can be usefully interpreted with mixed model approaches and may contribute to relatively weak parameter estimates (Bolker et al. 2009). Despite calls per minute being count data, the mean rate of calls per minute across colonies was not non-normal

(Shapiro–Wilk test, $p = 0.75$, $W = 0.95$); thus, we used a normal distribution for fitting the model.

To test the precision of acoustic activity as an index of abundance, we used leave-one-out cross validation (LOOCV; Efron 1983). LOOCV reported the error in predicting nest abundance for each of the 12 measurements of call rate if each had been left out of the initial model building. We report the mean absolute value of the prediction error.

We used linear regression to compare the change in call rate to the change in nest abundance (mean active nests and total nests) between the 2009 and 2010 breeding seasons at 5 colonies. Mixed model approaches were not necessary as each of the 5 sites was independent. Once again, we used LOOCV to measure prediction error in change in nest abundance for observations of change in call rate.

We considered that the spatial dispersion of nests could influence rates of acoustic activity in addition to changes in abundance. We calculated the distance of each nest from the sensor and determined whether the median distance to nests in the colony helped explain differences in acoustic activity among colonies by considering it as an additional fixed effect.

We compared the costs of acoustic monitoring with the costs of traditional colony censuses. We estimated the cost of acoustic monitoring, rather than directly accounting for costs because we were still refining methodologies. A traditional survey involves a minimum of 4 visits to each colony during the breeding season by a staff of up to 4 (technicians are required to contain chicks when visiting an active colony to prevent chicks swimming off and being depredated). Our traditional surveys were more extensive than required; thus, we conservatively estimated the cost of typical shorter and less frequent colony visits required to count and individually mark nests.

Results

Acoustic sensors sampled 4984 h in 2009 ($n = 5$ colonies) and 7836 h in 2010 ($n = 7$ colonies). In total, there were 1302 sensor days, or an average of 86 d of acoustic sampling per colony. We found and monitored 725 tern nests during 2009 and 2010, and colony size ranged from 15 to 111 total nests/colony site (Supporting Information). The maximum number of simultaneously active nests in a colony was 77. The seasonal mean numbers of active nests ranged from 1.3 to 29.6 active nests/d and were correlated with total nest abundance ($r^2 = 0.80$, $p < 0.01$, $n = 12$).

Performance of Call Detection

Overall, spectrogram cross-correlation identified 1,370,071 sounds as kerr calls. Mean calling activity ranged from 0 to 99.2 calls/min and averaged 10.8 calls/min.

The automated detector accurately classified 77.3% (95% CI 4.7) of calls across sites and years (measured as the percentage of sounds detected that were correctly classified as tern kerr calls by the software). We found no significant influence of colony site or year on accuracy, but there was weak evidence of an interaction of colony site and year (two-way ANOVA, site: $F[4, 9] = 1.26$, $p = 0.29$, $\beta_{(\delta = 0.1)} = 0.41$; year: $F[1, 9] = 1.83$, $p = 0.18$, $\beta_{(\delta = 0.1)} = 0.9$; site∗year: $F[4, 9] = 2.28$, $p = 0.06$, $\beta_{(\delta = 0.1)} = 0.36$).

Automatic detector sensitivity (percentage of sounds the software detected compared with all tern kerr calls available for detection on the recording) was 53.6% (95% CI 7.3). We found a significant influence of colony site and its interaction with year on detector sensitivity (two-way ANOVA, site: $F[4, 9] = 7.19$, $p < 0.01$; year: $F[1, 9] = 0.08$, $p = 0.77$, $\beta_{(\delta = 0.1)} = 0.87$; site∗year: $F[4, 9] = 3.22$, $p = 0.04$). Differences in sites explained 10.2% of the variation in detector sensitivity.

Acoustic activity and Nest Abundance Among Colonies

Within each breeding season, mean acoustic activity (calls per minute) varied among sites (2009, $F[6, 91] = 5.65$, $p < 0.01$; 2010, $F[4, 60] = 5.76$, $p < 0.01$). Mean active nest abundance explained 71% of the variation in the mean acoustic activity (calls per minute) among sites ($F [1, 2] = 40.7$, $p = 0.02$; Fig. 1a). The model predicted nest abundance with a mean magnitude of error of 5.1 mean active nests (LOOCV).

Similarly, the season's total nest abundance explained most of the variation in mean call rate when used as an alternative to the mean nest abundance (Fig. 1b). Total nest abundance explained 70% of the variation in mean call rate ($F[1, 2] = 26.70$, $p < 0.01$). The model predicted abundance with a mean magnitude of error of 22.1 total nests (LOOCV).

Acoustic Activity and Nest Abundance between Years

At the 5 colonies monitored over 2 separate breeding seasons, the change in mean active nest abundance between years explained 88% of the variation in mean calls per minute between years (Fig. 2; df = 3, adj. $R^2 = 0.88$, $p = 0.01$). When we used LOOCV to estimate prediction error for future observations of call rate, the mean magnitude of error was a change of 7.1 mean active nests.

When total nest abundance was included in the model instead of mean nest abundance, the change in total nest abundance between years explained 85% of the variation in change in mean call rate (df = 3, adj. $R^2 = 0.85$, $p = 0.02$). The mean magnitude of error was a change of 19.8 total nests.

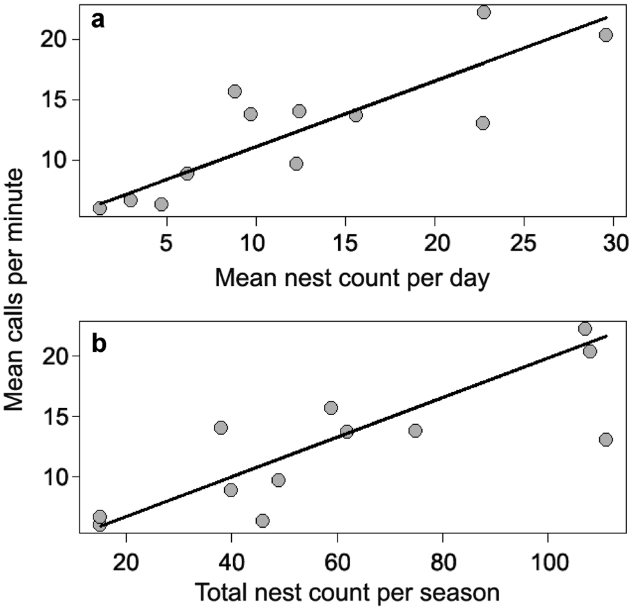

Figure 1. Seasonal mean acoustic call activity (calls per minute) relative to (a) mean nest abundance and (b) total nest abundance during the 2009 and 2010 breeding seasons of Forster's terns at colonies in San Francisco Bay, California. Black line is best fit from a linear mixed model incorporating nest abundance with site, year, and sensor type as random factors.

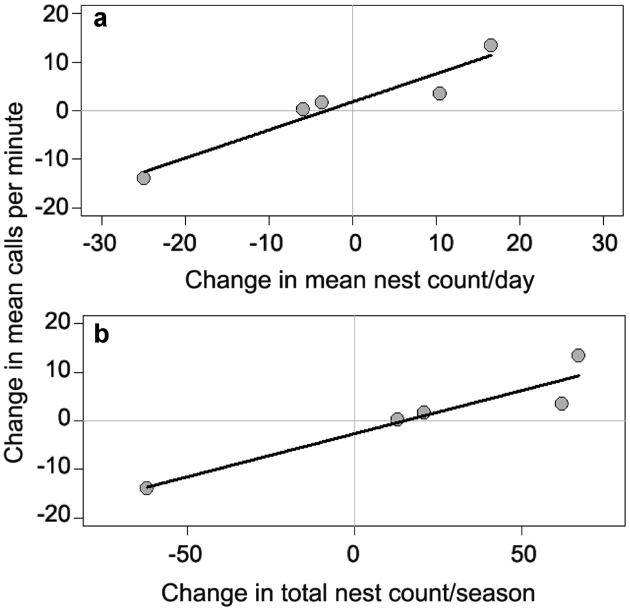

Figure 2. Change in seasonal mean of acoustic call activity (calls per minute) relative to (a) change in mean nest abundance and (b) change in total nest abundance from 2009 to 2010 at 5 Forster's tern colonies in South San Francisco Bay, California. Black line is linear best-fit line.

Acoustic Activity and the Spatial Dispersion of Nests

Fine scale spatial dispersion of nests in relation to the sensor did not explain differences in acoustic activity. The median distance to nests at each site was considered a fixed effect, but its influence was not significant (type III ANOVA of fixed effects, $F[1, 1] = 2.52, p = 0.36$).

Cost Comparison

The estimated cost of traditional nest census monitoring to measure abundance for 7 colonies was $9092 over 3 years; costs were relatively equal across years (year 1: $612/site, year 2: $343/site). An automated acoustic survey was estimated to cost $8,849 over 3 years. The first year cost automated acoustic monitoring was $966/site, due to the costs of sensors and detector development. Subsequent years cost $149/site (Supporting Information).

Discussion

Tern acoustic activity was strongly correlated with nesting abundance among colony sites (71% of variance explained) and with changes in nest abundance between years (88% of variance explained). These results suggest that for Forster's terns, acoustic activity can be a reliable index of relative abundance across colonies or between years and that automated acoustic surveys can be an effective method for measuring changes in relative abundance after management actions. Given that effective conservation actions like the removal of invasive species (Aguirre-Munoz et al. 2008) and social attraction techniques (Parker et al. 2007) can lead to >3-fold increases in seabird abundance over 10 y (Whitworth et al. 2013), this technique seems well suited to document effective conservation actions. We suggest that acoustic activity may be a reliable index of abundance for colonial species and outline some challenges and limitations for further consideration to drive improvements in automated acoustic monitoring so that its advantages may be fully realized.

Potential Challenges

Like other monitoring approaches, acoustic indices are sensitive to changes in signal availability (the rate birds vocalize when present) and signal detection (the rate vocalizations can be detected; Diefenbach et al. 2007). Factors influencing availability for detection in human surveys include life history, phenology, vegetation (Walsh et al. 1995), co-occurring species (Piatt et al. 1990), weather (Bourgeois et al. 2008), bird behavior (Harding et al. 2005), and behavior of human observers (Gutzwiller & Marcum 1997). For colonial seabirds, it is important to consider how demography or behavior influences signal availability when interpreting results from automated surveys. For example, changes in pair bonding or changes in colony attendance patterns (as a consequence of forage conditions) could drive differences in acoustic activity rates independent of differences in abundance. Interspecific interactions such as predation or competition, might lead to changes in calling rates that are unrelated to abundance. All these factors and their impact on acoustic monitoring are likely variable across colonial seabird species. For these reasons it is important to consider these factors in the context of the species and system being monitored. Many of these same challenges exist with traditional monitoring methods; however, they can be harder to confront without a permanent record of the sound environment.

Signal detection is easier to assess because unlike traditional surveys, the raw record of acoustic activity is retained. In our study, detector accuracy was not significantly influenced by site or year, but detector sensitivity differed by colony site (two-way ANOVA for sensitivity across sites, $F[4, 9] = 7.19, p < 0.01$). This suggests that site soundscape characteristics influenced the automated detection process. Vegetation and ambient noise affect human visual and aural detection (Barbraud & Gélinaud 2005; Simons et al. 2007) and likely automated detection. Automated approaches, like human observers are influenced by wind and other noise that reduces the signal:noise ratio. Because detector sensitivity rates varied by site and over time, we suggest developing an improved framework for accounting for this dynamically by understanding the relationship between soundscape properties and detection rates. Doing so could help account for the effect of other noise on signal ability and detection when predicting abundance. Additional sensors such as anemometers or light level meters might also help explain additional variance in acoustic activity.

Analyzing recordings with automated detection introduces new layered complexity. Like traditional surveys, there is the collection of field data (in this case raw acoustic data rather than human interpreted summary data). However, analyzing recordings efficiently requires developing automated detection and classification algorithms to replace decisions made in the field by observers. Automated detectors can be designed to be highly sensitive (e.g., detecting a high percentage of the available tern calls), but high detection rates come at the cost of generating many false positives (i.e., detecting many non-tern sounds with similar features). This can lead to a labor intensive, costly process of auditing automated results in search of relatively rare events. However, when sampling from environments where target calls are common, detectors with average sensitivity can effectively track relative activity levels. For this study, we employed a detector with moderate sensitivity (53% of known calls available were detected), minimizing false positives (23% of detections were incorrectly classified). We did not audit false (i.e., nontern advertisement calls) detections because our accuracy rate was consistent across sites and years.

The sensitivity of our detector was higher than that of similar studies in which researchers used the XBAT cross-correlation detector. The effectiveness of spectrogram cross-correlation varies depending on the target signal complexity and ambient sound environment. Charif and Pitzrick (2008) reported 22% sensitivity for detecting Cerulean Warbler (*Setophaga cerulea*) songs, while Swiston and Mennill (2009) reported 17% and 24% sensitivity for detecting woodpecker "double knocks" and calls, respectively. Alternatively, an approach based on hidden Markov models (SongScope by Wildlife Acoustics) reported detector sensitivity ranging from 56% to 69% for 3 seabird species (Buxton & Jones 2012). This wide range of detector performance emphasizes the importance of evaluating detectors in the context of monitoring goals, whether the goal is to capture rare events or track relative activity levels. If making comparisons of activity across sites, years, or species, it is important to evaluate the sensitivity and accuracy of automated detectors.

Comparison with Traditional Methods

The automated acoustic survey approach reduces a number of biases common in long-term monitoring programs, namely observer and temporal biases. Once deployed, acoustic sensor arrays can sample multiple sites simultaneously over long periods and be analyzed by a single observer with a single automated detector. Unlike human surveys, automated acoustic surveys retain original field recordings, facilitating reinterpretation and analysis of unanticipated questions. Observer differences that influence bird behavior are also reduced. Reducing researcher visits also reduces stress on breeding birds and the likelihood of nest abandonment.

These advantages come at comparable cost to traditional methods in our study, and there may be potentially larger savings in more remote colonies. Our estimated costs for 3-year monitoring programs were similar ($8849 for acoustic monitoring and $9092 for traditional nest censuses); however, our study colonies were easily accessed, whereas many of the world's seabird colonies are not. Acoustic sensors are relatively low-cost (SongMeters are currently, approximately, $700), and autonomous operation reduces the number and duration of expensive visits to remote field sites. The cost of designing and testing automated detectors for a single call type was roughly $350 (for 16 h of work at $22/h). Acoustic detection can be done on relatively low cost computers ($500–$1000), and free alternatives to MATLAB (required for XBAT) exist, such as seewave (Museum national d'Histoire naturelle, Paris, France [Sueur et al. 2008]) and Ishmael (CIMRS Bioacoustics, Corvallis, OR, U.S.A. [Mellinger 2001]). Importantly, the costs of traditional monitoring are unlikely to decrease in the future, whereas the costs of acoustic sensors, electronics, and analytical processing are likely to decrease over time.

Monitoring costs are highly species and colony dependent, influenced by species' remoteness and accessibility of their nesting colonies. We focused on a colonial species nesting atop the ground near a major population center, hardly the case for most threatened seabirds. Many seabird colonies cannot be accessed during the breeding season due to remote locations, disturbance sensitivity, and poor conditions. Thus, while our cost estimates for a 3-year program were similar, this represents a best case scenario for generally more expensive traditional seabird monitoring.

Extending Principles to other Colonial Species and Environments

It remains to be seen how effective this technique will be with species with different behavior from Forster's terns. Over 40% of threatened island-breeding seabirds vocalize primarily at night and nest in cryptic, inaccessible burrows, making them extremely challenging to monitor. Preliminary data from acoustic monitoring studies at Wedge-tailed Shearwater (*Puffinus pacificus*) breeding sites show a significant positive relationship between burrow densities and calling activity (M.M., unpublished data).

Furthermore, Forster's terns nest in predominately single species colonies, but many seabirds nest in mixed colonies with more complex sound environments. Interspecific interactions could mask relationships between a species' abundance and calling activity. Future extension of automated acoustic monitoring for colonial birds should attempt to explore these challenging sound environments and life histories.

Finally, Forster's tern colonies in this study represented relatively small colonies, with dozens of nests rather than hundreds. Larger and denser colonies may be problematic for signal detection (as overlapping calls increase, automated detectors are likely to detect fewer individual calls), and the relationship between vocal activity and abundance could change (the number of calls each bird makes could exhibit density dependence). Two potential solutions are to reduce microphone sensitivity of sensors (and reduce the effective sampling area) or calculate alternative metrics of acoustic activity (such as mean amplitude) that are less sensitive to overlapping calls.

Future Directions

The next step is extending this approach to more challenging soundscapes and colonial species of varied behaviors with multispecies and larger colonies. We have already identified some challenges in signal detection and availability, but we found a correlation between relative abundance and vocal activity. It appears that colonies present varied soundscapes, and explicitly measuring and incorporating these differences may further improve the utility of acoustic monitoring of relative abundance at larger spatial and temporal scales. Measuring the relative abundance of colonial animals with acoustic signals is hardly new, but the marriage of automated acoustic sensors and automated analysis fundamentally changes one's ability to monitor populations at a broader scale.

Our results suggest automated acoustic sensors present a low-cost and scalable tool for monitoring colonial terns. We encourage testing approaches with more challenging colonial species, particularly threatened seabirds, but also other sound producing, aggregated, or colonial species. Patterns in acoustic activity represent important biological information and can provide estimates of population metrics that are less costly and more replicable than from traditional methods. In the future, the cost of automated acoustic sensors and analysis are likely to decrease (Brandes 2008; Porter et al. 2009), leading to acoustic monitoring solutions that managers and conservationists can implement without engineering, programming, or signal detection expertise. Testing and improving the efficacy of these monitoring approaches will allow managers and conservation scientists to measure human impacts on wildlife populations and the outcomes of conservation actions at increasing scales.

Acknowledgments

We thank D. Spatz for her hard work in colonies and G. Baldassarre for his early inspiration. Funding provided by the Packard Foundation Marine Birds Program, Friends of the Long Marine Lab, the Frank M. Chapman Fund of the American Museum of Natural History, Michelle Wainstein, Conservation, Research and Education Opportunities (CREO), and a NSF Graduate Research Fellowship. The use of trade, product, or firm names in the publication is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supporting Information

Colony maps (Appendix S1), call rates and nest abundance (Appendix S2), and cost estimates (Appendix S3) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Acevedo, M. A., and L. J. Villanueva-Rivera. 2006. Using automated digital recording systems as effective tools for the monitoring of birds and amphibians. Wildlife Society Bulletin **34**:211–214.

Ackerman, J. T., J. D. Bluso-Demers, and J. Y. Takekawa. 2009. Post-fledging Forster's tern movements, habitat selection, and colony attendance in San Francisco bay. Condor **111**:100–110.

Aguirre-Munoz, A., et al. 2008. High-impact conservation: invasive mammal eradications from the islands of western Mexico. Ambio **37**:101–107.

Barbraud, C., and G. Gélinaud. 2005. Estimating the sizes of large gull colonies taking into account nest detection probability. Waterbirds **28**:53–60.

Barlow, J., and B. L. Taylor. 2005. Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. Marine Mammal Science **21**:429–445.

Blumstein, D. T., et al. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. Journal of Applied Ecology **48**:758–767.

Bluso-Demers, J. D., J. T. Ackerman, and J. Y. Takekawa. 2010. Colony attendance patterns by mated Forster's Terns *Sterna forsteri* using an automated data-logging receiver system. Ardea **98**:59–65.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J-S.S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution **24**:127–135.

Bourgeois, K., S. Dromzée, E. Vidal, and J. Legrand. 2008. Yelk ouan shearwater *Puffinus yelkouan* presence and behaviour at colonies: not only a moonlight question. Comptes Rendus Biologies **331**:88–97.

Brandes, T. S. 2008. Automated sound recording and analysis techniques for bird surveys and conservation. Bird Conservation International **18**:163–173.

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

Carey, M. 2009. The effects of investigator disturbance on procellariiform seabirds: a review. New Zealand Journal of Zoology **36**:367–377.

Celis-Murillo, A., J. L. Deppe, and M. F. Allen. 2009. Using soundscape recordings to estimate bird species abundance, richness, and composition. Journal of Field Ornithology **80**:64–78.

Charif, R., and M. Pitzrick. 2008. Automated detection of Cerulean Warbler songs using XBAT data template detector software. Preliminary Report. Cornell University Bioacoustics Research Program.

Croxall, J. P., S. H. M. Butchart, B. Lascelles, A. J. Stattersfield, B. Sullivan, A. Symes, and P. Taylor. 2012. Seabird conservation status, threats and priority actions: a global assessment. Bird Conservation International **22**:1–34.

Dawson, D. K., and M. G. Efford. 2009. Bird population density estimated from acoustic signals. Journal of Applied Ecology **46**:1201–1209.

Diefenbach, D. R., M. R. Marshall, J. A. Mattice, and D. W. Brauning. 2007. Incorporating availability for detection in estimates of bird abundance. Auk **124**:96–106.

Efron, B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. Journal of the American Statistical Association **78**:316–331.

Figueroa, H. 2007. XBAT. v5. Cornell University Bioacoustics Research Program.

Fitzpatrick, J. W., et al. 2005. Ivory-billed Woodpecker (*Campephilus principalis*) persists in continental North America. Science **308**:1460–1462.

Goyette, J., R. Howe, and A. Wolf. 2011. Detecting tropical nocturnal birds using automated audio recordings. Journal of Field Ornithology **82**:279–287.

Gutzwiller, K. J., and H. A. Marcum. 1997. Bird reactions to observer clothing color: implications for distance-sampling techniques. The Journal of Wildlife Management **61**:935–947.

Hall, J. A. 1998. Vocal repertoire of Forster's Tern. Colonial Waterbirds **21**:388–405.

Harding, A., J. Piatt, G. Byrd, S. Hatch, N. Konyukhov, E. Golubova, and J. Williams. 2005. Variability in colony attendance of crevice-nesting Horned Puffins: implications for population monitoring. Journal of Wildlife Management **69**:1279–1296.

Karanth, K. U., and J. D. Nichols. 1998. Estimation of tiger densities in India using photographic captures and recaptures. Ecology **79**:2852–2862.

Keith, D. A., T. G. Martin, E. McDonald-Madden, and C. Walters. 2011. Uncertainty and adaptive management for biodiversity conservation. Biological Conservation **144**:1175–1178.

Legg, C. J., and L. Nagy. 2006. Why most conservation monitoring is, but need not be, a waste of time. Journal of Environmental Management **78**:194–199.

McDonald-Madden, E., P. W. J. Baxter, R. A. Fuller, T. G. Martin, E. T. Game, J. Montambault, and H. P. Possingham. 2010. Monitoring does not always count. Trends in Ecology & Evolution **25**:547–550.

Mckown, M. W. 2008. Acoustic communication in colonial seabirds: individual, sexual, and species-specific variation in acoustic signals of *Pterodroma* petrels. University of North Carolina, Chapel Hill.

Mellinger, D. K. 2001. Ishmael 1.0 user's guide. NOAA Technical Memorandum OAR PMEL **120**:96349–98115.

Mellinger, D. K., and C. W. Clark. 2000. Recognizing transient low-frequency whale sounds by spectrogram correlation. The Journal of the Acoustical Society of America **107**:3518–3529.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution **4**:133–142.

Parker, M. W., S. W. Kress, R. T. Golightly, H. R. Carter, E. B. Parsons, S. E. Schubel, J. A. Boyce, G. J. McChesney, and S. M. Wisely. 2007. Assessment of social attraction techniques used to restore a Common Murre colony in central California. Waterbirds **30**:17–28.

Piatt, J., B. Roberts, and S. A. Hatch. 1990. Colony attendance and population monitoring of least and crested Auklets on St. Lawrence Island, Alaska. Condor **92**:97–106.

Porter, J. H., E. Nagy, T. K. Kratz, P. Hanson, S. L. Collins, and P. Arzberger. 2009. New eyes on the world: advanced sensors for ecology. BioScience **59**:385–397.

Sherley, R. B., T. Burghardt, P. J. Barham, N. Campbell, and I. C. Cuthill. 2010. Spotting the difference: towards fully-automated population monitoring of African penguins *Spheniscus demersus*. Endangered Species Research **11**:101–111.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth 2007. Experimental analysis of the auditory detection process on avian point counts. The Auk **124**:986–999.

Strong, C. M., L. B. Spear, T. P. Ryan, and R. E. Dakin. 2004. Forster's tern, caspian tern, and california gull colonies in san francisco bay: habitat use, numbers and trends, 1982–2003. Waterbirds **27**:411–423.

Sueur, J., T. Aubin, and C. Simonis. 2008. Equipment review: seewave, a free modular tool for sound analysis and synthesis. Bioacoustics **18**:213–226.

Swiston, K. A., and D. J. Mennill. 2009. Comparison of manual and automated methods for identifying target sounds in audio recordings of pileated, pale-billed, and putative Ivory-billed woodpeckers. Journal of Field Ornithology **80**:42–50.

Thompson, M. E., S. J. Schwager, K. B. Payne, and A. K. Turkalo. 2010. Acoustic estimation of wildlife abundance: methodology for vocal mammals in forested habitats. African Journal of Ecology **48**:654–661.

Wade, P. et al. 2006. Acoustic detection and satellite-tracking leads to discovery of rare concentration of endangered North Pacific right whales. Biology Letters **2**:417–419.

Walsh, P. M., D. J. Halley, M. P. Harris, A. Del Nevo, I. M. W. Sim, and M. L. Tasker. 1995. Seabird monitoring handbook for Britain and Ireland: a compilation of methods for survey and monitoring of breeding seabirds. Page 168. Joint Nature Conservation Committee, the Royal Society for the Protection of Birds, the Institute of Terrestrial Ecology, and the Seabird Group 1995.

Whitworth, D., H. Carter, and F. Gress. 2013. Recovery of a threatened seabird after eradication of an introduced predator: eight years of progress for Scripps's murrelet at Anacapa Island, California. Biological Conservation **162**:52–59.

Wood, K., D. Bynton, E. VanderWerf, L. Arnold, M. LeGrande, and D. Kuhn. 2002. The distribution and abundance of the band-rumped Storm-Petrel (*Oceanodroma castro*): a preliminary survey on Kaua'i, Hawai'i 2002. Pages 1–21.